Thèse   2015   **Open Access**

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

- -

Temporal reference: empirical and theoretical perspectives: converging evidence from English and Romance

- -

Grisot, Cristina

How to cite

GRISOT, Cristina. Temporal reference: empirical and theoretical perspectives: converging evidence from English and Romance. Doctoral Thesis, 2015. doi: 10.13097/archive-ouverte/unige:82583

This publication URL: https://archive-ouverte.unige.ch/unige:82583
Publication DOI: 10.13097/archive-ouverte/unige:82583

Temporal reference: empirical and theoretical perspectives

Converging evidence from English and Romance

Thèse de doctorat

Présentée à la Faculté des Lettres de l'Université de Genève pour l'obtention du titre de Docteur ès Lettres, mention Linguistique

le 2 Novembre 2015

par

Cristina Grisot

Membres du jury de thèse:

Président	Professeur Eric WEHRLI, Université de Genève
Directeur	Professeur Jacques MOESCHLER, Université de Genève
Membres	Professeur Jesús ROMERO-TRILLO, Université autonome de Madrid
	Professeur Louis DE SAUSSURE, Université de Neuchâtel
	Docteur Vladimir ŽEGARAC, Université de Bedfordshire
	Docteur Andrei POPOESCU-BELIȘ, Institut IDIAP, EPFL

Acknowledgments

Table of Contents

List of Tables

List of Figures

List of most frequent abbreviations

AOR: Aoristic
CA: Contrastive Analysis
E: Event moment
EN: English
FR: French
FUT: Future
IMP: Imparfait
IMPERF: Imperfective
IND: Indicative
L1: First/source language
L2: Second/target language
MT: Machine Translation
NLP: Natural Language Processing
PastCont: Past Continuous
PC: Passé Composé
PH: Historical Present
PL: Plural
PERF: Perfective
PRES: Present
PresCont: Present Continuous
PP: Past Perfect
PresPerf: Present Perfect

PRET: Preterit
PS: Passé Simple
PQP: Plus-que-parfait
R: Reference point
RO: Romanian
RT: Relevance Theory
S: Speech moment
SG: Singular
SL: Source Language
SMT: Statistical Machine Translation
SP: Simple Past
SPres: Simple Present
TL: Target Language
$X<Y$: X preceeds Y
$X>Y$: Y is posterior to X
$X \geq Y$: Y is posterior or simultaneous with X
$X=Y$: X and Y are simultaneous
1st (following a verb): 1st person
2nd (following a verb): 2nd person

Introduction

> It is important to keep the two concepts of time and tense strictly apart. The former is common to all mankind and is independent of language; the latter varies from language to language and is the linguistic expression of time-relations, so far as these are indicated in verb forms. (Jespersen 1961, 1)

This thesis deals with temporal reference – the location of eventualities (states and events) in time – in natural language. There are numerous ways in which temporal reference may be expressed, such as the grammatical categories of Tense and Aspect (generally referred to through the generic notion *verbal tense*)[1], inherent temporal features of the verb phrase known as lexical/situation aspect or Aktionsart, temporal adverbials and connectives such as *yesterday, before,* special particles such as the Mandarin Chinese aspectual particles *-le* and *-guo* and principles of human communication, among others. In tensed languages, research on temporal reference has primarily focused on Tense and, secondarily, on Aspect and Aktionsart. Cross-linguistic research from formal semantics in the past forty years pointed out that there are languages without the grammatical category of Tense (the so-called *tenseless* languages, such as the Mandarin Chinese and Yucatec – Mayan, Mexico) and *mixed tense* languages (with optional tense marking as well as untensed clauses such as Navajo, Japanese and Korean) as noted by Tonhauser (2015). These studies argued that Tense does not *fully determine* the temporal reference of a sentence but only locates eventualities relationally (Smith, 2008, 232).

Despite a long research tradition on verbal tenses in tensed languages, there is no general consensus among current theories, except on a certain number of basic points, such as the use of *temporal coordinates* (for example, Reichenbach 1947; Klein 1994) to describe verbal tenses and the notion of *temporal anaphor* (such as Partee 1973, 1984; Webber 1988; Kamp and Reyle 1993). The challenge in current researches on temporal reference is to propose a cross-linguistically valid model that holds for both tensed and tenseless languages mainly through decentralizing the role played by Tense and through increasing the attention given to the other 'ingredients' of temporal reference. Up to now, numerous semantic and pragmatic studies have been dedicated to individual verbal tenses in languages such as English, French and Spanish (see for example Moeschler et al. 1998; Saussure 2003 for French verbal tenses) without completely discriminating among types of temporal information coming from the categories of Tense, Aspect and Aktionsart. The direct consequence of approaches in which the three categories are not completely distinguished is the lack of a cross-linguistically valid model of temporal reference, which goes beyond language specificities and which permits consistent contrastive analyses. Likewise, keeping

---

[1] As a convention, in this thesis I use capitalised Tense and Aspect for grammatical categories and lower case letter to refer to individual verbal tenses in particular languages (e.g. the English Simple Past or Present Perfect, the French Passé Composé), where often the same verb stem expresses both Tense and Aspect. I refer to lexical aspect as Aktionsart. The words *tense* and *verbal tense* are used interchangeably and refer to individual verbal tenses in particular languages.

the centrality of Tense for tensed-prominent languages is without any doubt accurate but it provides a too narrow image of temporal reference in natural language. Taking into consideration aspect-prominent, tenseless and mixed-temporal languages is crucial for having an all-encompassing comprehension of temporal reference. This thesis makes a first step in this very challenging goal by investigating the categories of Tense, Aspect and Aktionsart and their role for expressing temporal reference in four tensed languages (English EN, French FR, Italian IT and Romanian RO) and by developping a theoretical model that deemphasizes the previously assumed centrality of Tense. Reference to tenseless languages and aspect-prominent languages will be made from a theoretical perspective and based on previous researches carried out by Comrie (1976, 1985), Smith (2003, 2007, 2008), Žegarac (1991), Bohnemeyer (2009), Bohnemeyer & Swift (2004), Lin (2003, 2006, 2012), among others. Reflections as theirs are necessary for developing a comprehensive model of temporal reference. Future work will be dedicated to develop the present model for aspect-prominent and tenseless languages through a rich and consistent empirical and experimental work.

The interest in temporal reference originates in an apparently simple empirical question: how can one improve the translation of verbal tenses by automatic translation systems? This question was asked in the framework of two inter-disciplinary research projects COMTIS and MODERN[2], which aimed at improving the quality of machine-translated texts in terms of their overall coherence. The coherence of a text depends on several cohesive ties, such as pronouns, discourse connectives and verbal tenses. The fact that verbal tenses are problematic for machine translation systems is illustrated in Table 0-1. The second column of the table provide examples of several types of errors made by this typical automatic system makes when translating verbal tenses from EN into FR (signalled by the * symbol). The third column provides examples of correct translations, in which the EN Simple Past (SP) is translated through a Passé Composé (PC), a Passé Simple (PS) or an IMP (Imparfait).

Regarding the type of errors, in general it is a question of wrong verbal tense choice combined or not with a wrong lexical choice, as well as difficulties with pronominal reference. In the Table 0-1, examples 1 and 2 illustrate the case of correct lexical choices but incorrect verbal tenses. Examples 3, 4 and 5 illustrate the case of correct verbal tenses but it is accompagned by an erroneous lexical choice. The example 6 illustrates the case of incorrect lexical choice and erroneous verbal tense choice. Moreover, the example in 4 is accompagnied by a problem of pronominal reference. There are also cases in which the automatic system provides a correct translation such as in 1 (the first clause), in 2 (the first clause) and in 5 (the second clause).

Table 0-1Examples of human and automatically translated sentences (EN-FR)

	EN Sentence: SP	FR Google Translate[3]	FR Human translation
1	John *was* in love. He *got married*.	John *était amoureux* (IMP). Il **se marie*. (PRES)	Jean *était amoureux*. Il *se maria/s'est marié* (PS/PC).
2	The train *left* London. One hour later, it already *entered* the Birmingham station.	Le train *a quitté* Londres (PC). Une heure plus tard, il **est* déjà *entré* dans la gare de Birmingham (PC).	Le train *quitta* Londres (PS). Une heure plus tard, il *entrait* déjà en gare de Birmingham (IMP).
3	John *ordered* coffee. It *was* too hot.	John **a ordonné* café (PC). Il **faisait* trop chaud. (IMP)	Jean *commanda/a commandé* du café (PS/PC). Il était trop chaud (IMP).
4	John *ordered* coffee. It *was* delicious.	John **a ordonné* café (PC). **Ce était délicieux*. (IMP)	Jean *commanda/a commandé* du café (PS/PC). Il *était délicieux* (IMP).
5	John *ordered* coffee. The waiter *refused*.	John **a ordonné* café (PC). Le serveur *a refusé* (PC).	John *commanda/a commandé* du café (PS/PC). Le serveur *a refusé.* (PC)
6	Last month France *dispatched* a special investigation team.	Le mois dernier, la France **dépêche* une équipe d'enquête spéciale (PRES).	Le mois dernier, la France *envoya/a envoyé* une équipe d'enquête spéciale (PS/PC).

In order to answer this empirical question, other related issues had to be addressed, such as which verbal tenses are problematic, what is the meaning of a verbal tense, how do humans process this temporal information in order to understand and translate it and does the translation of a tense depend on other temporal devices in the discourse. This thesis aims at answering these questions, on the one hand through revisiting the most relevant existent syntactic, semantic and pragmatic theories, neurolinguistic and computational approaches to Tense, Aspect and Aktionsart in discourse, and on the other hand, through my own empirical research on multilingual corpora and experimental data. I propose a theoretical reanalysis of these components of temporal reference and a predictive model applied to natural language processing and machine translation. The model defended in this thesis was developed based on empirical work carried out within both *data-based* (i.e. testing of theoretical predictions built based on the literature and *data-driven* (i.e. novel theoretical insights emerged from the analysis of the data) approaches.

The main proposal defended in this thesis can be summarized as follows:

Temporal reference should be defined *narrowly* and *broadly*. Narrow temporal reference represents the localization of an eventuality with respect to the moment of speech (before, at the moment of speech and after the moment of speech). Broad temporal reference represents the localization of an eventuality with respect to another eventuality (phenomenon classically treated at *temporal sequencing*). In order to have an accurate view of temporal reference in natural languages, the currently generic term *verbal tense* was decomposed into its constituent categories (Tense and Aspect applied to Aktionsart) and the role played by the category of Tense is deemphasized. Hence, narrow and broad temporal reference is linguistically expressed in tensed languages (such as the languages examined in this research) through Tense, Aspect, Aktionsart, as well as temporal connectives and temporal adverbials, among others.

Examples (1) and (2) illustrate the role played by the category of Tense for establishing

---

[3] Translation with the Google Translate system available at https://translate.google.com/ on 1.11.2014.

narrow temporal reference. In these two examples, the FR PC can express reference to past time and to present time, and this information is determined in the linguistic context, more precisely, due to the temporal adverbial *hier* 'yesterday', and respectively, *demain* 'tomorrow'. Examples (3) and (4) illustrate broad temporal reference, where eventualities can be temporally ordered as in former or can take place simultaneously as in the latter. These two interpretations can be rendered explicit by the temporal connectives *puis* 'then', and respectively *lorsque* 'while, when'.

(1) Hier, j'*ai fini* mon article.
     Yesterday, I finish.PC my article
     'Yesterday, I finished my article.'

(2) Demain, j'*ai fini* mon article.
     Tomorrow, I finish.PC my article
     'Tomorrow, I will have finished my article.'

(3) Jean *courut* à travers le parc. [*puis*] Il *arriva* à la grande statue.
     John run.PS through the parc. [then] He arrive.PS at the big statue.
     'John run through the parc. [then] He arrived at the big statue.

(4) Ce matin là, Jean *courait* à travers le parc. [*lorsque*] Deux chiens le *suivaient* en aboyant fort.
     That morning, John run.IMP through the parc. [while] Two dogs follow.IMP barking loudly.
     'That morning, [while] John was runing through the parc. Two dogs followed him loudly barking.'

Examples (5) and (6) illustrate the role played by Aspect for establishing broad temporal reference. In (5), the two situations referred to through the PS, classically described as expressing the perfective aspect, are temporally ordered whereas in (6), the two two situations referred to through the IMP, classically described as expressing the imperfective aspect, occur simultaneously. As above, these two interpretations can be rendered explicit by the temporal connectives *puis* 'then', and respectively *lorsque* 'while, when'.

(5) Il *descendit* les escaliers. [*puis*] Il vit un chat blessé.
     He go down.PS the stairs. [then] He see.PS a wounded cat.
     'He went down the stairs. [then] He saw a wounded cat.'

(6) Il *descendait* les escaliers. [*lorsque*] Il vit un chat blessé.
     He go down.IMP the stairs. [when] He see.PS a wounded cat.
     'He was going down the stair when he saw a wounded cat.'

Finally, examples (7) and (8) illustrate the role played by lexical aspect − Aktionsart − for establishing broad temporal reference. In the literature, it was argued that the type of situation referred to determines the temporal interpretation of the discourse (Dowty 1986). Accordingly, the interpretation of (7) which contains an event in the second clause is that of temporal ordering (rendered explicit through the connective *puis*) whereas the interpretation of (8) which contains a state in the second clase is that of temporal simultaneity (rendered explicit through the connective *lorsque*).

(7) Jean entra dans le bureau du président. [*puis*] Le président *se leva*.

'John entered the president's office. [then] The president got on his feet.'
(8)   Jean entra dans le bureau du président. [*lorsque*] Le président *était heureux.*
       'John entered the president's office. [while]The president was happy.'

These examples also show that categories of Tense, Aspect and Aktionsart underdetermine the speaker's intended temporal meaning of her utterance. This is illustrated from a cross-linguistic perspective in Table 0-2, which contais original examples from translation corpora. They exemplify that the meaning of the SP is not specific enough in order to have a one-to-one corresponce for each line of the table between the SP and the verbal tense used in FR for its translation.

Table 0-2 The translation of the English SP into FR

	EN Sentence: SP	FR Human translations: PC, PS, IMP, PRES
1	General Musharraf *appeared* on the national scene on October 12, 1999.	Le Général Moucharraf *est apparu* sur la scène nationale le 12 octobre 1999.
2	With significant assistance from the US, disarmament was orderly, open and fast. Nuclear warheads *were returned* to Russia.	Avec l'assistance non-négligeable des Etats-Unis, le désarmement a été méthodique, ouvert et rapide. Les ogives nucléaires *furent renvoyées* en Russie.
3	He *seemed* about seventeen years of age, and *was* of quite extraordinary personal beauty.	Il *paraissait* avoir seize ans, et il *était* d'une beauté absolument extraordinaire.
4	The fine little fellow, who *seemed* to have never known the meaning of fear, early revealed a keen and active mind.	Ce digne enfant, qui *paraît* n'avoir jamais connu la crainte, annonça promptement un esprit très vif.

Hence, the categories of Tense, Aspect and Aktionsart must be contextually worked out by making use of their *conceptual* and *procedural* types of encoded information, as it was suggested in the relevance theoretic framework. Guided by the need to have cognitive temporal coherence, the hearer treats information coming from several sources and their interrelations in a coherent manner.

The model is developed based and derived from empirical (translation corpora) and experimental work (offline experiments with linguistic judgement task). Translation corpora were used in order to investigate the translation of verbal tenses from English into French and from French into English. Furthermore, a multilingual translation corpus was built in order to the translation of the English SP into French, Italian and Romanian. In the exeperimental work carried out, hypotheses formulated based on the current literature were formulated and tested for the English SP, the French PC, PS, IMP and PRES, the Italian PC, PS, IMP and the Romanian PC, PS, IMP. The results of the experiments indicated that the category of Tense encodes both conceptual and procedural types of information whereas aspectual information is encoded at the conceptual level through Aktionsart and at the procedural level through Aspect. Approaching the ingredients of broad and narrow temporal reference in these terms based on experimental work is original and represents a contribution to the general state of the art.

Moreover, it is suggested that establishing temporal reference is a cognitive principle and that all humans apply it regardless of their language. Consequently, languages employ various means to convey it and they can be typologically classified according to what are

predominant linguistic and non-linguistic means at their disposal: tensed-prominent, aspect-prominent and tenseless languages in which aspectual information is crucial. The model developed in this research permits its developpement in further work for typologically different languages, and therefore, for natural language as a cognitive faculty specific to the human species.

Finally, describing temporal reference in terms of its ingredients and operationalizing their encoded conceptual and procedural types of information as medium-coarse grained features proved to be the appropriate approach in order to answer the research question regarding the automatic translation of verbal tenses. More precisely, the features proposed were not only successfully implemented for automatic procesessing but also their implementation in Natural Langue Processing and their application to Statistical Machine Translation produced significant improvements of the results of these systems. These ameliorations represent in the same time an empirical indirect but solid validation of the theoretical model defended in this thesis.

1 Theoretical and methodological preliminaries

1.1 At crossroads

This thesis aims at being a bridge between theoretical linguistics and computational linguistics, corpus linguistics, contrastive analysis and empirical linguistics. This research is innovative in that it proposes an interdisciplinary approach of *temporal reference* and its processing in discourse. On the one hand, temporal reference is approached from a theoretical point of view through a selective and analytical presentation of the very rich literature both in Romance and in English linguistic research, through anchoring this investigation in the relevance theoretic cognitive framework of language comprehension, and through suggesting an empirically valid theoretical model of temporal discourse coherence. On the other hand, temporal reference is investigated from an empirical point of view and using several types of methods and methodologies: translation corpora, contrastive analysis of EN, FR, IT and RO verbal systems, offline experiments with linguistic judgment task, application to natural language processing (NLP) and machine translation (MT).

The present study is framed by two Swiss SNF Sinergia projects on MT, COMTIS and MODERN. The COMTIS project aimed at improving the quality of translated texts in terms of their overall coherence. Research in discourse analysis (Halliday and Hasan 1976; Hobbs 1979; Mann and Thompson 1986; Sanders et al. 1992, 1993; Sanders 2005) showed that the coherence of a text depends on cohesive ties and coherence relations established intersententially. In Sanders' cognitive framework, it is argued that coherence relations identified in discourses 'are ultimately cognitive relations' (Sanders et al. 1992, 1). The proponents of the COMTIS project analysed the discourse-level phenomena that most influence the perceived coherence of a text, and used surface cues to label them automatically with intersentential dependency labels (ISD). The labelled source was handed over to a statistical machine translation (SMT) system equipped with translation models that are capable to process the ISD labels (COMTIS project proposal, March 2010). This ambitious task involved the integration of linguistic (semantic and pragmatic information), empirical (corpus work and experimental work) and automatic analysis and treatment of language. The MODERN project continues the research initiated in the COMTIS project and aims at detecting and modelling automatically cases where the translation of words and phrases depends on the translation of previous ones. The MODERN project uses jointly computational and experimental techniques and methodologies for investigating discourse relations.

Temporal reference and its processing in a discourse represents a problematic matter for MT systems. This thesis describes the conjoint research that was conducted in the COMTIS and MODERN projects for improving the translation of verbal tenses by MT systems. This research consists of a large linguistic background, corpus-based and experimental investigations for several languages, as well as computational models of temporal reference at the discursive level.

In this preliminary chapter, I wish to introduce theoretical and methodological issues

developed later in the following chapters. Since temporal reference in discourse expressed through verbal tenses and aspectual markers has been the subject of a very abundant literature, it is not the purpose of this thesis to be exhaustive, nor to speak about all linguistic and non-linguistic sources of temporal information in discourse. I will focus on Tense, Aspect and Aktionsart and I will confront my point of view with several previous linguistic theories. As for the existing literature on EN, FR, IT or RO individual verbal tenses expressing past time, I do not aim at providing fine-grained accounts but I will introduce them briefly and point to their most relevant usages with respect to the model developed in this thesis. I aim at investigating verbal tenses with quantitative analyses and retain in the final model only statistically significant results.

1.2   A cognitive pragmatic framework for language comprehension

Since Grice (1989) it is generally accepted that human communication is an inferential process driven by the desire to express and recognize intentions. The founders of Relevance Theory (RT) (Sperber and Wilson 1986/1995; Wilson and Sperber 1993, 2002, 2004) both refined and challenged Grice's ideas. RT proposes a model for human communication based on the notion of relevance. They claim that the expectations of relevance raised by an utterance are precise and predictable enough to guide the hearer towards the speaker's intended meaning (Wilson and Sperber 2004, 607). The speaker's intended meaning is inferred on the basis of the evidence provided. The hearer builds appropriate hypotheses about the explicit content via decoding, disambiguation, reference resolution, narrowing, loosening, saturation, ad hoc concept construction and free enrichment (Carston 2004).

As assumed in RT, the linguistic expressions that a speaker utters underdetermine the content that she communicates. This takes place not only at the level of implicatures but also the propositional contents she communicates explicitly (i.e. the explicature of the utterance). The hearer must therefore recover inferentially the speaker's intended meaning, and this takes place at the level of explicatures and implicatures. This interpretative process is guided by the expectation of relevance and the quest for cognitive effects. Linguistic expressions encode conceptual and procedural information (i.e. instructions for manipulating conceptual representations) that contribute and, respectively, constrain, the interpretative process (a more detailed discussion is provided in section 3.1.3).

As far as temporal reference is concerned, Tense, Aspect and Aktionsart underdetermine the content expressed by the speaker and their meaning must be contextually worked out. Currently, in the literature, it is assumed that verbal tenses encode procedural information that guides the hearer in the interpretation process (as discussed in section 3.1.3.5). In this research it is argued that Tense encodes both conceptual and procedural information, Aspect encodes procedural information and Aktionsart encodes conceptual information. Conceptual information represents most often a semantically incomplete *pro-concept*, which must be adjusted contextually in the form of an ad hoc concept. Speakers have easy conscious access to conceptual representations whereas procedural information is inaccessible through conscious thinking. In this thesis, I will provide empirical evidence that for some aspects validate and for other aspects challenge current theoreticl assumptions. Additionally, two quantitative measures are proposed for evaluating the type of meaning

encoded by linguistic expressions: *inter-judge agreement rate* measured in this research with the Kappa coefficient (as described in 4.3.2 from a methodological point of view, Chapter 6 regarding its use for evaluation of experimental results and Chapter 7 for its integration in a theoretical account) and the quantity of cross-linguistic variability in translation corpora (as described in section 5.3 regarding the analysis of parallel translations corpora and Chapter 7 for its integration in a theoretical account).

Establishing temporal reference at the discursive level involves determing relations among eventualities contained in the discourse segments. My suggestion is that temporal relations among eventualities fall under the competency of *discourse relations.* Sanders et al. (1992) point out that 'a coherence relationis an aspect of meaning of two or more discourse segments that cannot be described in terms of the meaning of the segments in isolation' (p. 2). Moreover, for them coherence relations are cognitive entities in that they play a central role for the construction of cognitive representations. Discourse relations can be expressed explicitly throught linguistic expressions encoding procedural information (i.e. instructions regarding the manipulaton of conceptual representations) or determined inferentially. In this thesis, I will be arguing that temporal relations existent among eventualities, as they are expressed by the procedural information encoded by Tense and Aspect, should be interpreted as pointing to a temporal *cognitive coherence relation.* This cognitive temporal relation is language independent and presents linguistic cues different from one type of language to another. Explicitly, it is triggered mainly by Tense, Aspect and Aktionsart in tensed languages and principally by Aspect and Aktionsart in tenseless languages.

1.3   Temporal reference and its ingredients

Formal models of natural language meaning takes as a basic working hypothesis the *principle of compositionality* (Katz and Fodor 1963; Partee, ter Meulen and Wall 1990) according to which the meaning of a complex linguistic expression is derived from the meanings of the morphemes (semantics), the way the morphemes combine (syntax) and the context in which the expression is uttered (pragmatics). Formal semantics (based on Montague's grammar) focused on computing the semantic value of sentences from the combination of its semantic constituents (propositions, predicates, arguments and quantificators) whereas pragmatic theories (such as RT, among others) targeted the computation of the meaning of an utterance in its cotext and context.

My proposal, developed in the RT framework, is that the speaker's intended temporal meaning of a sentence is computed contextually on the basis of the conceptual and procedural information provided by linguistic expressions and on the basis of contextual assumptions. Different sources of temporal information are treated in a coherent manner. Specifically, human mind tends to search for temporal coherence both at the cognitive and the discursive level. When information coming from different sources seems incompatible, it is adapted for meeting the cognitive needs of coherence. For example, the example in (9) could be treated as incoherent given that it contains a future time adverbial *tomorrow* and a past time verbal tense, the PC. However, this is not the case. The hearer interprets the PC based on the cotextual information provided by the temporal adverbial and establishes a location in the future (E>S) of the eventuality *finir* 'finish'. This is carried out by building the

ad hoc concept non-pastness, as it will be argued in Chapter 7.

(9) Demain, j'*ai fini* mon article.
     Tomorrow, I finish.PC my article
     'Tomorrow, I will have finished my article.'

However, in language not all elements are necessary for expressing temporal reference, which was primarily linked to Tense until research on a wide range of typologically different languages started 40 years ago. Languages are classified as *tensed* and *tenseless* with respect to their use or lack of use of the grammatical category of Tense (as discussed in section 2.1). Among tensed languages, a further distinction can be made between *tense-prominent* and *aspect-prominent* languages. Romance and Germanic languages are tense-prominent whereas Slavic languages are aspect-prominent. Prominence involves that both Tense and Aspect are used but one of the two is less developed than the other one (Bhat 1999). The category of aspect is usually used ambiguously to refer to both perfective/imperfective aspect, the so-called (*grammatical*) *Aspect* and to inherent features of situations[4], known as *Aktionsart* or *lexical aspect*.

Tense is defined as the 'grammaticalized marking of location in time' (Comrie 1985, 9) and it is expressed through morphemes realised on the verb. In generative grammar, Tense is part of the Inflection node Infl, where Infl is the functional head of the clause (Chomsky 1957, 1965, 1970, 1981, 1995), Stowell (1981, 2007), Pollock (1989), Belletti (1990), Zagona (1990, 1995), Guéron (1993, 2007, 2008), Giorgi and Pianesi (1997), among many others. In formal semantics, Tense was described as a temporal operator (Prior 1967, 1968), temporal anaphor (Partee 1973, 1984, 1989; Webber 1988; Kamp and Reyle, 1993; Lascarides and Asher 1993, to name but a few). Finally, in RT, Tense is a procedural marker constraining the interpretation process (Nicolle 1997, 1998; Moeschler et al. 1998; Saussure 2003, 2011). In this research, I suggest a *conceptualist* view of Tense. To be more precise, it encodes conceptual and procedural information, which contributes and respectively, constraints the interpretation process.

If Tense is defined as the grammatical category that relates the time of the situation described to some other time (the moment of speaking or a reference moment) (Reichenbach 1947), Aspect is described as "different ways of viewing the internal consistency of a situation" (Comrie 1976). In examples (10) - (14), the second verb presents the totality of the situation referred to (the entirety) without reference to its internal temporal consistency: a single unanalysable and indivisible whole. Verbal forms with this meaning have a perfective meaning and the grammatical verbal forms that express it are called *perfective* aspect. The forms referring to John's reading do not present the situation in the same way but rather there is explicit reference to its internal constituency. In this case, reference is made to an internal phase of John's reading, without giving explicit information about the beginning or the end of the situation. Verbal forms with this meaning have an imperfective meaning and the grammatical verbal forms that express it are called *imperfective aspect*.

---

[4] The terms *situation* and *eventuality* are generic terms used interchangeably to refer to all various possible aspectual classes and their names (state, action, event, process, activity, accomplishment, achievement). Unless it is necessary to distinguish among them, these two terms will be used in order to express *the situation/eventuality referred to* in the clause or in the sentence.

(10) John was reading when I entered.
(11) Ivan čital   kogda ja vošel.
    John read.IMPERF when I enter.PERF.
(12) Jean lisait   quand j'entrai.
    John read.IMP when   enter.PS.1stSG
(13) Gianni leggeva quando entrai.
    John read.IMP when   enter.PS.1stSG
(14) Ion citea   când   am intrat.
    John read.IMP when   enter.PC.1stSG

In example (15) in Mandarin Chinese, which is a tenseless language, temporal reference is expressed through a combination of Aktionsart and Aspect. Specifically, the punctual and telic situation *enter* indicates reference to past time whereas the imperfective *zheng zai* 'be+ing' is temporally anchored on the first clause.

(15) Dang wo jin lai de shi hou, Jean zheng zai du shu.
    when I enter   John *be+ing*   read
    'When I entered, John was reading'.

Hence, Aspect refers to the morphosyntactic features of verbs that present differently the temporal flow of situations, as being completed or ongoing. The difference between perfectivity and imperfectivity is not necessarily an objective difference between situations neither the speaker's objective perspective on the situation. It is possible for the same speaker to refer to the same situation, once with the perfective aspect and once with the imperfective. Her choice depends on her intention that is to present the situation as a whole and completed or to focus on an internal phase of an ongoing situation. The role played by Aspect for expressing temporal reference is discussed in section 2.3).

Aktionsart refers to inherent temporal properties of a situation. Based on these properties, situations can be temporally classified into several categories, more precisely *states*, *achievements*, *activities* and *accomplishments* (Vendler 1957,1967; Garey 1957; Verkuyl 1972, 1996; Dowty 1979, 1986; Mourelatos 1978, 1981; Parsons 1990; Smith 1991, 1997; Rothstein, 2004, among many others). These four aspectual classes can be described trough several distinctive features such as *homogeneity*, *dynamicity*, *telicity* and *boundedness*[5] (as discussed in section 2.4). Researches on tenseless languages and mixed-temporal languages (such as Smith 2005, 2006) indicated that boundedness and dynamicity play a fundamental role for determining temporal reference. My suggestion is that boundedness should be taken into account when investigating tense-prominent and aspect-prominent languages in order to guarantee an accurate model for temporal reference in natural language.

As for other sources for temporal information, there are linguistic sources, such as temporal adverbials[6], temporal connectives and pragmatic sources, such as world and

---

[5] Boundedness is inferred in most of the languages but it can also be expressed grammatically, as in the tenseless language Hausa (Chadic, West Africa) (Tonhauser, 2015 citing Musha 2013).

[6] Tonhauser (2015, 138) points out that not all languages have Tense but temporal adverbials play an important role for expressing temporal reference across all languages investigated in the literature. They are nevertheless optional in all languages. However, some languages lack particular kinds of adverbial

co(n)textual knowledge. As far as this thesis is concerned, I will make reference to these elements with respect to their interactions with Tense, Aspect and Aktionsart.

1.4   Data in linguistic research: corpus and experimental

Nowadays, one can observe the increasing aspirations of linguists to use robust and objective findings in addition to intuitive and subjective acceptability judgments or built examples. This is maybe due to a perceived dissatisfaction with methods from the so-called "armchair linguistics" (in the sense of Fillmore 1992). Empirical linguistics aims at having consistent data for supporting or challenging current theories, as well as proposing new models for the interpretation of linguistic phenomena.

McEnery and Wilson (1996) highlight that, broadly speaking, linguists have tended to favour the use of either introspective data (that is, language data constructed by linguists) or naturally occurring data (that is, examples of actual language usage). Nowadays, most linguists see these two types of data as complementary approaches, and not exclusive ones. Gibbs and Matlock (1999) and Gries (2002) argue that, although intuition may be poor as a methodology for investigating mental representations, linguists' intuitions are useful in the formulation of testable hypotheses about linguistic structure and behaviour. Introspective and corpus data were the two main sources of data for theoretical linguistics until the mid-1990s. After that time other sources have been considered, such as experimentation, language acquisition, language pathologies, neurolinguistics, etc. Kepser and Reis (2005) argue that linguistic evidence coming from different domains of data sheds more light on the issues investigated than data from a unique source. Multi-source evidence can either validate the theory or bring contradictory results, therefore opening new perspectives.

In this thesis, I use corpus data and data coming from offline experimentation where participants are required to do something with language they do not usually do (using units they usually interact with involving typical linguistic output) (Gilquin and Gries 2009, see section 4.3). Corpus Linguistics has flourished in the last fifty years mainly due the numerous advantages of corpus data (discussed more detailed in section 4.2.1.), of the availability of electronic corpora and of numerous tools to investigate them. Corpora have numerous advantages but also some limitations linked to their inherent features called *translationese* (Gellerstam 1996) and other limitations linked to possibilities of investigating speakers' processing of language. This aspect must be explored through experimentation. Experimentation is a rapidly growing source of data in contemporary research in linguistics. There are two types of experimental procedures: *online* and *offline* investigation of how the brain processes language. Online research can be carried out using methods such as electroencephalogram (EEG), eye tracking and neuroimaging among others (see section 3.3.1 for a description of recent findings from neurolinguistics with respect to Tense, Aspect and Aktionsart in several languages). Experimental design includes tasks such as lexical decision, picture naming, reading tasks, production tasks, acceptability, judgments and linguistic judgments among others.

Experimental methodology used in this thesis, that is offline processing based on linguistic

---

constructions and tenseless languages do not compensate the lack of Tense by using temporal adverbials to a higher degree than tensed languages (Bohnemeyer 2009).

judgement task, will be discussed in section 4.3.1. My suggestion is that corpus data and experimental data are complementary and necessary in pragmatic research, and may be used within various frameworks of linguistic description and analysis. This research is a plea for the necessity to develop linguistic models that are empirically covered. The data used in this thesis has two main sources: translation corpora and offline experiments with linguistic judgment task. Corpus data was randomly selected from four stylistic registers, in order to be able to generalize the conclusions. Experimental data consists of natural data (i.e. corpus data) and controlled experimental items built for the specific purposes of the experiments. Corpus data comes from four languages (English and three Romance languages) and it is analysed contrastively with the aim of proposing a cross-linguistic valid tertium-comparationis. Qualitative and quantitative analyses are carried out both for corpus data and for experimental data. A general mixed model is built to account for variation in the data and to predict the verbal tense used in target language. Both corpus and experimental work is replicable.

The research described in this thesis proposes a medium coarse-grained cross-linguistically valid model that is successfully applied to NLP and it improves the results of SMT systems in terms of their choices of verbal tenses and of lexical choices in a target language.

1.5   Human and automatic processing of temporal information

Human processing of temporal information both for comprehension and production is a topic that interests more and more psycholinguistics and neurolinguistics researchers. More recently, linguists also became interested in investigating the link between language as an object of study and the way in which the human brain processes language. Nowadays, linguists use experimental methodology targeting offline and online investigations of cognitive process in order to test and to improve existing linguistic theories, as well as, to develop computational models about the functioning of language unities at various levels. Until now, only a few studies investigated the way in which the human brain processes discursive cues expressing temporal information, such as Tense, Aspect and Aktionsart. For example, Bastiaanse and colleagues (2011) carry out a series of experiments with healthy and with aphasic speakers. They find that past-time and present-time references involve different neural processes, dissociation observed for both categories of participants. Dragoy and Bastiaanse (2013) found that past-time forms are more impaired than present time forms, and this occurs irrespective of the category conveying temporal location of eventualities (Tense or/and Aspect) and the type of linguistic expressions (inflexions, auxiliaries, free or bounded morphemes).

As for the automatic treatment of discursive cues expressing temporal information, it principally regards NLP and MT. MT systems came to life in the 1950s and they were *rule-based* systems. This means that they required a large amount of linguistic knowledge, precisely semantic and lexical information. Rule-based systems, like Systran[7], perform a word-by-word translation and use the rules of translation from a source language to a target

---

[7] Built by Peter Toma in 1968 and available at http://www.systranet.com/translate

language, often passing through an *interlingua* (i.e. a language-independent representation of meaning). Rule-based systems were almost completely disregarded from the research when *statistical* systems were designed. Carl and Way (2003) suggested using the previously translated sentences or texts in order to increase the knowledge of the MT system. This produced the so-called *example-based* systems. SMT systems make use of a statistical paradigm (Brown et al. 1990, 1991) which searches for the most likely translation *t* of a given source sentence by considering two constraints: (i) find the most likely translation *t* of a string of words using s statistically-learned *translation model* and (ii) find the string of words *t* that are most likely well-formed in the target language using a statistically-learned *language model* (as explained in the COMTIS project proposal, 2009). SMT systems need therefore large amounts of translation data, which can provide it with information about the translation and the language models. Nowadays, there are numerous translation corpora and monolingual data available for free usage on the Internet (Koehn 2005).

Despite their great results, SMT systems lack however to consider inter-sentential relations as those addressed in the COMTIS and MODERN projects. Human translated texts have discourse structure, which is used to express various aspects of meaning. In NLP (both text analysis and text generation domains), there were several attempts to modelise intersentential relations, such as *discourse parses* aiming at identifying stylistical register (Marcu 2000), the Penn Discourse Treebank (Prasad et al. 2004) and LEXCONN (Roze et al. 2010) based on the role of discourse connectives and models for inferring temporal relations such as TimeML markup language (Pustejovsky, Ingria, et al. 2005; Pustejovsky, Knippen, et al. 2005) and Li et al.'s (2001, 2004) model for processing temporal reference in Chinese. As for text generation, methods have been proposed correctly selecting verbal tenses and temporal connectives for EN (Dorr and Gaasterland 1995, 2002), for German (Grote 1998, 2003) and for generating tenses in EN by using lexical aspect (Olsen et al. 2000, 2001; Ye et al. 2006; Loáiciga and Grisot 2015), as well as for generating aspectual markers in Chinese (Ye et al. 2007).

The translation of verbal tenses is a very recent topic despite the fact that correct translation of Tense and Aspect is crucial for translation quality, given the fact that temporal information conveyed is essential for finding the interpretation intended by the speaker (Ye et al. 2007). The most known MT studies targeting tenses and aspectual categories were done for the pairs of languages EN-Chinese (Ye et al. 2006; Ye et al. 2007; Gong et al. 2012a, 2012b) and EN-FR in the COMTIS and MODERN projects (Meyer et al. 2013; Grisot and Meyer 2014; Loáiciga et al. 2014; Meyer 2014; Loáiciga and Grisot 2015).

1.6   Structure of the thesis

Chapter 2 explores briefly linguistic and non-linguistic means to express temporal reference in aspect-prominent and in tenseless languages (section 2.1). Then, it provides an overview of the semantics of the three linguistic parameters investigated in this thesis: Tense, Aspect and Aktionsart. Besides, it discusses certain discourse semantic theories that aimed at modelling the role played by Tense, Aspect and Aktionsart for discourse interpretation (sections 2.2, 2.3 and 2.4). Section 2.5 is dedicated to a monolingual discussion of the verbal tenses investigated in this thesis.

Chapter 3 aims at going beyond the semantics of Tense, Aspect and Aktionsart. This research is set in a cognitive pragmatic framework, precisely Relevance Theory with a focus on the conceptual/procedural distinction (section 3.1). Additionally, temporal reference is discussed within generative grammar (section 3.2), Neurosciences, NLP and MT (section 3.3).

Chapter 4 is dedicated to introducing and motivating the methodology used in this thesis coming from the following fields: Contrastive Analysis (section 4.1), Corpus Linguistics and NLP (sections 4.2 and 4.3). Experimental work carried out (section 4.3) is at the borderline between NLP (i.e. human annotation) and experimental designs (offline experimentation with linguistic judgment task). Finally, quantitative analyses of the data were carried out using descriptive and inferential statistics (section 4.4).

Chapters 5 and 6 describe the empirical work carried out in this research. Corpus work (description and analysis) was performed on bilingual translation corpora (EN and FR) and on multilingual translation corpora (EN, FR, IT and RO). Experimental work consisted of several offline experiments with linguistic judgment task, designed to test theoretical assumptions currently defended in the literature and in this research.

Chapter 7 incorporates the results of the empirical work carried out and proposes a cross-linguistically valid model for the interpretation of temporal reference and its ingredients (sections 7.1, 7.2 and 7.3). Finally, a reanalysis of the verbal tenses investigated in this thesis is proposed based on the model defended in this research (section 7.4).

Chapter 8 points to the main contributions of this thesis and suggests several recommendations for further research.

2 Semantics of Tense, Aspect and Aktionsart

2.1 On temporal reference in aspect-prominent and in tenseless languages

The link between Tense and broad temporal reference is the most obvious for languages such as EN, FR, IT and RO. Tense is the linguistic expression of temporal relations. Natural language employs several types of tense marking to locate situations in time. In some languages of the world, these markings express a past vs. non-past distinction with zero or several degrees of remoteness regarding the temporal distance from the time of speech. For example, Indo-European languages mark past time while Southeast-Asian languages are non-past marking languages (such as Mandarin and Cantonese Chinese, Burmese, Thai and Vietnamese). In some other languages, the inflectional marking expresses a future vs. non-future marking. For example, EN does not mark the future inflectionally whereas FR, IT, Spanish, Basque do mark it inflectionally (see Dahl and Velupillai 2013[8]).

Another issue of great importance for expressing and processing temporal information is aspectual marking. Linguists make the distinction between Aspect as a grammatical phenomenon and Aktionsart, which regards lexical or semantic categories. Aspect concerns the well-known perfective/imperfective distinction. This distinction plays an important role in many verbal systems and is commonly signalled by morphological means (rather than being expressed periphrastically such as the FR *être en train de*). This distinction is morphologically marked in Slavic languages, EN (the progressive morpheme *-ing*) and Mandarin Chinese and periphrastically marked in languages such as Spanish and FR. In languages such as RO, German, Hungarian and Japanese the perfective/imperfective distinction is not grammatically marked (Dahl and Velupillai, 2013). The association between Aspect and location of situation in time rests on the general interpretation of perfective verbs referring to a complete situation as expressing past time and imperfective verbs referring to an incomplete or ongoing situation as expressing present time.

Tensed languages can be classified in *tense prominent* and *aspect prominent* languages. According to Bhat (1999), this classification can be based upon the relative prominence that languages attach to one of three verbal categories, namely Tense, Aspect and Mood. The prominence of one category is signalled by its grammaticalizing to a greater degree than the others, and making it more obligatory, more systematic and more pervasive than other categories. The verbal system in Slavic languages is organised around the category of Aspect, as shown in Table 2-1 for Russian and in Table 2-2 for Serbian borrowed from Trnavac (2006).

Trnavac (2006, 24) notes that in Russian, the tense system is is aspectually constrained in the sense that perfective forms in the non-past (present) cannot get the interpretation of present time but of a future time, as in example (16), as opposed to imperfective forms, as in (17). The verbs in the perfective aspect appear in two tense forms (i.e. past and future), whereas the imperfective aspect allows the derivation of three tense forms.

---

(16) On pročita*et* knigu.
He read.PRES.PERF book.
'He will read the/a book.'

(17) On čita*et* knigu.
He read.PRES.IMPERF book.
'He reads/is reading the/a book.'

Table 2-1: Verbal system in Russian

	Name	Perfective aspect	Imperfective aspect
Past time	Past Tense	✓	✓
Present time	Present Tense	✗	✓
Future time	Future	✓	✓

In Serbian, perfective and imperfective aspects are morphologically expressed and occur with both past and non-past (present and future) verbal tenses, as in the following table.

Table 2-2: Verbal system in Serbian

	Name	Perfective aspect	Imperfective aspect
Past time	Past Tense	✓	✓
	Pluperfect	✓	✓
	Aorist	✓	✗
	Imperfect	✗	✓
Present time	Present Tense	✓[when, if]	✓
Future time	Future I	✓	✓
	Future II	✓	✓

There are four past verbal tenses: past tense (preterit), pluperfect, aorist and imperfect. The past tense and the pluperfect may occur with both imperfective aspects and perfective, as in (18) and (19) for the former and in (20) and (21) for the latter. The aorist verbal tense occurs only with the perfective aspect as in (22) whereas the imperfect verbal tense with the imperfective aspect (23). However, aorist, imperfect and pluperfect are not very common in modern Serbian language. When they do occur, they are regarded as stylistically marked replacements for certain uses of the general simple past. All non-past tenses may occur with both perfective and imperfective aspects. Perfective aspect used with present verbal tense does not refer to the moment of speech. They usually appear in temporal and conditional clauses.

(18) On je pita*o*.
He AUX ask.PRET.IMPERF
'He asked/was asking/has been asking'

(19) On je *u*pita*o*.
He AUX ask.PRET.PERF
'He asked/has asked.'

(20) On je bi*o* pita*o*.
He AUX be.PRET.IMPERF ask.PRET.IMPERF
'He had been asking.'

(21) On je bi*o* upita*o*.

He AUX be.PRET.PERF ask.PRET.PERF
'He had asked.'

(22) On *u*pita.
He ask.AOR.PERF
'He asked.'

(23) On pitaše.
He ask.IMPERFECTIVE.IMPERF
'He was asking.'

Other devices, such as lexical items (temporal adverbials or other lexically composite expressions), discourse anaphora and pragmatic reasoning can be used to express temporal relations. These temporal devices as well as grammatical and lexical aspect characterise the so-called *tenseless* languages. Lin (2012) gives a negative characterisation of tenselessness in opposition to Tense. Tense is the "grammatical expression of location in time" (Comrie 1985) and it is expressed typically by morphemes. Its main characteristic when compared to a temporal adverbial, for example, is that Tense is obligatory and it is always present in the sentence even if temporal information is expressed also by another device, such the adverbial *yesterday* or *tomorrow*. In syntactic theories, the Tense morpheme is assumed to occupy a functional T node, which is the head of the sentence (i.e. a string of words is not a sentence unless it contains a tensed verb). Tensed verbs are necessarily associated to a predication that denotes a situation, sentences being the syntactic realization of propositions, which describe the world (Puskás 2013). Tenseless languages, as Mandarin Chinese, Thai and Yukatek Mayan, present all other temporal devices except what has been defined as Tense (Binnick 1991).

In a tenseless language verbal stems are not obligatorily inflected for person, number, gender, tense or aspect and need not co-occur with a temporal, aspectual or modal marker, but they express temporal locations as precisely as tensed verbs in English do (Lin 2012, 670). Temporal information in Mandarin Chinese is inferred based on a very rich aspectual marking, temporal adverbials (such as *jiangyao* 'to be going to' for future and *cengjing* 'in the past, before' for past time reference, as well as *yǐjing* 'already' and *caí* 'just' for recency), cotextual information (for example, the shifted deictic *xiànzài* 'now' to refer to another time than present time) and larger context, including world knowledge (Smith 2008; see also Shi 2011 for a corpus-based investigation of the translation of Tense and Aspect from EN into Mandarin Chinese). According to Lin (2003), there are several aspectual viewpoints: two perfective morphemes (-*le* and -*guo*), a group of resultative and perfective verb complements and two imperfective morphemes (*zai* and −*zhe*).

Temporal reference to present time is expressed in stative sentences in Mandarin Chinese through an unmarked verb, adjective or nominal phrase that can be accompanied by temporal adverbials denoting present time. Dynamic verbs are not marked in sentences where they express present time reference. When a dynamic verb occurs in a sentence, it receives a generic or habitual interpretation. When a dynamic verb occurs accompanied by the progressive marker *zài,* it receives a present episodic reading. Nevertheless, dynamic achievements can also express past time reference or future time reference when accompanied by a future time adverbial.

Temporal reference to past time can be achieved by means of a temporal adverbial, a

perfective aspectual marker, an experiential marker or a zero form. Lin points to the fact that zero forms occur in sentences expressing both present and future time, arguing therefore against the existence of a null past time form. As for the perfective aspectual markers, Lin points out that they are not pure tense markers because they are required in every past time sentence and they encode a state change (which tense markers generally do not encode). As for lexical aspect, achievements and accomplishments are interpreted in the past.

Temporal reference to future can be achieved through the modal marker *huì* 'will', used to express a non-controllable prediction based on current information, the future-denoting expressions *jiāng* 'will' and *yào* 'want'. *Yào* has a volitional reading when the subject is animate, and a future meaning when the subject is inanimate or when the subject is not the agent of the event. Similarly to the expressions mentioned for reference to present and past time, expressions for future time reference can be used in sentences without reference to future.

According to Tonhauser (2015, 140) aspectual marking (i.e. both Aspect and Aktionsart) is implicated in temporal reference in Mandarin Chinese in two ways: (i) 'it mediates the temporal relation of the topic time (i.e. reference time in Reichenbachian terms) of one clause to that of the other, and (ii) 'leads to default inferences about the temporal location of the topic time relative to the utterance time' (i.e. moment of speech S in Reichenbachian terms). According to Smith (2008), in Mandarin Chinese, Aspect encodes the relation between reference time R and event time E (idea suggested initially in Tedeschi & Zaenen 1981), for example the *-le* perfective conveys that E=R and the *-guo* perfective conveys that E<R. The relation between S and R is pragmatically inferred from Aktionsart (i.e. bounded vs. unbounded situations). Smith and Erbaugh (2005) and Lin (2003, 2006, 2012, 681) suggested therefore a model to account for temporal location in Chinese based on Aspect and Aktionsart.

- Homogenous/unbounded/imperfective situations have a present time interpretation by default.
- Heterogeneous/bounded/perfective situations have a past time interpretation by default.

Lin (2012, and previous research) argued that sentences in tenseless languages have in their syntactic structure an aspectual functional head ASP, which plays the same role that the Tense head does in a tensed language (see section 3.2 for an extensive discussion on syntactic theories about Tense in tensed languages). The aspectual head ASP-P can be perfective or imperfective, as shown in Figure 2-1 for a perfective ASP-P.

Figure 2-1 Syntactic structure of an aspectual functional head ASP
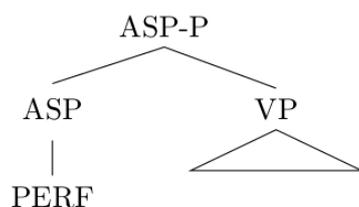
According to the model of *Default Aspect* (Bohnemeyer and Swift 2004), for sentences without an overt aspectual marker, the content of Aktionsart determines both Aspect and

temporal location with respect to S. More precisely, it is assumed that the Aspect of a telic situation is perfective, whereas that of an atelic eventuality is imperfective. Examples in (24) and (25) from Lin (2003, 262-263) illustrate this general assertion. According to Lin, if uttered out of the blue, the telic situation in (24) has perfective aspectual reference and, therefore, default past temporal location whereas the atelic situation in (25) has imperfective aspectual reference and, therefore, default present temporal location.

(24)　Ta dapuo yi-ge hua　ping.
　　　he break one　flower vase
　　　'He broke a flower vase.'

(25)　Ta hen congming.
　　　he very clever
　　　'He is very clever.'

Lin points out that these default interpretations can be reversed by overt expressions such as temporal adverbials, aspectual markers, modals or by a reference point R established in the context.

In a similar research flow, Smith (2008) developed a model for temporal reference in tenseless languages according to which temporal reference in sentences without temporal adverbials arises as a default inference from aspectual marking and universal pragmatic principles. These principles are the following:

- Bounded Event Constraint: Bounded eventualities (i.e. having an initial and a final endpoint) may not be located in the present because the bounds would go beyond the present moment (Kamp and Reyle 1993; Giorgi and Pianesi 1997).
- Simplicity Principle of Interpretation: Choose the interpretation that requires the least information added or inferred (cf Grice's Maxim of Quantity 1975; the Informativeness principle of Levinson 1983; the R-principle of Horn 1984).
- The temporal Schemata Principle: in a zero-marked clause[9], interpret boundedness according to the temporal features of the event or state entity.
- The Deictic Principle: in narratives, the speaker is the centre of linguistic communication and speech time S is the default orientation point. The *present* is located at S, the *past* precedes it and the *future* follows it (Reichenbach 1947; Klein 1994).

According to Smith, in tensed languages, these principles constrain temporal reference to present time whereas in tenseless languages, these principles underline the default pattern of temporal interpretation.

To sum up, in languages in which Tense is not grammaticalized, other devices are used to establish narrow and broad temporal reference. More precisely, Aspect and Aktionsart get the upper hand. Temporal information given by aspectual information can be confirmed or reversed by temporal adverbials, other aspectual markers, models and discursive temporal information. Lin raised the question of how reference time is determined by speakers of a tenseless language, given that it is Tense that has this role. Bohnemeyer (2009) argued in

---

[9] In Mandarin, zero-marked clauses (i.e. no overt grammatical aspect markers) are optional for event clauses and required for statives of all kinds (Smith 2008)

favour of a proposal suggesting that reference time resolution relies on universal "inference mechanisms of temporal anaphora". These universal mechanisms would therefore function both in tensed and tenseless languages. The consequence of this proposal is that in tensed languages, Tense facilitates the identification of the reference time, a task that is entirely left to pragmatic inference in tenseless languages.

At this point of the discussion, another question arises. It concerns the role played by each of the above-mentioned ingredients of temporal reference in discourse in tensed languages and the interactions among them. Lin's presentation of tenseless languages indicates that when one parameter is missing, the others take over and are able to express identic temporal information. One of my suggestions proposed in this thesis is that one could use the values of the other temporal parameters to predict the value of the missing one based on the assumption of *temporal coherence in discourse* (i.e. the coherence given by the values of the components of temporal reference: Tense, Aspect and Aktionsart). The idea is that it is possible to predict the value of a verbal tense based on the values of other temporal ingredients, specifically Aspect and Aktionsart (see section 6.1.8 for a general mixed model, where the dependent variable is the verbal tense and the independent variables are the contextual values of Aspect and Aktionsart among others). As for Tense, the functional head of the clause with scope over the entire sentence, I will argue that it encodes two types of information. The first is related to temporally locating eventualities regarding the moment of speech S. The second type of information regards location with respect to other eventualities passing through the reference time R (see Chapter 7 for a comprehensive discussion).

In the following sections, I address more closely three ingredients of temporal reference, as they are classically described in the literature: Tense, Aspect and Aktionsart. This non-exhaustive presentation points to some well-known works and builds a theoretical background in the semantics of these categories for the research carried out in this thesis. As for issues that go beyond the semantics of Tense, Aspect and Aktionsart, they will be addressed in Chapter 3.

2.2   Tense

Tense has played a central role in analyses of temporal reference ever since the beginning of the formal study of meaning in the early 1970s, where it was defined as a *temporal operator* (Prior 1967, 1968)[10]. As a temporal operator, Tense applies to the basic (i.e. untensed) form of a sentence and it shifts the evaluation time of that sentence to the past or to the future. However, the view of Tense as a sentential operator is too reductionist, mainly because it cannot account for the interactions among the temporal properties of the constituents of a

---

[10] Prior's tense logic offered an *internal* perspective on time (i.e. humans stand inside time, at the point of speech, which is the deictic center). The major debate in philosophy on the metaphysics of time is between the *A-theory* (known as the *tensed* theory) and the *B-theory* (known as the tenseless theory of time) (see Prosser 2013 and Ludlow 2013 for detailed discussions). Prior's logic is situated in the A-theory (or *A-series* of time as proposed by McTaggart 1908), which postulated that one time is *present* while other times are ordered degrees of *pastness* and *futurity*. *Pastness, presentness* and *futurity* are therefore properties of times and they change as time passes. In the B-theory, on the contrary, times are ordered through three relations: *being earlier than, later than*, or *being simultaneous with*. No time is objectively *past*, *present*, or *future* and the apparent passage of time is an illusion (Proser 2013).

sentence, as pointed out by Giorgi and Pianesi (1997, 21).

In a different framework, the so-called the *referential approach*, Tense refers directly to temporal entities and expresses temporal relations. Tense is a deictic category in that it relates entities to a deictic centre, which most of the times, is the moment of speech S (i.e. the *now* of the speaker). Described in these terms, reference to a temporal point seems to be both the meaning[11] and the function of Tense in discourse. It is on this matter that referential approaches to Tense focused[12]. Referential approaches have known theories developed by Reichenbach (1947) for the EN verb system and the Port-Royal Grammar (Arnauld and Lancelot 1660/1972) and Beauzée (1667/1974) for the FR verb system. These theories influenced later work by of numerous scholars such as McCawley (1971), Dowty (1979), Kamp (1979), Kamp and Reyle (1993), Partee (1973), Steedman (1997) and Hornstein (1990) to name but a few. These approaches assume that a verbal tense expresses a relation between two or three (in Reichenbach's system) coordinates: the moment of speech, the event moment and respectively, the reference moment.

2.2.1 *Tense and temporal reference*

Both the Port-Royal Grammar and Beauzée attempted a formalization of the meaning of FR verbal tenses addressing the intuitive idea that they express a relation between the moment of speech and the moment when the eventuality occurred. Arnauld and Lancelot (1660/1975) propose a system of two coordinates: the moment of speech and the event moment (in Reichenbach's terminology). These two coordinates can be linked by a relation of anteriority (for example, E<S for the PS in FR) or by a relation of simultaneity (for example, E=S for the FR PRES). This model, even if innovative for its time, had an important limit: several verbal tenses received the same formal description expressing past (such as the PS, PC and IMP in FR), present or future time. They need hence to be further discriminated based other criteria than the relation E/S. An example of plausible criterion is the *24 hours* rule proposed to distinguish between the PS and the PC. According to this rule, the PS expresses eventualities that took place 24 hours and more before the moment of speech (called a *definite past time*) and the PC expresses eventualities that took place less than 24 hours before the moment of speech (called an *indefinite past time*).

It was the French linguist Beauzée (1967) who offered a solution to Port Royal's problem. He suggested using a third coordinate called *comparison term* (reference point in Reichenbach's terminology) corresponding to *the moment from where the eventuality is considered*. This comparison term is the concretisation of the need to have secondary criteria to distinguish among several

---

[11] As it was argued in Relevance Theory, in this thesis I assume that language is underdetermined. This means that verbal tenses do not have "meanings" but rather "usages" defined contextually as a result of an interpretation procedure (for a developpement see Chapters 3 and 7).

[12] French linguists such as Guillaume (1929; 1971-1992) and Damourette and Pichon (1911-1936) investigated verbal tenses in FR in a different perspective, that is the expression of the speaker's psychological state (attitudes, beliefs). It has been claimed that verbal tenses give access to the way in which the speaker builds temporal representations of eventualities. This idea has been recycled and refined in latter cognitive theories, as pointed out by Sthioul (1998), Saussure and Sthioul (1998). Despite their innovative ideas, Guillaume and Damourette and Pichon's works have been cricized for the vagueness, rigidity and the ad-hoc explanations of their models. For example, Damourette and Pichon attached the speaker's psychological attitude to the morpheme *–ait* contained in FR by the IMP, Present Conditional and Past Perfect (Saussure and Sthioul, 1998).

verbal tenses. Beauzée's model gives an account of two pairs of coordinates *existence period/moment* (event moment S in Reichenbach's terminology) and comparison term R on one hand, and R and the moment of enunciation E on the other. These three coordinates can be in a relation of *anteriority, simultaneity* or *posteriority*. The combination of the pairs of coordinates and relations leads to nine tenses, which could be discriminated more specifically with the help of other secondary criteria.

Introducing a third coordinate allowed for a better discrimination between the PS and the PC, where the PS expresses an eventuality seen from the past and the PC an eventuality seen from the present. Beauzée focused on another important opposition, namely between the PS and the IMP. The distinction between the two verbal tenses being a problem of Aspect, he tried to express it through the comparison term, which can be either a point (i.e. a moment) or an interval. This idea lead to a first limitation of Beauzée's model, that is the imprecise nature of the comparison term and the need to identify this specific data in the cotext. Saussure (1998) noted that the comparison term can be interpreted either a mental projection of S, an aspectual point of perspective or a time interval concomitant to the event itself. A second limitation is the circular explanation of the distinction between the PS and the IMP: the PS provides a comparison point to the IMP and the IMP procures an interval of comparison for the PS (see Saussure 1998 for an extensive discussion). Despite these limitations, Beauzée contributed to the advancement of linguistic knowledge about the functioning of verbal tenses in discourse due to his proposal of a third temporal coordinate.

At the same time as the French flow of research on verbal tenses, Reichenbach (1947) proposed an abstract formalization of the EN verbal system. His framework includes three temporal coordinates used for temporal anchoring of eventualities. Reichenbach assumes that there is a time line (represented graphically from left to right) and argues that 'tenses determine time in reference to the time point of the act of speech, i.e. of the token uttered' called the *point of speech* S (1947, 288). His model includes (as Beauzée's did) the moment when the eventuality occurred called the *point of event* E and a third point called the *point of reference* R, which is a temporal point of view. The point of reference is a key notion in Reichenbach's model[13].

R is a parameter necessary for temporal anchoring pointed out by the semantics of the Past Perfect. Reichenbach (1947, 288) notes that:

> For a sentence as 'Peter had gone' we see that the time order expressed in the tense does not concern one event, but two events, whose positions are determined with respect to the point of speech. […] In the example the point of the event is the time when Peter went; the point of reference is a time between this point and the point of speech.

---

[13] Klein (1994) also proposes three parameters to explain the relationships between Tense and Aspect, namely, topic time TT, time of situation TSit and time of utterance TU. They correspond more or less to what Reichenbach called R, E and S, but there are some theoretical differences. In the sentence *The light was on* the TSit corresponds to the time at which the light was on, and the TT corresponds to the time for which such a claim was made. Both TT and TSit are different from the time when the utterance was made, which is time of utterance TU. According to Klein, TT precedes TU and TU is included in TSit, since it is possible that the light was one before, during and after the time of utterance. In Klein's words, 'TT is the time span to which the speaker's claim on this occasion is confined' (Klein 1994, 4).

Unfortunately, Reichenbach did not describe the nature of R in a detailed manner: it is given by the context (i.e. covert expression of R) or in the cotext by a temporal adverbial (i.e. overt expression of R) and it serves for temporal anchoring of one or more eventualities (p.289). In his words:

> In an individual sentence like the one given it is not clear which time point is used as the point of reference. This determination is rather given by the context of speech. In a story, for instance, the series of events recounted determines the point of reference, which in this case is in the past, seen from the point of speech. Some individual event lying outside this point are then referred, not directly to the moment of speech, but to this point of reference determined by the story.

Saussure (1998, 38) argues that this lack of specification on the nature of R leads to several possible interpretations. R could be either a *projection of S* and thus an *observation/evaluation point* situated on the time line. Saussure (1998) assumes that the addressee establishes the point of reference starting from contextual hypotheses, and if more specific information is provided, he either confirms or re-evaluates the initial calculation of R. Reichenbach did not specify whether R should be seen a point, as an interval or both but his analysis of extended tenses seems to indicate that the notion of *temporal extension* is linked to the speaker's aspectual viewpoint (i.e. Aspect) and not to R. Reichenbach's system is not designed to accommodate Aspect more than to consider that "in some tenses, an additional information is given concerning the time extension of the event" (pp. 290). He notes that in languages as FR, two verbal tenses are used to express this aspectual difference: IMP for extended events and PS ("passé défini") for events that are not extended. This aspectual difference corresponds to the *imperfective/perfective* distinction (see section 2.3.2 for an extensive discussion).

For Reichenbach, all three temporal coordinates are necessary for each verbal tense in order to establish temporal reference of one or more eventualities. However, S and E play a crucial role for defining the semantics of the so-called *simple* tenses (past, present and future). He assumed that:

- The present tense conveys that S and E are simultaneous;
- The past tense conveys that E precedes S;
- The future tenses conveys that E follows S.

Moreover, the three coordinates are linked through three possible temporal relations, i.e. *precedence*, *simultaneity* and *succession*. The anchoring procedure begins with the relation between R and S and continues to that between E and R. This leads to having *relative* and *absolute* tenses, where R coincides with S for the former and R is distinct of S for the latter. Positioning E, R and S on the time line, Reichenbach provides a logical taxonomy of verbal tenses, given in Table 2-3, where columns 3 and 4 provide the traditional terminology in EN and in FR (which are the same in IT and RO). This taxonomy assumes that R is punctual since the IMP is not accounted for.

Table 2-3 Description of EN and FR verbal tenses in Reichenbach's terms

Structure	Reichenbach's terminology	Traditional terminology in EN	Traditional terminology in FR
E<R<S	Anterior past	Past perfect *He had come.*	Plus-que-parfait *Il était venu.*
E=R<S	Simple past	Simple past *He came.*	Passé simple *Il vint.*
R<E<S; R<S=E; R<S<E	Posterior past	-	Mode conditionnel *Il viendrait.*
E<S=R	Anterior present	Present perfect *He has come.*	Passé composé *Il est venu.*
S=R=E	Simple present	Present *He comes.*	Present *Il vient.*
S=R<E	Posterior present	Simple future *He is going to come.*	Future proche *Il va venir.*
S<E<R; S=R<E; E<S<R	Anterior future	Future perfect *He will have come.*	Futur antérieur *Il sera venu.*
S<R=E	Simple future	Simple future *He shall/ will come.*	Future simple *Il viendra.*
S<R<E	Posterior future	- (latin *arbiturus ero*) *would* + verb	- (latin *arbiturus ero*)

Reichenbach's system uses R to account for the difference between the Present Perfect (periphrastic construction) and the SP (preterit). In both cases, E precedes S (eventuality took place in the past), but it is the position of temporal point of reference R and its relation to E and S that discriminates between the two tenses: for the SP, R coincides with E and precedes S, and for the Present Perfect, R coincides with S while E precedes them. This distinction is also made in FR between the PS and PC, where the simultaneity of R and S illustrates the relevance of the resulting state in the present for the PC.

According to Reichenbach, the EN Present Perfect is often used as an extended tense, with the specification that the duration of the event reaches up to S (pp. 292), as in (26) and (27). If the speaker does not intend to communicate the duration of the event then the SP is used, as in (28).

(26) I have seen him.
(27) I have known him for 10 years.
(28) I saw him ten years ago.

Reichenbach's system provides an interesting account also of the sequence-of-tense phenomenon (SOT). According to Reichenbach, when sentences are combined to form a compound sentence, the verbal tenses of the considered clauses are adjusted to one another through certain rules. He proposed two rules: (a) *the permanence of the reference point:* R is the same for all clauses as in (29), and (b) *the positional use of the reference point*: R is the carrier of the temporal position. When temporal location is given by an adverbial, it refers not to E but to R. In example (30), the adverb *yesterday* refers to both R and to E, which are simultaneous, whereas in (31) the adverb refers only to R.

(29) I mailed the letter when John came and told me the news.
(30) I met him yesterday.
(31) I had met him yesterday.

In example (29), the connective *when* signals that eventualities *mailing the letter, John's coming* and *John's telling the news* have the same reference moment R. However, eventualities are temporally sequenced: the event of the first clause preceds that of the second and of the third clause. If the temporal relation is explicitly specified in the sentence through connectives such as *before* or *after*, the rule of the permanence of R is replaced by the more general rule: the positional use of R. In example (32), R changes incrementally: R_1 in the first clause changes in R_2 in the second clause and finally in R_3 in the third clause.

(32)　　He was healthier when I saw him than he is now.

In EN, the SP is used in contexts where the compound form PC is used in other languages, such as FR and German as in examples (33), (34) and (35). According to Reichenbach, this is due to the strict adherence to the principle of the positional use of R in EN. In this way, the sentence in (33) is possible in FR even in the absence of a definite temporal adverbs, as in example (34), while in German the PresPerf would be used, as in (35). Reichenbach notes that a language is compelled to satisfy one of the two principles but not both (pp. 295).

(33)　　This is the man who drove the car.
(34)　　This is the man who drove the car yesterday.
(35)　　Dies ist der Mann, der den Wagen gafahren hat.
　　　　'This is the man who has driven the car'

Reichenbach's system has several limitations that received special attention and amendments in the literature (such as Comrie 1976, 1981, 1985; Hornstein 1990; Declerck 1986; see Giorgi and Pianesi 1997 for a discussion). However, they do not reduce the importance and the vast application of Reichenbach's model. Most of the critics made to Reichenbach's system concern the nature and the functions of the reference moment R. I will briefly discuss two of them in the following lines (see, for example, Hornstein 1990 for a comprehensive discussion).

A first problem pointed out in the literature concerns Reichenbach's suggestion that R and S are included in the semantics of all tensed constructions. This forced him to provide a complex description of simple tenses, such as E=R<S for the SP and E=R=S for the PRES. Only complex verbal tensed constructions provide evidence that R is distinct from S and E, as the case of past and future perfect. A second limitation is the lack of specification on the nature of R that permits several hypotheses about how Reichenbach conceived R and what its exact function is (limitation already identified also in Beauzée's system). This limitation has conducted to several proposals of improvement, such as Comrie (1981) who proposed to remove R for absolute tenses (present, past and future), to keep it for relative tenses[14] (such as PresPerf or Past Perfect) and to duplicate it for Past Conditional. Another proposal was Vetter's (1996), who considered R to be an aspectual point of perspective that would allow

---

[14] One of Comrie's amendments of Reichenbach's framework was to modify the distinction between *absolute* (deictic) and *relative* tenses. Deictic tenses have S as one of their arguments and relative tenses use an unanchored reference time instead of S. Both types have E as a second argument, thus deictic tenses convey the relation between S and E, while relative tenses convey the relation between R and E. Whereas S refers deictically to the moment of utterance, R is determined anaphorically in the context.

the perfective vs. imperfective distinction.

It's worth to note Bertinetto's (1986) proposal with respect to the "ambiguity" of R in Reichenbach's system. Reichenbach suggested that when a temporal adverbial is used, it signals the reference moment R. Bertinetto made two propositions: a temporal adverbial does not necessarily signal the reference moment, and R must necessarily be posterior to E. Therefore, temporal adverbials have two functions: (i) when the temporal adverbial expresses simultaneity with E, it has the function of *temporal localization* (TL), and (ii) when the temporal adverbial expresses posteriority with respect to E, it coincides with R. The second case can also be expressed in aspectual terms (i.e. Aspect): E<R conveys a meaning of perfectness (i.e. *compiutezza* in IT and *accompli* in FR).

According to Reichenbach's view on the relation between R and a temporal adverbial, examples (36) and (37) receive the same description, E<R<S, where the adverb *exactly* could be inserted in the first example whereas *already* could be inserted in the second one. In other words, R has a complex function: (i) temporal location of E with absolute tenses, as in (36), and (ii) signalling of a subsequent interval of time when the resultative state of R still hold with relative tenses, as in (37).

> (36)　Giovanni *uscì* a mezzogiorno.
> 　　　John go out.PS at noon.
> 　　　'John went out at noon.'
> (37)　Giovanni *era uscito* a mezzogiorno
> 　　　John go out.PC at noon.
> 　　　'John was out at noon.'

However, Bertinetto (1986, 47) argues against this interpretation and suggests two possibilities:

- Both in (36) and in (37) the temporal adverbial has the function of temporal localization and R is not explicitly expressed in these two utterances. Hence, R is implicitly determined in the context.
- The temporal adverbial in (36) has the function of temporal localization whereas in (37) it signals R.

His suggestion is that the PS in (36) does not require an R, whereas the PC in (37) does, and therefore, it is either implicitly determined in the context or it is provided by the temporal adverbial.

Moreover, Bertinetto speaks about a third closely linked notion, more precisely that of *temporal anchoring* (TA) as shown in examples (38) and (39). In the two examples, the second clause is temporally anchored on the first one. However, in terms of the temporal organization of the events, they are very different: temporal simultaneity in the former and temporal sequencing in the latter.

> (38)　Quando dormo bene, russo fragorosamente.
> 　　　'When I sleep well, I snore vociferously.'
> (39)　Quando dormo bene, lavoro meglio.
> 　　　'When I sleep well, I work better'.

In the light of these observations, Bertinetto makes the following suggestions (1986, 73):

- It is erroneous to consider that every temporal adverbial signals R.
- Three different notions used in the literature under the same label of "reference moment R" can be distinguished, namely *TL, TA* and the proper R.
- R points to the aspectual notion of perfectness 'compiutezza'.
- TL points to the situation of the E on the time line and it can be explicit or implicit depending on pragmatic reasons.
- Absolute verbal tenses (i.e. simple forms) do not require an R whereas relative verbal tenses (i.e. composed forms) do.
- Temporal evaluation of a series of eventualities can be made based on R or TL; when this is not possible, other linguistic and non-linguistic elements can provide the TA necessary for establishing relations of anteriority, simultaneity and posteriority.
- Some verbal tenses necessarily require a TA, such as the IMP.

Another revision of the initial Reichenbachian system is the one suggested by Reichenbach himself and then further discussed by Comrie (1985) and Hornstein (1990). It has been suggested that the relation among the three points should be split into two distinct relations, one between R and S, and one between E and R. The relation between E and S is never realized directly: it is inferred (see section 3.2 for a syntactic approach; see Moeschler et al. 2012 for a pragmatic model of verbal tenses in FR that suggests a three paired division). R is thus pivot information between E and S. For example, as Giorgi and Pianesi (1997, 88) argue, R permits explaining the incompatibility of the SP in (40)-(43) and the compatibility of PC in (44)-(47) with present time adverbials in EN, IT, FR, RO and in many other languages. The prediction is that if R is the temporal specification of S and not of E, then compound past forms described by Reichenbach as E<R=S are compatible with present time adverbials, while simple past form are not.

(40)   *Now I ate enough.
(41)   *Adesso *mangiai* abbastanza.
(42)   *Maintenant je *mangeai* assez.
(43)   *Acum *mâncai* destul.
(44)   Now I have eaten enough.
(45)   Adesso *ho mangiatto* abbastanza.
(46)   Maintenant j'*ai mangé* assez.
(47)   Acum *am mâncat* destul.

Notwithstanding the criticisms, Reichenbach's description of the semantics of verbal tenses is topical and continues to represent the basis on which researches involving temporal information build their models. This thesis makes no exception. In the model defended in this research, Reichenbachian temporal coordinates are a means to express conceptual and procedural information encoded by Tense. E and S operationalize the pro-concept TIME, which can be contextually specified as the ad hoc concept *pastness* (E<S) or *non-pastness* (E≥S). Pro-concepts are semantically incomplete and are contextually worked out. Contextual information consists of linguistic (typically temporal adverbials, prepositional phrases and temporal pragmatic connectives among others) and non-linguistic (such as world knowledge) sources. This suggestion was validated in an offline experiment with elicitation task, where

participants had to provide the tensed form of an infinitive verb based on the context. Participants had similar results both for built examples (the source of contextual temporal information is controlled) and for natural data (no control of the source of contextual temporal information). They had high inter-judge agreement rates pointing to the easiness of evaluating this type of temporal information (see Experiment 7 for a detailed description and section 7.2.1.1 for a theoretical interpretation of the results).

As far as R is concerned, the relation E/R has a procedural nature. Specifically, it is linked to the instruction encoded by Tense to temporally relate eventualities with respect to one another. As pointed out by Reichenbach, the semantics of the Past Perfect, as in (48), indicates that it puts in relation two temporal entities: the moment when the police arrived (i.e. E) and the moment for which the claim was made (i.e. R). The latter is between E and S. R could be, for example, rendered explicit by the continuation 'when the ambulance did'.

(48)   The police had arrived at the place of the accident. (when the ambulances did)

In this case, the R in the Past Perfect is linked to its instruction to temporally link two eventualities: the event of arriving of the police and the event of the arriving of the ambulances. Consequently, I make the hypothesis that this instruction is encoded by the Tense category and, hence, it is applicable for all other verbal tenses.

Finally, temporal coordinates E, S and R are variables that must be contextually saturated, as well as the relations among them (i.e. precedence, simultaneity and posteriority) and this takes place at the level of the explicature of the propositional content (as discussed more in detail in section 7.2.1.1).

2.2.2 *Tense, temporal reference and discourse structure*

Formal discourse semantics approaches investigated the semantics of verbal tenses building their models on the previous referential approaches (as discussed in section 2.2.1) and with regards to the questions of discourse coherence and discourse relations (Halliday and Hassan 1976). In this section, I will briefly recall the *Discourse Representation Theory* (DRT: Kamp 1981, Kamp and Rohrer 1983, Kamp and Reyle 1993) and the *Segmented Discourse Representation Theory* (SDRT: Lascarides and Asher 1993, Asher 1993). These approaches have their roots in formal semantics (which proposes rigorous and detailed descriptions of the meaning-form relation and where each sentence is interpreted individually) and logic (a science of inference that goes to back to Aristotle). Discursive theories aim at describing discourse as being composed of complex structures that are logically related.

Discourse Representation Theory

DRT (Kamp 1979, 1981, Kamp and Reyle 1993) represents a logical approach to discourse and evaluates the representation of discourse in relation to the representation of reality in terms of truth conditions. DRT is a formal semantics theory that considers the articulation between the truth conditions of sentences (which, according to a model-theoretic semantics based on Montague Grammar[15], depend on the connection between the meaning

---

[15] As referred to by Partee (1973) and explained by Kamp and Reyle (1993)

of the expression uttered on one hand, and on its factuality on the other) and the phenomenon of language interpretation (Kamp and Reyle 1993, 23). In proposing DRT, Kamp and Reyle assume that the interpretation of sentences and texts is constructed in the form of abstract structures that they call *discourse representation structures* or *DRSs.* DRSs are logically related and built by applying certain rules, called *DRS Construction Rules.* A DRS consists of *discourse referents* (entities that a piece of discourse is about, functioning as variables) and *conditions* that apply to these referents.

The main idea of this theory regarding the semantics of coherent multi-sentence discourse and text uttered by the same speaker is that each new sentence of a discourse is interpreted in the context provided by the sentences preceding it, that is a representation structure (van Eijck and Kamp 1997). DRSs are linguistic units larger than single sentences but their representation is made sentence by sentence while maintaining the semantic cohesiveness of the discourse or text (Kamp and Reyle 1993). Semantic cohesiveness is provided by various kinds of *cross-reference* that connect coherent pieces of discourse (Kamp and Reyle 1993, 59). The processing of a piece of discourse is incremental; specifically, in the process of understanding, the interpreter must relate or connect the new sentence to the information structure he has already obtained from the preceding ones. The "old" information structure will be "updated" in the light of the interpretation process, and the updated information structure becomes the new context for the processing of a following sentence, until the entire discourse has been interpreted.

This representational and dynamic approach of meaning at the discursive level was influenced by psycholinguistic *procedural approaches* of the meaning of linguistic expressions (Bras 2008 citing Sanford and Garrod 1981, Fodor 1983). Meaning is seen as instructions for incrementally building mental representations of discourse. The procedural nature of the meaning of some linguistic expressions, as well as their role for discourse interpretation, was also debated in RT and its followers (see section 3.1.3.2). I will consider more in detail the procedural/conceptual distinction applied to verbal tenses (see section 3.1.3.4), as well as its integration in the model proposed in this thesis (see section 7.2).

The initial problem that motivated DRT is the interpretation of nominal and temporal reference in discourse. DRT offers a concrete tool for interpreting anaphoric expressions through the idea of updating contexts by introducing new elements that can serve as antecedents. Kamp (1981) and Kamp and Reyle (1993) provide rules of processing sentences in a representation structure. The main important feature of these rules is that they impose formal constraints on the availability of discourse referents for anaphoric linking. Specifically, the available discourse referents are those of the current structure and those from structures going backward step by step or from an encompassing structure.

For example, in example (49), the arrival occurs at some indefinite time on a specific day in the past. Mary's entering the house is linked to the time of the arrival. The interpretation of (49) involves establishing an event discourse referent for the arrival event and linking it to a reference time discourse referent that points at an interval just after the time of arrival. The processing of the second sentence introduces an event that is constrained to be included in the reference time interval and has the property to shift the reference time discourse referent from just after the time of arrival to just after the time of the entering of the house.

(49)   Mary arrived during the day. She let herself into the house.

Constructing a DRS for one sentence basically consists of introducing a new discourse referent for the described eventuality, setting the temporal relation between this discourse referent and the time of utterance, introducing a discourse referent *t* for the time denoted by the adverb in the case when the sentence contains a temporal adverb, and finally, setting the temporal relation between this discourse referent (a constraining information) and the described eventuality (Kamp and Reyle 1993, 514). An additional step specifies the type of the described eventuality (state or event).

Regarding the construction of DRSs for sequences of sentences, and thus temporal sequencing such as in (50), Kamp and Reyle (1993, 521) argue that 'the eventuality described by a non-initial sentence is interpreted as standing in some specific relationship to some other event introduced by an earlier sentence or to some earlier introduced time *t*.' They make use of Reichenbach's temporal coordinates, more specifically of the reference point R in the form of a new condition α that represents a time or an event that is already present in the DRS. Following this idea, interpreting the second sentence (with a progressive verb form and describing a state) involves setting the α from the first sentence as included in the reference point of the second one, hence leaving R unchanged. For the processing of the third sentence (with a simple past tense and describing an event) the relation is not that of inclusion but that of succession: the reference point of the current interpreted sentence succeeds temporally the reference point of the preceding ones.

(50)   A man entered the White Hart. He was wearing a back jacket. Bill served him a beer.

The discursive contribution of verbal tenses and of sentences containing a verbal tense consists thus of introducing temporal discourse referents (states or events) and temporal relations that the discourse referents have with the surrounding cotext, making use of Reichenbach's coordinates. Reichenbach's analysis envisages verbal tenses as anaphoric devices through the introduction of the reference point R. Reichenbach's analysis led to the apparition of several formal semantic discursive theories that envisage the interpretation of verb tenses as being temporally related to the preceding sentences (Kamp 1979, Hinrichs 1986; Kamp and Rohrer 1983, Partee 1984).

Kamp and Rohrer (1983) combine the DRT structure and textual function for investigating the meaning of the FR PS, PC and IMP. They argue that the choice of a tense form depends on the "function that the sentence in which it occurs has in a text…the factors which determine the use of IMP or PS can only be explained at the level of discourse representation" (1983, 253). They also underline that the reference point R is established by the context and includes the antecedent discourse.

Kamp and Rohrer argue that verbal tenses encode information about how to establish temporal reference of an eventuality related to a reference point and related to other eventualities in a discourse. In a text containing a succession of sentences whose main verb is in the PS or PC, the order of the sentences corresponds to the order of the events. The same principle can be applied for interpreting a succession of events in a complex sentence. The PS and the IMP thus encode interpretation rules: a PS introduces a new event representation with a reference point that succeeds the reference point of the previous

sentence, whereas an IMP introduces a new state representation, which covers a period that includes the reference point of the event introduced previously by a PS or a PC. Kamp and Rohrer therefore propose a predictive model for interpreting the French PS (time moves forward) and IMP (time stagnates) and they put forward some of the exceptions of the rules, explained in terms of complexity of temporal indexicality and the role of temporal adverbs for building DRs, as well as the notions of *temporal* and *personal perspectives* on eventualities.

In conclusion, DRT proposes a semantic interpretation of discourse (in terms of mental representations) with a focus on the notions of reference (nominal and temporal) and contextual interpretation. Within DRT, only sentential syntax and compositional semantics of the DRSs affect the interpretation of temporal anaphora. Lascarides and Asher (1993) underline that in DRT forward movement of time is encoded in the logical form of the clauses through the forward movement of their reference times, while statives do not encode this information.

One of the limits of DRT is the fact that the semantic rules provided are too specific and limit the empirical cover of the model explained (Kamp and Rohrer 1983). Lascarides and Asher (1993) propose counterexamples such as '*Max fell. John pushed him.*' to the DRT's analysis, where the temporal order of events mismatches their textual order, and thus rules for constructing the logical form yield a DRS with wrong truth conditions. These limits have been considered in SDRT.

Segmented Discourse Representation Theory

SDRT is a dynamic semantic theory that considers the logical relations between segments of a text and analyses them in terms of *cohesion* (Halliday and Hasan 1976) and *coherence relations* that structure the discourse (Hobbs 1979,1985; Mann and Thompson 1987; Sanders et al. 1992, 1993).

SDRT was developed as a continuity of DRT in the beginning of the 1990s (Asher 1993, Lascarides and 1993; Asher and Lascarides 2003) and proposes a certain number of refinements, related to the problems mentioned by Kamp and Rohrer (1983) but not only them. SDRT, offering a formal account of the hypothesis that discourse has a hierarchical structure upon which interpretation depends, was received and settled rapidly as the most elaborated semantic alternative to pragmatic models of temporal interpretation at the discoursive level. SDRT addresses temporal relations among discourse segments (sentences in SDRT and mental representations in DRT) based on the context (consisting of cotext and world knowledge in SDRT, compared to only cotext in DRT).

Lascarides and Asher (1993, 1) point out the limits of DRT by identifying the problem of temporal relations (examples (51)-(54)), which do not depend only on sentential syntax and compositional semantics. They argue that in (51), the order in which the events are described matches their temporal order, in (52) descriptive order mismatches temporal order, in (53) the event and the state temporally overlap, and in (54) they do not. If DRT explains forward movement and stagnation of time, it does not give an account of backward movement and of the difference of interpretation between (53) and (54). SDRT focused on these cases, showing that temporal relations must be calculated on the basis of semantic content, knowledge of causality and language use, as well as sentential syntax and compositional semantics

(Lascarides and Asher 1993).

> (51)　Max stood up. John greeted him.
> (52)　Max fell. John pushed him.
> (53)　Max opened the door. The room was pitch dark.
> (54)　Max switched off the light. The room was pitch dark.

SDRT proposes to segment a discourse in *segmented discourse representation structures* (SDRSs) that is, representations of discourse segments. Each SDRS consists of DRSs that are representations of the minimal discourse constituents (the sentences) and are linked through discourse relations of an inferential nature. The minimal discourse constituents, also called *elementary constituents,* are propositions (expressing the semantic content of the sentence) that can be evaluated in terms of truth-values. Elementary constituents can be grouped hierarchically to form *complex constituents* that have the same structure as a SDRS.

As far as the discourse relations are concerned, SDRT distinguishes between *coordination* and *subordination* relations. Coordination relations concern labels of the same level and subordination relations concern labels of an inferior level. Discourse relations are either identified based on a specific linguistic marker occurring in the constituent treated at that moment or inferred based on the semantic content of the constituent and the discourse context through a *nonmonotonic inference*[16] using *commonsense entailment*[17] (Asher and Morreau 1991). As far as examples (51) and (52) are concerned, Lascarides and Asher (1993) argue that they can be distinguished based on a defeasible causal knowledge that exists between falling and pushing and which is lacking for standing and greeting. They argue that this type of logic is a suitable system of inference for modelling Gricean pragmatic maxims and the world knowledge necessary for temporal interpretation.

SDRT (Lascarides and Asher 2003) proposes five temporal discourse (also called *rhetorical*) relations (modelled after those proposed by Hobbs 1985) that determine the hierarchical structure of discourse and represent constraints according to which discourse segments can be linked together: *narration, explanation, elaboration, background* and *result*. Narration involves sentences where textual order matches temporal order; result involves sentences where textual order matches temporal order with a causal link between the events; explanation involves a causal link between the events but no temporal order; and background involves sentences where events and states overlap with no causal link.

It is not the purpose of this thesis to describe in detail the logic underlying SDRT, such as principles of nonmonotonic inference (Defeasible Modus Ponens, the Penguin Principle, the Nixon Diamond) or monotonic inference (for a complete presentation, see Lascarides and Asher 1993). As Saussure (2003) underlines it, the power of SDRT consists of a complete system of rules that can produce each of the proposed discourse relations, the awareness and

---

[16] Monotonic operations correspond to classical or standard logic formalism (calculation of predicates) while nonmonotonic operations correspond to non-classical logics (such as linear logic). A nonmonotone conditional operator (">") as in A > B is interpreted "if A is true then, normally, B is also true" (Bras 2008, 37). This logic uses the notion of *nonmonotone consequence* that gives account of plausible but revisable or cancellable inferences (where the most specific information has priority for choosing the discourse relation).

[17] Commonsense entailment is a conditional logic for nonmonotonical reasoning, where defeasible consequences do not equate with the conditional connective ">".

modelization of the fact that the human mind must make a decision between two possible interpretations by cancelling the default interpretation and favouring the most specific one. Saussure also points out some of the limits of this theory related to a lack of correspondence between their model and real linguistic situation on one hand, and its lack of cognitive likelihood on the other hand. Firstly, he remarks that the used rules of logic can have an extreme degree of specificity that renders the model very complex and even ambiguous, and for this reason, they should be replaced by general pragmatic principles which are activated contextually and based on the addressee's world knowledge. The second difficulty of SDRT identified by Saussure (2003) is the default narration relation. He argues that for the narration relation, or temporal sequencing in other words, to take place, a minimal set of conditions is required, such as a conceptual relation and occurrence of verb tenses that instruct for progressing of time in that specific context. Saussure argues that discourse relations seem to be the consequence and not the reason of temporal interpretation. He thus proposes a procedural pragmatic framework where verb tenses encode instructions on how to temporally interpret discourse.

SDRT thus proposes a discourse modelization based on the property of coherence given by the structure of discourse. Discourse structure is based on discourse relations holding between discourse segments that have been investigated in other paradigms than SDRT.

To sum up, the semantic and discourse semantics of Tense and temporal reference focus on modelizing the meaning of verbal tenses through temporal coordinates and their contribution to the interpretation of a piece of discourse.

2.3   Aspect

2.3.1 The category of Aspect

Aspect, known as *grammatical aspect*, refers to the possibility of using grammatical forms (i.e. verbal forms) to express the way in which the speaker wants to represent the internal temporal structure of a situation, in other words, her viewpoint of the situation referred to (Declerck 2005, 28). Viewpoint may be expressed through suffixes, auxiliaries or a combination of the two, such as the EN progressive aspect expressed through *be*+V+*ing*. Not all languages use all of these means in an identic manner. For example, Aspect is expressed morphologically especially in Slavic languages.

A speaker may use a special verb form to refer to a situation in its entirety (*perfective* aspect), refer to its beginning (*ingressive* or *inchoative* or *inceptive* aspect), its ending (*egressive* or *terminative* aspect) or refer to it as ongoing (*progressive* or *continuous* aspect). The latter types exemplify the *imperfective* aspect. Additionally, the speaker may choose to view a situation as actualizing only once (*semelfactive* aspect) or as being a situation consisting of a series of the same situations (*iterative* or *repetitive* aspect). Finally, the speaker may express a situation as being a habit (*habitual* aspect) (Declerck 2005).

At this point of the discussion, two terminological distinctions should be made: *perfective* (vs. imperfective) and *perfect* (vs. aorist or preterit). The former distinction represents the grammaticalized expressed of the speaker's viewpoint with respect to the internal temporal constituency of a situation. The latter distinction is rooted in Ancient Greek and Latin,

where it represented a combination of Tense and Aspect. The perfect aspect presents an eventuality as being already completed at the time of reference, which is S for the PresPerf and another past eventuality for the Past Perfect. The aorist is the unmarked form of a past tense, which combines, in general, with the perfective aspect. However, in German and English, the preterit does not always imply perfective aspect and can also have imperfective interpretations. In Romance languages, the preterit is the simple (non-compound) past tense and the perfect is the compound past. It is argued in literature (e.g. Squartini and Bertinetto 2000) that the compound past (PC) in Romance languages passes through a process of aoristicization, changing from a pure perfect to an aorist (see section 2.5.1.2.3).

2.3.2 *Perfective and imperfective aspects*

In this thesis, I deal only with the perfective/imperfective distinction. The perfective[18] aspect indicates the viewpoint of a situation as a single whole, without internal structure, with highlighted boundaries. The imperfective expresses the viewpoint on the internal structure of the situation or on a moment other than the initial or the final ones. The perfective and imperfective aspects are morphologically expressed in Russian[19]. Perfective verbs can be derived from imperfectives by means of prefixation, such as *delat' > **s**delat' (to do)*, root change: *pos**y**lat' > poslat' (to send)*, or stress change *razrez**á**t' > razr**é**zat' (to cut)*. Suffixation is used when an imperfective verb is being derived from its perfective counterpart, such as *dat' > da**va**t' (to give)*. A number of verbs are used both perfectively and imperfectively, such as *obeschat' (to promise)* (Dragoy and Bastiaanse 2013).

Hence, Aspect can be expressed through *derivational* morphology (as in Russian), *inflectional* morphology (as for example, the IMP form[20] in FR, IT and other Romance languages) or periphrastic constructions (as the FR *être en train de*). The consequence of the former case is that aspectual counterparts are not inflected forms of the same verb but different lexical items (i.e. have different entries in the dictionary). In Russian, Aspect has a direct relationship with Aktionsart[21], specifically, it has been argued that it *restricts* grammatical

---

[18] The *perfective* aspect suggested in studies of Slavic languages is called *boundedness* aspect by Allen (1966). This multiple usage of the same term might bring to confusion. In this thesis, boundedness represents bounded and unbounded representations of telic and atelic situations as they are actualized contextually (see section 2.4.2 ; see also Declerck 2005, 72).

[19]According to Comrie (1976, chapter 5), aspectual oppositions are expressed in natural languages in two ways: morphologically (synthetic means) and syntactically (analytic means). For example, the progressive aspect is expressed syntactically in Yoruba through an adverbial phrase, in English with the construction copula+verb+predicate, and in FR with the paraphrase *etre en train de*. As far as morphological means are concerned, there are languages where there is a clearly identifiable marker of aspect (e.g. the invariable affix –*zhe* for progressive in Chinese and -*mi* for the imperfective in Persian), languages where the distinction is made through different stems of the verb (e.g. Arabic, where there are distinct stems are the verbal endings indicating person, number and gender are distinct for the two aspects) and languages where verb stems are the same for both aspects (aspectual oppositions are expressed through prefixation and less often through suffixation as in modern Slavic languages).

[20] It is a classical assumption that the IMP in Romance languages is inherently imperfective. This assumption was challenged by several scholars, such as Blücher (1974). In this thesis, I will be arguing that Aspect and Tense are binary features and four combinations of them are possible and they occurring with different frequencies (see Chapter 7).

[21] Reference to the past (the suffix –*l*) and future can be used with both perfective and imperfective verbs, but only imperfective verbs can be used to express reference to the present (Dragoy and Bastiaanse 2013).

marking, influencing both the distribution and the meaning of Aspect. For example, the choice of aspectual prefixes is motivated by the lexical aspect of the verb (Tatevosov 2002).

The notions of Aspect and temporal location are only partially covering one another (see for example, ter Meulen 1997, and section 2.4.3). According to Bertinetto (1986) aspectual marking does not provide information about the temporal location of eventualities with respect to one another but rather it represents the speaker's viewpoint on the eventuality expressed. The imperfective aspect provides an *instant of focalization*. More precisely, it explicitly brings the focus on an instant included in the open time interval when the eventuality takes place. The perfective aspect, on the contrary, refers to a closed time interval and no other instant can be focalized than the final boundary (or more rarely, the initial boundary) of the eventuality.

Aspect is the grammatical expression of the speaker's viewpoint (contrary to Aktionsart which represents temporal information inherent to the VP). The speaker makes use of one or another form in order to express her standpoint regarding the eventuality. For example, in (55), the first verb in PS, expressing the perfective aspect, and the second one in IMP, expressing the imperfective aspect, refer to the same past time event. However, each of the two forms provides the reader a different viewpoint: from the exterior in the former, and from the interior in the latter.

(55) Quel mattino, Giovanni *andò* a scuola come al solito. Ma mentre *andava*, si avvide di una cosa sconvolgente: era uscito in pantofole. (Bertinetto 1986, 80)
‘That morning, John went to school. But while he was going, he noticed a disturbing thing: he had left in slippers. ’

Numerous misconceptions and misuses of the notion *perfective* aspect lead to an important confusion among linguists and therefore their descriptions of individual languages (as pointed out by Comrie 1976; Žegarac 1991). Firstly, there is the assumption that the perfective vs. imperfective aspects indicate situations of *short* vs. *long* duration. The EN sentence (56) can be translated into Russian either trough the perfective in (57) suggesting (subjectively) a short period, through a perfective form in (58) suggesting a (subjectively) long period of time, or through the imperfective in (59) which is neutral (Comrie 1976, 16-17). Another example is the distinction between the FR PS in (60) and IMP in (61), where there is no objective or subjective differentiation with respect to the period of time. Rather, the former expresses the period of thirty years as a single complete whole, whereas the latter focuses on the internal structuring of the reign expressing that at any point during the thirty years he was reigning.

(56) I *stood* there for an hour.
(57) Ja *postojal* tam čas.
      He stay.PERF.SUBJECTIVE for an hour
(58) Ja *prostojal* tam čas.
      He stay.PERF.SUBJECTIVE for an hour
(59) Ja *stojal* tam čas.
      He stay.IMPERF for an hour
(60) Il *régna* trente ans.
      He reign.PS for thirty years

'He reigned for thirty years.'

(61)   Pendant son marriage avec Lady Ann, il *régnait* trente ans.
During his marriage to Lady Ann, he reign.IMP for thirty years
'During his marriage to Lady Ann, he reigned for thirty years.'

Secondly, perfective aspect was associated to *limited*, *punctual* or *momentary* duration, the imperfective expressing *unlimited* duration. Sentences in (56)-(61) show that both perfective and imperfective forms can be used to express limited periods such as *an hour* or *thirty years*. Comrie argues that the "punctuality" interpretation is due to the fact that the perfective aspect does not give direct expression to the internal structure of a situation but presents it as a single unit. Moreover, Žegarac (1991, 43) points out that the perfective in (62) and (63) in Serbian indicates that the eventuality *preplivati* 'swim across' took place in ten minutes whereas the eventuality *stići* 'arrive' occurred ten minutes after some point in time. These examples illustrate very well the interaction between Aspect and Aktionsart.

(62)   *Preplivali* su reku za deset minuta.
'They swam across the river in ten minutes.'
(63)   *Stigly* su za deset minuta.
'They arrived in ten minutes.'

Thirdly, a frequent characterisation of perfectivity is that it indicates a *completed* action. The term "completed" is misused for "complete" in the sense that the former focuses on the ending point of a situation (Comrie 1976, 18). The perfective denotes a complete situation, with a beginning, middle and end, without focusing on either of these. And this is the case when it is explicitly contrasted with an imperfective form, which expresses a situation in progress. The perfective can be used to express the beginning of a situation when it is combined with stative verbs (lexical aspect), such as the Russian *ponimat* ("understand"). In (64), the perfective *ponjal* means "come to understand, grasp" (Comrie 1976, 19).

(64)   Nakonec on *ponjal*, v čem delo.
'At last he grasped what was up.'

Fourthly, the perfective is associated with a *resultative* interpretation, indicating the successful completion of the situation. Similarly to the completion interpretation, the resultative one is identified when the perfective in (65) is contrasted with the imperfective form in (66) (Comrie 1976, 20).

(65)   Ja *ugovoril* ego.
I persuade.PERF him
'I succeeded in persuading him.'
(66)   Ja ugovarival ego.
I persuade.IMPERF him
'I tried to persuade him.'

The interpretations of *perfective* in terms of *completion* and *resultative* meanings are due to a focus on the final stage of a situation that arises in the opposition to the *imperfective*. In other words, the perfective stands in opposition to the imperfective, being the unmarked member

of the binary opposition perfective/imperfective. Most of the descriptions of the perfective do not correspond to an inherent meaning of the perfect but to it's functioning in opposition to the imperfective.

With respect to imperfectivity, there are languages that have a single category to express it and there are other languages where imperfectivity is divided into a number of distinct semantic distinctions shown in Figure 2-2, and finally languages where the existent category corresponds to only a part of the meaning of imperfectivity (Comrie 1979, 25).

Figure 2-2 Classification of aspectual oppositions
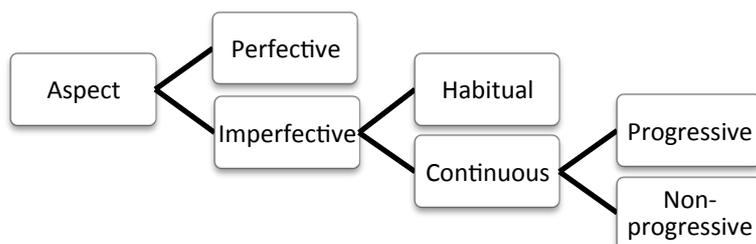
Tense prominent languages, such as EN and Romance languages, do not have morphologically expressed Aspect (except the EN progressive –*ing*[22]). Aspect can be inferred from the use of a certain verbal tense, which can be related to one or the other of the two aspects. If the verbal tense is unmarked for the use of Aspect, then the lexical aspect of the VP is used to determine it (Trnavac 2006). In Romance languages, the distinction between these two aspects is mainly inferred through the distinction between the PS and the IMP. The FR PS is used in perfective contexts[23] whereas the IMP is used in imperfective contexts, such as in examples (67) and respectively (68). Examples (69) and (70) illustrate the IT IMP and PS. As the temporal adverbial signals it, both tenses make reference to past time (E<S). The difference between the two utterances is therefore an aspectual one. In the former the event is presented as in progress at the moment of reference ('around 5') whereas in the latter the event is completed and it has a resultative state holding at the moment of reference ('around 5').

(67) Tout à coup, Jean *tomba*.
Suddenly, John fall.PS
'Suddenly, John fell.'

(68) A cette époque, Jean *tombait* souvent.
At that time, John fall.IMP often
'At that time, John used to fall often.'

(69) Ieri, verso le 5, Giovanni *andava* a scuola.
Yesterday, around 5, John go.IMP to school
'Yesterday around 5 John was going to school.'

---

[22] In this thesis, I do not deal with the progressive in EN and its opposition to the preterit form. However, this question has preoccupied scholars (see for example the discussion in Parsons 1989) proposing accounts of the semantics of the progressive, accounts which emphasize the importance of Aktionsart and of temporal adverbials (such as the *imperfective paradox*, Dowty 1979).

[23] For a detailed presentation of the perfective and imperfective indicators provided here and their correlation with the PS or the IMP, see Vetters (1996, section 2.3).

(70) Ieri, verso le 5, Giovanni *è andato* a scuola.
Yesterday, around 5, John go.PC to school
'Yesterday around 5 John went to school.'

The EN verbal system allows expressing the habitual aspect only for the past tense as in (71) and the progressive as in (72). The simple form, the SP, makes no distinction of aspect and allows the expression of habitual as in (73) but excludes the progressive. FR verbal system makes a clear distinction between perfective and imperfective for past tense, and this corresponds to the PS in (74) and the IMP in (75). The IMP expresses both habitual and progressive aspect. The construction *être en train de* is a supplementary means to express the progressive aspect, and it can be used for past and present time as in (76) and (77). FR, as IT, Spanish or Russian, has a general imperfective periphrase corresponding to habitual and progressive aspects in EN. RO, on the contrary, does not have an imperfective periphrase.

(71) John *used to* work here.
(72) John *was working* when I entered.
(73) John *worked* there.
(74) Jean *lut*.
John read.PS
'John read.'
(75) Jean *lisait* quand le facteur est arrivé.
John read.IMP when the postman arrived
'John used to read/was reading when the postman arrived.'
(76) Jean *était en train* de lire quand le facteur est arrivé.
John be.IMP+V+*ing* read when the postman arrive.PC
'Jean was reading when the postman arrived.'
(77) Jean *est en train* de lire et ne veut pas te parler.
John be.PRES+V+*ing* et does not want to talk to you
'Jean is reading and he does not want to talk to you.'

As far as the continuous aspect is concerned, languages present two categories of continuity in time: progressive as in (78) and non-progressive as in (79). There are languages where the two types of meaning must be expressed by the means of progressive and nonprogressive forms, such as EN, and others where the use of the specifically progressive form is optional, such as IT and FR. This means that the nonprogressive form does not exclude progressive meaning, as the translation into EN of the IT sentence illustrates in example (79).

(78) Gianni sta cantando.
'John is singing.'
(79) Gianni *canta*.
'John sings/John is singing.'

A series of indicators of the perfective and imperfective aspects have been suggested for tense prominent languages, as shown by the following examples in FR. Vetters (1996) argues that there are several types of perfective indicators occurring in perfective contexts which are incompatible with the IMP: (i) temporal indicators explicitating the end of the situation such

as *jusqu'à 8 heures* 'until 8 o'clock' or the beginning and the end of the situation such as *du matin jusqu'au soir* 'from morning until night', as in (80) and (81), (ii) temporal indicators explicitating the total duration of the situation such as *en 50 minutes* 'in 50 minutes', as in (82), (iii) the repetition of the situation signalling that it is completed such as *trois fois* 'three times', as in (83), (iv) temporal indicators explicitating the punctual realization of the situation such as *à l'instant* 'instantaneously', as in (84), and (v) the change of state or position accompanied by temporal indicators such as *le jour d'après* 'the next day', *quelques secondes plus tard* 'a few seconds later', as in (85). In (86), the IMP is possible but it is used in its *narrative* or *breaking* interpretation (Tasmowski- De Ryck 1985) (see section 2.5.1.2.2 for a detailed presentation of this usage).

(80)   Le 5 juin 1989, Jules *attendait/*attendit* jusqu'à 5 heures.
       'On the 5th of June 1989, Jules *was waiting/waited until 5 o'clock.'
(81)   Le 5 juin 1989, Jules *étudiait/*étudia* du matin jusqu'au soir.
       'On the 5th of June 1989, Jules *was studying/studied from morning until night.'
(82)   Le 5 juin 1989, Jules *rentrait/*rentra* chez lui en 50 minutes.
       'On the 5th of June 1989, Jules *was coming back/came back home in 50 minutes.'
(83)   Le 5 juin 1989, Jules *sonnait/*sonna* trois fois à la porte.
       'On the 5th of June 1989, Jules *was ringing/rung tree times at the door.'
(84)   À l'instant, Paul *trouvait/*trouva* la solution.
       'Instantaneously, Paul found the solution.'
(85)   Quelques secondes plus tard, Luc *fut* sous le chapiteau.
       'A few seconds later, Luc was under the tent.'
(86)   Quelques secondes plus tard, Luc *était* sous le chapiteau.
       'A few seconds later, Luc was-IMP under the tent.'

As for imperfective indicators signalling that the situation is in progress, there are also several categories: (i) background situations introduced by *quand* 'when', *pendant que* 'while', *pendant* 'during', as in (87) and (88), (ii) situations interrupted by other events preventing them of being completed, as in (89), (iii) telic situations accompanied by temporal adverbials expressing a long period such as *pendant la guerre* 'during the war', *pendant sa jeunesse* 'during his youth', as in (90), and (iv) non-specified repetition of a situation accompanied by temporal adverbials such as *souvent* 'often', *toujours* 'always', *en général* 'usually', *regulièrement* 'regularly', as in (91).

(87)   Nous *fûmes/*étions* à l'étude quand le proviseur entra.
       'We *studied/were studying when the teacher came in.'
(88)   Elle remonta à sa chambre, et pendant que je l'*embrassai/*embrassais*, ell edit (…).
       'She climbed back up in her room, and while I *kissed/was kissing her, she said (…).'
(89)   Il se *noya/*noyait* quand l'agent le sauva.
       'He *drowned/was drowning when the agent saved him.'
(90)   Quand il était jeune, Jean *prit/*prenait* son café avec moi.
       'When he was young, John *drank/was drinking his coffee with me.'
(91)   A cette époque, Jean * *étudia*/*étudiait* toujours du matin jusqu'au soir.
       'At that time, John *studied/was always studying from morning until night.'

As indicated in (86) and in (92) below, one cannot talk about a one-to-one mapping. The

narrative IMP is used in a perfective context. There are therefore numerous scholars who have criticized this classical distinction between perfective PS and imperfective IMP as discussed by Vetters (1996) for FR (see also section 2.5.1.2), and several modern models were suggested in order to explain this lack of one-to-one correspondence (such as the neutrality of the IMP or the proposal made in this thesis, according to which the tensed verbal form consists of both Tense and Aspect, whose values combine, see Chapter 7).

(92)   Tout à coup, Jean *tombait*.
       Suddenly, John fall.IMP
       'Suddenly, John fell.'

As far as RO is concerned, it is only recently that the Romanian Academy introduced the category of Aspect (GLR edited by V. Guțu-Romalo 2005 in 2 volumes). As pointed out by Margan (2009), traditionally Romanian grammars consider that aspectual distinctions are lexicalized, as in aspectual verbs (*a începe* 'to begin', *a înceta* 'to begin', *a se pune pe* 'to start doing something'), aspectual words (*deja* 'already', *tot* 'still', *mereu* 'always, *în fiecare zi/lună* 'every day/month) and aspectual prefixes (*a reciti* 'to read again'). GLR proposes the category of *aspect* (which includes both Aspect and Aktionsart), which is 'specific to the verb and which points to the structure of the time interval when the situation described by the verb takes place' (2005, vol. 1, 449). GLR makes the following aspectual distinctions in terms of : [±perfectivity], [±durativity], [±genericity], [±iterativity] and [±inchoativity], as illustrated in the following examples from Margan (2009, 52):

(93)   El *a scris*.PC. (perfective)
       'He wrote.'
(94)   El *scria*.IMP. (imperfective)
       'He was writing.'
(95)   El *scrie*.PRES. (durative)
       'He writes.'
(96)   El *intră*.PRES. (punctual)
       'He enters.'
(97)   El *ascultă*.PRES muzica anilor 70. (determinate)
       'He listens music from the 70s.'
(98)   El *ascultă*.PRES muzică. (generic)
(99)   'He listens music.'
(100) El *a scris*.PC o scrisoare saptămâna trecută. (unic)
       'He wrote a letter last week.'
(101) El *a scris*.PC două scrisori saptămâna trecută. (iterative)
       'He wrote two letters last week.'
(102) El *a scris*.PC scrisori în fiecare zi. (repetitive)
       'He wrote letters every day.'
(103) El *începe*.PRES să scrie. (inchoative)
       'He starts to write.'
(104) El *continuă*.PRES să scrie. (continuative)
       'He continues to write.'
(105) El *termină*.PRES de scris. (egressive)
       'He finishes to write.'

According to GLR, among these distinctions, RO expresses grammatically only the first one: perfective and imperfective. All other distinctions are expressed lexically. Aspect in RO is expressed only with past and future time verbal tenses, and it is associated with 'the interpretation "anterior to the moment of reference R", which is different than S' (GLR, 2005, 449). The category of Aspect is, therefore, dependent on Tense since it can be expressed only when R≠S. I will come back to the category of Aspect in RO in section 2.5.1.4, when I will discuss individual verbal tenses.

To sum up, Aspect consists of a binary distinction expressing a viewpoint on the situation. If in Slavic languages this distinction is expressed morphologically, in Romance languages it is most often associated with the PS and IMP used in perfective and imperfective contexts. Žegarac (1991, 50) points out that the lack of the perfective-imperfective grammatical distinction in one language has been erroneously taken as indicating the lack of lexical meaning characterising one or the other member of the distinction. He supports his claim with Ferdinand de Saussure's words:

> Les langues slaves distinguent régulièrement deux aspects du verbe: le perfectif représente l'action dans sa totalité comme un point en dehors de tout devenir; l'imperfectif la montre en train de se faire sur la ligne du temps. Ces categories font difficulté pour un français parce que sa langue les ignore: si ells étaient prédéterminées, il n'en serait pas ainsi[24]. (1967, 161-162)

To fill in this gap, French scholars identified a series of indicators of the perfective and imperfective aspects. Nevertheless, the great variety of these indicators and the lack on one-to-one mapping represent an important drawback for their actual utility for applicative purposes, such as NLP and MT. One could make the hypothesis Aspect could be a relevant criterion that would explain the cross-linguistic variation of verbal tenses. It seems that numerous languages grammatically encode the [±perfectivity] feature (Dahl and Velupillai 2013), which can be considered a parameter with two values: positive and negative.

The question that arises at this point of the discussion regards the means that would make possible the application of the perfective/imperfective distinction in contrastive studies. Let's imagine that one or more languages expressing grammatically this distinction are contrasted to one or more languages that do not express it through the same grammatical means. If the contrastive analysis is carried out based on translation corpora then the target language can be used to infer features of the source language. This is the principle that stands behind the *translation spotting* and *cross-linguistic transfer* methods (see section 4.2.3 for a detailed presentation). Precisely, if a text written in a language where the distinction is not grammatically expressed is translated into a language where this distinction is grammatically expressed, then identifying the aspectual information in the target language makes it possible to transfer it backwards to the source language. My argument is that this makes it possible to have an abstract perfective/imperfective distinction, which is detached from concrete lexical and language-specific means, such as those suggested by Vetters (1996) for FR. Moreover, I will be arguing that operational contrastive analyses of languages, be it from typologically

---

[24] 'Slavic languages regularly distinguish two aspects of the verb : the perfective represents the action in its totality as a point outside of every becoming ; the imperfective presents it in progression on the time line. These categories are difficult for a French person whose langue neglects them : if they were predetermined, it wouldn't be as it is.'

different languages or languages from the same family, can be carried out only if they consider abstract and language-independent features.

In the following sections, I will discuss Aktionsart and its interaction with Aspect aiming at emerging an abstract and cross-linguistically valid feature. I will also discuss the role played by Aktionsart, Aspect and their interaction for the temporal interpretation of a discourse[25].

2.4   Aktionsart

2.4.1   The category of Aktionsart

Aktionsart[26] refers to temporal information intrinsic to situation types. It is also called *ontological* aspect because it refers to ontological features used to describe situations, such as *stativity, durativity, homogeneity*, etc. (see Declerck 2005 for a detailed discussion of ontological features and their application to EN). Aktionsart is the expression of these inherent features of a situation represented by a verb phrase outside of its marking for Aspect and Tense. This is due to the fact that in many cases, Tense and Aspect modify and override inherent temporal features of a situation.

In the literature, there have been several propositions of taxonomies of Aktionsart. These taxonomies make use of ontological features such as stativity, durativity, homogeneity, agentivity and telicity (see section 2.4.2 for a selective discussion). Among these classifications, the most known and distinctive are Lyons' (1977) and Vendler's (1957, 1967)[27]. Lyons' fourfold distinction distinguishes among *states, actions, processes* and *events* and it makes use of the ontological features of dynamicity, homogeneity and agentivity[28]. Vendler's taxonomy distinguishes among *states, activities, accomplishments* and *achievements*. His classification was suggested for EN verbs and makes use of the ontological features of durativity and telicity. Using the compatibility with the progressive criterion, a two-fold classification may be done: states vs. non-states. For FR verbs, Garey (1957) makes use of the telicity feature to propose a two-fold classification: telic situations (states and activities) vs. atelic situations (accomplishments and achievements). Vendler's four-folded taxonomy was

---

[25] Boogaart (1999), investigating the role played by Aspect and Aktionsart for determining temporal ordering of eventualities in English and Dutch, notes that Aspect does not *determine* temporal interpretation of a discourse and that it allows both temporal sequencing and temporal simultaneity. In Boogaart's model, there are three factors that influence temporal interpretation of a discourse containing reference to past time: Aktionsart (states vs. events), Aspect (perfective vs. imperfective) and discourse-type (narrative vs. non-narrative). He suggests that discourse-type and pragmatic incompatibility (i.e. an interpretation is pragmatically incompatible with an utterance if it is not supported either by world knowledge, or by the cooperative principle and maxims in the sense of Grice).

[26] Known as *mode d'action* in FR (Vetters 1996) and *azione verbale* in IT (Bertinetto 1986).

[27] It is also worth mentioning Parsons' syntactic and semantic features of events (1990). In his subatomic semantics, English sentences contain three main elements, namely, subject, verb and tense, which constrain the event. In this thesis, I will however not make more detailed reference to Parsons' approach.

[28] A state is a kind of situation which is conceived as *existing* (rather than being done or taking place) and which is homogenous throughout its duration. Situations that are not static are called *dynamic*. Within the class of dynamic situations, *actions, events* and *process* may be distinguished. Actions are carried out under the control of an agent (e.g. *John dug a hole*) whereas processes and events are not. Moreover, events may be both punctual and durative whereas processes are only durative (Declerck 2005).

proposed also for IT (Bertinetto 1986) and for RO (Stoicescu 2010, Novakov and Lazović 2009). In this thesis, Aktionsart is operationalized as the [±boundedness] feature, representing the actual realization of situation types in the context.

Research on tenseless languages and mixed-temporal languages (such as Smith 2005, 2006) indicated that boundedness and dynamicity play a fundamental role for determining temporal reference. Smith (2005) proposed the Bounded Constraint Principle (as discussed in section 2.1) for tenseless languages, which gives rise to default temporal interpretations. In the following sections, I will discuss Vendler's typology of aspectual classes, their distinctive features and the linguistic tests used to distinguish them, as well as the role played by boundedness in tenseless and mixed-temporal languages to express temporal reference. Finally, I will speak about the interaction between Aspect, Tense and Aktionsart and the discursive function of Aktionsart.

2.4.2 *Aktionsart and [±boundedness]*

Vendler discusses the relation between verbs and time, a relation that can be expressed through Tense on the one hand, and through the use of a verb, on the other. In his words, it is 'the particular way a verb presupposes and involves the notion of time' (1957,143). He proposes that EN verbs[29] can be grouped into four 'time schemata' or aspectual classes: activities, accomplishments, achievements and states[30] and distinguishes them through their restrictions with time adverbials, verbal tenses and logical entailments.

- Activities: *run, push a cart*
- Accomplishments: *run a mile, draw a circle*
- Achievements: *recognize, reach the top, spot the plane, win the race*
- States: *love, know, like*

Vendler's classification of aspectual classes presents a first distinction based on the criterion of the *compatibility with the progressive*. This criterion provides a coarse-grained classification of Aktionsart as *states* on the one hand and as *events* on the other. Specifically, events (i.e. accomplishments, achievements and activities) are compatible with the progressive whereas states are not. However, Žegarac (1991, 195) points out that the number of verb states in EN incompatible with the progressive is reduced, as shown by the following plausible examples:

(106)   Peter is being polite.
(107)   John is living in Muswell Hill.
(108)   Mary is loving the fruit salad.

A finer-grained distinction among the first three types of events can be drawn based on their compatibility with temporal adverbials: activities combine with *for* adverbials as in

---

[29] Scholars (Dowty 1979; Verkuyl 1972; Comrie 1979) have argued that Vendler's approach was too simplistic and that lexical aspect applies to a verb phrase (verb and objects) rather then the verb alone, since the objects can modify the aspectual class. For example, *sing* is an activity and *sing a song* is an achievement.

[30] Mourelatos (1978) argues that Vendler's scheme is too narrow and proposes an ontological typology. For him, all verb predicates are *situations*. Situations can be divides between *states* and *actions* (occurrences). Actions are divided between *activities* (processes) and *events* (performances). Events include *developments* (accomplishments) and *punctual occurrences* (achievements).

(109), accomplishments combine with *in* adverbials as in (110), and achievements, which are punctual, combine with *at* adverbials as in (111).

(109)   He run in the forest *for* thirty minutes.
(110)   He run five miles *in* one hour.
(111)   He knocked at my door *at* 4 a.m.

Generally speaking, the linguistic tests applied for distinguishing aspectual classes (Dowty 1979, 55-60) are:

- non-stative tests to distinguish between states and non-statives VPs:
  - states fail to be used with the progressive, in the imperative, in pseudo-cleft constructions, with adverbs such as *deliberately*, *carefully*, *reluctantly*[31];
  - when achievements pass the non-stative tests, it is due to a change in interpretation: the focus is on the development of the process and they are recategorized as activities;
- use of the *for*-adverbials and *in*- adverbials test
  - states and activities take *for*-adverbials
  - accomplishments and achievements take *in*- adverbials
- entailment tests with the progressive
  - x is V[32]-*ing* entails x has V-*ed* for activities but not for accomplishments[33] (i.e. *if one stops pushing a cart it still means it pushed it*)
  - the test does not apply to states and achievements

However, Stoicescu (2010) argues that the progressive test does not function in RO since all four aspectual classes may co-occur with the IMP as in examples (112)-(116).

(112)   Ion *iubea* muzica.
        John love.IMP the music.
        *John was loving the music.
        'John used to love music.'
(113)   Ion alerga.
        John run.IMP
        'John was running.'
(114)   Ion *săpa* şanţul.
        John dig.IMP the ditch
(115)   'John was digging the ditch.'
(116)   Ieri la ora 5 Ion *găsea* inelul.
        Yesterday at 5, John find.IMP the ring
        'Yesterday at 5, John was finding the lost ring.'

---

[31] Other tests for states are proposed by Lakoff (1965) as pointed out by Žegarac (1991). States are incompatible with: the imperative, the complements of the verbs *persuade* and *remid*, the *do-something* construction and use with *instead of.*

[32] V stands for the verb

[33] These cases were described as the *imperfective paradox* (Dowty 1972; Parsons 1990, Moeschler and Reboul 1994; Reboul 1996).

She suggests that *telicity[34]* is a more appropriate criterion for a coarse-grained distinction. *Telicity, dynamism* and *durativity* are ontological features that are used to distinguish among aspectual classes. Dynamic situations 'require a continual input of energy (Vendler 1967, 13). The term durativity refers to the fact that 'the given situation lasts for a certain period of time' (Vendler 1967, 41). Telic situations have a change of state, which becomes the outcome, or the goal of the eventuality. Telic eventualities have a natural final endpoint, which is an intrinsic boundary. Atelic eventualities have arbitrary final endpoints.

Garey (1957) classified FR verbs also based on the notion of telicity. Telicity concerns the realisation of the inherent goal of the action expressed by the verb. For example, *to swim* is an atelic verb because it is realised as soon as it begins while *to arrive* is a telic verb because the action expressed has an inherent goal that must be reached for the action to have taken place. Telicity is a criterion that distinguishes between states and activities on the one hand (atelic) and accomplishments and achievements on the other hand (telic). As for IT, Bertinetto (1986) suggests a fine-grained classification of verbal action using several criteria. However, he acknowledges the classification suggested by Vendler and Garey based on the telicity criterion. Precisely, telic situations are *azione transformativo* ('achievements[35]') and *azione risultativo* ('accomplishments'), whereas atelic situations are *azione continuativo* 'activities' and *azione stative* 'states'.

Boundedness is closely related to telicity, as pointed out by scholars like Declerck (1979,1989, 1991a, 1991b, 2005) and Depraetere (1995a, 1995b). Telicity and boundedness are the two faces of the same coin, i.e. lexical reference[36]. If telicity evokes the potential actualisation of a situation, boundedness represents the actual realisation of the situation in a context. Situations are telic or atelic, and they can be realized contextually as bounded or unbounded. For example, *running a mile* is a telic situation. It can be expressed in an utterance as bounded as in (117) or unbounded as in (118). These examples indicate that telicity is an inherent feature of eventualities that is not sensitive to linguistic context. Boundedness on the contrary is sensitive to context, such as the tense of the verb and grammatical aspect, past perfective in (117) and present imperfective in (118).

(117)   Max ran the one-mile race.
(118)   Max is running the one-mile race.

Depraetere (1995a) comments that '(a)telicity has to do with whether or not a situation is described as having an inherent or intended endpoint; (un)boundedness relates to whether or

---

[34] Vender did not use the term *telicity*. Jespersen (1948) made the distinction in meaning between telic and atelic situations (in his words, *conclusive/nonconclusive*). In this thesis, telicity points to virtual boundaries of a situation whereas boundedness refers to actual boundaries.

[35] These are equivalences established in broad terms since Bertinetto's classification makes use of several criteria. As far as my purpose in this thesis is concerned, I will not go further into the details. For a discussion of these criteria and of the classes suggested, see Bertinetto (1986, section 2.2.).

[36] The notion of reference goes back to Frege and was used in linguistics by Milner (1982) for referential expressions. It was developed by Reboul (1994), Moeschler (1994) and Moeschler et al. (1998) and applied to temporal and lexical reference. They suggest the notions of *virtual* and *actual temporal* and *aspectual* reference. According to them, when a sentence is uttered, the corresponding utterance receives an actual temporal reference corresponding to the localisation of the eventuality in time. As for Aktionsart, telicity represents the virtual lexical reference whereas boundedness represents the actual lexical reference of a situation.

not a sitation is described as having reached a temporary boundary' (pp. 2-3). A situation is bounded if it is presented as having reached a temporal boundary, irrespective of whether the situation has an intended or inherent endpoint, as in examples (119) to (121). A situation is unbounded if it is presented as not having reached a temporal boundary, as in examples (122) to (124).

(119)   I met John at 5 o'clock.
(120)   Judith played in the garden for an hour.
(121)   Julian lived in Paris from 1979 until May 1980.
(122)   I have lived in Paris.
(123)   She lives on the corner of Russel Square.
(124)   She is writing a nursery rhyme.

A situation has two main boundaries, the left one expressing the beginning and the right one expressing the end. Telicity indicates only the right boundary, i.e. the end of the process. Boundedness indicates one (beginning or end) or both boundaries. In discourse, other linguistic markers such as temporal adverbials serve at marking the boundaries, such as *since, from, as soon as* for the left boundary, *until, till* for the right boundary, and *from… until* for both boundaries. Boundaries are important for marking the limits of a situation in time and have thus an influence on the temporal structure of the discourse. Generally speaking, telic verbs take *in*-adverbials and express non-homogenous and bounded VPs (accomplishments and achievements) and atelic verbs take *for*-adverbials and express homogenous and unbounded VPs (states and activities). Depraetere (1995a) discusses factors that influence the classification of situations as an accomplishment, achievement, activity or state, such as NPs, PPs, Tense and Aspect. She argues that NPs affect telicity (i.e. a NP can turn an atelic situation into a telic one, as in the pair of sentences (125) and (126)). My suggestion is that it is boundedness that is affected rather than telicity. For example, Aspect influences boundedness, as shown in the pairs of examples (126) and (127): in the former there is an atelic unbounded situation which turns into an atelic bounded situation due to the perfective aspect.

(125)   Petrol was leaking out of the tank.
(126)   The petrol was leaking out of the tank.
(127)   The petrol leaked out of the tank.

Finally, another discriminating criterion is the notion of *homogeneity*, described by Vendler (1957,145-146) as follows:

> …running and its kind go on in time in a homogenous way; any part of the process is of the same nature as the whole. Not so with running a mile or writing a letter; they also go on in time, but they proceed towards a terminus, which is logically necessary for their being what they are.

Let's take sentences in (128) and (129) on the one hand, and in (130) and (131) on the other hand. If it is true that someone has been running for an hour, then it is true that he has been running for every period within that hour. The same is true for loving someone. In this case the situation takes place in a homogenous way. In case of running a mile an hour, the

mile was not reached in the first quarter of that hour but only at the end. Running a mile consists of several internal phases oriented towards the natural end. Reaching the top of the mountain is a punctual occurrence with no internal phases, thus the feature of homogeneity is strictly speaking not applicable.

(128)   Max ran for an hour.
(129)   Max loves Mary.
(130)   Max ran a mile in an hour.
(131)   Max reached the top of the mountain.

If the compatibility with progressive and homogeneity criteria are applied, the outcome is that activities and states are homogenous and atelic while accomplishments are telic and non-homogenous[37].

In this thesis, Aktionsart was operationalized as the [±boundedness] feature. The linguistic tests used to distinguish between bounded and unbounded eventualities are *in/for* adverbials, homogeneity and entailment with the progressive. For example, the eventuality 'writing the long letter' in (132) is bounded as shown by its compatibility with *in* adverbial, its lack of homogeneity (the writing of the letter took place in several phases and each phase is different than the others) and its lack of entailment with the progressive (if the president stopped in the middle of the writing, the letter would not have been written).

(132)   John entered the president's office. The president wrote a long letter.
(133)   John entered the president's office. The president sat behind his desk.

On the contrary, the eventuality 'sitting behind the desk' in (133) is unbounded, as shown by its compatibility with *for* adverbials ('for an hour/ ten minutes'), its homogeneity (the sitting behind the desk does consist of different phases but the president has been sitting for the whole time) and its entailment with the progressive (if the president stopped sitting at a certain moment, he can say that he has sit).

The interaction between Tense, Aspect and Aktionsart

The interaction between Tense, Aktionsart and Aspect has received much attention in the literature (Gary 1957; Moens 1987; Dowty 1979; Comrie 1979; Parsons 1989; Smith 1991 to name but a few). There are cases when the interpretation of one category depends on the other, and there are also cases of incompatibility. Garey (1957), for example, describes the interrelations between Aspect[38] and Aktionsart for FR verbs as in Table 2-4. He explains that telic situations expressed with imperfective aspect are interpreted as the action of directing oneself towards a goal without knowing if the goal was attained, whereas those expressed with perfective aspect are interpreted as the action of attaining the goal previously targeted. Atelic situations expressed with imperfective aspect are interpreted in terms of the existence in time of that situation, without saying anything about its beginning or its end,

---

[37] With respect to the properties of dynamism and durativity, the four aspectual classes can be described in the following terms: (i) states are not dynamic whereas activities, accomplishments and achievements are; and (ii) states, activities and accomplishments are durative whereas achievements are not (Smith 1997).

[38] Garey's analysis takes for granted that the FR PS is perfective and the IMP is imperfective.

whereas those expressed with perfective aspect are interpreted as the affirmation of the existence in time of an action, including its cessation.

Table 2-4 Aktionsart and Aspect: interrelations

	Imperfective	Perfective
Telic	*Pierre arrivait.*	*Pierre est arrivé.*
Atelic	*Pierre jouait.*	*Pierre a joué.*

The perfective aspect in (135) and (136) depicts an atelic situation as bounded whereas the non-perfective PRES in (134) depicts the situation as unbounded.

(134) John too loves Mary.
(135) John too has loved Mary.
(136) At that time, it was clear that John too had loved Mary.

Moreover, perfective forms referring to telic situations entail the attainment of the ending point of that situation as in (137) (Dowty 1979; Comrie 1979). This principle does not apply to atelic situations, such as *push a cart* or *sing songs*, where the sentence does not entail the realization of the ending point of the situation as in example (138). The imperfective forms do not carry such implications neither for telic as in (139) nor for atelic situations as in (140). The imperfective applied to atelic situations entails a different kind of information, subject matter called the *imperfective paradox*. This is not the case for telic situations, as making a chair.

(137) Il fabriqua/a fabriqué une chaise.
      He make.PS/PC a chair
      'He made/has made a chair.'
(138) Il *poussa/* il *a poussé* un chariot.
      He push.PS/PC a cart
      'He pushed/has pushed a cart.'
(139) Il *fabriquait* une chaise.
      He make.IMP a chair
      'He was making a chair.'
(140) Il *poussait* un chariot.
      He push.IMP a cart
      'He was pushing a cart.'

Comrie (1989) and Smith (1986) observed that in EN states are incompatible with the progressive, as in (141), whereas in Russian, the perfective applies only to telic situations. FR on the contrary does not impose restrictions regarding the combination between lexical and grammatical aspect (I will discuss the classical characterisation of the FR PS as being perfective and the IMP as being imperfective in section 2.5.1.2).

(141) *She *was being* tired.

Tense plays a significant role for determining the Aktionsart of a sentence (Moens 1987). Example (142) in the SP points to a single event and it is a telic bounded situation whereas (143) is interpreted as a habitual state of affairs and it is an atelic unbounded situation (Moens 1987, 54). Depraetere (1995a) argues that it is because the PRES triggers a habitual

reading that the situation is classified as atelic and unbounded. He therefore suggests that any factor which induces habitual reading can affect a situation's classification in terms of (un)boundedness and (a)telicity as in (144) from Depraetere (1995a, 12).

(142)   John *wrote* a good book.
(143)   John *writes* a good book.
(144)   He *went* to London five times.

The strong relationship between Aktionsart and verb inflection has been observed also for Russian (Dragoy and Bastiaanse 2013). They note that Russian children strongly prefer to use perfectives to refer to past time, and imperfectives to refer to the present as suggested by Gagarina (2004). Moreover, the acquisition of Aspect is dependent on children's developing ability to distinguish aspectual lexical categories as shown by Stoll (1998). Dragoy and Bastiaanse (2013) underline that the Russian Aspect is built upon lexical aspect, and that the lexical nature of the verb semantically guides time reference assignment in children. More precisely, "situations with defined boundaries (e.g. punctual events) expressed through perfective verbs naturally refer to the past, and situations focused on internal structure (e.g. on-going activities) as expressed through imperfective verbs refer to the present time frame" (p. 116).

According to Dahl (1985), these correlations are often observed across languages: past and perfective inflections are generally associated with telic and bounded situations (predicates that presuppose an inherent endpoint of the eventuality), while present and imperfective inflections are associated with atelic and unbounded situations (predicates that describe eventualities without an endpoint). For example, in a recent study on language acquisition, Stoicescu (2010) investigated these correlations in Romanian children aged between 1;5–2;2. She found that in 70% of the cases, atelic situations (states and activities) were used with the PRES verbal tense, whereas more telic situations are used with the PC (76%). These patterns decrease with age. For example, after 2;2 years, the correlation telic situations/PC decreases to 50% of the predicates. The correlation atelic situations/PRES starts decreasing at the age of 1;10[39] (Stoicescu 2010, 189). Stoicescu suggests a possible explanation for the correlation between [±telicity], [±perfectivity] and [±pastness], that is, the notion of [±*boundedness*]. In her words:

> Telicity, perfectivity and pastness involve the notion of boundedness. It is possible that children operate with this single concept when employing past morphology. Similarly, atelicity, imperfectivity and present tense all involve the notion of unboundedness. Working only with two representations and applying them at several level of the language seems like a good strategy to relieve pressure on the linguistic system. (Stoicescu, 2010, 190).

---

[39] Stoicescu (2013) points out that these mismatches observed in children older than 2;2 and in adults are dealt through *coercion*, an idea previously suggested by de Swart (1998) for FR verbal tenses. Stoicescu notes that RO verbal tenses are aspectually sensitive (similarly to FR IMP and PS, as suggested by de Swart) and they select either atelic or telic predicates. In case of mismatches, coercion operators trigger a recategorization to the necessary aspectual class (de Swart 1998). However, aspectual shifts are cognitively costly, and are likely to be avoided. Thefore, speakers produce structures where Aktionsart and Aspect match (Stoicescu 2013).

To sum up, inherent temporal information of VP's can be categorized in four classes: states, activities, achievements and accomplishments. It seems that a coarser-grained distinction can be made depending on the language. For EN, Vendler (1967) suggests a classification according to the compatibility with the progressive: accomplishments and activities accept the progressive whereas states and achievements do not. For FR (and Romance languages in general), Garey (1957) proposes a classification regarding the expression of inherent ending boundaries: states and activities are atelic whereas accomplishments and achievement are telic. Aktionsart influences the interpretation of Aspect and also the general temporal interpretation of a discourse, an observation that I will motivate in section 2.4.3.

2.4.3 Aktionsart and discourse structure

Scholars such as Jespersen (1965), Dry (1981, 1983), Dowty (1986) and ter Meulen (1997) among many others made the hypothesis that the aspectual classes of the VP determine or at least play an important role in determining the temporal relationships between sentences in a discourse. In this section, I will briefly describe Dowty's Temporal Discourse Interpretation Principle (1986), Smith's approach of temporal information in tenseless languages (2005, 2006) and ter Meulen's dynamic interpretation of time (1997).

Dowty (1986) builds his model on narrative texts and argues (p. 37) that there is temporal progression with accomplishments and achievements, as in (145) and (146), and lack of temporal progression with activities and states, as in (147) and (148).

(145) John entered in the president's office. The president walked over him.
(146) John entered the president's office. The president woke up.
(147) John entered the president's office. The president sat behind a huge desk.
(148) John entered the president's office. The clock on the wall ticked loudly.

In his words,

> If a sentence in a narrative contains an accomplishment or achievement predicate but no definite time adverb, that sentence is understood to describe an event occurring later than the time of the previous sentence's event (…narrative time "moves forward" in the second sentence)…If on the other hand the second sentence of the sequence has a stative or an activity predicate, the state or process it describes is most usually understood to overlap with that of the previous sentence: narrative time does not "move" in the second sentence. (Dowty 1986, 37)

Dowty proposes a first exception to this rule, that of the progressive. Specifically, when a progressive form is used, the sequence is interpreted as lacking temporal advancement, no matter what the aspectual class of the VP, whether it is an activity as in (149) and (150).

(149) John entered the president's office. The president was looking out the window.
(150) John entered the president's office. The president was writing a letter.

Another exception is that of some lexical stative verbs (e.g. *stand, sit, realize*) that are ambiguous between a *stative* and an *inceptive* interpretation. In the inceptive interpretation, they behave as achievement VPs and determine the temporal progression in discourse as in

(151). Other stative verbs can receive an inceptive interpretation through adverbials such as *suddenly* or *in a moment* leading to temporal progression as in (152).

> (151) John entered the president's office. The president realized why he had come.
> (152) John sat in his chair going over the day's perplexing events again in his mind. Suddenly, he was asleep.

Dowty's proposition (1986) is that temporal information in discourse depends on sentence semantics (that includes determining aspectual classes) and pragmatic principles. Dowty claims that the temporal relationships between sentences of a discourse are determined by three factors:

- Semantic analysis of aspectual classes using the interval semantics model (Taylor 1977, Dowty 1979). The main idea of the model is that recursive semantic clauses are to be stated in terms of the notions of truth of a sentence with respect to an interval of time. The truth of a sentence with respect to a given interval I is independent of the truth of that same sentence with respect to either subintervals of I, or moments within I or superintervals of I.
- The Temporal Discourse Interpretation Principle (TDIP)
- Gricean conversational implicatures and the "common sense" principle

Dowty (1979, 1986) argues that it is the aspectual class of the whole sentence (rather than any of its constituents) that is relevant to the temporal effect on discourse interpretation. Dowty's idea is that the aspectual class of a phrase or a sentence is determined in a mechanical and completely explicit way by the lexical aspectual class of its main verb, NPs adverbials, tenses and other constituents through compositional semantic rules. Sentence in (153) is an example for the computation of the aspectual class of the sentence: *walk* is an activity, *walk to the station* is an accomplishment and the whole sentence is stative because of its progressive form.

> (153) John was walking to the station.

Based on this observation, Dowty proposes the TDIP for temporally interpreting successive sentences in a discourse. The TDIP postulates that the reference time[40] R of a sentence in a sequence of sentences is to be interpreted to be consistent with the definite time adverbials occurring in the sentence (if there are any), or otherwise with a time immediately preceding the reference time of the previous sentence. Dowty points out that time distance between the R points of the two sentences is determined by pragmatic principles, such as the hearer's understanding of the nature of the events related, the overall degree of detail in which events are being described and common knowledge about the usual temporal relations among events.

According to the interval semantics model, when a sentence with an accomplishment or achievement interpretation is true at an interval I, it is false at all subintervals and it is false at all superintervals of I (by entailment). TDIP thus predicts thus that for sentences with an accomplishment or achievement reading, the sequence of sentences must be interpreted as

---

[40] For Dowty (1982), reference time R and speech time S are contextual parameters of the utterance.

non-overlapping intervals. Consequently, there is temporal progression.

The case of states and activities is different in this respect. Again, according to the interval semantics model, when a sentence with a stative/activity interpretation is true at an interval I, it is true at all subintervals of I. The model makes no prediction for the superintervals of I, thus allowing for sentences with a stative interpretation to be true at all superintervals of I also as in (154).

(154)   John was asleep from 1pm to 2 pm; in fact, he fell asleep at noon and did not wake up until 3pm.

TDIP thus predicts that for sentences with a stative/activity reading the sequence of sentences must be interpreted as overlapping intervals. Consequently, there is no temporal progression as in (155). The situation is different in (156), where the causal relation between the first and the second sentence cancels the overlapping interpretation of the *look out the window* activity.

(155)   John entered the president's office. The clock ticked loudly.
(156)   John asked were the children were. Mary looked anxiously out the window. Their coats lay on the lawn, but they were not in sight. (Dry 1978)

As far as the interpretation of sequences of sentences containing the progressive or expressing iterative or habitual aspect, Dowty (and Dry 1983) assume that they behave as stative sentences. The TDIP predicts no temporal progression, as in (157) and (158).

(157)   John entered the president's office. The president was writing a letter.
(158)   John entered the president's office. They played football together on Sundays.

Kozlowska (1998) and Moeschler (1998) gave arguments against Dowty's hypothesis that aspectual classes *determine* the temporal structure of a discourse. Sentences (147) and (148) have a temporal progression interpretation if the verbal tense is changed, as shown in the FR examples in (158) and (159), where a PS form is used corresponding to the inceptive reading of *to sit* and *to tick* (Kozlowska 1998, 117). Dowty himself points out that the effect of the aspectual class of temporal interpretation can be cancelled by an inceptive reading introduced for example by an adverbial such as *suddenly* as already showed in (152).

(159)   Jean entra dans le bureau du president. Le president *s'assit* derrière un énorme bureau.
        'John entered the president's office. The president sat behind a huge desk.'
(160)   Jean entra dans le bureau du president. L'horloge murale *marcha* bruyamment.
        'John entered the president's office. The clock ticked loudly.'

There are some cases where the temporal interpretation predicted by the TDIP does not apply, such as cases where the second sentence in a discourse describes the same situation but in a more detailed manner, as in (161), cases where a simultaneous interpretation is inferred from the context, as in (162), cases where the second sentence describes subevents of the situation expressed in the first sentence, as in (163) and cased where a progressive expresses the speaker's subjective viewpoint, as in (164).

(161) John knelt at the edge of the stream and washed his face and hands. He washed slowly, feeling the welcome sensation of the icy water on his parched skin. (Dowty, 1986, 58 citing Dry 1983)

(162) At the signal, every one went to work at once. Mary searched the room for any of the items of the list that might be there. John went next door to do the same in Bill's apartment. (Dowty, 1986, 58)

(163) Pedro dined at Madame Gilbert's. First there was an hors d'oeuvre. Then the fish. After that the butler brought a glazed chicken. The repast ended with a flaming desert. (Dowty, 1986, 58 citing Kamp)

(164) In the darkness, John felt his way up the stairway of the dilapidated ild house. Halfway, there was a loud cracking noise under his feet, and suddenly he was falling through space. (Dowty, 1986, 55)

Dowty points out that the TDIP may be considered as describing the 'default' cases of discourse interpretation and it is applicable when the discourse does not provide other sources of temporal information having priority, such as time adverbials, entailments and implicatures regarding the ordering of events.

Smith (2005, 2006) suggested an aspectual model of discourse interpretation for tenseless and mixed-temporal languages. She proposed a model for Mandarin Chinese developed in the DRT framework (Kamp and Reyle 1993), which has syntactic, semantic and pragmatic components and also makes use of contextual information. The syntactic component is the functional ASP-P node of the syntactic structure. The two semantic components are Aspect (perfective, imperfective and neutral viewpoints) and Aktionsart (realized by the verbs and its arguments). The pragmatic component is represented by the default inferences about temporal location in time.

 The viewpoints introduce the reference time R, and the event moment E and their relation, as well as boundedness information, which is represented by conditions relating the situation time interval to the entity E (information specified in the construction rules). Specifically, perfective viewpoints introduce a bounded eventuality, imperfective viewpoints introduce an unbounded eventuality and, finally, neural viewpoints provides information that the situation is visible only partially (i.e. no information regarding boundaries). Moreover, lexical aspect conveys boundedness in zero-marked sentences containing a neutral viewpoint: punctual and telic eventualities (i.e. accomplishments and achievements) are bounded whereas ongoing events (i.e. activities) and states are unbounded.

The third temporal coordinate involved in temporal reference, the moment of speech S, is introduced automatically into the Discourse Representation Structure (DRS) for each clause (Kamp and Reyle 1993). The relation of R to S is done through pragmatic inference. More precisely, by default, bounded situations are located previously to S (i.e. in the past) and unbounded situations are located simultaneously to S (i.e. in the present). These default inferences may be overridden by additional information. The pragmatic principles that underlie Smith's account of temporal reference are the Deictic Principle, the Bounded Event Constraint and the Simplicity Principle of Interpretation (as discussed in section 2.1).

Ter Meulen's Dynamic Interpretation of Tense and Aspect (1995/1997) is a discourse semantics approach of temporal reference based on the role played by aspectual classes, and used for dynamically interpreting sentences in a discourse. Her suggestion is that aspectual

classes and their aspectual properties determine how the events are temporally related in a discourse and that this aspectual information 'controls the dynamics of the flow of information about described change encoded in text' (1997, 6). She redefines the well-known aspectual classes as *holes* (i.e. activities such as *drive around, pour, damage*), *filters* (i.e. accomplishments such as *walk a mile, drive home, land*) and *plugs* (i.e. achievements such as *arrive, finish, begin*). Similarly to the DRT and SDRT frameworks, ter Meulen assumes that sentences are interpreted at the level of the discourse: each sentence is interpreted based on the information provided by the previous ones. For example, if a given sentence is interpreted as describing an event as a *hole*, then the information expressed in the following sentence is interpreted as being part of that event 'as if information it conveys flows through the hole' (p. 7). When a sentence is interpreted as a *filter* than it restricts the information in the following sentence to describe another simultaneous situation. Finally, when a sentence is interpreted as a *plug* it blocks any information about a simultaneous situation. Hence, the context is constrained to redirect its temporal direction by interpreting the next sentence as describing another later event.

Ter Meulen points out that factors such as Tense, Aspect, NPs, PPs and the arguments of the verbs (as previously discussed by Depraertere 1995a) interact with verbs when it comes to determining their aspectual class, and therefore, their function as holes, filters or plugs. Moreover, causal connections or other knowledge of the world can modify and overrule these general semantic principles. Temporal reasoning, a form of logical reasoning, requires that supposed true premises trigger supposed true conclusions if the argument is valid. The temporal information manipulated in logical reasoning can come from three sources: (i) the descriptive content of the utterance, (ii) aspectual classes, and (iii) perspectival information (i.e. provided by grammatical aspect). In ter Meulen's model, these types of temporal information are modelled as *ordered representations of information* obtained based on rules provided by Dynamic Aspect Trees (DATs). Two other important elements in the study of temporal reasoning in ter Meulen's model are *temporal adverbials* and *verbal tenses*. For example, events described by simple past tense clauses and interpreted as filters and plugs affect the perspective by shifting the temporal vantage point. In example (165), the third sentence is interpreted as a different event occurring after the event from the first sentence. On the contrary, the event that caused the perfect state in the second sentence must precede both the simple past event from the first and the one from the third sentence. From (165), one can validly infer (166) (as pointed out by ter Meulen 1997, 15).

(165)   The car hit the fence. The driver had been killed. The police arrived.
(166)   The driver was killed before his car hit the fence and before the police arrived.

In conclusion, this section dealt with Aktionsart and its utility for expressing temporal reference both in a tensed language (English in Dowty and ter Meulen's models) and in a tenseless language (Mandarin Chinese in Smith's model). Dowty's and ter Meulen's models claim that aspectual classes determine the temporal structure of the discourse in English. This seems to be applicable for tenseless languages, where Aktionsart plays a fundamental role for temporal reference. According to the argument that a comprehensive model of temporal reference in natural language meaning must give account for both tensed and tenseless languages, I suggest that Aktionsart is only one of the several factors that should be

taken into consideration when investigating human and automatic processing of temporal reference (see section 6.1.7 for the empirical basis, and Chapter 7 for the theoretical account of the reanalysis proposed in this thesis).

Up to now, this chapter has given an account of the semantics of temporal reference and its ingredients mainly in tensed (sections 2.2, 2.3 and 2.4) and briefly in tenseless languages (section 2.1). The investigated ingredients, namely Tense, Aspect and Aktionsart, were assumed to *play a significant role* and even *to determine* temporal reference in discourse. Despite the fact that their various interactions were pointed out in the semantic accounts described in this chapter, the precise contribution of each of these ingredients has either been maximized or minimized. Moreover, their contributions to establishing temporal reference in discourse have been combined in the generic notion of *verbal tense* (i.e. Tense and Aspect applied to Aktionsart).

In the following section entitled 'Verbal tenses cross-linguistically', I will address another topic classically linked to temporal reference: the description of several individual verbal tenses in the four languages investigated in this thesis.

2.5 Verbal tenses cross-linguistically

2.5.1 Monolingual descriptions

In this section, I will provide monolingual descriptions of three verbal tenses expressing most frequently past time reference and a verbal tense expressing most frequently present time reference. The choice of verbal tenses is based on bilingual corpus-based work (see section 5.1) that revealed an important translation divergence in the EN to FR direction of translation. More precisely, the EN SP is translated into FR most frequently through a past time tense (PS, IMP or PC) or through a present time tense (PRES). Multilingual corpus corpus-work confirmed the choice of the same verbal tenses in IT and RO.

The languages considered in this thesis mark the categories of Mood, Tense and Aspect synthetically (by inflection) and analytically (by periphrases) on the verb. According to traditional grammars, Romance languages have four moods: the *indicative*, the *subjunctive*, the *conditional* and the *imperative*. RO presents another paradigm deriving from the epistemic future called *presumptive*. They present a temporal-aspectual system for all moods but the most complex one belongs to the indicative mood: *present*, *past* (the *simple past*, also called aorist or preterit, the *compound past*, the *imperfect* and the *pluperfect*) and *future* forms (the *future*, the *future perfect* and the *future in the past*). As for the EN verbal system, the indicative mood is the most developed. Subjunctive and conditional interpretations may be expressed through the preterite form V+-*ed* and the second form of irregular verbs. Table 2-5 provides the names and the abbreviations of the verbal tenses considered in this thesis, or more precisely the source SP and the target verbal tenses into FR, IT and RO.

Table 2-5 Selected verbal tenses in EN, FR, IT and RO

Time reference	Past			Present
Language	Preterit/Aorist	Compound Past	Imperfect	
EN	Simple Past SP			
FR	Passé Simple PS	Passé Composé PC	Imparfait IMP	Présent PRES
IT	Passato Semplice PS	Passato Composito PC	Imperffeto IMP	Presento PRES
RO	Perfectul simplu PS	Perfectul compus PC	Imperfect IMP	Prezent PRES

In this section, I will introduce the existent descriptions of the usages and values of these tenses as follows: section 2.5.1.1 is dedicated to the source verbal tense, the EN SP, section 2.5.1.2 is dedicated to FR target tenses, section 2.5.1.3 to the IT target tenses, and finally, section 2.5.1.4 is dedicated to RO target tenses. I aim at showing that the lack of a common framework makes it impossible to compare the verbal systems of the four languages considered in this thesis. This comparison is necessary for identifying the features to be included in a model that explains and predicts the cross-linguistic variation of the translation of the considered verbal tenses.

2.5.1.1   English

2.5.1.1.1   General remarks

As Huddleston and Pullum (2006) point out, the verb paradigm in EN is fairly simple, containing simple and complex forms as in Table 2-6.

Table 2-6 Paradigm of the English verb

Paradigm		
Plain form	V *walk*	V *walk*
Simple forms	Preterit *walked*	Present *walk/walks*
Complex forms	V+ing *was walking*	V+ing *is walking*
	V+have *had walked*	V+have *has walked*
	V+have+be+ing *had been walking*	V+have+be+ing *has been walking*

The plain form of the verb is used in three syntactically distinct clause constructions: *imperative, subjunctive* and *infinitive*. Simple forms represent Tense strictly speaking, or more precisely, the location of eventualities with respect to the moment of speech: E<S for the preterit and E=S for the present[41]. The preterit is expressed through the inflectional morpheme –*ed* for regular verbs and through irregular forms of irregular verbs, and it is called the *simple past*. The simple form expressing present time, called *simple present*, is identic to the plain form for the 1st and 2nd person singular and plural, and for the 3rd person plural. For the 3rd person singular, it is formed by adding the inflectional suffix –*s*. In EN, both the simple past and the simple present are marked forms (i.e. marked through inflectional

---

[41] However, both the preterit and the simple present form may express reference to other times than the canonical one (past time for the preterit and present time for the present form) and have atemporal and non-temporal interpretations.

morphemes).

Complex forms are formed through the combination of Tense and Aspect, or more precisely, the *–ing* suffix and the auxiliary *be* marking the progressive aspect and the auxiliary *have* used to express the perfect aspect. Forms expressing reference to past time are called *past continuous* (past progressive), *past perfect* (past perfect) and *past perfect continuous* (past perfect and progressive). The *present perfect* (present perfect) and *present perfect continuous* (present perfect and progressive) are mixed forms that can express reference both to past time and to present time. Finally, the *present continuous* (present progressive) expresses reference to present time (E=S). In EN, the perfect and the progressive are marked forms.

2.5.1.1.2 Simple Past

The SP has temporal and non-temporal usages[42]. In this thesis, I will deal only with temporal usages. Classical descriptions of the SP (Quirk et al. 1985; Comrie 1985; Leech and Svartvik 2002; Radden and Dirven 2007) present it as the "the deictic time preceding speech time" (Radden and Dirven 2007, 218) that has a main temporal meaning in reference to *past* time. Hence, the EN SP is usually described as representing an action or state as having occurred or having existed at a past moment or during a past period of time that is definitely separated from the actual present moment of speaking or writing.

Radden and Dirven (2007, 218) note three properties of the SP: *focus on the past time, detachment from present* and *definiteness.* Also for Quirk et al. (1985), the SP combines two features of meaning in reference to past time: the event/state must have taken place in the past, with a gap between its completion and the present moment, as in (167), and the speaker must have in mind a definite time at which the event/state took place, as in (168) and (169):

(167)   I *stayed* in Africa for several months (→ I am no longer in Africa)
(168)   Freda *started* school last year/in 1950.
(169)   Prices *slumped* last winter/yesterday.

As for the combination of the SP with Aspect, it can express both perfective and imperfective aspect, as in (170) and (171) from Huddlestum and Pullum (2006). The former example has a perfective interpretation: it reports a promise made in the past. The latter example can be interpreted perfectively or imperfectively. In the former case, the sentence denotes a single act of mowing the lawn located as a whole in the past time. In the latter case, Sue habitually or regularly mows the lawn, and this state of affairs holds at the moment that is being referred to.

(170)   I *promised* to be back for lunch.
(171)   Sue *mowed* the lawn.

---

[42] Non-temporal usages are *modal* (I wish they *lived* nearby), *optative* (It's time we all *took* a rest.), *conditional* (If I *were* to go to Dubai, I can get things like electrical goods) and *politeness* (I *wanted* to ask you a little about The Exorcist). However, in a pragmatic perspective, Saussure and Sthioul (2005) propose that the French IMP used in sentences such as *Je voulais te demander quelque chose au sujet de l'Exorciste*, should be treated as *temporal metarepresentation*, i.e. interpretative usage of the IMP in which the hearer represents her own thought in the past. It produces an attenuation effect because the truth of the situation holding at S is not asserted but only implicated.

Aktionsart and types of situations also play a role for the interpretation of the SP. Specifically, the distinction between states and events gives rise to three interpretations (Leech and Svartvik 2002): *state* in (172), *single event* in (173) and *set of repeated events* (i.e. *habit*) in (174). According to Leech and Svartvik, the 'habit' interpretation combines event interpretation and state interpretation. The *state* interpretation can be specified by adding an adverbial of duration as in (175) whereas the *habit* interpretation can be specified by adding an adverbial of frequency or duration as in (176).

(172) Napoleon *was* a Corsican.
(173) Columbus *discovered* America.
(174) Paganini *played* the violin brilliantly.
(175) Queen Victoria *reigned* for sixty-four years.
(176) He *played* the violin every day from the age of five.

The SP may be accompanied by an overt indicator of time (Quirk et al. 1985). The element of definite meaning (a past event/state) must be recoverable through inference from immediate or larger context, or general world knowledge. Comrie (1985, 41) emphasizes that the SP "only locates the eventuality in the past, without saying anything about whether the situation continues up to the present or into the future". As we have noted above, one of the properties of the SP is *detachment from present*. This is due to a conversational implicature (Grice 1975) based on Grice's maxim of manner, explained as follows by Comrie (1985, 41-42):

> Statements about the present moment are more relevant than those about other times, so that the use of a form explicitly locating a situation in the past suggests that that situation does not hold at the present, otherwise the present tense would be used[43].

The SP may be used in relation to an immediate situation, which has a definite character, as in (177), in a domestic situation where it is known that the front door is locked at bedtime every night. Situational definiteness given by general knowledge explains the use of the SP in historical or biographical statements that have specific people, places or objects as their topics, as in (178). The use of the Present Perfect in the preceding sentence provides a context for mentioning the time, and so it allows a SP in the second sentence, as in (179).

(177) *Did* you *lock* the front door?
(178) Byron *died* in Greece.
(179) They *have decided* to close down the factory. It *took* us completely by surprise.

Radden and Dirven (2007, 219) also note the use of the SP to express bounded past situations, presented as a series of events, typically in narratives, as in (180). The individual events from example (180) are temporally ordered (signalled by their coordination and the conjunction *and*) and are thus interpreted as being successive. Labov and Waletzky (1967)

---

[43] However, as suggested by Louis de Saussure (University of Neuchâtel) at the defence of this thesis (2015), this should be true for all past experiences of any language. In French, for example, the PS shares this propriety with the SP but not the PC in all its usages (see section 2.5.1.2). As for Romanian, the PS may be used to express recent past (corresponding to the PresPerf in EN and PC in FR) (see section 2.5.1.4).

argued that two sentences, which are interpreted as being temporally successive, form a narrative text. The first event is deictically situated in the past time related to the speech moment S while the other is related to the first one.

(180)   I *grabbed* his arm and I *twisted* it up behind his back and when I *let go* his arm there *was* a knife on the table and he just *picked* it up and *let* me have it and I *started* bleeding like a pig. (Labov and Waletzky 1967, quoted in Radden and Dirven 2007, 219)

According to Quirk et al. (1985, 187) the SP also has special uses that occur in certain contexts, such as (a) in *indirect speech,* where there is a transfer from the past tense of the reporting verb to the verb of the subordinate clause (known as *back shifting* or *harmony of tenses*), as in (181), or *forward shifting,* as in (182), where the sentence containing speech or thought in the future contains reported speech referring to present time.

(181)   A: *Did* you say you have/*had* no money? B: Yes, I am completely broke.
(182)   My wife will be sorry that she *missed* seeing you this evening.

One point that arises from these traditional descriptions is that they simply depict intra-linguistically the meanings of the SP, namely the main usage that specifically means "true before speech time" (Riddle 1986, 267).

The competition between SP and Present Perfect

In EN, there is an important competition between the SP and the PresPerf for expressing reference to past time. The PresPerf is a compound form that expresses the *perfect* aspect and makes reference to past time. More specifically, the PresPerf locates an eventuality in the past (E<S), and this is expressed from a reference moment that is in the present (R=S). The SP on the other hand expresses the preterit (i.e. aorist) aspect and locates eventualities prior to S (E<S). This is expressed from a reference moment that is in the past (R=E).

Traditionally, grammars of EN provide a list of adverbials which are compatible with only one of the two verbal tenses, and a list of adverbials compatible with both, as provided in Table 2-7 (Leech and Svartvik 1975). Adverbials compatible only with the SP point to the moment or period of time that finished in the past, whereas adverbials compatible only with the PresPerf point to the period leading up to the present or recent past time.

Table 2-7 Adverbials in relation to SP and PresPerf

SP	PresPerf	SP and PresPerf
I *saw* him yesterday (evening)	I *have*n't *seen* him since Tuesday	I *saw*/*have seen* him today
last night	last week	this week
a week/month ago	I met you	this month
in the morning	so far	recently
on Wednesday	up to now	He always/never *forgot*/*has*
in June	lately	*forgotten* my wife's birthday.
in 1974		
at 4 o'clock		
the other day		

Klein (1992) names the impossibility of occurrence of the PresPerf with a definite

temporal adverbial such as *at 4 o'clock* the *Present Perfect puzzle*. Giorgi and Pianesi (1997) point out that there are [+Present Perfect puzzle] languages, such as EN, Norwegian, Danish and Swedish and [-Present Perfect puzzle] languages, such as Romance languages, German, Dutch and Icelandic. Kamp and Reyle (1993) argued that in languages that are subject to the Present Perfect puzzle, their perfect form expresses only the last stage of a complete event, which are *preparatory stage, culmination point* and *result stage.* This characteristic explains their incompatibility with past time adverbials[44] as in (183) and their lack of usage in narratives[45] as in (184).

(183)   *Mary *has arrived* at 5.

(184)   *Mary *has arrived* and *has started* to cook. She then *has turned* on the TV and she *has watched* a movie.

Languages that are not subject to the Present Perfect puzzle, such as Romance languages, express more than the *result stage* (Vișan 2006). This is what makes them compatible with a past time adverbial as in (185), and what explains their possible usage in narratives as in (186).

(185)   Marie *est arrivée* à 5 heures.
        Mary arrive.PC at 5 o'clock
        'Mary arrived at 5 o'clock.'

(186)   Marie *est arrivée* et *a commencé* à cuisiner. Elle *a* ensuite *allumé* la télé et *a regardé* un film.
        Mary arrive.PC and begin.PC to cook. She then turn.PC the TV and watch.PC a movie
        'Mary arriveed and began to cook. She then turned on the TV and watched a movie.'

The main difference pointed out in grammars with respect to the competition between the SP and PresPerf is the absence, and respectively, the presence of a link between the past time referred to and the present time. More precisely, the SP implies a gap between past and present time (i.e. the two moments are disconnected), whereas the PresPerf implies that the eventuality expressed, be it a state as in (187), a habit as in (188), or an event as in (189), continues at the present time pointing to the *resultative* eventuality holding at S. In (190) and (191), the PresPerf makes reference to an indefinite eventuality located in a period leading up to the present (Leech and Svartvik 1975, 66).

(187)   That house *has been* empty for ages.

(188)   He *has attended* lectures regularly.

(189)   The taxi has arrived.

(190)   *Have* you ever *been* to Florence?

(191)   All my family *has had* measles (in the last year).

As for the usage illustrated in (190) and (191), there is a tendency in American EN to prefer the SP, as in (192).

---

[44] Spanish and Catalan are however subject to a constraint called by Comrie (1985) the *hodiernal restriction*, also known as the *24 hours rule* (Vișan 2006; Aménos-Pons 2011).

[45] With the exception of the Dutch complex past, which is compatible with definite past time adverbials but it cannot be used in narratives (Boogaart 1999, Vișan 2006).

(192) *Did* you ever *go* to Florence?

In this thesis I deal only indirectly with the PresPerf. Explicitly, the competition between the SP and the PresPerf to express reference to past time takes into consideration the part of meaning shared by the two verbal tenses, that is, E<S. However, as pointed out by Reichenbach (1947), the two verbal tenses differ with respect to the position of the reference points: R=E for the SP and R=S for the PresPerf. This difference was tested and validated in an annotation experiment, which is described in section 6.1.6.

To sum up, in EN, information regarding Tense and Aspect is expressed in the same verbal form. There are simple and compound forms. Tense expresses reference to past, present and future times: in other words, it locates an eventuality as being prior, simultaneous or posterior to the moment of speech S. As for Aspect, EN expresses the progressive aspect through the *–ing* morpheme which applies to past and present verbal tenses, as well as verbal forms making reference to future time. With respect to past time reference, there is a competition between the aorist (SP) and the perfect (PresPerf) form. In this thesis, I will be arguing that these two verbal tenses share conceptual information (E<S) and behave differently with respect to their procedural information (E/R).

2.5.1.2 French

FR verbal tenses expressing past time have been extensively studied and described by scholars, among whom are Benveniste (1959, 1966), Kamp and Rohrer (1983), Comrie (1985), Vetters (1992, 1996), Moeschler et al. (1998), Reboul and Moeschler (1998) and Saussure (2003) to name but a few. Each of the FR verbal tenses described briefly in this section were abundantly discussed in the literature, both regarding their descriptive usages and also regarding their numerous interpretative usages. My aim is not to provide an exhaustive presentation: for the purposes of this thesis, I will briefly and selectively recall some of the main approaches discussing the meaning and the usages of the PS, IMP and PC. A brief section will be dedicated to the PRES, where the focus will be on cases where the PRES expresses present or past time reference. The distinction *descriptive* vs. *interpretative* usages will be used to make reference to usages where the verbal tenses considered express reference to the time intuitively expected (i.e. past time for the PC, PS and IMP and present time for the PRES) contrasted to usages where they express reference to other times than those intuitively expected (i.e. present or future time for the PC, PS and IMP and past or future time for the PRES). The *descriptive* vs. *interpretative* distinction finds its roots in Relevance Theory (Sperber and Wilson 1986/1995, 228-229[46]). In these pages, however, I do not adopt their technical definition of the two terms, but use the *interpretative* notion to describe cases where the hearer is brought to interpret an unexpected usage of a verbal tense.

---

[46] In their words. 'any representation with a propositional form, and in particular any utterance, can be used to represent things in two ways. It can represent some state of affairs in virtue of its propositional form being true of that state of affairs ; in this case the representation is used descriptively. Or it can represent some other representation which also has a propositional form – a thought for instance- in virtue of a resemblance between the two propositional forms ; in this case the first representation is used interpretatively'.

2.5.1.2.1 Passé Simple

Scholars have described the PS within several approaches known as the *classical* approach, the *aspectual* approach, the *anaphoric* approach, the *textual* approach and the *pragmatic* approach[47]. In the following lines, I will briefly discuss the first four them and pay more attention to the last one.

Firstly, within the classical approach, the PS expresses a past event completed in the past with no connection to present time (Grevisse 1980, 873; Wagner and Pinchon 1962, 413). The focus on the accomplishment of the event in the past is the feature that distinguishes the PS from the PC, the second one expressing a link to the speaker's or a third person's present time (see below in section 2.5.1.2.3). Scholars have argued that the PS provides an objective interpretation of the situation described. The PS is also distinguished from the IMP, which presents a past situation as not accomplished (except for the narrative IMP, see below in section 2.5.1.2.2).

Secondly, the aspectual approach is mainly developed based on the classic approach and insists on the aspectual and not temporal distinction between the PS, and the IMP (Martin 1971, 93-94). This approach assumes the perfective aspect of the PS providing a global view of the event and the imperfective aspect of the IMP as offering an interior view of the event in progress[48].

Thirdly, in the anaphoric approach, Kamp and Rohrer's (1983) main argument is that the interpretation of verbal tenses depends on the temporal relations that they establish between discourse segments. They argue that the PS is used in contexts where time progresses and events are temporally ordered, as in (193). They base their analysis on the three coordinates proposed by Reichenbach (S, R and E), pointing out that sentences with a PS introduce a new R moment in the discourse that is prior to the event moment E, while sentences with an IMP adopt the existing R (introduced by the precedent sentence with a PS), as in (194). This description has numerous exceptions, as Kamp and Rohrer themselves and other scholars pointed out, as in examples (195) and (196).

(193) Pierre entra. Marie t*éléphona.*
'Peter entered. Mary *made* a phone call.'
(194) Pierre entra. Marie *téléphonait.*
'Peter entered. Mary *was calling*.'
(195) Marie *chanta* et Pierre l'*accompagna* au piano.
'Mary sung and Peter accompanied her at the piano'
(196) L'été de cette année-là *vit* plusieurs changements dans la vie de nos héros. François *épousa* Adèle, Jean-Louis *partit* pour le Brésil et Paul s'*achet*a une maison à la campagne.
'The summer of that year saw several changes in our heroes' lives. François married Adele, Jean-Louis left to Brazil and Paul bought a house in the countryside.'

---

[47] This section is based on Tahara (2000, 2004) who provides a detailed presentation of the various approaches of the PS in FR. For other discussions see also Vetters (1996). For a pragmatic account, see Saussure (1998, 2003).

[48] There have been several attempts to question the perfective aspect of the PS and the imperfective aspect of the IMP such as Guenthner, Hoepelman and Rohrer (1978) and respectively, Molendijk (1990). For counter-arguments, see Vetters (1996).

Fourthly, the textual approach considers verbal tenses with respect to their function in a discourse. For Benveniste (1966), the difference between the PS and the PC is a stylistical difference: the PS is used in the written language of *histoire* 'history' whereas the PC (as well as all other verbal tenses) occurs in the *discours* 'discourse' (written or spoken)[49] register. Weinrich (1973) discusses the opposition between PS and IMP in the *histoire* (that he calls *monde raconté* 'narrated world') register in terms of emphasis and background information. Explicitly, the PS is used to express emphasized information whereas the IMP is used uniquely for background information.

Finally, pragmatic approaches aim at integrating semantic and pragmatic principles in the analysis of the PS and its discourse function. Pragmatic approaches make use of the assumptions suggested by Grice (1967) and developed in RT (Sperber and Wilson 1986/1995; Wilson and Sperber 1993) about language comprehension (see sections 3.1.2 and 3.1.3). Briefly, language is underdetermined and it must be contextually worked out through inferences. The comprehension process is guided by the communicative and cognitive principles of relevance.

Among the pragmatic approaches, one can distinguish between two trends. According to the first one, verbal tenses have inferential descriptive and interpretative usages computed based on the instructions encoded by a verbal tense and contextual information (Moeschler et al. 1998; Luscher and Sthioul 1996; Luscher 1998, Sthioul 1998; Saussure and Sthioul 1999; Tahara 2000; Saussure 2003; Saussure and Sthioul 2005). Descriptive usages of the PS are described in terms of a basic semantic description using Reichenbachian coordinates E, R and S, or more precisely, E=R<S. This description corresponds to the procedural information encoded by the PS, namely to locate the eventuality previously to S passing through an R which is simultaneous with E. The temporal location of an eventuality must therefore be calculated contextually, and this is an inferential process. Following Saussure (2003), I will from now on call this trend the *procedural pragmatics* approach. In the following lines, I will describe briefly the second pragmatic trend, and I will come back to discussing the PS from in the procedural pragmatics approach.

The second trend aims at reducing the role of pragmatic (non-linguistic and cognitive) factors for determining the meaning of a verbal tense, and therefore at increasing the semantic input (Kleiber and Riegel 1989, 1991; Kleiber 1994; Vetters 1996). Vetters (1996) speaks about a *pragma-semantic* approach. He argues that the PS/IMP opposition can be explained in the pragma-semantic approach using a model with three levels (1996, 142):

- The opposition perfective/imperfective is semantic, therefore descriptive and truth-conditional;
- The rules for temporal interpretation identified by Kamp and Rohrer (1983) are instructions encoded by the two verbal tenses;
- The communicative principle of relevance guides the pragmatic interpretation of the sentences.

As far as the analysis of the PS in these terms, Vetters argues that it is aspectually *non imperfective* (be it perfective or inchoative, as suggested by Guenthner, Hoepelman and

---

[49] For Benveniste, the main difference between *discours* and *histoire* is that discours refers to a situation of communication involving a speaker and a hearer who interact.

Rohrer, 1978) and that it encodes instructions for forward temporal sequencing (called *linearity* by Vetters). He suggests that cases where the PS is used without a forward temporal sequencing, as in examples (195) and (196) above, are interpreted following the principle of optimal relevance. In his words, "the PS expresses temporal progression except when it is used in contexts where the linear interpretation would be more costly than a non-linear interpretation" (1996, 150), the higher cognitive cost being due to world knowledge. For example, in sentence (196), all the events are temporally located during *the summer of that year* and their order is not specified. The hearer assumes that the speaker does not intend a sequential interpretation and that the utterance is worth processing despite the lack of a temporal specification of the order.

In the procedural pragmatics approach, Saussure (2003, 222) argues that the PS encodes by default an instruction for temporal progression. This instruction is blocked when the hearer does not have sufficient contextual information for interpreting the utterance, as in (197), and in cases of temporal encapsulation, as in (198) and (199) from Saussure (1998, 249).

(197) François *épousa* Adèle. Paul s'*achet*a une maison à la campagne.
'François married Adele. Paul bought a house in the countryside.'

(198) Une terrible tempête *fit* rage. Quelques tuiles *tombèrent*. Un arbre du jardin *fut arraché*.
'A terrible storm raged. Some tiles fell. A tree was torn from the garden.'

(199) Ce samedi *marqua* le début de la relation de Paul et Marie. Ils *déjeunèrent* ensemble. Ils se *promenèrent* sur les berges. Le soir, ils *s'embrassèrent* pour la première fois.
'That Saturday marked the beginning of Paul and Mary's relation. They had lunch together. They went for a walk on the riverbank. In the evening, they kissed for the first time.'

The PS may occur in contexts with backward temporal sequencing, but only accompanied by an appropriate connective, such as *dès que* 'as soon as', as shown in examples (200) and (201) from Saussure (2003, 223). Without the connective, the PS imposes temporal progression (i.e. the convicted fainted before the reading of the sentence).

(200) Le condamné s'*évanouit* dès que le juge *lut* la sentence.
'The convicted fainted as soon as the judge read the sentence.'

(201) Le condamné s'*évanouit*. Le juge *lut* la sentence.
'The convicted fainted. The judge read the sentence.'

Consequently, Saussure proposes two possible descriptions of the semantics and the pragmatics of the PS (Saussure 2003, 228), which are interpretative procedures[50]. Explicitly, in version 2, as in the left side of Figure 2-3, there are *two semantic procedures* specific to the PS (i.e. to locate E before S via a R simultaneous to E, and to increment R if possible, marking temporal progression) and *two pragmatic procedures* (i.e. if there is a connective or a conceptual relation requiring backward temporal progression, allow it). However, hearers apply the default procedures unless they are blocked by the contextual information regarding the connectives and conceptual rules. Hence, in version 2 as in the right side of Figure 2-3, based

---

[50] For the exact algorithm to follow, see Saussure (2003, 228).

on the argument that the value of R must be computed contextually, this step is independent of the semantic procedure and is included in the pragmatic interpretation, which therefore includes three steps.

Figure 2-3 Interpretation of the PS: Version 1 and 2
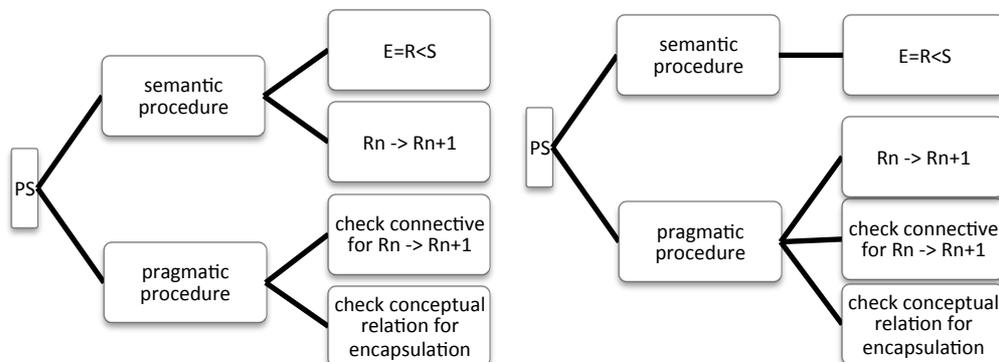
In the same procedural pragmatics approach, Sthioul (1998), Tahara (2000) and Saussure (2003) discuss *descriptive* and *interpretative* usages of the PS. In RT, utterances point to world representations, i.e. they represent hypotheses, thoughts, beliefs, etc. about the world (or the fictional world in novels). In this case, utterances are used descriptively. There are also cases when an utterance is used to represent the thought or belief of third party different than the speaker's at the moment of speech S. In this case, utterances are used subjectively (Saussure 2003, 130). As far as verbal tenses are concerned, Saussure argues that semantic and pragmatic temporal procedures combined with contextual assumptions may trigger interpretative usages. According to Tahara (2000), the PS has descriptive and interpretative usages that present the features provided in Table 2-8.

Table 2-8 Descriptive and interpretative usages of the PS

Descriptive usages of PS	Interpretative usages of PS
- Perfective $E=R<S$ $R_n \rightarrow R_{n+1}$ - Neutral perspective - Emphasized information	- Inchoative or Perfective $E=R<S$ $R_n \rightarrow R_{n-1}$ or R_{n+1} - Subjective perspective - Emphasized information

Descriptive usages correspond to the classical description of the PS. As for the interpretative usages, the PS can be inchoative, as in (202) and in (203) from Sthioul (1998, 217 and 218). Interpretative PS can also be perfective instructing for backward temporal sequencing as in example (204) borrowed from Vuillaume (1990, 10). In all these examples, the PS presents the situation from a subjective perspective identified by the hearer based on contextual assumptions (i.e. the moment when Paul perceives the cold in the first example and sees the monster in the second example, and the moment signalled by the temporal deictic *today* corresponding to the character's and not the speaker's *today* in the third example).

(202)   Paul *sortit*. Dehors, il *fit* bigrement froid.

'Paul got out. Outside, it was fantastically cold.'

(203) Paul leva les yeux. Un monstre *se tint* devant lui.
'Paul looked up. A monster was standing in front of him.'

(204) Le malheur diminue l'esprit. Notre héros eut la gaucherie de s'arrêter auprès de cette petite chaise de paille, qui jadis avait été le témoin de triomphes si brillants. Aujourd'hui personne ne lui *adressa* la parole; sa présence était comme inaperçue et pire encore. (Stendhal H. de., *Le rouge et le noir*).
'Misfortune lessens the spirit. Our hero had the clumsiness to stop next to this small straw chair, which was long ago the witness of such brilliant triumphs. Today nobody talked to him, his presence was as if unnoticed and even worse.'

To sum up, according to the procedural pragmatic approach, the PS encodes a general interpretative procedure consisting of semantic and pragmatic procedures. The interpretation process is guided by the communicative principle of relevance and the hearer uses linguistic and non-linguistic information in the inferential process.

2.5.1.2.2 Imparfait

The IMP has a *temporal* value when the process refers to past time (realis) and a *modal* value when the situation is projected in irrealis (Riegel et al. 1994). In this thesis, I am interested in the temporal interpretation of the IMP.

As noted in section 2.5.1.2.1, traditionally scholars often described the IMP in opposition to the PS as being a tense of background information (Weinrich 1973), aspectually unaccomplished and imperfective, which needs a hosting event previously presented (Guillemin-Flescher 1981) as shown in example (205). Most of the FR scholars agree that the IMP is an anaphoric tense (Ducrot 1979; Kamp and Rohrer 1983; Tasmowski-De Ryck 1985; Molendijk 1990; Kleiber 2003; Berthonneau and Kleiber 1993, 1994; Vetters 1996) that must be related to an existing situation.

(205) Pierre entra. Marie *téléphonait*.
Peter enter.PS. Mary call.IMP
'Peter entered. Mary *was calling*.'

These features situate the IMP in opposition with the PS, which marks a breaking between the moment of speaking S and the global image of the situation happening before S. The IMP gives an interior perspective on the situation, which allows the distinction between what has effectively happened and what has not happened yet. Martin (1971, 70) argued that the IMP puts in opposition at a certain moment '*la partie accomplie du processus avec la partie inaccomplie*' ('the accomplished part of the process with the unaccomplished one') as in (206). The PS, on the contrary, as discussed in section 2.5.1.2.1, considers the situation globally without analysing inherent parties, though a temporal complement can mark the beginning or the end of the situation, as in the following examples from Riegel et al. (1994):

(206) Coupeau eut un accident. Il *sortait* du village.
Coupeau have.PS an accident. He get out.IMP of the village
'Coupeau had an accident. He was getting out of the village.'

(207) Après son accident, Coupeau *se mit* à boire.

After his accident, Coupeau start.PS to drink

'After his accident, Coupeau started to drink'

(208) Gervaise *attendit* le retour de Lantier jusqu'à l'aube.

Gervaise wait.PS come back Lantier until daybreak

'Gervaise waited for Lantier to come back until daybreak'

Temporal reference in an utterance containing an IMP is generally calculated taking into account three observations (Sthioul 1998, 207). Firstly, temporal reference is fixed in relation to an existing reference period R. Accordingly, a sentence containing an IMP cannot be interpreted in isolation, as in (209). The anchoring reference period can be provided by a temporal adverbial, as in (210) or another event, as in (211).

(209) ?Marie *buvait* un café.

Mary drink.IMP a coffee

'Mary was drinking a coffee.'

(210) Hier à huit heures, Marie *buvait* un café.

Yesterday at o'clock Mary Mary drink.IMP a coffee

'Yesterday, at eight o'clock Mary was drinking a coffee. '

(211) Paul entra. Marie *buvait* un café.

Paul enter.PS. Mary drink.IMP a coffee

'Paul entered. Mary was drinking a coffee. '

Secondly, the reference period is prior to S, as shown by the compatibility of a past temporal adverbial in example (212), and the incompatibility with a present time adverbial in (213) or future time adverbial in (214).

(212) Il y a une heure, Paul *lisait* le journal, et ça n'est pas prêt de changer.

An hour ago, Paul read.IMP the newspaper, and this is not going to change soon.

'An hour ago, Paul was reading the newspaper, and this is not going to change soon.'

(213) *Au moment où je vous parle, Paul *lisait* le journal.

*At the moment, Paul read.IMP the newspaper

(214) * Dans une heure, Paul *lisait* le journal.

* In an hour, Paul read.IMP the newspaper

Thirdly, the period when E holds is larger than the reference period, as in example (212) where it continues up to present, and it cannot be smaller than the reference period, as in example (215) from Ducrot (1979). On the contrary, this is possible with the PS or the PC as shown in example (216). The IMP therefore presents the situation as unbounded (R included in E) and locates it prior to S. The consequences of R being included in E are the impossibility to have achievement implicatures even for telic situations, as in (217), and the interpretation that the event expressed with the IMP includes the event expressed with the PS or PC, as in (205) or (211).

(215) L'année dernière, Paul *habitait* à Paris (*mais seulement en mai).

Last year, Paul live.IMP in Paris (*but only in May)

'Last year, Paul was living in Paris (*but only in May).'

(216) L'année dernière, Paul *habita/a habité* à Paris, mais seulement en mai.

Last year, Paul live.PS/PC in Paris (but only in May)

'Last year, Paul lived in Paris (but only in May).'

(217) Pendant la reunion, Marie *buvait* un café, qu'elle n'a d'ailleurs jamais fini.
During the meeting, Mary drink.IMP a coffee, which by the way, she never finish.PC
'During the meeting, Mary was drinking a coffee, which by the way, she has never finished.'

Saussure and Sthiul (2005, 105) summarise these observations by suggesting that the basic semantic features that can be attributed to the IMP regardless of its discoursive context are the dislocation of the referential anchoring of S and the inclusion of this reference point within the eventuality denoted by the verb.

These observations correspond to the descriptive usages of the IMP. Probably the most known and discussed exception to the description of the IMP in these terms is the so-called *breaking* or *narrative* IMP[51] (Tasmowski-De Ryck 1985; Vetters 1996, Comrie 1976) that has features completely opposed to the first type of IMP, as illustrated in example (218) and (219).

(218) Comme elle avait été à l'opéra, une nuit d'hiver, elle rentra toute frissonnante de froid. Le lendemain elle toussait. Huit jours plus tard elle *mourait* d'une fluxion de poitrine.
Since she go.PQP to the opera, one winter evening, she come.PS back all shivering. The day after, she cough.IMP. Eight days later, she die.IMP of phthisis
'Since she had gone to the opera, one winter evening, she came back all shivering. The day after, she was coughing. Eight days later, she died of phthisis.'

(219) Le lendemain, il *partait*.
The next day, he leave.IMP
'The next day, he left.'

Therefore, scholars investigating the IMP were constrained to suggest a model that would explain both the existence of the imperfective and that of the narrative IMP[52]. In the pragma-semantic approach, Vetters (1996, 142) argues that the IMP is in opposition to the PS regarding aspectual information and the instructions for temporal sequencing (as I noted in section 2.5.1.2.1) where the IMP is imperfective and instructs the hearer to relate the situation to another past situation with a meronymic[53] relation. He does not include simultaneity in the procedural meaning of the IMP since the narrative IMP does not express it. The interpretative process is finalized at the pragmatic level under the guidance of the principle of optimal relevance. As for the narrative IMP, it is characterized as it follows (Vetters 1996, 128):

- It instructs for temporal progression
- It can be replaced by the PS
- It is promoted by an anteposed temporal adverbial

---

[51] Vetters (1996, 128) points out that the narrative IMP was identified for other languages as well, such as in most of the Romance languages, in EN (Klum 1961,190) and in ancient Greek (Kiparsky 1968,40).

[52] In a different framework, Moledijk (2002) reduces the semantics of the PC, PS and IMP to a series of logical-temporal relations allowed by these verbal tenses in a sequence of sentences, the semantics of the IMP being the relation of temporal simultaneity.

[53] Bethonneau and Kleiber (1993, 73) argue that the relation between a situation expressed with the IMP and another past time situation is similar to associative anaphora, where a part is linked to the whole.

- With states, it expresses an inchoative meaning (as in (220))

    (220)   Quelques secondes plus tard, Luc *était* sous le chapiteau.
            A few seconds later, Luc be.IMP under the tent
            'A few seconds later, Luc was under the tent.'

As for the aspectual value of the narrative IMP, there are two cases. The impossibility to have the interpretation of a unique and entire event in (221), which is possible with the PS as in (222), shows that the IMP remains imperfective. He second possibility is that it can remain undetermined regarding aspectual information due to the Principle of optimal relevance. In other words, the hearer can correctly interpret an utterance with a narrative IMP without determining its aspectual value, and this interpretation produces sufficient cognitive effects.

    (221)   Le lendemain, il *travaillait* de 5h à 8h.
            The next day, he work.IMP from 5 to 8am
            'The next day, he was working from 5 to 8am.'
    (222)   Le lendemain, il travailla de 5h à 8h.
            The next day, he work.PS from 5 to 8am
            'The next day, he worked from 5 to 8am.

According to Vetters, the pragma-semantic approach explains both descriptive and interpretative usages of the IMP without being obliged to make use of a split analysis between the two types of IMP. For example, in all its usages the narrative IMPs in (223)[54] needs a reference situation recoverable from the context (i.e. the IMP is acceptable with difficulty unless there is a logical link between the two events, as shown in (224)).

    (223)   Le commandant se jeta sur l'interphone et hurla qu'il *avait* à parler à M. Chisnutt. Trois minutes plus tard, M. Chisnutt *se présentait* chez le commandant.
            'The commandant threw himself of the intercom and screamed that he had to speak to Mr Chisnutt. Three minutes later, Mr Chisnutt showed up to see the commandant.'
    (224)   * ?Le commandant se jeta sur l'interphone et hurla qu'il avait à parler à M. Chisnutt. Trois minutes plus tard, M. Brown se présentait chez le commandant.
            'The commandant threw himself of the intercom and screamed that he had to speak to Mr Chisnutt. Three minutes later, Mr Brown showed up to see the commandant.'

In the procedural pragmatics approach, Saussure (2003) (also Saussure and Sthioul 2005) systemised these observations and proposed a general procedure of the interpretation of the IMP. Based on previous work (Saussure and Sthioul 1999), he argued that the IMP instructs the hearer to build an *unsaturated P variable interior to the event*, which will be saturated contextually either by the reference moment R (corresponding to descriptive usages of the IMP) or by a moment of consciousness C (corresponding to interpretative usages of the IMP). It is thus in the process of assignation of temporal reference that the hearer builds a subjective perspective on the situation.

As far as the narrative IMP is concerned, he suggested that it occurs when the hearer infers based on contextual information either achievement implicatures (blocked in the

---

[54] From Vetters (1996, 144) inspired from Tasmowski-De Ryck (1985, 66)

descriptive usages of the IMP) or forward, as in (225), and backward, as in (226), temporal sequencing. He emphasizes that the narrative IMP is not interchangeable with the PS because it provides a view on the event from the interior, whereas the PS instructs to view the process as a whole. In (225), the adverb *déjà* ('already') suggests the speaker's subjective perception of the situation from the interior and it occurs with the narrative IMP. The same utterance is not acceptable with the PS as in (227), which imposes a view from the exterior.

(225) Le train quitta Londres. Une heure plus tard, il *entrait* déjà en gare de Birmingham. (Sthioul 1998, 213)
The train left London. One hour later, it enter.IMP already in Birmingham station.'
'The train left London. One hour later, it was already entering in Birmingham station.'

(226) Judith ne reconnut pas le "joyeux colporteur" qui la *quittait* quelques semaines plutôt. Klum (1961, 258)
Judith did not recognize the "happy peddle" who break up.IMP with her three weeks before.
'Judith did not recognize the "happy peddle" who broke up with her three weeks before.'

(227) Le train quitta Londres. Une heure plus tard, il *entra* ?déjà en gare de Birmingham.
'The train left London. One hour later, it enter.PS already in Birmingham station.'

To sum up, the IMP encodes a general interpretative procedure consisting of the instruction to build an unsaturated variable P included in E, which will be saturated contextually either by R or by a C moment. As for temporal sequencing, the IMP does not provide a directional instruction (in its descriptive usage). Its interpretation as forward or backward temporal sequencing instructions (in its interpretative usages) is due the pressure of contextual information.

2.5.1.2.3 Passé Composé

In a cross-linguistic typological analysis, Squartini and Bertinetto (2000) investigate the usage of the PC and the SP in Romance languages. The main hypothesis for explaining the variation in usages across Romance languages is the process of *aoristicization*. According to Harris (1982), the aoristicization process consists of a change from a purely perfect (the PresPerf in EN) to an aoristic, passing through several steps, of which the third corresponds to what is known in the French literature as the *accomplishment PC* and the fourth to the *anteriority PC*.

- The PC is restricted to present states resulting from past actions, and is not used to describe past actions themselves, however recent;
- The PC occurs in durative or repetitive contexts (similar to the EN PresPerf and the Present Perfect Continuous);
- The PC expresses the archetypal PresPerf value of past action with present relevance
- The PC expresses the aoristic function, while the PS is restricted to formal registers

Squartini and Bertinetto (2000) argue against distinct steps in the aoristic drift and for a continuum from perfect to aorist. Romance languages and dialects would then be situated on this continuum as in Figure 2-4. Portuguese is the only language that presents an opposite pattern, as the PC is less used than the PS for expressing past time reference. In all other

languages and vernaculars, the PC is more frequent than the PS, a scalar orientation that is at a maximum in northern IT and FR vernaculars.

Figure 2-4 Scalar orientation of Romance languages in the aoristicization process

Portuguese	Spanish	Occitan Catalan	Standard IT	Standard FR	Standard RO	Northern IT and FR vernaculars
perfectal--------------->----------------------->------------------------->-------------------------aoristic						

The third and the fourth step were identified in the FR literature as the accomplishment and the anteriority usages of the PC. Squartini and Bertinetto argue, as I will show later on, that in IT there is an important difference between central and northern parts of Italy where the PC is used more frequently than the PS, and in the southern part of the country where the situation is reversed. In RO on the other hand, the PC is more advanced in the aoristic drift, being the most frequent tense used to express past time reference. In this section, I will describe traditional and pragmatic approaches to the FR PC.

Traditionally, the PC is described as a "tense with two faces" (Martin 1971) because of the possibility to express both past and present time. When describing the PC, scholars suggested monoguist and ambiguist analyses. Monoguist analyses consist of a focus either on the past time reference (i.e. the anteriority PC as in (228) and (229) such as Brunot, 1922), on the present time reference (i.e. the accomplishment PC as in (229)-(232) such as Guillaume 1929) or on both usages unified and undistinguished (e.g. Reichenbach 1947). Anteriority PC provides information about E being prior to S, whereas the accomplishment PC allows achievement inferences about a resultative state relevant at the moment of speech S. As for the third type of analysis, Reichenbach assumes a one-to-one correspondence between the PresPerf and the PC, which are both characterized by the concomitance between R and S (i.e. E<R=S). As Luscher and Sthioul (1996, 198) point out, however, Reichenbach's analysis is problematic for examples (233) and (235) where the PC is translated through a PresPerf. They are perfectly acceptable in (234) and (236), where the PC is translated through an SP.

(228) Une fois, j'*ai conduit* sans le permis de conduire.
         'Once, I drove without driving licence.'
(229) Victor Hugo a écrit *Les Misérables.* (Luscher and Sthioul 1996, 206)
         'Victor Hugo wrote Les Misérables.'
(230) Policier: Votre permis de conduire, s'il vous plait?
         Chauffeur : Je l'*ai oublié* à la maison.
         'Policeman: You driving license, please?
         Driver: I forgot/ have forgotten it at home.'
(231) Isabelle *est sortie.* (Saussure 2003, 232)
         'Isabelle has gone out.'
(232) Il *a plu.* (de Saussure 2003, 232)
         'It has rained.'
(233) Hier, il *a plu.* (Luscher and Sthioul 1996, 199)
         *'Yesterday, it has rained.'
(234) Yesterday, it rained.

(235) Le 21 janvier 1976, le Concorde *a atteri* à Rio. (Luscher and Sthioul 1996, 199)
    *'On the 21 of January 1976, the Concorde has landed in Rio.'
(236) On the 21 of January 1976, the Concorde *landed* in Rio.

Riegel et al. (1994, 301-303) point out that the accomplishment PC in (237) is opposed to PRES, which expresses a situation in the moment of speech, as in (238). In this case, the PC marks the resultative state, especially with perfective verbs conjugated with *être*, as in (239). In a structure where the PC is followed by the present tense, it marks the precedence of the former and has a habitual meaning, as in (240) where *quand il a déjeuné* ('when he has had lunch') means *après avoir déjeuné* ('after having had his lunch').

(237) Nous *avons emporté* de quoi faire à manger.
    We bring.PC what to cook.
    'We have brought what to cook.'
(238) Nous *emportons* de quoi faire à manger.
    We bring.PRES what to cook.
     'We are bringing what to cook.'
(239) Il *est parti*.
    He go.PC
    'He is gone.'
(240) Quand il *a déjeuné*, César fait la sieste.
    When he have lunch.PC, Cesar take a nap.PRES
    'When he has had lunch, Cesar takes a nap.'

Ambiguist analyses, on the contrary, pleaded for the existence of an ambiguous PC where only contextual information can disambiguate among its possible interpretations (such as Vet 1980, Luscher and Sthioul 1996 among others). Vet (1980) suggested describing the PC with two reference points: a main reference point expressing simultaneity to S and an auxiliary reference point expressing anteriority to S. His second suggestion is that the analysis depends on the lexical aspect of the situation: transitional (i.e. telic) vs. non-transitional (i.e. atelic) situations. Precisely, telic situations allow an anteriority interpretation of the PC accompanied by past time adverbial as in example (241) and an accomplishment interpretation with a present time adverbial as in (242). Atelic situations allow only accomplishment interpretations, as in (243), where the PC is incompatible with a present time adverbial (from Luscher and Sthioul's 1996 discussion of Vet's analysis).

(241) Hier, Chantal *est sortie.*
    Yesterday, Chantal go.PC out.
    'Yesterday, Chantal went out.'
(242) En ce moment, Chantal *est sortie.*
    Today, Chantal be.PRES out
    'Today, Chantal is out.'
(243) *L'enfant *a* maintenant *pleuré.*
    The child cry.PC now
    'The child has just cried.'

In the procedural pragmatics approach, Luscher and Sthioul (1996) argue that Vet's

analysis in terms of "two semantics of the PC" (1996, 202) is not convincing and suggest a pragmatic analysis consisting of a unique semantic content or a *base value* and two pragmatic or *contextual* usages. The base value consists of the event moment E prior to S as in (229). In its base value, the PC shares semantic information (i.e. reference to past time: E<S with the other simple and composed tenses, such as PS, IMP, and PQP. As for the two pragmatic usages, the distinction is given by the position of the reference moment R: in the first usage, the PC points to the event that took place in the past (R=E), as in (233), whereas in the second usage, the PC points to the resultative state relevant in the present time (R=S) as in (242).

Luscher and Sthioul (1996) propose a complex procedure for interpreting the PC that consists of a general procedure for recuperating R that is applicable to all tenses, as shown in the left panel of the Figure 2-5, and two sub-procedures specific to the PC as shown in the right panel. The temporal interpretation of a sentence involves determining the R point from the context, such as a temporal adverbial from the current sentence or from the previous sentences, or through inference based on the temporal interpretation of previous sentences and world knowledge. The procedure of recuperating R is the same for all tenses, where E can be previous, simultaneous or posterior to R, or where R is previous, simultaneous or posterior to S. The first sub-procedure for interpreting the PC instructs the hearer to instantiate a P moment so that E<P<S. The second sub-procedure for interpreting the PC instructs the hearer to recuperate a resultative state, either lexically (i.e. for example *get out* entails *be out*) or through inference (i.e. *having eaten* implicates *not be hungry*). Saussure (2003) argues that the resultative state is a product of conceptual relations holding between eventualities. The main idea is that accomplishment usages of the PC communicate that the event is perceived from S and that the same event produced a resultative state true at S.

Figure 2-5 Preliminary sub-procedures for interpreting the PC
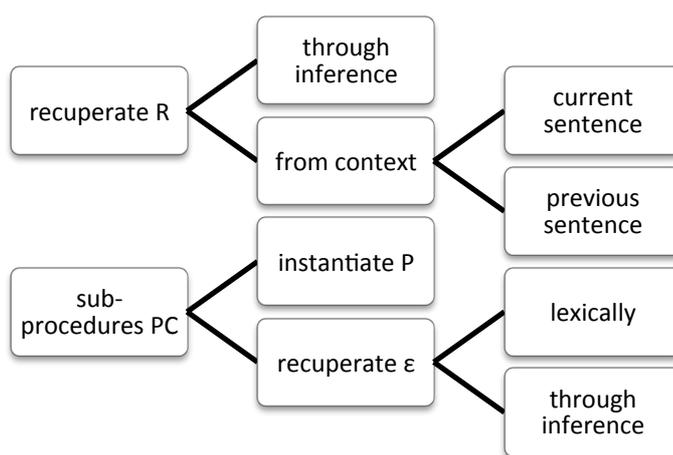
The complete procedure for interpreting the PC proposed by Luscher and Sthioul (1996) (reasserted in Luscher 1998) is provided in Figure 2-6. The PC presents a basic semantics according to which E is previous to S. The hearer is instructed to instantiate a P so that E<P<S. P is saturated based on contextual information through pragmatic inferences. It can be saturated as a reference moment R that is either simultaneous to E or simultaneous to S. The former case corresponds to the anteriority usage, whereas the latter to the

accomplishment usage. As far as the latter usage is concerned, the hearer can further infer either a right bounded (e.g. *be out* in example (242)) or a right unbounded (e.g. *the novel that was written* as in the second interpretation of (229)) resultative state true at S.

Figure 2-6 Final procedure for interpreting the PC
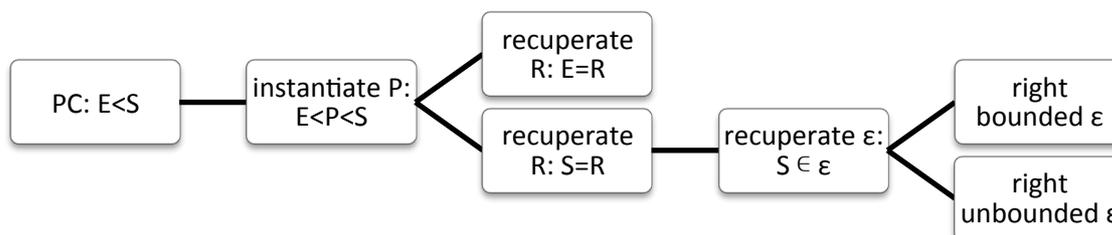
Saussure (2003) argues that the PC, contrary to the PS, does not provide information regarding temporal progression and that both backward and forward temporal sequencing inferences are possible, as in (244) where time progresses from the first to the second event and regresses from the second to the third. The adverb *en plus* 'besides' illustrates that there is no temporal order imposed by the PC. He also points out that the PC can refer to future time when it is accompanied by a future temporal adverbial as in (245). This is an interpretative usage of the PC, where it does not refer to a fact but to a thought: the speaker imagines herself at a moment S' (i.e. that is two months after S), when she can assert *I finished my thesis* (S<E<S').

(244)   Le concierge *a sorti* sa clef, il *a quitté* les lieux, et en plus il *a fermé* la porte.
        'The concierge took out his key, he left and besides he locked the door.'
(245)   Dans deux mois j'*ai fini* ma thèse.
        In two months, I finish.PC my thesis.
        'In two months, I will have finished my thesis.'

To sum up, the PC is a verbal tense locating E<S via an R which can be R=E or R=S. In the usage E=R<S, the PC is similar but not identical to the PS and the IMP. It can be distinguished of the PS based on the instructions and constraints on temporal progression encoded by the PS. Similarly, the PC can be distinguished of the IMP based on the difference of viewpoint, i.e. perfective for the PC and imperfective for the IMP.

2.5.1.2.4   Présent

The PRES is opposed to the PS, PC and IMP based on the relation E/S. The PRES expresses a relation E=S whereas the PS, PC and IMP express a relation E<S. The description of the semantics of these verbal tenses in these terms applies to what has been called *descriptive* usages. Similarly to the PS, PC and IMP, the PRES can refer to other times than present time in its interpretative usages. In this section, I will not provide an exhaustive presentation of all the usages of the PRES. I am particularly interested in the cases when the PRES expresses reference to past time, a usage called *historical present* (HP).

Traditionally, the PRES is used to express eventualities that take place in the moment of speaking, for habitual and timeless statements (general truths as maxims, proverbs, and theorems). Riegel et al. (1994) point out that the PRES can place the situation in any epoch,

past or future, and even in all epochs (omnitemporal value). As a simple form, the PRES expresses the process while ongoing, without taking in consideration it's delimitations or duration. The limits and the duration are given by the semantics of the verb as in (246) with a punctual situation, and in (247), with an activity. The omnitemporal value, i.e. the *permanent present*, occurs in definitions, as in (248), in general truths (considered by the speaker to be true at any epoch), as in (249), and in proverbs or maxims, as in (250).

(246)   La bombe *explose*.
         'The bomb blows up.'
(247)   Elle *marche* au milieu de la rue.
         'She walks in the middle of the street.'
(248)   Une haquenée *est* un petit cheval ou une jument, de taille moyenne.
         'A hackney is a small horse or a mare, of a middle size.'
(249)   Le soleil *se lève* à l'Est.
         'The sun rises in the East.'
(250)   Tous les matins du monde *sont* sans retour. (Riegel et al. 2002, 300)
         'All mornings from the whole world never come back.'

A PRES utterance may also express the iterative aspect with an appropriate temporal adverbial, as in the following examples. Sometimes sentences without a temporal adverbial may remain ambiguous between a present action and a habitual activity.

(251)   Elle *regarde* la télévision parfois/souvent /tous les soirs.
         She watch.PRES TV sometimes/often/every evening
         'She watches TV sometimes/often/every evening.'
(252)   Claire *joue* au tennis.
         Claire play.PRES tennis
         'Claire plays tennis.'

The PRES can also express reference to other times than the present, and this occurs in its interpretative usages. For example, it can make reference to past or future times (i.e. R is located before or after S) with the help of a temporal adverbial or based on contextual knowledge. The utterance is related to S but the event is shifted in the past, as in (253), or in the future, as in (254) and (255):

(253)   Je *sors* à l'instant du lycée.
         I get out.PRES from the high school
         'I have just gotten out from the high school.'
(254)   Elle *part* demain pour le Pérou.
         She leave.PRES tomorrow to Peru
         'She leaves tomorrow to Peru.'
(255)   J'*arrive* dans cinq minutes.
         I arrive.PRES there in five minutes
         'I will be arriving in five minutes.'

Temporal adverbials may express a shorter or a longer period of time from the past or the future. They can also mention an initial or a final border of the process. In example (256), the temporal adverbial marks the initial border, while the final one is indefinite. In example

(257), the initial border of the period of time from the past is specified, while the final border remains unspecified. On the contrary, in (258), the process is oriented towards future starting with the initial border marked by *désormais* 'from now on'.

(256) Il *neige* depuis vingt-quatre heures.
It snow.PRES for twenty-four hours
'It has been snowing for the last twenty-four hours.'

(257) Je me *lève* à cinq heures depuis vingt ans.
I wake up.PRES at five o'clock for the last twenty years
'I have been waking up at five o'clock for the last twenty years.'

(258) Désormais, je me *lève* à cinq heures.
From now on, I wake up.PRES at five o'clock
'From now on, I will wake up at five o'clock.'

Another and probably the most investigated interpretative usage is the *historical* or the *narrative* PRES (HP). It is used to make reference to real or fictional past events, in an independent phrase or in a whole paragraph. In contrast to the PRES, which expresses an immediate past with the appropriate temporal adverbials, the HP shifts the event into the past, as in example (259). The HP may alternate with the PS to give a particular vivacity to the narration or to the IMP, where the HP expresses main events, and the IMP expresses secondary events.

(259) En 1789, le peuple de Paris *prend* la Bastille.
In 1789, people from Paris take.PRES the Bastille.
'In 1789, people from Paris took the Bastille.'

From a procedural pragmatics perspective, Luscher (1998) proposes for the PRES a similar procedure as for the verbal tenses expressing past time described in sections 2.5.1.2.1, 2.5.1.2.2 and 2.5.1.2.3. He suggests that the PRES has a unique semantics identified in the descriptive usages, which is preserved in the interpretative usages. Similarly to the PC, when interpreting a HP utterance, the hearer is instructed to instantiate a moment of perspective P so that P=S, as in examples (246) to (252).

In some cases, identifying P=S does not correspond to the situation described, as in (260) where the speaker has already arrived, and in (261), where the speaker has not arrived yet (from Luscher (1998, 203)).

(260) Tu es là depuis longtemps? Non, j'*arrive.*
Are you here for a long time? No, I arrive.PRES
'Are you here for a long time? No, I have just arrived.'

(261) Commencez sans moi, j'*arrive.*
Begin without me, I arrive.PRES
'Begin without me, I am arriving.'

He points out that the hearer's assumption is that the speaker used the PRES and not another possible form (*venir de* corresponding to a recent past and respectively, immediate future) so that her interlocutor could infer a set of specific inferences using the instruction P=S. For interpreting the utterances in (260) and (261), the hearer must build a moment of

conscience S' different than S so that S' is included in E. The hearer must instantiate S' to a moment that produces the largest cognitive effect. The interpretation of (260) is that it corresponds to the speaker's thought at the moment when he was arriving, which occurred in the recent past, whereas for (261), the preferred interpretation is that it corresponds to the hearer's perception when the speaker will be arriving in the next few minutes. The same process occurs in (262), where the hearer builds a posterior moment of conscience S' corresponding to the speaker's thought about E.

(262)   Dans dix ans, je *suis* à la retraite.
        In ten years, I be.PRES retired
        'In ten years, I will be retired.'

As far as the HP is concerned, the interpretative process is similar[55]. Because the hearer is constrained to consider the semantics of the PRES S=P, he interprets the utterance as being the thought of an external observer occurring at a moment of conscience S' given by the temporal adverbial. An alternative analysis is proposed in Moeschler (2014), who suggests that the usages of HP may be characterized through three pragmatic features: [±narrative], [±subjective] and [±explicit]. The HP presents five of the six possible combinations of these features. Hence, he proposes a minimal basic reichenbachian semantics shared by the PRES and the PH combined with different groupings of pragmatic features.

To sum up, the FR PRES has a basic semantics consisting of E=P, where P is an unsaturated variables. In descriptive usages of the PRES, the hearer instantiates P=S, whereas in interpretative usages, the hearer builds a moment of conscience S' different than S and included in E. S' corresponds to the speaker's thought or perception regarding E. Interpretative usages arise due to the incompatibility between the semantics of the verbal tense and the available contextual assumptions[56].

To conclude, I dealt in these pages with traditional and semantico-pragmatic approaches of verbal tenses used for translating the EN SP into FR, namely three verbal tenses expressing past time (PC, PS and IMP) and the PRES. In FR, information regarding Tense and Aspect is expressed in the same verbal form. There are analytic forms (i.e. forms

---

[55] A different approach of the HP is provided by Schlenker (2004). Following Banfield (1982) and Doron (1991), he suggests that the notion of *context of speech* should be splitted in two sub-types: *context of thought* and *context of utterance.* For Schlenker, the context of thought is the point at which the thought originates and it includes a thinker, a time of thought and a world of thought. The context of utterance, on the other hand, is the point at which the thought is expressed and it includes a speaker, a hearer, a time of utterance and a world of utterance. He argues that in free indirect discourse (FID) and in narrations in the HP this distinction is particularly relevant. Schlenker's claim is that (i) in ordinary discourses, the context of utterance and the context of thought are identical and correspond to the actual context of speech; (ii) in FID, the context of utterance and the context of thought are different, the actual context being the context of utterance and (iii) in sequences in narrative PRES, the actual context is the context of thought and the context of utterance is presented as having its time coordinate in the past. Moreover, he argues that tenses and pronouns depend of the context of utterance while other indexicals depend on the context of thought. Tenses and pronouns are variables whose domains of reference are determined by grammatical features they carry, such as gender, person and tense.

[56] In a more recent pragmatic approach, Saussure (2013) argues that the so-called interpretative usages of verbal tenses in FR can be accounted for through general pragmatic principles: one of the components of the temporal representation (the deictic point, the reference point, or the eventuality moment) is modified under the pressure of contextual consistency or relevance.

consisting of an auxiliary followed by the main verb and its flexional endings) and synthetic forms (i.e. forms consisting only of the main verb and its flexional endings). Tense expresses reference to past, present and future times, in other words, it locates an eventuality prior, simultaneous and posterior to the moment of speech S. The aspectual distinction perfective/imperfective is expressed grammatically only for past[57]. In this thesis, I will be arguing that Tense and its expression by verbal tenses in FR (similarly to IT and RO) encode conceptual information (i.e. the relation E/S) and procedural information (the relation E/R). The four verbal tenses described in this section can coarsely be classified with respect to their conceptual information: the PC, PS and the IMP on the one hand, and the PRES on the other hand.

In sections 2.5.1.3 and 2.5.1.4, I will describe IT and RO verbal tenses used for translating the SP in these languages. I will conclude this chapter with section 2.5.2, where I will argue that a common framework for these four languages is necessary in order to have a language-independent comprehension of the functioning of temporal reference in discourse.

2.5.1.3   Italian

As mentioned at the beginning of section 2.5.1.2, IT verbs mark Tense and Aspect synthetically and analytically. IT contains regular and irregular verbs. Regular verbs are classified in three main classes according to their endings[58] as shown in Table 2-9. Irregular verbs present irregularities most often in the PS and the past participle forms. These irregularities consist of an alternation between three weak and three strong (i.e. form with stress on the stem) forms for the PS, such as the verb *prendere* 'to take'. All verbs that have a strong PS retain weak forms in the 2SG, 2PL and 1PL PRES IND, except *dare* 'to give', *stare* 'to stay' and *essere* 'to be', which have strong forms throughout (Lepschy and Lepschy 1998, 150).

Table 2-9 Inflectional classes of verbs in IT

Class	Infinitive ending for main class	Endings for subclasses	Example (INF; 1SG, PRES IND)
Class 1	-are	/	*comprare, compro* 'to buy'
Class 2	-ere	/	*credere, credo* 'to believe'
Class 3	-ire	-o -isc-o	*dormire, dormo* 'to sleep' *finire, finisco* 'to finish'

The analytical PC consists of the auxiliary *essere* 'to be' or *avere* 'to have' and the past participle form of the verb. As a rule, transitive verbs take the auxiliary *avere* and intransitives *essere*. Impersonal verbs and reflexive verbs take the auxiliary *essere* (Lepschy and Lepschy

---

[57] As far as the Future is concerned, it is not considered as a tense. Lyons (1997, 677) writes 'Futurity is never a purely temporal concept; it necessarily includes an element of prediction or some related notion'. Moreover, Enç (1996) suggests that the uncertainty of prediction drive it close to modality (see for example, Jaszczolt 2009 for an approach of temporality as epistemic modality).

[58] This is a coarse-grained classification. IT verbs present also other fine-grained specificities, such as the retaining of the velar sound at the end of the stem for verbs in –*care* and –*gare* for the PRES and FUT IND: *gioco* 'I play', *giochi, giocherà* and *pago* 'I pay', *paghi, pagherà* (for a more detailed presentation, see Lepschy and Lepschy 1998, chapter 16).

1998, 144).

2.5.1.3.1 Passato Composito

The use of the Passato Composito (PC) (also called *Passato Prossimo*) and the Passato Semplice (see section 2.5.1.3.2) varies in different parts of Italy. In the north, the PS is rarely used in spoken Italian, while in the south, it is more widely used than the PC. In central Italy, a distinction is made between the two tenses, and it is observed in literary Italian.

The PC is described as having two types of usages: *deictic* and *non-deictic*. The PC is described in opposition to the PS as being used to express a past time eventuality which is linked to the present time, i.e. E<R=S. These deictic usages of the PC have been traditionally called as follows[59] (Bertinetto 1986, 415-419):

- *current relevance* PC
- *experiential* (Comrie 1976) or *existential* perfect (McCawley 1971)
- *inclusive* PC (Jespersen 1948/1961)

The first case is illustrated in examples (263)-(266) from Lepschy and Lepschy (1998, 228-229). Bertinetto (1986) includes in this first case the so-called *notizia fresca* PC 'fresh news PC' as in (267).

(263) Perché sei così arrabbiato con lui? Perché mi *ha dato* un calcio.
      Why are you so angry against him? Because he give.PC me a kick
      'Why are you so angry against him? Because he kicked me.'

(264) Mio fratello *è partito* due ore fa.
      My brother leave.PC two hours ago
      'My brother left two hours ago.'

(265) Negli ultimi dieci anni *abbiamo cambiato* casa sette volte.
      In the last ten years we move.PC seven times
      'In the last ten years we have moved seven times.'

(266) Dante ci *ha dato* nella "Comedia" la maggiore opera della nostra letteratura.
      Dante give.PC us with his "Comedy" the greatest work in our literature
      'Dante has given us with his "Comedy" the greatest work in our literature.'

(267) La sai l'ultima? *È arrivato* Gianni!
      Do you know the latest news? Arrive.PC John
      'Do you know the latest news? John has just arrived!'

The second case is illustrated in (268), where the PC expresses an eventuality that covers S. Bertinetto (1986, 418) argues that this case could be considered as an extreme case of current relevance, where not only the resultative state but also the eventuality itself continues at S and maybe even beyond S. He points out that inclusive usages of the PC are restricted as far as lexical aspect is concerned. More precisely, inclusive interpretations of the PC can occur only with non-telic durative situations, as with situations such as the one in (268), and other types of situations which become statives under the scope of negation, as in (269).

(268) Finora, Gianni *ha vissuto* in questa casa.

---

[59] Deictic usages of the IT PC correspond to the accomplishment FR PC described in section 2.5.1.2.3 and to deictic usages of RO PC described in section 2.5.1.4.1.

Until now, John live.PC in this house

'Until now, John lived in this house.'

(269) Le donne di questo posto non *hanno* sempre *portato* gonne corte.

Women in this position not wear.PC always skirts short

'Women in this position have not always worn short skirts.'

Finally, in (270), the eventuality is part of the life experience of the speaker and it is therefore considered as being linked to the moment of speech S. Bertinetto points out that when the PS is used instead of the experiential PC, as in the pair of examples (271) vs. (271), there is an implication that the referred period of time is completed. This implication could be explicated with temporal adverbials such as *tra il 1968 e il 1973* 'between 1968 and 1973' or *durante la sua vita* 'during his life'. The PC in (271), on the contrary, does not trigger this type of implication.

(270) *Sei* mai *stato* a Parigi?

Be.PC ever to Paris

'Have you ever been to Paris?'

(271) Luca *fu* tre volte in Francia.

Luca be.PS three time in France

'Luca was three time to France.'

(272) Luca *è stato* tre volte in Francia.

Luca be.PC three time in France

'Luca has been three time to France.'

In non-deictic usages[60] the reference moment R is disjoined of S. In these usages, the PC expresses a relation of anteriority of an eventuality with respect to another past eventuality mentioned in the context, as in (273) and (274) from Bertinetto (1986, 421)

(273) Ti avevo detto che *è finito* il latte; parché non mi stai mai ad ascoltare?

I had told you that the milk finish.PC; why don't you ever listen to me

'I had told you that the milk is finished; why don't you ever listen to me?'

(274) La casa è crollata dopo che tu *sei uscito.*

The house collapsed after you get out.PC

'The house collapsed after you got out.'

The PC can also express an anteriority relation with respect to a future eventuality, as in (275) in the subordinate clause, and in (276) in the main clause. In these examples, the PC expresses a reference moment R posterior to S. Procedural-pragmatic studies for the FR PC (such as Saussure 2003) explained that in the futurate PC, the speaker projects herself at a moment S' (i.e. one that is one month after S in (275) and the next day in (276)), when she can make the assertion (S<E<S').

(275) Soltano fra un mese sarà possibile capire chi *ha avuto* ragione tra noi due.

Only in a month it will be possible to understand who of the two of us be.PC right

'Only in a month it will be possible to understand who of the two of us was right.'

---

[60] Non-deictic usages of the IT PC correspond to the anteriority FR PC described in section 2.5.1.2.3 and to non-deictic usages of RO PC described in section 2.5.1.4.1.

(276)   Domani ho finito.
          Tomorrow finish.PC
          'Tomorrow I will have finished.'

Bertinetto (1986, 1996) argues that:

> La caratteristica fondamentale che oppone il PC al PS, e che accomuna al PC tutti i tempi compsti, è la possibilità di attivare per mezzo di eventuali avverbiali temporali un Momento di Riferimento (MR), ossia un punto di valutazione che segue sulla liea del tempo il Momento dell'Avvenimento (MA), e rispotto al quale l'avvenimento stesso vienne rappresentato come compiuto.[61] (Bertinetto 1996, 386).

He argues that the semantics of the PC requires instantiating the three temporal coordinates E, R and S on the time line, where R and S generally coincide. The compatibility of the PC with a temporal adverbial explicitating R indicates that R is part of the semantics of the PC. R can refer to the moment of speech, as shown in (277), to an adverbial indicating the lapse of time between E and R, as in (278), or in cases when there is reference, to no specific moment (i.e. omnitemporal value), as in (279).

(277)   A quest'ora (adesso, etc.) Gianni *è arrivato.*
          At this time (now, etc.) John arrive.PC
          'At this time John has arrived.'
(278)   Gianni *è partito* da due giorni.
          John leave.PC since two days
          'John left two days ago.'
(279)   Una persona che *ha studiato* non deve comportarsi cosí.
          A person who study.PC cannot behave like this
          'A person who studied cannot behave like this.'

These three types of examples are incompatible with the PS, which imposes a temporal location of the eventuality before the moment of speech E<S. The following section is dedicated to the PS.

2.5.1.3.2 Passato Semplice

Traditionally, the PS (also called *Passato Remotto*) is described in the literature as a deictic tense expressing a past time eventuality disconnected from the moment of speech, i.e. E<S. Bertinetto (1986) describes the PS as having the aoristic aspect, that is, it expresses the eventuality as completely ended. He argues that aoristic tenses do not present in their semantics a reference moment R, contrary to perfective tenses such as the PC. Consequently, the PC can be used in temporal (i.e. it allows reference to past and future) and in atemporal sentences (i.e. the omnitemporal value), whereas the PS expresses necessarily a relation of anteriority of the eventuality with respect to the moment of speech S. Bertinetto

---

[61] 'The main feature opposing the PC and the PS is the possibility with the PC to activate through temporal adverbials a reference moment R, that is a point of evaluation which follows the event moment on the time line and at which the eventuality is represented as being completed.' (my translation)

(1986, 430) points out[62] that in examples (280) vs. (281), the PS imposes a temporal and definite interpretation (i.e. an identifiable woman who lost her son, where E<S) whereas the PC allows an atemporal and indefinite interpretation (i.e. a hypothetical situation where a woman who could lose her son at an imaginary moment).

(280) Per consolarmi, cercai di pensare ad una madre che *persi* il proprio figlio.
       To comfort me, I tried to think of a mother who lose.PS her son
       'To comfort me, I tried to think of a mother who lost her son.'
(281) Per consolarmi, cercai di pensare ad una madre che *ha perso* il proprio figlio.
       To comfort me, I tried to think of a mother who lose.PC her son
       'To comfort me, I tried to think of a mother who lost her son.'

This description corresponds to the so-called descriptive[63] usage of the PS. There are cases, however, where the PS may produce different interpretations. Firstly, it can have a non-deictic usage as in (282), where it behaves like a PQP (from Bertinetto 1986, 431). Secondly, the PS is used to express atemporally in saying and proverbs (i.e. the so-called gnomic usage) as in (283).

(282) Ritornando dal viaggio che *feci/avevo fatto*, trovai una montagna di posta.
       Coming back from the journey I do.PS/PQP, I found a mountain of mail
       'Coming back from the journey I made/had made, I found lots of mail.'
(283) Cosa fatta in fretta non *fu* mai buona.
       Things made in haste not be.PS ever good
       'The things done in haste have never been good.'

Bertinetto (1986) and more recently Squartini and Bertinetto (2000) argue that in IT, the PS and PC, being perfective, are more similar than different, and this becomes more visible when compared to the IMP. This is mainly due to the aoristic drift undergone by the PC. I will come back to the distinction between the perfective PC and PS, on the one hand, and the imperfective IMP, on the other hand, in section 2.5.1.3.3.

2.5.1.3.3 Imperfetto

IT IMP has temporal and modal values. As for the FR IMP, in this research I am interested in the temporal values. According to Bertinetto, the IT IMP is a 'clearly imperfective' verbal tense (1986, 345). It presents all three aspectual oppositions recognised in the literature, namely, progressive, continuous and habitual (Comrie 1976), though the continuous aspect seems to be most representative. The progressive aspect of the IMP is shown by the contrast between examples in (284) and (285), where the IMP expresses that the eventuality of having breakfast started before the moment when the news arrived, whereas in (286), the PS expresses that the eventuality of having breakfast started exactly at the same moment when the news arrived. The habitual aspect is shown in (286), where there

---

[62] Bertinetto notes that this opposition between PC and PS in IT is similar to the relation between the definite and indefinite articles.

[63] The notions of *descriptive* and *interpretative* usages have been proposed for FR verbal tenses, as discussed in section 2.5.1.2.

is no information about the total duration of the eventuality without it being explicitly marked by an adverbial or by contextual information. The same holds for the continuous aspect in (287) from Bertinetto (1986, 347, 349).

(284) Quando arrivò la notizia, Andrea *faceva* tranquillamente colazione come ogni mattina.
When the news arrived, Andrea make.IMP calmly breakfast as every morning
'When he heard the news, Andrea was having breakfast as every morning.'

(285) Quando arrivò la notizia, Andrea *fece* tranquillamente colazione come ogni mattina.
When the news arrived, Andrea make.PS calmly breakfast as every morning
'When he heard the news, Andrea had breakfast as every morning.'

(286) Tino *pedalava* ogni giorno per due ore.
Tino pedal.IMP every day for two hours
'Tino used to pedal/was pedaling every day for two hours.'

(287) Cosa *facevi* ieri dalla 2 alle 4? *Dormivo.*
What do.IMP yesterday from 2 to 4? Sleep.IMP
'What were you doing yesterday from 2 to 4? I was sleeping.'

Bertinetto (1986, 352) points out that the aspectual information expressed by the IMP is linked to the notion of *indetermination* evaluated with respect to the continuation of the eventuality beyond the interval considered, with respect to the delimitation of the interval considered and with respect to the number of iterations. This indetermination is most often resolved due to contextual knowledge.

The main temporal interpretations of the IMP are simultaneity in the past and the narrative IMP. The interpretation as simultaneity in the past is linked to the notion of *temporal anchoring* (TA, see section 2.3.2). Both the progressive and the continuous IMP require temporal anchoring, which cannot be provided uniquely by a temporal adverbial as in (288) and (290) respectively. Examples (289) and (291) on the contrary show that temporal anchoring can be established with respect to another eventuality. This is linked to the fact that temporal adverbials do not necessarily signal a reference moment R (as discussed in section 2.2.1).

(288) ?Ieri *giocavo* a carte.
Yesterday play.IMP cards
'Yesterday, I was playing cards. '

(289) Ieri a quest'ora *giocavo* a carte; come passa il tempo!
Yesterday, at this time I play.IMP cards; how time fly.PRES
'Yesterday, at this time I was playing cards; time flies.'

(290) ?La settimana scorsa *mi vedevo* un film dopo l'altro.
Last week watch.IMP a movie after another
'Last week I was watching a movie after another.'

(291) La settimana scorsa, mentre tu passavi tutto il tempo sui libri, *mi vedevo* un film dop l'altro.
Last week, while you pass.IMP all your time on books, I watch.IMP a movie after another
'Last week, while you were passing all your time on books, I was watching a movie after another.'

Bertinetto points out that the IMP can carry out the interpretation of simultaneity in the past independently of the occurrence of explicit markers such as *at the same time, when* and *simultaneously,* markers that are necessary for the PC or PS, as shown in examples (292) and (293) from Bertinetto (1986, 357). In the absence of an explicit marker or an appropriate context, the PC has an inceptive and sequential interpretation as in (294).

(292) Quando Luca *è caduto*, Marco *faceva* le scale assieme a lui.
When Luca fall.PC, Marco walk.IMP on the stairs with him
'When Luca fell, Marco was walking on the stairs with him.'

(293) Quando Luca *è caduto*, Marco *ha fatto* le scale assieme a lui al tempo stesso.
When Luca fall.PC, Marco walk.PC on the stairs with him at the same time
'When Luca fell, Marco walked on the stairs with him at the same time.'

(294) Quando Luca *è caduto*, Marco *ha fatto* le scale assieme a lui.
When Luca fall.PC, Marco go.PC the stairs with him
'When Luca fell, Marco went down the stairs with him.'

Nevertheless, the IMP can also be used to express temporal sequencing, and this can be observed with the habitual IMP in (295) and when the context triggers it as shown by the contrast between (296) and (297) from Bertinetto (1986, 358, 359).

(295) Il professore *si alzava* ale 7 e un quarto, *si rasava*, *raccoglieva* le sue cose, e *scendeva* al bar per fare colazione.
The professor wake up.IMP at a quarter past seven, he shave.IMP, gather.IMP his things and go down.IMP at the café to have breakfast
'The professor woke up at a quarter past seven, he shaved, gathered his things and went down at the café to have breakfast. '

(296) *Suonavano* le 8 ed egli si *alzò*.
Ring.IMP 8 and he wake up.PS
'The alarm clock rang at 8 and he woke up.'

(297) *Suonavano* le 8. Egli si *alzò*, si *lavò*, si *vestì*.
Ring.IMP 8. He wake up.PS, wash.PS, dress.PS
'The alarm clock rang at 8. He woke up, washed himself and got dressed.'

As for the narrative IMP, it represents an interpretative usage of the IMP attested in all Romance languages (Savić 1979 cited by Bertinetto, 1986) and which received particular attention in FR (see section 2.5.1.2.2). Classically, the narrative IMP is characterized by contrast to the IMP, mainly by the perfective aspect triggering an interpretation of perfectivity of the eventuality, by the temporal sequencing of the eventualities expressed, and by the presence of a temporal adverbial which sets the reference moment R, as in (298) from Bertinetto[64] (1986, 383).

(298) L'indomani, a mezzogiorno in punto, egli *usciva* dalla città.

---

[64] There are scholars like Blücher (1974, cited by Bertinetto, 1986) who argue that the FR IMP is aspectually neutral. It would then retrieve its aspectual value (i.e. imperfective or perfective) from the context. Bertinetto argues against this position by suggesting that the IMP is essentially imperfective and that the narrative IMP was initially used in contexts where an imperfective IMP (i.e. a descriptive IMP) would have been expected. The narrative IMP can be inserted directly in the narrative flow and produce 'descriptive pauses'.

The next day at noon precisely, he get out.IMP from the town
'The next day at noon precisely, he got out from the town.'

However, Imbs (1960, cited in Bertinetto 1986, 393) gives examples of narrative IMPs from Flaubert's *Bouvard et Pécuchet* whose perfective aspect can be questioned. Bertinetto therefore suggests that the aspectual nature of the narrative IMP can only be determined contextually, and this is due to the following idea:

> […] la forte tensione imperfettiva che questo Tempo possiede nelle sue accezioni standard, e le possibili neutralizzazioni aspetuali cui esso va incontro in certi particolari contesti […] ma normalmente esso non giunge fino ad annullare del tutto le connotazioni aspetualli primarie.[65] (Bertinetto 1986, 393-394)

This idea explains the fact that the narrative IMP can rarely be replaced by a PS without any loss, and more specifically, a 'temporal dilation of the event' that triggers a focalization on that event during the interpretation process. It is a psychological dimension rather than a discursive one[66].

Bertinetto points out that the 'narrative' interpretation is therefore mainly due to both linguistic and pragmatic factors, and this occurs only when the context is taken in consideration. The narrative IMP is used as a tool to focalize on the eventuality expressed, an interpretative effect that does not occur with the PS as in (299).

(299) L'indomani, a mezzogiorno in punto, egli *uscì* dalla città.
       The next day at noon precisely, he get out.PS from the town
       'The next day at noon precisely, he got out from the town.'

To sum up, in this section I have proposed a non-exhaustive list of the temporal usages of the IT IMP. The IMP is one of the three verbal tenses expressing past time in their descriptive usages analysed in this thesis. The following section is dedicated to the PRES, which is the fourth most frequent verbal tense used to translate the EN SP into IT.

2.5.1.3.4 Presente

The two main categories of usages of the PRES are those regrouping non-deictic (standard usages) and deictic usages. Among deictic usages, one can distinguish between reference to present (descriptive usages) and reference to past and future time (interpretative usages). Similarly to the FR PRES, I am at proposing a non-exhaustive discussion of the PRES (for a detailed discussion, see for example Bertinetto 1986).

Non-deictic usages of the PRES are characterized by a maximally generic relation

---

[65] '…collision between the strong imperfective force that this tense has in its standard usages, and the possibility of aspectual neutralization in certain contexts […] but normally this cannot completely cancel the primary aspectual connotations.' (my translation)

[66] Bertinetto exemplifies this idea with the following passage from the novel *La cognizione del dolore* by C.E. Gadda: "*Ebbe* per il dottore, che non vedeva da tempo, espressioni cordiali ma brevi; e gli *demonstrava* la sua stima. Con garbo native *diede* senz'altro per inavvertiti i quattro millimetri di barba…". This passage describes an important moment in the existential developpement of the protagonist. The use of the narrative IMP instead of the PS transfers what is said from the discoursive level to the psychological level.

between S and E, where S is represented as an extended period of time. It occurs with adverbials such as *now* in a generic reading or *today* in a generic reading, as in (300).

(300)   Adesso, i giovani *bevono* tanto.
        Today, teenagers drink.PRES a lot
        'Today, teenagers drink a lot.'

According to Bertinetto (1986, 329-331) the PRES is used non-deictically in atemporal statements as in (301), definitions, proverbs, gnomic statements, omnitemporal assertions as in (302).

(301)   La verità *è* un bene supremo, ma non sempre.
        The truth be.PRES a supreme good, but not always
        'The truth is a supreme good, but not always.'
(302)   Parigi si *trova* in Francia.
        Paris be.PRES in France
        'Paris is in France.'

Descriptive deictic usages correspond to cases where the PRES expresses simultaneity to the moment of speech, precisely E=R=S, and therefore reference to present time. Bertinetto (1986, 325) points to the fact that the PRES has deictic usages when it occurs with temporal adverbials such as *al momento attuale* 'at the present moment' and *in questo preciso istante* 'in this very moment' only with durative situations, as in (303). When used with non-durative verbs, it expresses a comment about a situation in progress (i.e. reporting PRES) as in (304). When the PRES is used with non-durative verbs, it expresses the inceptive aspect, as shown in example (305) compared to (306), where the latter illustrates the impossibility to have a progressive interpretation. It can also have a reiterative interpretation as in (307).

(303)   In questo preciso istante, Carlo *dorme*.
        In this very moment, Carlo sleep.PRES
        'In this very moment, Carlo is sleeping.'
(304)   In questo preciso istante, Clara *esce*.
        In this very moment, Carla get out.PRES
        'In this very moment, Carlo gets out.'
(305)   Ora *piange*; lo sapevo!
        Now cry.PRES. I know.IMP
        'Now he cries. I knew it!'
(306)   ??Adesso Edoardo *piange*.
        Now Eduard cry.PRES
        'Eduardo is crying now.'
(307)   In questo momento, Gaetano *raccoglie* le biglie che Monica ha sparso per terra.
        In this moment, Gaetano pick up.PRES the marbles that Monica scatter.PC on the floor
        'In this moment, Gaetano picks up the marbles that Monica scattered on the floor.'

The PRES can also express habitual situations, as in (308), interpreted as 'every time I go to the mountains, I feel another like person', and in (309).

(308)   In montagna *mi sento* un altro.
At the mountains I feel.PRES another person
'At the mountains I feel like another person.'

(309)   Amedeo *viaggia* sempre in prima classe.
Amedeo travel.PRES always in the first class
'Amedeo always travels in the first class.'

Interpretative deictic usages of the PRES are usages where it expresses reference to other times than the present[67]. The PRES can locate the eventuality in the future. as in examples (310)-(314) where the posteriority of R with respect to S is expressed explicitly through an adverbial or inferred in the context, as in the last two examples.

(310)   *Resto* a casa nel pomeriggio.
Stay.PRES at home this afternoon
'I'll stay at home this afternoon.'

(311)   *Parto* domani.
Leave.PRES tomorrow
'I am leaving tomorrow.'

(312)   Fra un anno *mi trasferisco* a Milano.
In a year I move.PRES to Milan
'In a year's time I will move to Milan.'

(313)   Adesso *esco*.
Now get out.PRES.
'I'm getting out in a second.'

(314)   *Vengo* subito.
Come.PRES at once
'I'll come at once.'

As for reference to past time, two types of usages can be distinguished. The first is the *narrative*[68] PRES (also called *historical present* HP), which is used in written and oral narratives, as in (315) and (316) from Bertinetto (1986).

(315)   In armonia con questo giudizio, Andreotti *compie* con regolarità, a Firenze, dove era nato il 15 marzo 1924, gli studi medi...
According to this opinion, Andreotti carry out.PRES regularly in Florence, where born-PC on 15th of March 1924, his medical studies
'According to this opinion, Andreotti carried out regularly in Florence, where he was born on 15th of March 1924, his medical studies.'

(316)   Ieri *vado* al cinema, e chi ti *trovo*? Cinzia e Mario, naturalmente!
Yesterday, go.PRES to cinema, and who find.PRES? Cynthia and Mario, naturally
'Yesterday I went to the cinema, and who did I find there? Cynthia and Mario, naturally.'

---

[67]Other usages are the hypothetical PRES, the modal PRES and the prophetic PRES among others (see Bertinetto 1986).

[68] In a finer-grained analysis, Bertinetto (1986, 335) distinguishes between *narrative* and *dramatic* PRES on aspectual basis. More precisely, only narrative PRES can be interpreted as habitual and accepts the progressive periphrase. Here I consider both usages under the label *narrative* PRES.

The second is the *recent past* PRES, used most often with actions implying no duration and a resultative state[69], such as *arrive, get out* or *say* as in (317).

(317) Sii paziente con Teresina, *esce* da un collegio di suore.
       Be patient with Teresina, get out.PRES from a nuns school
       'Be patient with Teresina, she has just gotten out from a nuns school.'

As far as the aspectual nature of the PRES is concerned, it is a verbal tense that presents an eventuality from both imperfective and perfective points of view. The imperfective PRES can be easily be replaced by progressive periphrases *stare* + gerund and *stare a*+ infinitive 'be + *-ing'*, whereas the perfective PRES occurs mainly in the deictic non-standard usages when it expresses reference to past time.

To sum up, in IT, information regarding Tense and Aspect is expressed in the same verbal form. Tense locates an eventuality prior, simultaneous and posterior to the moment of speech S. The aspectual distinction perfective/imperfective is expressed grammatically only for past time verbal tenses. Verbal tenses can have deictic usages expressing the relation E/S, and/or anaphoric expressing the relation E/R. In this thesis, I will be arguing that Tense and its expression by verbal tenses in IT, similarly to FR and RO encode conceptual information (i.e. the relation E/S) and procedural information (the relation E/R). The four verbal tenses described in this section can coarsely be classified with respect to their conceptual information: the PC, PS and the IMP on the one hand, and the PRES on the other hand.

2.5.1.4   Romanian

Similarly to IT, RO verbs mark Tense and Aspect synthetically and analytically. As for Aspect, in RO, the distinction perfective/imperfective is expressed only for past and future verbal tenses, which have the relation R≠S (GLR, 2005). For the PRES, which is the unique form to express reference to present time (E=R=S), the aspectual distinction perfective/imperfective can be expressed only contextually through pragmatic means. There are three past time verbal tenses that express the feature [+perfective], namely the PS, the PC and the PQP, and one verbal tense, the IMP, which expresses the feature [-perfective] (Margan 2009).

Unfortunately, the Romanian literature investigating the semantics and pragmatics of RO verbal tenses is poorer than the French literature for example. This is mainly due to two reasons. The first is the lack of a database for contemporary Romanian. The immediate consequence is that most of the studies are carried out either on built examples or on Old Romanian using historical texts. The second reason is that it is only the last fifteen years that local Romanian researchers started to be interested in pragmatic theories of language[70] and in having an empirical basis for their research, coming most often from language acquisition (such as for example Stoicescu 2009, 2010 for grammatical and lexical aspect; 2012 for

---

[69] Called performative verbs (Bertinetto 1986, 274)

[70] It is worth mentioning the work of Nadia Vișan in the discourse semantics DRT framework (2006 among others), as well as several studies in syntax by researchers such as C. Dobrie-Sorin, G. Soare, E. Soare and P. Monachesi among a few others.

temporal deixis). As for Romanian linguists living and working abroad, their task is complicated by the lack of a database available for consultation and the lack of studies and language grammars published with international publishers. In this thesis, I propose an analysis of four RO verbal tenses based on a relatively reduced translation corpus and on offline experimentation with two Romanian judges. However, the research is carried out in a multilingual perspective aiming at identifying a language independent and cross-linguistically valid *tertium comparationis* for verbal tenses. With this thesis, I hope to have contributed to advancing our knowledge of the pragmatics of verbal tenses in RO. It should, however, be followed and developed by more consistent empirical and experimental work targeting Romanian language in a monolingual context.

RO verbs are regular and are classified in five inflectional classes, characterized by the infinitive endings, as in Table 2-10. In contrast to other Romance languages, RO inherited the first four classes from Latin and developed the fifth class. Except to the second class in −*ea*, all other classes divide into more subclasses depending on other endings that are inflectionally relevant (for a more detailed presentation, see Zafiu 2013). There is also a series of verbs which belong to one of the five classes but which present irregularities in the realization of the root, of the affixes and/or of the syncretic forms[71].

Table 2-10 Inflectional classes of verbs in RO

Class	Infinitive ending for main class	Endings for subclasses	Example
Class 1	-a	-ez (1SG, PRES IND) +ez(1SG, PRES IND)	*învăța, învăț* 'to learn' *lucra, lucrez* 'to work'
Class 2	-ea	╱	*plăcea, place* 'to like'
Class 3	-e	-ut (PAST PART) -s (PAST PART) -t (PAST PART)	*face, făcut* 'to do' *întoarce, întors* 'to turn' *rupe, rupt* 'to tear'
Class 4	-i	-esc (1SG, PRES IND) +esc (1SG, PRES IND)	*ieși, ies* 'to come out' *citi, citesc* 'to read
Class 5	-î	-ăsc (1SG, PRES IND) + ăsc (1SG, PRES IND)	*coborî, cobor* 'to go down' *hotărî, hotărăsc* 'to decide'

Before discussing the values and usages of the verbal tenses considered in this thesis, I will give an example of the conjugation paradigm of a verb from Class 1 for the indicative mood, *a cânta* 'to sing', provided in Table 2-11. This will facilitate the reader's comprehension of the examples provided later on during the discussion of the usages of the RO verbal tenses.

Table 2-11 Paradigm of conjugation for the indicative mood

Tense	Aux +	Root+ Endings (1SG)
PRES		cânt
IMP		cânt**am**
PC	am	cânt**at**
PS		cânt**ai**
PQP		cânta**sem**
FUT	voi	cânta
FUT PERF	voi fi	cântat
FUT PAST	aveam să cânt	

---

[71] Such as the Class 1 (and the −ez subclass) *a vrea* 'to want' which has specific syncretic forms INFINITVE=3SG PRESENT INDICATIVE=3SG PRESENT SUBJUNCTIVE=3PL PRESENT SUBJUNCTIVE (for other cases and sources of irregularities, see Zafiu 2013, 23).

Synthetic verbal tenses consist of inflectional suffixes added to the root, marking tense, person and number. For the PRES, verbs take different thematic suffixes and inflectional endings depending on the class and subclass of the verb. The PS is formed of the root followed by suffixes and the inflectional endings marking tense, which are different for all classes and subclasses of verbs. The IMP forms consist of the root (usually the same as for the 1st and 2nd person PRES), a specific tense suffix (*-a* or *−ea*) and another inflectional ending marking person and number. The pluperfect (PQP) consists of the perfective root (identical to the root of the participle and the SP), followed by the more general tense suffix *−se* and a thematic suffix. Analytical tenses PC and FUT[72] are formed of auxiliaries and the verb. The PC is formed with the auxiliary *a avea* 'to have[73]' and the invariable participle of the verb.

2.5.1.4.1 Perfectul Compus

The PC, in its descriptive usages, presents the following features: [E<R], [R=S], [+perfective] and [+realis] (Margan 2009). In other words, the PC presents an eventuality as being anterior to S (E<S) from a moment of reference that is in the present (R=S), as being accomplished and situated in the speaker's reality. The PC does not specify the temporal distance between E and S, as shown in (318) and (319) from Zafiu (2013, 58), where both short and long periods are possible.

(318) Dan *a venit* acum cinci minute și te așteaptă.
Dan come.PC five minutes ago and he wait.PRES for you
'Dan came five minutes ago and he has been waiting for you.'

(319) Basarb I *a trăit* acum șapte sute de ani.
Basarab I live.PC seven hundred years ago
'Basarab I lived seven hundred years ago.'

The PC has deictic usages[74], as in examples in (318) and (319), and, in certain contexts, it functions as an anaphoric tense, as in examples (320)-(322) from Zafiu (2013, 58). In (320), the PC makes anaphorically reference to a past time R (i.e. a moment situated a week before the moment of speech). In (321), the PC makes reference to a future time R (i.e. a moment situated somewhere in the future). Finally, in (322), the PC makes reference to a recent past time R (i.e. a moment situated the day before the moment of speech at lunch). Zafiu points

---

[72] The FUT presents several competing periphrastic forms, called the standard future (auxiliary *a vrea* 'to want' followed by the bare infinitive form of the verb), the regional future (reduced form of the auxiliary into *oi* followed by the bare infinitive) and two colloquial futures (*o să* future formed of the particle *o* followed by the subjunctive form and the *am să* future formed of the auxiliary *a avea* 'to have' followed by the subjunctive form). The future perfect (FUT PERF) corresponds to the standard future and consists of the auxiliary *a fi* 'to be' in the FUT followed by the participle form of the verb. The future in the past (FUT PAST) is formed with the auxiliary *a avea* 'to have' at the IMP followed by the subjunctive form of the verb. RO differs from western Romance languages and resembles Balkan languages with respect to the periphrastic FUT using a volitional verb (Zafiu, 2013).

[73] Within Romance languages, RO is similar to Spanish and Portuguese in that the PC is formed only with the auxiliary *to have*, whereas IT and FR use an alternation between *to have* and *to be* (Zafiu, 2013).

[74] It was Comrie (1976) who proposed that deictic tenses express the relation between E and S (where R refers deictically to the moment of speech S), whereas anaphoric tenses express the relation between R and E (where R is determined anaphorically in the context).

out that in these examples, the PC replaces the PQP, the FUT PERF and the IMP, respectively.

(320) *A găsit* ieri scrisoarea. *Am pierdut*-o acum o săptămână.
He find.PC the letter yesterday. He lost.PC a week ago
'Yesterday, he found the letter. He had lost it a week ago.'

(321) O să merg la Ploiești și, când *am terminat* treaba, o să mă întorc.
I go.FUT to Ploiești, and, when I finish.PC, I go.FUT back
'I will go to Ploiești, and when I will have finished the business, I will come back.'

(322) Te-*am căutat* ieri la prânz. Unde-*ai fost*?
I look.PC for you yesterday at lunch. Where be.PC
'I looked for you yesterday at lunch. Where have you been?'

The PC can be used to express anticipation (a future value) as in example (323), where the speaker expresses her intention to accomplish an action very quickly in the immediate future (from 2013, 58), and very recent past, as in (324) from Sporiș (2012, 70).

(323) Gata, am plecat.
Ready leave.PC
'I am done and I'm off.'

(324) *A intrat* chiar acum pe poarta liceului.
Enter.PC right now through the highschool's gate
'He has just entered through the highschool's gate.'

From a procedural pragmatics approach, these usages of the PC are interpretative usages, where the utterance refers to a thought: the speaker imagines herself at a moment S' (i.e. one that is a few moments after S), when she can assert *I am done and I'm off.* This usage occurs also in FR (see section 2.5.1.2.3).

Vişan (2006) points out that the RO PC, as well as the PC in all Romance languages, does not perform the Present Perfect puzzle (Klein 1992, see section 2.5.1.1.2 for the PresPerf in EN). In other words, the RO PC may co-occur with a definite past adverbial as in (325). The RO PC is still a perfect due to its link to S (R=S) and its usage to express resultativity, as shown in (326) from Vişan (2006). At the same time, the PC is may be used in narratives, as shown in (327).

(325) Ion *a plecat* ieri la ora 5.
John leave.PC yesterday at 5 o'clock
'John left yesterday at 5 o'clock.'

(326) Mi-*ai spart* capul!
You crack.PC my skull
'You've cracked my skull!'

(327) Marin și Ion *au plecat* de acasă devreme și s-*au întors* seara târziu. După ce *au intrat* în casă, *au făcut* de mâncare, *au luat* cina si s-*au culcat*.
'Marin and John left.PC home very early and came back.PC very late in the evening. After they entered.PC the house, they prepared.PC the dinner, ate.PC and went.PC to bed.'

The PC in standard RO is fairly advanced in the aoristicization process (the so-called *aoristic drift*), more than IT and in FR (Squartini and Bertinetto 2000). Previous corpus-based studies showed that the PS has a more reduced usage in written cultivated texts, literary style and narratives in RO than in FR and IT (Squartini and Bertinetto citing Savić 1979; Călăraşu 1992). Squartini and Bertinetto (citing Călăraşu 1992) note than the analysis of a contemporary Romanian epistolary novel showed that the PS is completely absent while it is used in its FR translation. As for newspaper texts, Squartini and Bertinetto (citing Savić 1979) also note the absence of the PS[75]. These results are similar to those found in the analysis of the multilingual corpus described in section 5.3. Vişan (2006) therefore suggests that the RO PC is semantically richer than its EN counterpart Present Perfect, and it ranks highest in the aoristicization process. She furthermore suggests that the PC and PS in RO have identical functions in narratives and that they are interchangeable. When they alternate in the same text, this takes place without necessarily imposing a change of perspective, as shown in (328) from Vişan (2006, 65). This behaviour is due to the great advancement of the PC in the aoristic drift.

(328) Vorbind aşa, *au ajuns* aproape de Tecuci, unde poposiră la o dumbravă.
       'As they were speaking, they arrived.PC close to Tecuci, where they stopped.PS in a glen.'

Vişan points out that RO data reveals that the PC is frequently used in both spoken and written RO and that 'the narrative value of the PC alternates witht he narrative PS' (Vişan 2006, 66). Her usage of the term *narrative* makes reference to Smith's (2003) notion of *discourse modes*. In this thesis, I will be arguing that narrativity is not a discourse mode, but an inherent feature of Tense, i.e. a binary procedural feature.

Regarding aspectual information expressed by the PC, the PC presents the situation from a perfective viewpoint, i.e. it expresses a completed situation as in (329) from Zafiu (2013, 59). It can be accompanied by a durative adverbial as in (330) and by iterative temporal adverbials as in (331).

(329) *A citit* cartea.
       She read.PC the book
       'She read the book.'
(330) *A citit* din carte timp de trei ore.
       She read.PC from the book for three hours
       'She read from the book for three hours.'
(331) *A venit* în vizită în fiecare zi.
       She come.PC to visite every day
       'She came to visit every day.'

The PC can be used both in contexts with temporal progression as in (332) and with backward temporal sequencing as shown in (333) from Zafiu (2013, 59) and in (334) from Vişan (2006, 63). As seen in all three examples, explicit temporal connectives are possible but

---

[75] There is however a variety of Romanian spoken in the regions called Oltenia and Muntenia, where the PS is used very often and it expresses temporal proximity (short periods prior to S). The PC is used to express more distant situations (Squartini and Bertinetto 2000).

not necessary to express temporal progression or reverse order. Vișan (2006) makes the hypothesis that temporal ordering is triggered by the sequence of tense forms themselves.

(332) *Am ajuns* acolo. *Am văzut* dezastrul. *Am chemat* poliția.
I get.PC there. I see.PC the disaster. I call.PC the police
'I got there. I saw the disaster. I called the police.'

(333) *Am ajuns* acolo pe la prânz. *A fost* o zi teribilă.
I get.PC there around noon. It be.PC a terrible day
'I got there around noon. It was a terrible day.'

(334) În acea zi *a plouat* mult. Cu toții erau iritați. Maria l-*a lovit* din greșeală pe Mihai. Ion *a căzut*. Martin l-*a împins* și *a căzut* și el.
That day it rain.PC a lot. We be.IMP all out of sorts. Maria hit.PC Mihai by mistake.
Ion fall.PC. Marin push.PC him and then he fall.PC too
'That day it rained a lot. We were all out of sorts. Maria hit Mihai by mistake. Ion fell.
Marin pushed him and then he fell too.'

In a procedural pragmatics perspective, one could make the hypothesis that the RO PC, similarly to the FR PC, does not encode instructions about temporal sequencing and allows both readings under the pressure of contextual relevance. In Chapter 7, I will suggest a reanalysis of the RO PC, PS, IMP and PRES in pragmatic terms, based on the corpus and experimental work described in this thesis.

2.5.1.4.2 Perfectul Simplu

The RO PS is an absolute tense, presenting the following features: [E<S], [E=R], [+perfective] and [+realis] (Margan 2009). In other words, it locates an eventuality prior to S and simultaneously to R, it presents it from a perfective viewpoint and it locates it the reality of the speaker.

Because of its advancement in the aoristic shift, the PC became the preferred tense for expressing past time reference. Consequently, the PS is currently being used in RO as a narrative tense in literary fiction and as a recent past, in southern regional varieties. Zafiu (2013, 59) notes that the narrative PS is used with 3rd person pronouns and expresses impersonal remarks (i.e. there is no explicit speaker who commits to what was said). It designates situations prior to the present, without indicating any relation with the moment of speaking S, as in example (335). When used in fiction, the SP cannot be subordinated to verbs of declaration, as shown in (336), and it contrasts with the PC in direct and indirect speech, as shown in (337).

(335) Monstrul o   *văzu*   pe prințesă.
The monster see.PS.3SG the princess
'The monster saw the princess.'

(336) **Spuse*   că *fu*   acasă.
Say.PS.3SG that be.PS.3SG at home.
''He said that he was at home.'

(337) Am văzut casa, *spuse* el.
See.PC the house, say.PS
'I saw/have seen the house, he said.'

As for the PS used in southern varieties, it expresses recent past (i.e. eventualities that took place during the same day) and can be used for all persons, contrary to the literary SP that can only be used for the 3rd person, as in example (338).

(338)　Unde *fuseşi* de dimineaţă? Mă *dusei* la moară.
　　　　Where be.PS this morning. I go.PS to the mill
　　　　'Where were you this mornig ? I went to the mill.'

Similarly to the PC, the PS expresses the eventuality from a perfective viewpoint, i.e. it expresses a completed situation, as in (339). It can be accompanied by durative (as in (340)) and iterative temporal adverbials (as in (341)).

(339)　*Citi* cartea.
　　　　read.PS.3SG the book
　　　　'She read the book.'
(340)　*Citi* din carte timp de trei ore.
　　　　read.PS. 3SG the book for three hours
　　　　'She read from the book for three hours.'
(341)　*Veni* în vizită în fiecare zi.
　　　　Come.PS.3SG to visit evevry day
　　　　'She came to visit every day.'

As indicated in this section, the PS in RO is an infrequent verbal tense, being replaced both in oral and written discourses by the PC (except in the south-east of Romania). As suggested in section 2.5.1.2.3, this is due to the aoristicization process undergone by the perfect in Romance languages (FR, IT and RO). This process is very advanced in RO, where the PC is more frequent than in IT and RO (see section 5.3.2 for a quantitative analysis).

2.5.1.4.3　Imperfectul

The IMP is an anaphoric tense expressing reference to past time[76] and it presents the following features [E<S], [E=R], [-perfective] and [+realis] (Margan 2009). In other words, it locates an eventuality prior to S and simultaneously to R, it presents it from an imperfective viewpoint, and it locates it in the reality of the speaker. The reference moment can be explicitly marked by an adverbial, as in (342), or by an adjunct temporal clause, as in (343), or it may be implicit and therefore inferable from the context, as in (344).

(342)　Atunci *erai* mai înţelegător.
　　　　At that time be.IMP more understanding
　　　　'You were more understanding at that time.'
(343)　Ana *dormea* când a sunat telefonul.

---

[76] The IMP has modal values when it refers to irrealis. Interpretative usages of the IMP include counterfactual information, the hypocoristic IMP and the politeness IMP. As far as I am aware, there does not exist a pragmatic model that would explain the temporal and modal usages of the RO IMP. I will make the hypothesis that the procedural pragmatics model suggested for FR (see section 2.5.1.2.2) could be applied for RO. I will discuss this matter in chapter 7.

Ana sleep.IMP whe the phone ring.PC

'Ana was sleeping when the phone rang.'

(344) Am ieșit pe terasă. Soarele *răsărea*.

Come.PC on the terrace. The sun rise.IMP

'I came out on the terrace. The sun was rising.'

Classically, the IMP is described as presenting eventualities from an imperfective viewpoint. Zafiu (2013, 60) argues that the IMP is better described as a means of marking the aspect in the past rather than a proper tense. The IMP expresses durative (continuous) and iterative aspect, as in (345) and (346) respectively, from Zafiu (2013, 60-61).

(345) Afară *ploua* iar tu *stăteai* în casă și *citeai*.

Outside rain.IMP and you stay.IMP indoors and read.IMP

'It was raining outside and you stayed indoors and read.'

(346) *Deschidea* și *închidea* ușa de mai multe ori pe zi.

Open.IMP and close.IMP the door several times a day

'He was opening and closing the door several times a day.'

Similarly to the FR IMP, the RO IMP is described as providing a null temporal instruction, being a tense of 'background'. Zafiu points outs that under the pressure of contextual information, such as in popular epic poetry, the IMP can express forward temporal sequencing instructions.

2.5.1.4.4 Prezentul

Traditionally, the RO PRES is described as being an absolute deictic tense with numerous temporal values. Prototypically, the PRES has the following features: [E=S] and [+realis] (Margan 2009). In its descriptive usages, the PRES expresses temporal simultaneity among three temporal coordinates: E, R and S. More precisely, the PRES indicates that the eventuality takes place in an interval of time that is simultaneous with S, as in (347) from Zafiu (2013, 55). The PRES can place the situation in any epoch, past or future, and even in all epochs (*omnitemporal* or *atemporal* value) as in (348).

(347) Acum *locuiesc* aici.

'I live here now.'

(348) Triunghiul *are* patru laturi.

'A triangle has four sides.'

The PRES does not encode the aspectual distinction perfective/imperfective. As a simple form, the PRES expresses the situation while happening, without taking into consideration it's delimitations or duration. The limits and the duration are provided by the semantics of the verb or through adverbials in the context. Hence, sentences with the verb in the PRES may express imperfective and continuous as in (349), iterative as in (350) and perfective with an achievement (punctual situation) in (351) (from Zafiu (2013, 57)). She points out that contemporary RO, like FR (Bertinetto 2000), does not have grammaticalized periphrases with a progressive meaning, contrary to other Romance languages such as IT. The distinction between progressive and non-progressive PRES must therefore be made based on

other lexical items or contextual information.

(349) *Așteptă* în stradă.
       He wait.PRES in the street
       'He is waiting in the street.'

(350) El își *verifică* mesageria telefonică de trei ori pe zi.
       He check.PRES his voicemail three times a day
       'He checks his voicemail three times a day. '

(351) Deodată, fotografia îi *cade* din mâini.
       Suddenly, the photograph drop.PRES from his hands
       'Suddenly, the photograph drops from his hands.'

As for the interpretative usages of the PRES, it can express reference to future (the *futurate* PRES or to past time (the HP). The PRES expresses future time when the reference time is not S but an immediately subsequent interval or time (as in promises or predictions). The reference time may be explicit, expressed with future adverbials, as in (352), or implicit, inferable from the context, as in (353), from Zafiu (2013, 56). Authors studying the futurate PRES in Romance languages (Rebotier 2009, Salvi and Vanelli 2004, Manea 2008) note that the futurate PRES is used in RO and IT more often than in FR.

(352) Măine *plec* la Ploiești.
       Tomorrow I leave.PRES to Ploiești
       'Tomorrow I am leaving to Ploiești.'

(353) *Citesc* și eu toate textele primite.
       Read.PRES also I all received texts
       'I am reading too all the received texts.'

The HP occurs in informal conversation and in fiction (Zafiu 2013) and the past time reference is provided by the context, as shown by the incompatibility of the PRES with an indexical past time adverbial in (354). It can be given in temporal sequences the alternation with past time verbal tenses, as in (355), or by temporal adverbials whose past time interpretation is based on contextual knowledge, as in (356), from Zafiu (2013, 56).

(354) *Ieri *plec* la Ploiești.
       Yesterday leave.PRES to Ploiești
       'Yesterday I left to Ploiești.'

(355) Ieri *am fost* la Ploiești. *Am mers* cu trenul. În compartiment, *văd* o figură cunoscută.
       Yesterday go.PC to Ploiești. I go.PC by train. In the compartment, I see.PRES a familiar face
       'Yesterday I went to Ploiești. I went by train. In the compartment I saw a familiar face.'

(356) Cuza *moare* în 1873.
       Cuza die.PRES in 1873
       'Cuza died in 1873.'

In subordinate clauses, the PRES functions as an anaphorical tense: its reference moment is not S but the temporal interval expressed by the matrix verbal tense, as in the following two examples:

(357)   Acum un an mi-*a spus* că *se simte* cam obosit.
          A year ago he tell.PC that he feel.PRES rather tired
          'A year ago he told me that he was feeling rather tired.'

(358)   Vom vedea ce tren *luăm*.
          See.FUT which train take.PRES
          'We will see which train we take.'

To sum up, in RO similarly to FR and IT, information regarding Tense and Aspect is expressed in the same verbal form. There are analytic forms and synthetic forms. Tense expresses reference to past, present and future times, in other words, it locates an eventuality prior, simultaneous and posterior to the moment of speech S. The aspectual distinction perfective/imperfective is expressed only for past verbal tenses. Verbal tenses can be deictic, expressing the relation E/S, and/or anaphoric, expressing the relation E/R. In this thesis, I will be arguing that Tense and its expression by verbal tenses in RO, similarly to FR and IT, encode conceptual information (i.e. the relation E/S) and procedural information (the relation E/R). The four verbal tenses described in this section can coarsely be classified with respect to their conceptual information: the PC, PS and the IMP on the one hand, and the PRES on the other hand.

2.5.2 A cross-linguistic valid framework?

Monolingual descriptions of verbal tenses in EN, FR, IT and RO point out several problematic issues. The first is related to the dissimilar manner in which classical grammars and studies describe verbal tenses and their usages. For example, FR literature addresses verbal tenses in terms of their main and secondary usages, also called descriptive and interpretative usages. EN literature points to the reference of a verbal tense to past, present or future time and grants an essential role to aspectual information, such as the progressive morpheme *-ing* and lexical aspect (states vs. events). Moreover, IT and RO literature continue to describe verbal tenses in terms of their deictic and anaphoric usages.

The second issue regards the level of analysis adopted in classical and more recent research of verbal tenses in these four languages. Research on FR verbal tenses is clearly more advanced than that on all other three languages, consisting of points of view proposed in formal semantics, and semantico-pragmatic and procedural pragmatics frameworks.

The third issue is the use of the generic term *verbal tense* for referring to the categories of Tense and Aspect, which are applied to Aktionsart. Using a generic term is a source of confusion with respect to the precise origin of a certain type of information. For example, the IMP in Romance languages is considered as being an imperfective verbal tense, and therefore, it presents an event as being incomplete and the period when E holds is larger than the reference period (as discussed in section 2.5.1.2.2 for the FR IMP). The question that arises at this point is what is the source of this temporal information. The most probable answer is the category of Aspect. Consequently, this temporal information identified for the IMP and triggered by Aspect could be identified and analysed cross-linguistically even if the verbal form of the IMP does not exist, as is the case for Mandarin Chinese for example. My suggestion is that a cross-linguistically valid framework should primarily be based on meaning features and, only secondly, on formal features.

This dissimilar description is problematic for a systematic contrastive analysis of verbal tenses in EN, FR, IT and RO. According to the CA methodology, a neutral and cross-linguistically valid *tertium comparationis* is needed in order to compare verbal tenses and establish degrees of similarities and differences. To propose such a model, translation corpora were investigated and a series of theoretical assumptions currently suggested in the literature and their application to EN, FR, IT and RO data were tested experimentally. The results of the empirical research are provided in Chapters 5 and 6. The cross-linguistically valid model defended in this thesis, as well as a reanalysis of verbal tenses in EN, FR, IT and RO are provided in Chapter 7.

2.6 Conclusive remarks

This chapter has given an account of the semantics of temporal reference and its ingredients, as well as its expression in tenseless and tensed languages. Tensed languages can be classified in two main categories: tense-prominent languages (such as EN and Romance languages) and aspect-prominent languages (such as Slavic languages). The particularity of tense-prominent languages is that Aspect is not expressed morphologically, as is the case for Slavic languages (except for the EN *–ing*), and that a verbal tense represents a mixed form of Tense and Aspect. In tense-prominent languages, Tense therefore plays a primary role for expressing temporal reference, and it interacts with both Aspect and Aktionsart.

In this chapter, I have given an account of Tense, Aspect and Aktionsart in which they were considered individually and with respect to the role they play in the structuring of a discourse. The discussion was organised around semantic and discourse semantics theories. The various theories described in this chapter aimed at determining the *meaning* of verbal tenses, where meaning is determined at the level of a clause, disregarding contextual information. DRT and SDRT are discourse semantics theories investigating relations that link discourse segments. Discourse relations in these frameworks are semantic relations in that their meaning is compositional.

More recent pragmatic accounts, and specifically cognitive pragmatics, as well as accounts suggested in other frameworks (such as generative syntax, neurolinguistics, natural language processing and machine translation), will be addressed in Chapter 3. In the next chapter, it will be argued that a semantic approach of Tense, Aspect and Aktionsart is not appropriate for several reasons. The first is that elements such as the speaker's intention with respect to the temporal location of an eventuality and the speaker's perspective on the eventuality expressed should be integrated in the analysis. The second is the inferential nature of the *meaning* of a verbal tense, which is determined at the contextual level, making use of contextual information and world knowledge. The relevance theoretic suggestion that language, thus verbal tenses, underdetermine the speaker's communicated content at the level of both explicatures and implicatures is, in my view, an accurate approach for investigating verbal tenses. Consequently, I will be speaking about *usages* and not *meaning(s)* of a verbal tense. More specifically, it is argued in Moeschler et al. (1998) that a verbal tense has contextual usages. More recently, Moeschler et al. (2013) and Grisot and Moeschler (2014) argued that contextual usages are determined based on the contextual value of the conceptual and procedural information encoded by the verbal tense.

Chapter 3 will also incorporate neurolinguistic and syntactic views regarding Tense, Aspect and Aktionsart. It will be argued that conceptual distinction between pastness and non-pastness is neurolinguistically motivated, and that Tense and Aspect are contained in the Inflection-Phrase, which is the functional head of a sentence. Finally, I will address several computational models and the manner in which they treat temporal information at the discursive level.

3 Beyond the semantics of Tense, Aspect and Aktionsart

3.1 Cognitive pragmatic approaches

3.1.1 Introduction

The main assumption of pragmatic theories (Austin 1957; Searle 1969; Grice 1967; Sperber and Wilson 1986/1995, 2012; Horn 1984, 1992, 2004, 2006, 2007; Levinson 2000)[77] is that semantic interpretation of language is underspecified and it must be enriched through pragmatic process. The way in which these pragmatic processes work, the delimitation of the border between semantics and pragmatics as well as the terminology used depends on the different pragmatic trends. Tense, Aspect, Aktionsart, discourse markers and temporal adverbials expressing temporal relations are linguistic expressions that undergo a process of pragmatic enrichment. Temporal relations in discourse receive a different account depending on the pragmatic trend considered. Temporal relations are approached in the Gricean framework as conversational implicatures (see section 3.1.2) whereas in RT they are pragmatically determined aspects of what is said (Wilson and Sperber 1998) (see section 3.1.3). In RT, Tense is a procedural expression (vs. conceptual) (see section 3.1.3.2). In this thesis I adopt the relevance theoretic account of temporal relations and suggest a dual description of Tense, as encoding both conceptual and procedural information (see section 7.2).

3.1.2 Gricean account

Grice (1967, 1975, 1989) challenged the classical view that pragmatics is concerned only with the nonconventional or contextual meaning, such as irony and metaphor. He moved the focus from the *conventional* vs. *nonconventional* distinction to the *truth-conditional* vs. *non-truth-conditional* distinction. Truth-conditional meaning is expressed by *what is said* and belongs to the domain of semantics while non-truth-conditional meaning is expressed by *what is implicated* (i.e. implicature)[78] and belongs to the pragmatic domain. Grice sets thus a fixed

---

[77] The origin of pragmatics is Austin and Searle's theory of speech acts (Austin 1957; Searle 1969) and Grice's *William James Lectures* (1967). Speech act theory was developed as a formal theory of *illocutionary logic* (Searle and Vanderveken 1985). Grice's assumptions about language gave rise to two main pragmatic frameworks, the theory of *generalized conversational implicature* (Levinson 2000; Horn 2004) and Relevance Theory (Sperber and Wilson, 1986/1995). The former is called the neo-Gricean trend and the latter is called post-Gricean trend. Gricean pragmatics is based on the *principle of cooperation, conversation maxims* (quantity, quality, relation and manner) and the notion of *implicature*. Neo-Gricean pragmatics is a simplified version of Gricean pragmatics, based on two principles (the Q-principle recalling the maxim of quality and the I-principle or the maxim of minimization). Relevance Theory introduced the cognitive aspect in the study of language, suggesting that communication is an ostensive-inferential process involving the *principle of relevance*.

[78] Implicatures are a subtype of inferences. Inferences are logical processes where conclusions are derived from premises according to a set of rules. One can distinguish among logic, semantic and pragmatic inference. Logic inferences are valid if the truth of the premises entails the truth of the conclusion. Logic inferences are based on deductive or inductive rules. One example of semantic inferences is the semantic presupposition.

border between the two domains. On of the consequences of this position is that implicated meanings do not contribute to the truth-conditions of utterances. A first distinction proposed by Grice is between *conventional* implicatures and *conversational* implicatures. The schema in Figure 3-1 summarizes Grice's theory of meaning (Moeschler 2012a, 416, citing Sadock 1978, 283).

Figure 3-1 Types of implicatures
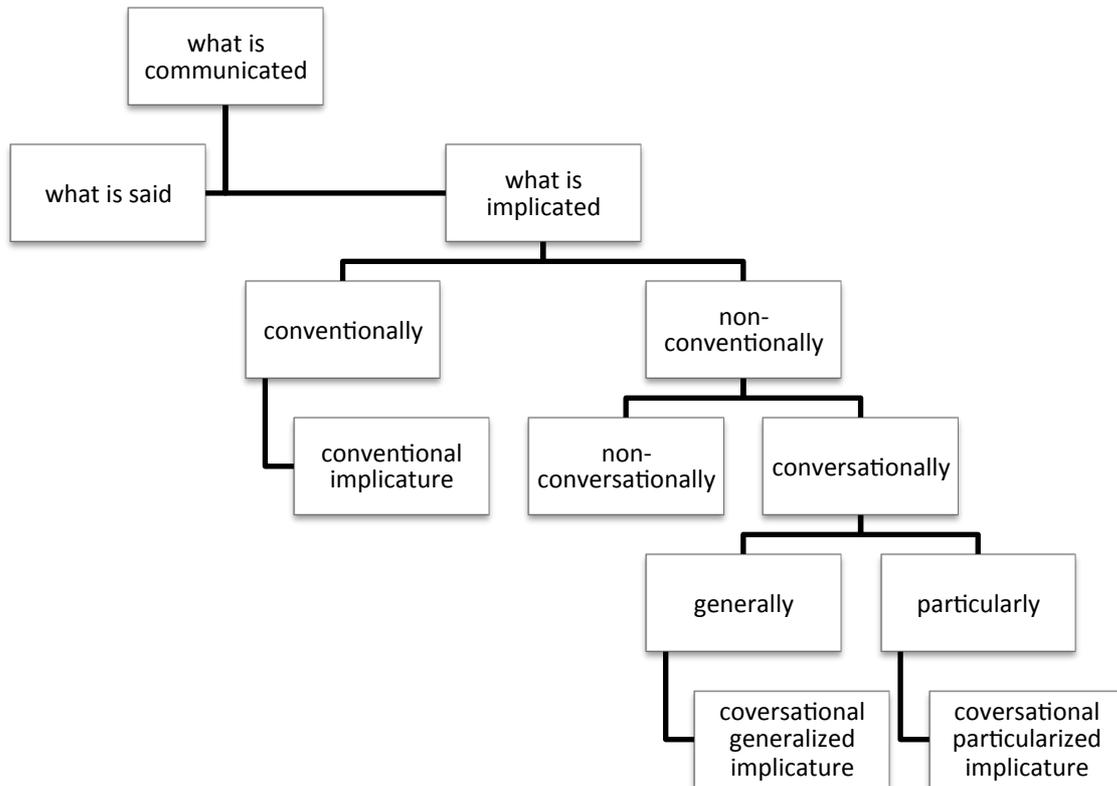
Conventional implicatures are triggered by specific expressions. In (359), the speaker implies that his friend and his colleagues are not honest people.

(359) How is your friend doing? Oh quite well, I think; he likes his colleagues, and he hasn't been to prison yet.

Conversational implicatures occur in discourse and are the result of the application of conversational maxims or the conversation principle. They can be triggered by specific words (i.e. *generalized conversational implicature*) as in (360), where the meaning of *and* is the temporal meaning 'and then', or not (i.e. *particularized conversational implicature*) as in (361), where the speaker implicates that the woman John is meeting is not his wife/sister.

(360) I took out the key and opened the door.
(361) John is meeting a woman tonight.

Semantic presuppositions can be defined as a statement Q presupposed by a statement P that is true in all cases when P is true and in all cases when P is false. Pragmatic presuppositions (type of pragmatic inferences), on the other hand, can be false and are non-truth-conditional aspects of the meaning. Gricean and neo-Gricean assume the inferences are based on inductive rules while post-Griceans (RT) adopt the idea that inferences are non-demonstrative abductive (the premises dot not guarantee the conclusion).

Grice (1975, 57-58) and Sadock (1978) propose a list of six criteria to test for conversational and conventional implicatures (see Moeschler 2012a, 416-417 for a detailed presentation of the six criteria). According to these criteria, conversational implicatures are *calculable* (originate in a working-out procedure), *cancellable*, *non-detachable*, *non-conventional*, *carried out not by what is said but by the speech act* and *indeterminate* (do not have a precise content attached). On the contrary, conventional implicatures are *non-calculable, non-cancellable, detachable, conventional, carried out by what is said* and *determinate.* According to Sadock (1978), and as pointed out by Moeschler, these conditions are neither necessary nor sufficient for testing for implicatures mainly because they are linked to each other.

As shown in (360), temporal (and causal) relations are interpreted as conversational implicatures[79] (the application of the orderliness maxim). Contrary to what has been suggested by Grice among others, temporal and causal relations in sentences (362)-(365) (Wilson and Sperber, 1998; Wilson 2011) are not necessarily triggered by the connective *and* since they arise whether it occurs or not in the sentence[80].

(362)   John dropped the glass and it broke.
(363)   John dropped the glass. It broke.
(364)   Peter left and Mary got angry.
(365)   Peter left. Mary got angry.

One of Grice's assumptions was that logic operators such as & 'and' and their correspondents in natural language (the connective *and*) are semantically equivalents. A consequence of this assumption is that if the order of the two conjuncts is reversed, the truth conditions of the utterance do not change. Cohen (1971) pointed out that Grice's treatment of temporal and causal relations as conversational implicatures (thus not-truth-conditional) is not appropriate. Sentence in (366) illustrates that the temporal ordering of the two eventualities is part of the truth-conditions of the utterance. And this is what makes the disjunction in (366) not redundant.

(366)   It's always the same at parties: either I get drunk and no-one will talk to me or no-one will talk to me and I will get drunk.

Based on these observations, RT (specifically, Carston 1988) proposes that temporal and causal interpretations of such sentences should be analysed as pragmatically determined aspects of *what is said*, hence encoded by linguistic expressions. In other words, there are

---

[79] I will not discuss the Neo-gricean account of temporal relations, which is similar to the Gricean one, specifically temporal relations being implicatures triggered by the connective *and* (Atlas & Levinson 1981, Levinson 1983, 1987, 1989; Horn 2004).

[80] Since temporal relations arise also in the absence of the connective *and,* I did not speak about it in this thesis. The reader may see Wilson (2011) and Blochowiak (2014, 2015b) for interesting discussions regarding the puzzles concerning the connective *and* defined as five types of problems: sequencing, interval, cause-consequence, unspecified sequence and Horn's problem. Grice's solution was for these temporal interpretation triggered by *and* was the maxim of orderliness. Dowty's solution (1986) was the Temporal Discourse Interpretation Principle. Blochowiak proposes a solution in the relevance theoretic framework by suggesting a more fine-grained notion of contextual assumptions (i.e. the Relevance Nomological Model, see Blochowiak 2014) and by discussing the usages of *and* with respect to two oppositions: *extensionality* vs. *intensionality* and *homogeneity* vs. *heterogeneity*.

aspects of what is linguistically encoded that are pragmatically determined. This remark is linked to the proposal in RT of the *inferential model* of communication (consisting of two phases: decoding and inference; see section 3.1.3.1) and to the conceptual vs. procedural distinction of types of encoded information (Blakemore 1987; see section 3.1.3.3). The following section, 3.1.3, is dedicated to a series of theoretical aspects defended in RT and to relevance theoretic accounts of verbal tenses and temporal relations.

3.1.3 Relevance theory

3.1.3.1   General remarks

Relevance Theory (Sperber and Wilson 1986/1995, 1998, 2004, 2012) is a cognitive pragmatic theory of language comprehension. The *cognitive* characterization is due to several hypotheses assumed by RT:

- processes implied in pragmatic interpretation are not specific to language but are localized in the central system of thought;
- the essential feature of human communication is the expression and recognition of intentions (Grice 1989);
- the search for relevance is a basic feature of human cognition.

The first hypothesis finds its roots in the theory of modularity of mind (Fodor 1983; Sperber 2005; Caruthers 2006). According to this theory, the mind is built of *modules* and the *central system of thought*. A module is a dissociable functional component with a specific job for the functioning of the whole system. Fodor's definition of a module (the so-called fodor-module) is *a domain-specific innately specified processing component, with it's own proprietary transducers, and delivering 'shallow' outputs* (as expressed by Caruthers 2006, 3). In other words, each module is restricted to the type of content it takes as input that constitutes its domain (for example, visual inputs for the visual system). Each module contains its specific transducers allowing the caption and the transport of the input. The outputs of modules are *shallow* in the sense that they generate *fragments of information* of various sorts, which will be used in further processing. For example, the visual module might deliver a fragmentary representation of a chair. The recognition of the object as a chair and the belief that the chair exists is due to further processing, specifically in the central system of thought (as suggested in Sperber and Wilson 1986). According to Fodor modules are innate and localized in specific structures in the brain (not always biologically localizable but rather distributed across a set of dispersed neural systems). Modules work automatically, in the sense that they are not under voluntary control, and their functioning is mandatory. They generate outputs extremely fast in comparison with other holistic or non-modular systems, such as for example the conscious decision-making system. As far as the processing of the module is concerned, it is described as encapsulated and inaccessible. The former means that the processing system of a module cannot take place using information existent outside of that system. It functions only using the domain-specific information of the module (visual input for the visual module, auditory input for the auditory module, etc.). The internal processing of a module is inaccessible to other modules. The only accessible component is the output.

The central system of thought is complex and nonspecialized, working out data received from different modules and reasoning. RT assumes that linguistic data is treated in a

specialized module providing fragmentary information that is the logic form of an utterance. Its output is transmitted as input to the central system of thought, where pragmatic processes take place in order to yield the full interpretation (I will come back later to what the full interpretation of the sentence is). According to RT, all modules are not involved in the same way in human communication (Wilson 2011). While certain modules are connected to communication only remotely such as the visual system, others play a crucial role in linguistic communication, such as those modules involved in inferential comprehension, social cognition, language production or parsing. RT's hypothesis concerning two types of encoded information by linguistic expressions (conceptual vs. procedural information) is based on Fodor's description of modules. According to RT, procedural expressions explicitly activate domain-specific procedures (i.e. concerns inputs of a particular sort) whereas conceptual expressions are linked to concepts, which are components of the language of thought (Fodor 1975, 1983).

The second hypothesis led to suggesting an inferential model of communication that included both the *code* model (since Aristotle it is believed that communication is achieved through coding and decoding messages) and the *inferential* model (according to Grice, communication is achieved by producing and interpreting evidence about speaker's intended meaning[81]).

Another of Grice's central claims recycled in RT is that utterances automatically create expectations that guide the hearer towards the speaker's intended meaning. Grice defines these expectations in terms of the *conversation maxims* and the *cooperative principle.* RT does not adopt Grice's maxims and the cooperative principle but make the hypothesis that "the expectations of relevance raised by an utterance are precise and predictable enough to guide the hearer towards the speaker's meaning" (Wilson and Sperber 2004, 607). Moreover, "utterances raise expectations of relevance… because the search for relevance is a basic feature of human cognition, which communicators may exploit" and this is expressed in RT through the *Cognitive Principle of Relevance:*

- ▪ Human cognition tends to be geared to the maximization of relevance " (Wilson and Sperber 2004, 610)

An input (a sight, a sound, an utterance, a memory) is relevant when it connects with existing or background knowledge in order to produce a *positive cognitive effect*, i.e. bringing new information, developing existing information or correcting existing information. These cognitive effects are positive if they help the hearer to create true representations of the world. There are numerous potentially relevant stimuli but humans will search for the most

---

[81] In RT language comprehension is a two-phase inferential process that includes *a decoding phase* (provides input to the next phase and takes place in the linguistic specialized module) and *a central inferential phase* in which a linguistically encoded logical form is contextually enriched and used to construct a hypothesis about the speakers' intention (in the central system of thought). RT makes a difference between two types of intention. The *informative* intention is the speaker's intention to bring to the audience of some piece of information (a set of assumptions) whereas communicative intention is the speaker's intention to make known her informative intention. Consequently, the hearer must identify the communicative intention in order to get to the communicative intention, and thus to the message the speaker wants to transmit. This happens through inferring the message based on the evidence the speaker provides. In RT, communication is thus *ostensive* and *inferential*: the speaker manifestly shows her communicate intention and the hearer makes an inference about the speaker's informative intention.

relevant stimulus. According to RT, in equal situations, the greater the positive cognitive effect achieved by processing an input, the greater its relevance will be. Relevance is thus weighed in terms of *cognitive effects* and *processing efforts:*

- Other things being equal, the greater the positive cognitive effect achieved by the least processing efforts, the greater the relevance of that input to the individual at that time.

RT argues that in communication, the speaker makes use of *ostensive stimuli* designed to attract the hearer's attention and to focus on the communicator's intended meaning. Ostensive stimuli generate precise and predictable expectations of relevance not raised by other inputs. Therefore, by using an ostensive stimulus the speaker encourages the hearer to presume that it is relevant enough to be worth processing. This hypothesis is expressed through the *Communicative Principle of Relevance* and, respectively, the *Presumption of optimal relevance*:

- Every ostensive stimulus conveys the presumption of its own optimal relevance. (Wilson and Sperber, 2004, 612)
- The ostensive stimulus is relevant enough to be worth the audience's processing effort. It is the most relevant one with communicator's abilities and preferences. (Wilson and Sperber, 2004, 612)

According to the RT, in the interpretation process, the hearer follows a path of least effort for finding the cognitive effects needed. The interpretation process stops when the hearer's expectations of relevance are satisfied. It consists of taking into consideration *what is said* and *what is implicated* (as suggested by Grice). For Grice, the explicit/implicit distinction refers to the difference between the truth-conditional content of utterance and the non-truth-conditional content, where the latter depends solely on pragmatics. Relevance theoretic framework assumes a quite different position. There are two kinds of assumptions communicated by a speaker: explicatures and implicatures defined as it follows (Carston 2004, 635, citing Sperber and Wilson 1986).

- An assumption communicated by an utterance U is explicit [hence an explicature] if and only if it is a development of a logical form encoded by U. [in case of ambiguity, a surface form encodes more than one logical form]
- An assumption communicated by U which is not explicit is implicit [hence an implicature]

Explicatures are developments of the logical form through decoding and through pragmatic enrichment into a full propositional form of the utterance, as in (368), which is the explicature of (367). The star assigned to the word Mary indicates that a particular referent has been assigned to the name "Mary". The explicature consists of more precise and elaborated information, such as reference assignment, narrowing of concepts *get* and *unit*, enrichment of the meaning of words like *enough*, and adding the cause-consequence relation between the two segments[82]. On the other hand, the sentence (369) is an independent assumption inferred as a whole from (368) and a further premise concerning the relation between Mary's recent failure at university and her current state of mind (Carston 2004).

---

[82] This is an important point of divergence between relevance theorist and neo-gricean pragmatic frameworks. Neo-griceans have followed Grice in considering these aspects of communicated meaning as implicatures (see Carston 2004, sections 4-6).

(367) X: How is Mary feeling after her first year at university?
        Y: She didn't get enough units and can't continue.

(368) Mary* did not pass enough university course units to qualify for admission to second-year study, and as a result, Mary* cannot continue with university study.

(369) Mary* is not feeling very happy.

As Moeschler (2012a) points out, implicatures have a very different status in RT than in Grice's theory and in neo-gricean theories[83]. Their principal characteristic is that they can be false, since they are formulated on contextual premises. Implicatures are defined in RT as a type of cognitive effect (i.e. addition of new information, strengthening of an old information, suppression of an old information) and are the result of *non-demonstrative inferences* (whose premises are the logical form of the utterance and a set of contextual assumptions). The results of these inferences are called *implicated conclusion* and they arise during a general procedure of comprehension, as stated in Wilson and Sperber (2004, 613):

- Follow a path of least effort in computing cognitive effects: test interpretative hypotheses (…) in order of accessibility.
- Stop when your expectations of relevance are satisfied (or abandoned).

According to RT, the interpretation process[84] consists of several subtasks that take place in parallel. The logical form encoded by an utterance containing incomplete conceptual representations is treated in the inferential process in three ways (Wilson and Sperber 2004, 615):

- Constructing an appropriate hypothesis about explicit content (explicatures) via decoding, disambiguation, reference resolution and other pragmatic enrichment processes (narrowing, loosening, saturation, free enrichment, ad-hoc concept construction[85])
- Constructing an appropriate hypothesis about the intended contextual assumptions (implicated premises)

---

[83] RT implicatures belong to Gricean notion of particularized conversational implicatures, whereas neo-griceans mostly worked on generalized conversational implicatures (Moeschler 2012a, 425).

[84] RT is a theoretical model for language interpretation (i.e. recuperating the speaker's intended meaning). It is an experimentally testable model, and there are numerous studies in experimental pragmatics that present offline and online experiments targeting hypotheses developed in the RT framework. As far as production is concerned, answers must be found in psycholinguistics and neurolinguistics.

[85] Carston (2004) discusses the pragmatics aspects of explicatures (pragmatically determined aspect of *what is said*). Disambiguation concerns the selection of sense for polysemantic words (such as *bank*), the candidates being supplied by the linguistic system itself. Reference resolution concerns referent assignment to deictics, overt indexicals and referential expressions. Saturation concerns pragmatic developments of the logical form required by covert indexicals (such as *better, same, too, enough*) and is under linguistic control. Free enrichment is not triggered by a linguistic expression and it concerns aspects of the interpretation of the utterance that are relevant for the implicatures. The utterance *I've had a shower* contains the idea of *today* that comes through free enrichment and which is considered in the implicature *I don't need to have another shower now/today*. In neo-Gricean pragmatics, these aspects of utterance meaning are generalized conversational implicatures. Ad-hoc concept construction concerns the pragmatic adjustment of the concepts encoded in the utterance. The utterance *He was upset but he was not upset* said by the defence lawyer of a man who murdered his wife, is not a contradiction. The hearer understands that the man was upset but not upset to the point to kill his wife. The two interpretations of *upset* correspond to two concepts of upsetness resulted through narrowing from the ad-hoc concept UPSET. In this thesis, I use the notion of ad-hoc concept for interpreting Tense (see section 7.2.1).

- Constructing an appropriate hypothesis about the intended contextual implications (implicated conclusions)

Wilson and Sperber point that there is no sequential order in which these subtasks of the comprehension process take place due to the fact that comprehension is an online process. They take place in parallel and the resulting hypotheses are, if needed, revised or elaborated as the utterance unfolds. Thus, explicatures and implicatures (consisting of implicit premises and conclusions) are constructed through a process of "mutual parallel adjustment with hypotheses about both being considered in order of accessibility" (Wilson and Sperber 2004, 617).

Implicatures in RT are of two types: *weak* and *strong* implicatures. Strong implicatures are determinate in content (opposed to the notion of conversational implicatures which do not have attached a precise content) as in example (370) from Moeschler (2012a, 425), where Axel implicates that he does not want to brush his teeth and then go to bed. Weak implicatures, on the contrary, are less determinate, and are left to the responsibility of the hearer. The hearer may draw an important number of weak implicatures.

(370)   Jacques: Axel, please go and brush your teeth.
        Axel: Dad, I'm not sleepy.

Finally, as far as the semantics/pragmatics division of the content of an utterance, RT defends the position that some types of pragmatic meaning (such as the temporal interpretation of *and*) are truth-conditional. Precisely, they are part of the explicature of an utterance as shown by Wilson and Sperber (1998). To sum up, what Grice called *generalized conversational implicatures* and *conventional implicatures* are in RT defined as being part of the lexicon, and are not the result of any particular contextual device (Moeschler 2012a, 427). I will come back to this in sections 3.1.3.2 and 3.1.3.3, where I will discuss Moeschler's (2012a, 2012b, 2013, 2015) fine-grained distinction of semantic and pragmatic levels of meaning (section 3.1.3.2) and the relevance-theoretic *conceptual/procedural* distinction (section 3.1.3.3).

3.1.3.2   Levels of meaning

In recent work Moeschler (2012a, 2012b, 2013, 2015a, 2015b, 2015c) has argued that meaning contains complex layer representations including semantically determined meaning (that is presupposition and entailment) and a set of pragmatically determined meanings (that is explicature and implicatures) (2013, 88).

Moeschler (2015c) proposes the following criteria to distinguish between semantic presuppositions and implicatures (arguing that pragmatic presuppositions are actually implicatures): *explicitness, inferentiality, context-dependence, truth-functionality, speaker's commitment* and *behaviour with negation*.

Entailments and presuppositions

Moeschler (2013) provides the classical definitions of entailment and semantic presupposition:

- P entails Q
  - If P is true, then Q is true.
  - If P is false, then Q is true or false
- P presupposes Q:
  - Iff P entails Q
  - Iff not-P entails Q

Example (371) shows that if P is true, then Q is true. Examples (372) and (373) show that if P is false, then Q can be true (i.e. he bought another breed of dog) or, respectively false (i.e. he bought a cat) from Moeschler (2013, 88).

(371)　Nath has a Chow. He bough a new dog.
(372)　Nath does not have a Chow; he bought a Labrador.
(373)　Nath does not have a Chow; he bought a cat.

As for presupposition, examples (374) and (375) show that if P is true or respectively, false, then Q is true, where Q given in (376) is a presupposition.

(374)　Sam knows that he failed.
(375)　Sam does not know that he failed.
(376)　Sam failed.

Moreover, entailments and presuppositions are semantic contents that can neither be made explicit as in (377) and (379) nor denied under descriptive negation as in (378) and (380) from Moeschler (2013, 89).

(377)　?Nath has a Chow, so he has a dog.
(378)　??Nath has a Chow, but he has no dog.
(379)　?My daughter is in Japan, so I have a daughter.
(380)　??My daughter is in Japan, but I have no daughter.

To sum up, presupposition and entailment are truth-functional, not explicit and not inferential. Presupposition may be contextually dependent whereas entailment is not context dependent, i.e. it does not require contextual premises. They both involve the speaker's commitment because they cannot be denied without producing a contradiction.

Explicatures

In RT, explicatures correspond to Grice's category of generalized conversational implicatures[86] (called *informational amplifications of utterances* by Levinson 2000). Explicatures are enriched forms of the propositional contents and are truth-functional (Sperber and Wilson 1986; Wilson and Sperber 2004; Cartson 2002): an explicature is true or false iff the

---

[86] As I will be arguing in this thesis, temporal and causal relations holding among eventualities treated as generalized conversational implicatures by Grice are considered as explicatures triggered by procedural information encoded by Tense (also Wilson and Sperber 1998). As for scalar implicatures, treated as generalized conversational implicatures by Grice, they are considered by Moeschler (2013) as implicatures.

proposition expressed by the utterance is true or false. The main idea is that a proposition cannot be true when its explicature is false, and the explicature cannot be true when the proposition is false.

Explicatures, contrary to entailments and presuppositions, can be made explicit. This happens either as basic explicatures as in (382), which is the development of (381) enriched with a temporal variable, or as higher-order explicatures as in (383).

(381)   It's raining.
(382)   It's raining, I mean, right now.
(383)   Can you take down the garbage? It's not a question; it's an order.

Explicatures are be negated as in (384). However, if the explicature of an utterance is made explicit, it cannot be further denied, as in (385).

(384)   It is raining but not right now.
(385)   *It's raining I mean, right now but not right now.

To sum up, explicatures represent explicit content, inferential and truth-functional. They are context-dependent requiring contextual premises for their computation and are pragmatically determined. They involve the speaker's commitment because they cannot be denied once they have been made explicit.

Implicatures

In RT, implicatures correspond to Grice's category of particularized conversational implicatures. According to Moeschler (2013, 90), implicatures can be true or false if the utterance is true, and true or false if the utterance is false. They do not have truth-functional contents, are different than presuppositions and are cancellable content.

For example, scalar implicatures are triggered by quantitative scales, such as *<4 children, 3 children, 2 children>*. In such a scale, two relations can be identified: a semantic relation of entailment ENT: *3 children* entails *2 children* and a pragmatic relation of scalar implicature SI: *3 children* implicates *not 4 children.* The assertion in (386) gives the entailment in (387) and the scalar implicature in (388) from Moeschler (2013, 91). The scalar implicature can be cancelled without contradiction as in (389) and (390).

(386)   Anne has three children.
(387)   Anne has two children.
(388)   Anne does not have four children.
(389)   Anne does not have three children, she has four.
(390)   Anne has three children, and even four.

The negation used in (389) to negate the assertion is a *metalinguistic* negation, shown by the fact that the scalar implicature in (388) is false because it is contradictory to the entailment in (387). Metalinguistic negation[87] requires a corrective clause whereas *descriptive* negation

---

[87] For more technical details (such as scope) regarding the difference between descriptive and metalinguistic negation, see Moeschler (2013).

(classical negation) does not, as shown by the contrast between (391) and (392).

(391)   Anne does not have three children (she has two).
(392)   Anne does not have three children * (she has four).

Metalinguistic negation also applies to presupposition. The presupposition cannot be false under the descriptive negation as in (393) but it can be false under metalinguistic negation, as in (394).

(393)   Mary did not stop smoking [presupposes: Mary smokes]
(394)   Mary did not stop smoking, since she never smoked.

In other words, metalinguistic negation does not have the same pragmatic effects when it cancels a presupposition or an implicature. Therefore, Moeschler suggests dividing metalinguistic negation in two types, which gives the following simplified categorization of negation (2013, 94):

- Descriptive negation: COR[88] entails NEG
- Metalinguistic negation 1: COR entails NEG (P) & NEG (Q)
- Metalinguistic negation 2: COR entails POS

This means that with metalinguistic negation 1, both the assertion P and the presupposition Q are negated, whereas with metalinguistic negation 2, both the assertion and the implicature are cancelled and both positive contents are entailed.

To sum up, implicatures are not explicit content, are inferential, context-dependent and pragmatically determined. They are not truth–functional and can be negated under metalinguistic negation 2. They do not therefore involve the speaker's commitment. Entailments cannot be negated whereas explicatures can be negated under descriptive negation. Finally, presuppositions can be negated under metalinguistic negation 1.

In conclusion, meaning contains complex layer representations ranging from semantic and logical entailment and presuppositions to pragmatically determined explicatures and implicatures. Using the procedural/conceptual distinction (suggested in RT by Blakemore 1987, see section 3.1.3.3), Moeschler (2015) suggests, in relation to connectives, that conceptual meaning is associated to logical entailments whereas procedural information is activated at two levels: explicatures (and therefore truth-conditional) and implicatures (and therefore non truth-conditional).

Table 3-1 summarizes the different layers of meaning as discussed in this section, starting from Grice's distinctions and as they were developed in a relevance-theoretic framework (Sperber and Wilson 2004; and mainly in Moeschler's recent work: 2012a, 2012b, 2013, 2015a, 2015b, 2015c).

---

[88] POS stands for the positive counterpart of the negative sentence NEG and COR is the corrective sentence.

Table 3-1 Overview of levels of meaning

Grice	RT (2004)	Conceptual/ procedural information	Inferential status	Truth-functionality status	Test
'What is said'	Entailment	Conceptual: logical entailment	Not inferential	Truth-functional	Does not apply
Conventional implicature	Presupposition	Semantic presupposition	Not inferential	Truth-functional	MetNeg1 COR -> NEG(P)& NEG(Q)
Generalized conversational implicature	Explicature	Procedural 1	Inferential	Truth-functional	Descriptive negation COR -> NEG
Generalized and particularized conversational implicature	Implicature	Procedural 2	Inferential	Non truth-conditional	MetNeg2 COR -> POS
Non-conventional implicature	General inference	General inference	Inferential	Non truth-conditional	/

I will make reference to these levels of meaning with respect to temporal reference and its components in my reanalysis developed in Chapter 7.

3.1.3.3   Conceptual vs. Procedural information

As Escandell-Vidal et al. (2011) argue, the conceptual/procedural distinction was first meant as a solution for the semantics/pragmatics division of labour and it has remained an important explanation for the contribution of linguistic meaning to utterance interpretation. A speaker is not expected to render more difficult than necessary his addressee's task in obtaining a relevant interpretation. Therefore, procedural meanings are instructions encoded by linguistic expressions that specify paths to follow during the interpretation process involving manipulating of conceptual representations in order to access the most relevant context. Saussure (2011, 61-62) points out that procedural expressions encode *specific* paths to follow for obtaining *specific* inferences. The first consequence is that it is not impossible for the hearer to get to the intended inference in the absence of the procedural expression but this would happen (though it is not guaranteed) at a higher cognitive cost. The second consequence is that there also exist more *general* inferences that are not encoded specifically by linguistic expressions. This is the case of inferences obtained through general ways of pragmatic reasoning starting from conceptual-encyclopaedic information.

Since its proposal (see Blakemore 1987[89], 2002; Wilson and Sperber 1993), conceptual and procedural information represent semantic information encoded by some types of linguistic expressions. The topic of discourse connectives was and remained a rich source for the contemporary research on the conceptual/procedural distinction: Blamekore (1987) on *but*, Blakemore (1988) on *so*, Blakemore (2000) on *nevertheless* and *but*, Blass (1989) on several

---

[89] In the French literature, a very influential work was that of Ducrot (notably Ascombre and Ducrot 1983) who suggested similar ideas in the framework of argumentation and polyphony. Ducrot spoke about *instructional expressions* (such as *puisque* 'since' and *mais* 'but') and his model aimed at modeling their argumentative function.

particles in Sissala, Ifantidou (2000) on the Greek particle *taha*, Moeschler (2002) on French *et* "and" and *parce que* "because", Zufferey (2012) on French *puisque, parce que* and *car* "because" among many others. Other phenomena that have been investigated with respect to the conceptual/procedural distinction and their role for discourse processing are mood and modality (Wilson and Sperber 1988; Ifantidou 2001 on evidentials; Ahern 2010 on speaker attitude) and verbal tenses (Ahern and Leonetti 2004 for the Spanish subjective; Nicolle 1997, 1998; Moeschler et al. 1998, 2012; Saussure 2003, 2013 for FR verbal tenses; Leonetti and Escandell-Vidal 2003 on the Spanish imperfective; Aménos-Pons 2010 on Spanish past tenses, 2011) to name but a few.

Many works are dedicated to the conceptual/procedural distinction from a theoretical point of view, which aimed at defining the two types of information and at proposing qualitative features. Wilson and Sperber (1993, 151) argue that conceptually encoded information contributes either to explicatures (to the proposition expressed and to higher-level explicatures) or to implicatures while procedurally encoded information represents constraints either on explicatures (to the proposition expressed and to high-level explicatures) or on implicatures, as shown in Figure 3-2 (Wilson and Sperber, 1993).

Figure 3-2 Types of information conveyed by an utterance
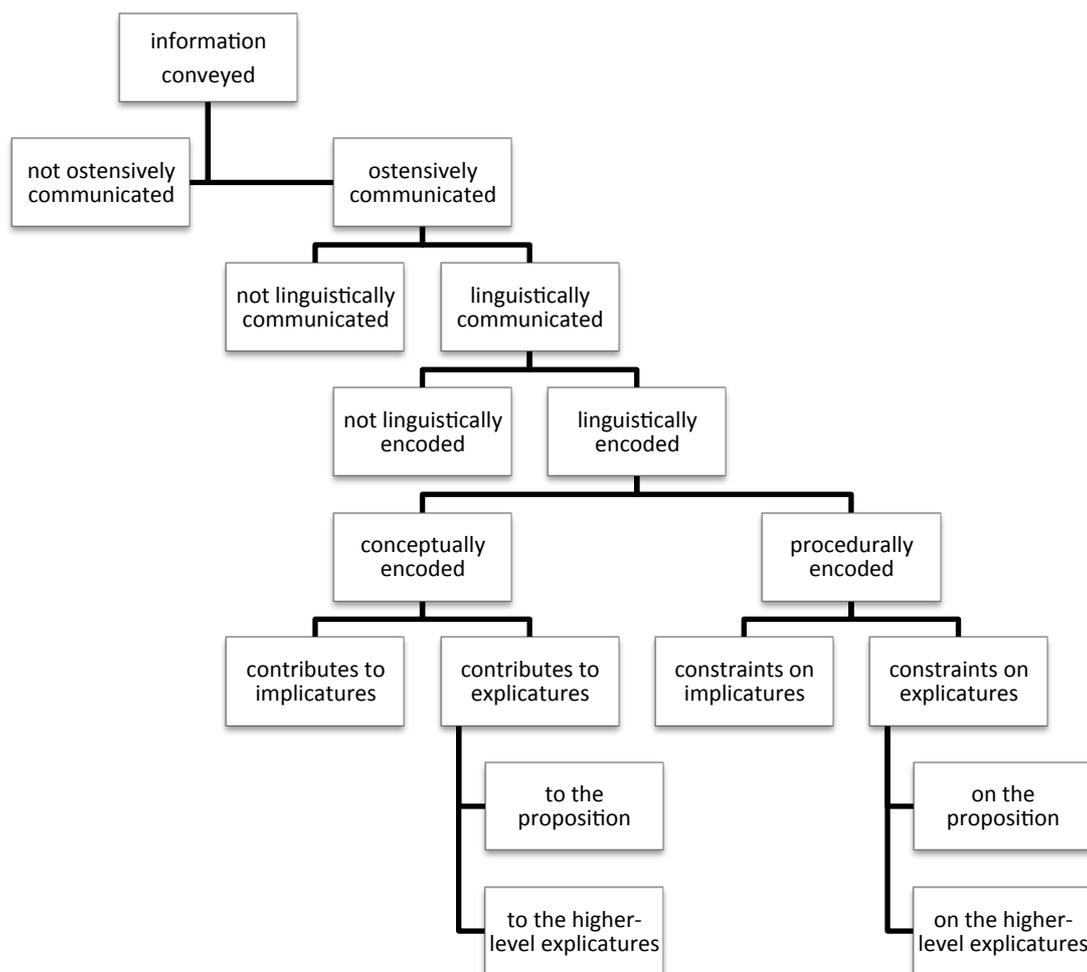
Figure 3-2 presents, in a schematic manner, a number of the key notions the relevance-theoretic account of meaning. The ostensive character of communication is related to the

notion of relevance. Some messages or part of messages can be communicated through paralinguistic clues as tone of voice or non-linguistic cues as facial expression. They are not linguistically encoded. This was Grice's main idea when he suggested the distinction between *what is said* (linguistically encoded) and *what is implicated* (not linguistically encoded). RT challenged this strict distinction assuming that some aspects of what is said are pragmatically determined through a combination of decoding and inference. For example, temporal and causal implicatures of an utterance (called *connotations* by Wilson and Sperber 1993) connected or not by the connective *and* are not treated as implicatures in RT but as pragmatically determined aspects of the proposition, contributing to its truth-conditions and falling under the scope of logical operators and connectives.

Linguistically encoded information is of two types: procedural and conceptual. The main idea is that during the interpretation process, the hearer builds conceptual representations and uses encoded procedures for manipulating them. A conceptual representation differs from other types of representations in that it has logical properties and truth-conditional properties. Sentence in (395) has the logical form (396) and the propositional form (397). Wilson and Sperber (1993) argue that the logical form recovered through decoding and the propositional form recovered by a combination of decoding and inference are conceptual representations.

(395)   Peter told Mary that he was tired.
(396)   x told y at t_i that z was tired at t_i.
(397)   Peter Brown told Mary Green at 3.00 pm on June 23 1992 that Peter Brown was tired at 3.00 pm on June 23 1992.

As far as procedural information is concerned, Wilson and Sperber (1993) argue that it represents constraints on the inferential phase of comprehension, as in example (398), which can be interpreted as in (399) and in (400). Wilson and Sperber (1993, 158) argue quoting Blakemore (1987, 1992) that the connectives *so* and *after all* do not contribute to the truth conditions the utterances, but constrain the inferential phase of comprehension by indicating the type of inference the hearer is expected to go through.

(398)   Peter's not stupid. He can find his own way home.
(399)   Peter's not stupid; so he can find his own way home.
(400)   Peter's not stupid; after all he can find his own way home.

It seems that the key idea to distinguish the two types of information is *contribute* (conceptual) or *constrain* (procedural) the construction of explicatures and of implicatures. Unfortunately, these two notions are quite vague and cannot easily be used as discriminating criteria. A first attempt has been made to use the contributing/constraining distinction in relation to the truth-conditional evaluation of a proposition. But the picture is not black and white: Wilson and Sperber (1993) distinguish four possible combinations: (a) conceptual and truth-conditional (most regular content words such as manner adverbials *seriously* and *frankly*), (b) conceptual and non-truth-conditional (illocutionary adverbials such as *seriously, frankly, unfortunately*), (c) procedural and non-truth-conditional (discourse connectives *so* and *after all*) and (d) procedural and truth-conditional (personal pronouns *I* and *you*). Discourse connectives constrain the construction of implicatures by guiding the search of intended

contexts and contextual effects. Pronouns impose constraints on explicatures by guiding the search for the intended referent appertaining to the proposition expressed. As far as conceptual information is concerned, it can contribute or not to the truth-conditions of the proposition expressed. They can have though their own truth-conditions (be false or true), they can therefore be contradicted, negated and used in entailment.

Another attempt to explain the contributing and constraining notions was to explain them relating to their function for determining the intended inferences. Saussure (2011) points out that procedural information encoded by expressions such as *but* (Blakemore 1987) linking two propositions P and Q excludes a variety of possible inferences that can hold between two P and Q, and guides the hearer toward the intended specific inference. It is in this way that procedural information constraints the inferential phase of communication and achieves better relevance by eliminating the unintended potential interpretations. Conceptual information on the other hand, through the rich encyclopaedic entry, opens a large array of possible assumptions, therefore contributing to the inferred premises and conclusions grasped through general inference. Though this account of the contribute/constrain division is interesting from a theoretical point of view, it is limited to using it as a discriminating criterion. My suggestion is that quantitative measures that allow objective comparisons and evaluations are needed. I will come back to this later.

The classical view on the conceptual/procedural distinction assumes that there is a clear-cut distinction between what is conceptual and what is procedural. This led to the assumption that there is a perfect mapping between conceptual/procedural information and lexical/grammatical categories. It is assumed that lexical categories (nouns, verbs, adjectives bearing descriptive content, *-ly* adverbs) encode concepts, whereas grammatical, or functional, categories encode various kinds of constraints on inferential processes. Nicolle (1997) argued against this position suggesting that a single expression can encode both procedural and conceptual meaning. Escandell-Vidal, Leonetti and Ahern (2011, 24) note that this could be the case for third person pronouns (*he, she*) encode the instruction to identify a highly accessible referent (Ariel 1994) but also include some conceptual information about the referent, such as male/female and animate. Hence, the accessibility requirement is common for the whole class of third person pronouns, whilst the conceptual information varies from pronoun to pronoun. For example, the case of the pronoun *it* remains problematic for this approach because it can refer both to objects and animate beings, without discriminating the gender such as for *dog* or for *baby*. Moreover, the gender distinction refers to grammatical gender and not to actual gender as illustrated by the cross-linguistic difference between *the people* in EN and *la gente* (feminine) in Italian[90].

Moeschler (2015) argues, with respect to connectives, that they are complex linguistic units conveying both conceptual and procedural information. He illustrates through an analysis of close connectives in FR (*parce que* 'because', *donc* 'therefore' and *et* 'and') that they have conceptual and procedural contents, which trigger different levels of meaning. More precisely, they all share a causal conceptual content however the set of entailments are not identical: P and Q for *parce que* and *et*, and only P for *donc*. Moreover, the causal meaning is an explicature (not defeasible) with *parce que*, and an implicature (defeasible) with *et* and *donc*.

---

[90] As pointed out by Louis de Saussure (2015) during the defence of this thesis (2015).

In the last twenty years, there has been a growing interest in establishing discriminating features of procedural in opposition to conceptual information and in applying the distinction to linguistic expressions. The first attempts to define and characterize conceptual vs. procedural information included qualitative features such as truth-conditional vs. non truth-conditional, representational vs. computational. One very significant contribution to the discussion is Sperber and Wilson's (1993) hypothesis of the cognitive foundations of the distinction. They characterise conceptual vs. procedural information in terms of accessible to consciousness vs. inaccessible to consciousness, easily graspable concepts vs. resistant to conceptualization, capable of being reflected on vs. not available through conscious thought (Wilson and Sperber 1993, Wilson 2011). These features of conceptual and procedural information find their roots, on one hand, in the parallel that has been made between natural language and the 'language of thought' and, on the other hand, in the 'massive modularity hypothesis' (Sperber 2005; Carruthers 2006). Sperber and Wilson (1998, 172-173) suggest that the constituents of a language are systematically put in relation with other objects, such as constituents of other language, with states of the user of the language, or with possible states of the world. Based on these remarks, Wilson (2011, 10) indicates that:

- Conceptual expressions in natural language are systemically linked to concepts, which are constituents of language of thought;
- Procedural expressions in natural language are systematically linked to states of language users;
- Sentences of the language of thought are systematically linked to possible states of the world;

Wilson argues that according to the second hypothesis, procedural expressions have the function to put the user of the language *into a state in which some of these domain-specific cognitive procedures are highly activated* (2011, 11). The output of the high-activated procedures functions as ostensive cues being more likely to be selected by the hearer during the comprehension process. In Wilson's words, expressions like *dog* or *think* encode conceptual representations (constituents of language of thought), which are accessible to consciousness and capable of being reflected on, evaluated and used in general inference. By contrast, procedural expressions such as *but* and *or* activate domain-specific procedures belonging to fodor-modules (encapsulated and inaccessible), are inaccessible to consciousness and resistant to conceptualisation.

Other studies (such as the papers published in the collection edited by Escandell-Vidal and colleagues, 2011) focused on procedural information regarding discriminating features and criteria, its status as semantic or pragmatic information and its application to reference, tense, modality, discourse markers and prosody. Saussure (2011) proposes a methodological criterion to distinguish between what is conceptual and what is procedural. In his words, an expression is procedural as it triggers inferences that cannot be predicted on the basis of a conceptual core to which general pragmatic inferences (loosening and narrowing) are applied. In his view, expressions that encode (at least apparently) both procedural and conceptual information (such as third personal pronouns, verbal tenses and some FR pragmatic connectives as *ensuite* 'then') should be considered as procedural. He argues that procedural information:

> *… either takes conceptual information as a parameter* as with *she*, and therefore *the conceptual information is simply under the dependence of the procedure,* or the conceptual meaning has no motivation anymore and is just a relic of ancient versions of that word (the *of ensuite*) (p. 65, original italics).

Escandell-Vidal and Leonetti (2011) propose rigidity as the major feature of procedural information. Their hypothesis is that conceptual information is flexible while procedural information is rigid.

> …Conceptual representations are flexible and malleable, which means that they can be enriched, elaborated on and adjusted in different ways to meet the expectations of relevance. (…) We claim that instructions, in contrast, are rigid: they cannot enter into the mutual adjustment process, nor can they be modulated to comply with the requirements of conceptual representations, either linguistically communicated or not. The instructions encoded by an item must be satisfied at any cost for the interpretation to succeed. (Escandell-Vidal and Leonetti, 2011, 86)

This idea was also suggested for FR markers expressing temporality by Moeschler (2000a, 2003), arguing that procedural information, encoded by temporal connectives for example, is stronger than conceptual information, encoded by aspectual classes (i.e. Aktionsart) for example.

Escandell-Vidal and Leonetti argue that despite the rigid character, instructions can give rise to a series of different interpretative effects. This is due to the different contextual assumptions and the data varying from one context to another. The main consequence is that cases of mismatch between the information coming from conceptual and procedural content will be solved following the procedural constraints on interpretation.

Moeschler et al. (2012) suggest that another feature of conceptual vs. procedural information is easily translatable vs. translatable with difficulty. Their suggestion is based on the assumption that conceptual information encoded by linguistic expressions is linked to conceptual representations, which are constituents of the language of thought. As language of thought exists beyond specific languages, this should facilitate the translation from one language to another. Procedural information encoded by linguistic expressions is systematically put in correspondence with states of the language user, where domain-specific procedures may be more or less highly activated in different circumstances. This user-specific character of procedural information is expected to produce difficulties in translation.

Escandell-Vidal and Leonetti (2011, 84-85) suggested a series of theoretical assumptions about procedural information, as it follows:

- Instructions are operational: they specify a set of algorithms or logical operations.
- Instructions operate on conceptual representations.
- Instructions can operate at two different levels: that of syntactic computation and that of interpretation.
- Linguistic items can encode concepts and instructions. Conceptual representations are linked to encyclopaedic knowledge, but instructional meanings lack such connections.
- A strong connection was initially established between the lexical vs. functional (i.e. grammatical) distinction and the conceptual/procedural distinction. Recent work suggests, however, that functional words can also encode conceptual information (such as connectives as suggested by Wilson 2015 and Blochowiak 2014, 2015a; and

verbal tenses as suggested by Grisot and Moeschler 2014, and this thesis).

- The conceptual/procedural distinction regards encoded meaning and it is therefore a semantic distinction. Precisely, it concerns pragmatically determined aspects of *what is said*.

The description of conceptual and procedural information in these terms is without any doubt accurate and reflecting the linguistic and cognitive reality of language users. Despite these numerous qualitative features of conceptual/procedural information, no study has suggested, as least as far as I am aware of, a description in quantitative terms. I think that quantitative criteria would have at least one advantage, that of objective comparison:

- comparison among expressions in order to know, for example, if one is more accessible to consciousness than another, and if it is, how much (in other words, if there is a black-and-white distinction or a continuum[91])

- comparison among expressions of the same type (procedural or conceptual), such as the FR *parce que* and *puisque* ('because' or 'since') in order to know if the activate similar or even identic procedures and to what degree of activation (i.e. low-medium-high)

- comparison among studies that investigated the same linguistic expression in order to evaluate them on other basis than whether or not the reader found the explanation plausible.

In this thesis, I suggest two quantitative measures that might be useful for evaluating conceptual and procedural information encoded by linguistic expressions. They are established based on the previous qualitative features suggested in the literature and discussed in this section.

The first quantitative measure is *inter-judge agreement rate* measured with the Kappa coefficient[92] (Carletta 1996) and it is used in this research for investigating the conceptual and procedural content of verbal tenses. My hypothesis is that evaluating conceptual information (which is accessible to consciousness, capable of being reflected on and evaluated and used in general inference, as described by Wilson and Sperber 1993) results in high values of the Kappa coefficient reflecting an important agreement between judges, other than that expected to occur by chance. On the contrary, evaluating procedural information (which is inaccessible to consciousness, resistant to conceptualization and not available through conscious thought) triggers low values of the Kappa coefficient (see more on general methodological issues in section 4.3.2 and for the its use in this thesis in Chapter 6).

The second measure is the amount the variability observed in translation corpora. This quantitative feature makes use of Moeschler's et al. (2012) suggestion that conceptual information is easily translatable whereas procedural information is translatable with difficulty. This suggestion is linked to the fact that conceptual information represents concepts that are constituents of the language of thought and therefore, language-

---

[91] See Moeschler (2002a, 2002b) for a proposition in this vein. He suggests that linguistic expressions can be of several types regarding the conceptual/procedural distinction: only procedural, only conceptual, averagely procedural (and averagely conceptual), strongly conceptual (and weakly procedural) and, finally, strongly procedural (and weakly conceptual).

[92] The Kappa coefficient is a statistic measure that corrects for expected chance agreement between two judges when confronted to judgments of sentences according to a given set of guidelines.

independent. Based on this observation, it is expected that translating conceptual information triggers reduced variability in the target language(s) whereas translating procedural information triggers high variability. In this thesis, this criterion was applied to the translation of the SP into three Romance languages (see section 5.3 for the results of empirical work) and it was found that translating conceptual information past vs. non-past triggers little cross-linguistic variation: in more than 70% of the cases, verbal tenses expressing past time are used in the target languages whereas the PRES is used in less than 8%. Besides, translating procedural information encoded by Tense and Aspect triggers important cross-linguistic variation: verbal tenses expressing past time and presenting different combinations of procedural types of information, that is IMP, PC and PS, are used with comparable frequencies (see Chapter 7 for a detailed discussion).

3.1.3.4   Pro-concepts and contextual adjustment

Wilson (2011, 10) argues that conceptual expressions in natural language are systematically linked to concepts, which are constituents of the language of thought. In earlier work Sperber and Wilson (1998) describe the relevance theoretic account of the mapping between the mental and the public lexicon. Specifically, they assume that mental representations consist of *mental concepts*, which are relatively stable units. A mental concept entertains causal and formal (semantic or logical) relationships with external objects (i.e. words in a language) and relationships with other mental concepts. Sperber and Wilson (1998) argue against a one-to-one mapping between words in a language and mental concepts. Consequently, there may be:

- Concepts for which there is no word in a given language (one might expect that some languages do express them or they can be expressed my means of a phrase) (none-to-one).
- Words lacking a conceptual counterpart (one-to-none) such as for example 3rd personal pronouns.
- Different words that correspond to one concept (many-to-one) such as synonyms.
- One word corresponding to several concepts (one-to-many) such as homonyms.

The lack of one-to-one mapping can be explained by the existence of words in a language appertaining to all grammatical categories that do not encode a 'full-fledged concept but what might be called a pro-concept. The semantic contribution of pro-concepts must be contextually specified for the associated utterance to have a truth-value' (Sperber and Wilson 1998, 184). The authors make the claim that pro-concepts are so common that 'all words behave as if they encoded pro-concepts' (Sperber and Wilson 1998, 185). This is due to the fact that their semantic meaning must be contextually worked-out, i.e. they have to be fine-tuned through inferential processes so as to create an *ad hoc* concept (Barsalou 1987, Carston 2002a; Wilson and Carston 2007).

This view of concepts was initially adopted for 'open' classes of words (nouns, verbs, adverbs and verbs) as in example (401) from Wilson and Carston (2007, 235), where the hearer is brought to build a fine-tuned ad hoc concept *drinking alcohol* thought lexical narrowing (i.e. the word conveys a more specific sense than the encoded one). Another

means of contextual adjustment is lexical broadening[93] involving the use of a word to convey a more general sense that the encoded one as in (402) from Wilson and Carston (2007, 235).

(401) I am not *drinking* tonight.
(402) That bottle is *empty*.

The hearer makes hypotheses about this type of content at the level of explicatures (i.e. truth-conditional content). This occurs mainly because:

> 'The meaning encoded in a linguistic expression underdetermines the content the speaker communicates, not only at the level of her implicatures but also the propositional content she communicates explicitly (i.e. the explicature of the utterance) (Carston 2010b, 156).

Lexical pragmatics was further developed for 'closed' class of words, such as discourse connectives (Zufferey 2010) and interjections (Padilla Cruz 2009). In his reanalysis of interjections, Padilla Cruz (2009) suggests a contrary position to the previously suggested procedural account (Wharton 2003) and he argues that:

> If interjections encode some conceptual content, that content would not amount to proper full concepts, but to pro-concepts that have to do with emotions, feelings, attitudes or mental states, which have to be contextually specified or fleshed out. (Padilla Cruz 2009, 257)

Padilla Cruz gives arguments in favour of a conceptualist account of interjections, among which there are:

- We cannot use many interjections indistinctively or interchangeably in the same context, as there seem to be fixed conditions for their use;
- Psycholinguistic findings pointed to the fact the choice of an interjection is planned ahead along with the other words of the sentence pointing to the assessment of some constant content;
- Interjections are closely linked to adjectives, with which they share similar content;
- Prosody, representing procedural information, constrains the fine-tuning of the meaning of an interjection.

He suggests that many interjections can be associated with the very broad concept FEELING. However, due to repeated and constant activation through narrowing of the concept related to the domain of a particular emotion, feeling, attitude and want, some interjections become specialized for activating the concept of HAPINESS whereas others for activating the concept SADNESS (Padilla Cruz 2009, 259). Finally, some interjections such as *oh*, do not undergo a process of specialization and continue to be associated to broader pro-concepts.

As far as temporal reference is concerned, it is generally assumed that Tense is a grammatical category encoding instructions about how to locate eventualities with respect to

---

[93] According to Wilson and Carston (2007) there are several types of broadening, namely approximation, hyperbolic extension, metaphorical extension and category extension (i.e. use salient brand names for a broader category), among others. Narrowing and broadening make use of the encyclopaedic properties of a concept, where at least one property is shared between the pro-concept and the ad hoc concept. The enrichment process is carried out using the encyclopaedic properties of the concept, contextual information and pragmatic expectations and principles of relevance.

S (see section 3.1.3.5 for a detailed discussion). In this thesis, I would like to suggest an alternative view of the content encoded by Tense, which is a *conceptualist* view. Precisely, Tense could encode a very broad pro-concept TIME. Each verbal tense in a language is constantly used in activation to reference to past or non-past (distinction also recognised in neurolinguistics, see section 3.3.1) and it becomes, therefore, specialized for these types of temporal interpretation. This does not prohibit a verbal tense to make reference to another time or to no time at all, if contextual information directs it. In other words, temporal information expressed in a proposition is not recovered by linguistic decoding alone. Pragmatic inference plays a role in the recovery of the speaker's meaning with respect to temporal information. This takes place on two levels: ad hoc narrowing of the TIME concept and relating eventualities with respect to one another (see chapter 7 for the theoretical model suggested in this thesis). In the remaining of this chapter, I will discuss existent procedural accounts of Tense (sections 3.1.3.5, 3.1.3.6) and of Aspect (section 3.1.3.7)

3.1.3.5 Verbal tenses as procedural expressions

The hypothesis that verbal tenses encode procedural information comes, on the one hand, from a series of works on temporal reference conducted by the so-called Geneva School of Linguistics[94], and on the other hand, from Nicolle (1997, 1998) in the RT framework. Reboul (1994) and Moeschler (1994, 1996) make the parallel between nominal and temporal reference and suggest that verbal tenses refer to temporal entities or eventualities in a discourse (similar analogy previously made by Partee 1973 between verbal tenses and pronouns). Eventualities can be temporally located in relation to the moment of speech S and in relation to each other. Temporal location of an eventuality in relation to S is dealt with by the procedural meaning of a verbal tense whereas temporal location of eventualities in relation to each other is not encoded in the meaning of a verbal tense (Aménos-Pons 2011). Tense markers impose constraints on the determination of temporal reference and thus they encode procedural information (Nicolle 1998, 4).

Wilson and Sperber (1999) and Sperber and Wilson (1998) argued that temporal relations holding among eventualities are truth-conditional aspects of a proposition and part of the explicatures of that proposition. In other words, they are not implicitly communicated. Moreover, Moeschler (2000a, 2003) proposed a model of temporal interpretation of discourse based on inferences that the hearer has to make about temporal location of eventualities in relation to each other, called model of directional inferences (MDI).

In this section, I will pay special attention to the procedural approaches of verbal tenses (more precisely, of category of Tense as it is distinguished in this thesis). I will discuss the status of temporal and causal connotations of relations holding among eventualities and the

---

[94] At the beginning of the eighties, the label "Geneva School" was given to a series of publications on discourse and conversation that applied basic principles of syntactic analysis to the domain of discourse (Roulet et al. 1985, Moeschler 1985). In the beginning of the nineties, two different directions could be identified in the Geneva School: (i) a general discourse-oriented framework of language based on the modular hypothesis (Roulet 1997) and (ii) a radical pragmatic perspective on discourse sequencing and discourse interpretation (Moeschler 1993, 1996) (see detailed presentation in Moeschler 2001). In the late nineties, Moeschler and colleagues worked on temporal reference and proposed pragmatic models for individual FR verbal tenses and connectives. The results of their research were mainly published in Moeschler et al. (1998) and in the collection *Cahiers de Linguistique Française*, no. 22 (2000).

MDI in section 3.1.3.6.

RT's main assumption regarding tense markers is that the meaning of a verbal tense is underdetermined. Consequently, to yield the speaker's intended meaning, a verbal tense must always be contextually enriched through inference consistently with the principles of relevance. N. Smith (1990) points out that a verbal tense can only locate temporal reference in an underspecified way. Establishing actual temporal reference takes place through contextual enrichment according to expectations of optimal relevance. Smith's assumption was that the various connotations associated to a verbal tense do not correspond to different meanings of that tense but to different interpretations of a unique meaning combined with various contextual assumptions.

Nicolle (1998) followed Smith's assumption and proposed that tense marking is procedural information. In his words,

> Tense markers, in those languages which have them, may be characterize as merely imposing constraints on the determination of temporal reference. Similarly, markers of modality may be viewed as encoding constraints on the existential status of situations and events. Conversely, it is difficult to see how markers of tense and modality could be characterized conceptually. Take for instance, example (1) [Mary has eaten] and (2) [Mary has climbed the Everest]; the "present perfect" does not encode information about itself but about the events described in (1) and (2), say, that the event [Mary eat] and [Mary climb Everest] are meant to be represented as occurring at some time in the past whilst having present relevance. As a result of these considerations, grammatical markers of tense and modality may be characterized as exponents of procedural encoding, constraining the inferential processing of conceptual representations of situations and events.

The discussion of verbal tenses involves their function to establish temporal reference through the positions of temporal coordinates (speech point S, reference point R and event point E[95]) and determining temporal sequencing of eventualities[96]. Saussure (2011) argues that a verbal tense is a procedural marker in that it specifies the computations that should be made on hearer's mental representations of eventualities. The output of the computation is a contextual value in the form of an inference. The procedure encoded by a verbal tense demands to the hearer to find the most relevant and consistent with contextual assumptions configuration of temporal coordinates S, R and E in order to locate an eventuality before, simultaneous or previous to S.

Saussure (2011, 2013) argues in favour of the procedural nature of verbal tenses and their underdetermined meaning based the case the complexity of the interpretation of certain FR tenses. He explores cases when tenses do not refer to time or refer to other times than those referred in most of the cases. To be more precise, he pays special attention to narrative and

---

[95] The description of Tense in these terms is due to Reichenbach (1947), who suggested a formalization of the semantic meaning of verbal tenses (see section 2.2.1).

[96] Most of the works in RT framework (besides the Geneva School) did not investigate the procedural content of the morpheme Tense, bur rather explored the link between Tense and temporal sequencing for discursive interpretation. Wilson and Sperber (1998) discussed temporal and causal connotations of conjoined propositions (as well as the case of and) and suggested that they are pragmatically determined aspects of what is said, thus part of explicatures. I will discuss these problems in section 3.1.3.6 and describe how the model proposed in this thesis accounts for them in section 7.2.1.2.

background uses of the FR IMP and future time reference with the FR PC among others. Saussure (2011, 67) suggested a methodological criterion for distinguishing conceptual and procedural expressions: *an expression is procedural when it triggers inferences that cannot be predicted on the basis of an identifiable conceptual core to which general pragmatic inferential principles are applied.* According to this methodological criterion, these distinctive possible interpretations cannot be accounted for, unless they are inscribed in the procedure of the tense. In other words, there are no identifiable conceptual cores of the IMP and PC respectively, that can predict their distinctive interpretations. According to Saussure, there are three aspects that play a role for the existence of distinctive interpretations of certain verbal tenses:

- Constraining contextual assumptions
- A contextual saturation of temporal coordinates R and S
- Communicative principle of relevance

Saussure argues that the IMP in (403) and (404) changes its behaviour (similar to the EN progresses, excepting it doesn't imply dynamicity) under contextual constraints of boundedness and temporal sequencing in (403) by virtue of relevance. As for the third party subjective perspective on the eventuality, the C-point as in (404), it is a semantic procedural constraint. Since the sentence carries the presumption of its own relevance, its interpretation must be consistent with the pertaining contextual assumptions. And this happens through a pragmatic modulation of the temporal interpretation associated to the IMP. The situation is similar for the interpretation of the PC in (405), where the representation of the eventuality is pragmatically shifted into the future, from where it is conceived as past. This shift occurs under the pressure of a future temporal adverbial positioning the projected point R corresponding to a third party's viewpoint.

(403) A huit heures, Marie *trouvait* ses clés et sortait.
At eight, Mary find.IMP her heys and leave-IMP
'At eight, Mary found her keys and left.'

(404) Le train quitta Londres. Une heure plus tard, il *entrait* déjà [surprise] en gare de Birmingham.
The train leave.PS London. Une hour later, it enter.IMP already [surprise] in Birmingham station
'The train left London. One hour later, it had already entered in Birmingham station.'

(405) Dans un an, j'*ai fini* ma thèse.
In a year, I finish.PC my thesis
'In a year, I will be done with my thesis.'

Another example of the procedural nature of tense markers is that of the analysis of the simple and compound past forms in Spanish (PS vs. PC). Aménos-Pons (2011) argues that the distinctive possible interpretations of the Spanish PC (resultative, existential, universal, hodiernal past, hot news) provided in Table 3-2 can be accounted for only is the "meaning" of the PC is a procedural one (Aménos-Pons 2011, 241).

Table 3-2 Interpretations of the PC in Castilian Spanish

Type of reading	Description	Example
Resultative	Change of state located between E and S	Los precios *han subido* mucho. Ahora es impossible comprar nada. 'Prices have increased a lot. Now it is impossible to buy anything.'
Existential	Single or multiple occurrences of an eventuality, contained in E	*Ha viajado* muchas veces a Europa. 'He/she has travelled many times to Europe.'
Universal	Eventuality coextensive with E	*He vivido* treinta años aquí y conozco bien este país. 'I have lived here thirty years and I know this country well.'
Hodiernal past	Eventuality contained in E, which is located within the interval S	Hoy, Luisa *ha salido* del trabajo a las ocho. 'Today, Luisa has left her work at eight.'
Hot news	Eventuality contained in E, relevant at S	!!Luisa se *ha divorciado* el mes pasado!! 'Luisa has got divorced last month!!'

He defines the procedural meaning of the Spanish PC as it follows:

- The hearer must represent an eventuality of any type as bounded, locate it in the past and consider some kind of relation between E and S (through an R connected to S)

Aménos-Pons comments that the relation E<S is common to all interpretations but has a changing nature. Both resultative and existential interpretations convey the idea that an event has taken place, causing a result state that is thought to hold at S (though the result state has a different source for each of the two interpretations). In hodiernal and hot news interpretations, the eventuality is seen as closely connected to S without any result state represented (closeness being chronologically established for the former and speaker-based for the latter). Finally, in universal interpretations no result state is represented, since the eventuality is still going on.

Another parameter that influences the interpretation of the PC in Spanish is Aspect, (perfective vs. imperfective viewpoint) and Aktionsart (telic vs. atelic eventualities) (see sections 2.3 and 2.4). Perfective tenses provide bounded representations of eventualities and thus they tend to associate with telics. Aménos-Pons argues that being perfective, the PC instructs the hearer to build a bounded representation of the eventuality, regardless of its type (telic or atelic). As far as the role of the lexical aspect is concerned, telics are particularly compatible with resultative interpretations of the PC. This is due to the fact that they have natural endpoints and involve a change of state. Aménos-Pons points out that the relation between resultative interpretations and telics is not systematic (there are cases of resultative interpretation arising with atelics and interpretations without a result state with telics). He suggests that this is due to the flexible conceptual nature of lexical aspect, which accepts contextual adjustment if required according to the criterion of consistency with the principle of relevance.

Aménos-Pons underlines that the procedural meaning of an indicative verbal tense influences it's chances to occur in narratives. Narratives require temporal location of eventualities in relation to each other in a chronological manner. In the case of the PC,

specifically temporal location of a bounded eventuality in relation to S and its resulting state holding at S, minimizes the chances of the Spanish PC to occur in narratives. The Spanish PS on the other hand, which does not involve a direct relation between E and S, becomes preferred in narratives. The procedural meaning of the Spanish PS is described in the following terms (Aménos-Pons 2011, 248):

- The hearer must represent an eventuality of any type as bounded and locate it in the past (through an R disconnected from S)

The PS is also a perfective tense, envisaging therefore both telic and atelic eventualities as bounded. This illustrates that grammatical aspect has the same effect on eventuality type for both verbal tenses.

Aménos-Pons' main hypothesis is therefore that the PC and the PS instruct the hearer to build bounded representations of eventualities thought an R connected to S for the former, and an R disconnected from S for the latter. The PC and the PS do not encode information about relations between eventualities. Temporal relations, as well as other discourse relations, result through pragmatic inference as a by-product from the contextual computation of the procedural meaning of a tense. They are calculated only when it is contextually appropriate commanded by the expectations of relevance (Aménos-Pons 2011, 251).

Aménos-Pons' explanation about the procedural meaning of verbal tenses in Spanish gives rise to a very intriguing conclusion: *procedural information encoded by a verbal tense provides aspectual information*. A similar explanation is given by Saussure (2003), who assumes that for treating temporal reference expressed by a verbal tense the human brain 'applies an aspectual strategy for building cognitive representations of eventualities' (Saussure 2003, 179; see also the discussion in section 3.1.3.6). In the view that I defend in this thesis, this aspectual information should be provided by the category of Aspect and temporal information with respect to S, R and E should be provided by Tense.

In my understanding, this type of model for interpreting verbal tenses has two main reasons. The first is the treatment of a verbal tensed form as a unique and generic category, without separating the subordinate categories, which are Tense and Aspect, as well as their interactions with Aktionsart. The second is the overwhelming emphasis given to the category Tense in Western European studies of Romance languages. Treating verbal tenses in this way prompts ambiguous temporal and aspectual interpretations of verbal tenses and prevents scholars from being able to contrast verbal tenses cross-linguistically[97] and to have

---

[97] The issue of representation of temporal information in a cross-linguistic perspective is also discussed in Jaszczolt's *Default Semantics* framework (2005, 2009, 2012). Default Semantics (DS) is a semantic theory of meaning based on two theoretical assumptions. The first assumption, shared with post-Gricean and neo-Gricean theories, is that linguistic meaning is underdetermined and that there exists a *pragmatic mechanism of completion* of the meaning. Consequently, there are pragmatically determined aspects of truth-conditional meaning. This corresponds to what is called *explicatures* in RT, *impliciture* (Bach 1994, 2001), *presumptive meanings* (Levinson 1995, 2000), *truth-conditional pragmatics* (Recanati 2004, 2010 among others), *default meanings* (DRT, SDRT) and *default interpretations* in DS. The second assumption adopted in DS, and shared with DRT, is that the theory of meaning of utterances and discourses is a *compositional semantic* theory (Jaszczolt 2005). DS integrates the formalism and the semanticization of meaning suggested in DRT, but it distances itself from it by assuming that compositionality is not sought at the level of sentence but at the level of higher level semantic representation of meaning called *merger representation*. In short, merger representations reflect the fact that various sources of information with equal strength contribute to the

an accurate understating of how temporal reference is expressed in natural language. Finally, any analysis of verbal tenses should also provide answers to questions regarding the status of *eventuality type* because of their very rich inter-relations, the status of *grammatical aspect* (perfective and imperfective, Comrie 1976) and that of the way in which these types of meanings are related to each other. Currenlty, in the literature it was suggested that eventuality type has a conceptual nature (Moeschler 1994), that grammatical aspect encodes procedural information (Žegarac 1991; Leonetti and Escandell-Vidal 2004) and that these types of information are hierarchically organized (Moeschler 2000a, 2003; see section 3.1.3.6 for a discussion of this hierarchy).

3.1.3.6 Tense, temporal and causal relations

Wilson and Sperber (1998) argue that the treatment of temporal and causal connotations of utterances involves considering three interrelated issues: the *interval* problem, the *cause-consequence* problem and the *sequencing* problem. The interval problem, as in examples (406) and (407), where the hearer assumes different time intervals: almost instantaneously in the former and a much larger interval in the latter. The cause-consequence problem, as in examples (406) and (408) where the hearer assumes that the glass broke because it was dropped as in the former and that Mary got angry because Peter left as in the latter.

(406) John dropped the glass. It broke.
(407) They planted an acorn. It grew.
(408) Peter left. Mary got angry.

The source of the cause is different in these two examples; it is conceptual in the former and speaker-based in the latter. Finally, the sequencing problem, as in (409) where the hearer would assume that Peter left before Mary got angry.

(409) I took out my key. I opened the door.

As far as the interval issue is concerned, Wilson and Sperber (1998) point out that it also applies to single sentences such as (410) and (411). If through the verbal tense the eventuality is located at some point within an interval stretching back from the moment of speech S, the hearer's task is to choose among a series of logical possibilities *within the last few minutes, within the last few hours, within the last few days, weeks, months*, etc. Wilson and Sperber argue that the hearer's choice affects the truth-conditions of the proposition and its cognitive effects. If the negation test is applied for verifying the truth-conditional status, as in (412) and (413), the claim that the speaker has not had breakfast may be true within the last few minutes or hours but false if the time interval refers to the last few weeks. Regarding the hearer's cognitive effects resulted from the processing of the sentence, they are greater for narrower than for larger intervals.

---

overall meaning of an utterance. Pragmatic means are used only for capturing some defaults or situated inferences. Furthermore, merger representations of utterances combine in a dynamic way (i.e. idea adopted from DRT: meaning in discourse is the function of the parts and the structure of the discourse) for determining the meaning of multi-utterance discourses.

<blockquote>
(410)   I have had breakfast.

(411)   I have been to Tibet.

(412)   I have not had breakfast.

(413)   I have not been to Tibet.
</blockquote>

Wilson and Sperber claim that the logical structure of the proposition is completed through the hearer's choice of the interval. This information is part of the explicature of the sentence. In the search for optimal relevance, the hearer narrows the interval based on contextual assumptions and encyclopaedic knowledge (or ready-made schema, such as take the key and immediately use it to open the door or have the breakfast each morning) to a point where he has an interpretation consistent with the communicative principle of relevance. In cases when no ready-made schema exists, such as in (414), the hearer might either make the hypothesis that the two events are unrelated and happened simultaneously or, in a very specific context (such as detective story for example), make the hypothesis that John used the handkerchief to open the door in order to avoid leaving fingerprints. In this case, there is an expectation of relevance for justifying later on the usage of the handkerchief.

<blockquote>
(414)   John took out his handkerchief and opened the door.
</blockquote>

As far as the cause-consequence problem is concerned, causality is an important part of human cognition[98], allowing language user to predict the consequences of their own actions and those of others. RT assumes, as Wilson and Sperber (1998) point out, that cause-consequence schemas are highly accessible to mind for the interpretation of sentences like (408). According to the communicative principle of relevance, the information that Peter left must contribute to the overall cognitive effects. If this is the case, all other possible interpretations will be discharged.

The sequencing problem is one major theoretical issue that has dominated the field on temporal information in discourse for many years. Scholars have focused on temporal relations holding between eventualities and questioned how temporal relations are inferred in discourse. Both discourse semantics (DRT and SDRT, see section 2.2.2) and pragmatic theories (Grice, Neo-Griean and RT frameworks) suggested an account for processing temporal relations. I will briefly recall and discuss these accounts in what follows.

SDRT assumes that discourse relations are inferred on the basis of world knowledge, discourse and linguistic knowledge. Let's consider two temporal discourse relations: Narration and Explanation (Asher 1993). The Narration discourse relation is characterized by a forward temporal inference (i.e. time advances) whereas Explanation is characterized by a backward temporal inference (time goes backward). According to discourse semantic theories, the knowledge of discourse type is crucial for determining these temporal relations. Labov and Waletzky (1967) argued that two sentences containing a preterit (as discussed in section 2.5.1.1.2), which are interpreted as being temporally successive, form a narrative text. The first event is deictically situated in the past time (i.e. E<S) whereas the other is temporally anchored on the first one. Hence, Labov and Waletztky argued that Narration is highly preferred in narrative discourses whereas Explanation is preferred in non-narrative

---

[98] See for example Hume (1738-1740), Davidson (1967, 1980), Talmy (1988) and more recent discussions as Moeschler (2007a), Reboul (2007), Blochowiak (2009, 2014), among others.

discourses.

(415)   I *grabbed* his arm and I *twisted* it up behind his back and when I *let* go his arm there *was* a knife on the table and he just *picked* it up and *let* me have it and I *started* bleeding like a pig. (Labov and Waletzky 1967, quoted in Radden and Dirven 2007, 219)

Linguistic knowledge is provided by temporal connectives and verbal tenses. Kamp and Rohrer (1983) (in the DRT framework) argued that FR verbal tenses expressing past time encode time direction necessary for inferring discourse relations. The PS encodes a forward temporal inference, the plus-que-parfait encodes a backward inference and the IMP encodes an inclusive temporal inference. Unfortunately, this idea has numerous counterexamples discussed by Kamp and Rohrer (1983, 260) themselves as in (416), Moeschler (2000a, 2000b), Saussure (1997, 2000a, b) and Tahara (2000) for the PS, Saussure and Sthioul (1999, 2005) for the IMP among others.

(416)   Bianca *chanta* et Pierre l'a*ccompagna* au piano.
        'Bianca sung and Peter accompanied her on the piano.'

SDRT specifies that world knowledge (WK) is stronger and has priority over linguistic knowledge (LK) for determining the inferred temporal discourse relation, as in (417). According to Kamp and Rohrer, the PS encodes a forward inference, which should hold in both (417) and (418). But in (418), there is a backward inference based on the causal relation that comes from world knowledge. The PC, on the contrary, accepts both forward and backward inferences, in (419) and (420), as suggested by Saussure (1997, 2000b).

(417)   Max *poussa* Jean. Il *tomba*.
        Max push.PS John. He fall.PS
        'Max pushed John. He fell.'
(418)   Jean *tomba*. Max le *poussa*.
        John fall.PS. Max push.PS
         'John fell. Max poushed him.'
(419)   Max a poussé Jean. Il est tombé.
        Max push.PC John. He fall.PC
        'Max pushed John. He fell.'
(420)   Jean est tombé. Max l'a poussé.
        John fall.PC. Max push.PC
        'John fell. Max poushed him.'

The weakness of this approach is circularity: discourse type is defined based on the usage of the appropriate verbal tenses corresponding to the intended temporal inference (forward or backward) and discourse relation is inferred based on the discourse type. Moeschler (2000b) points out that the most appropriate explanation for establishing temporal inferences in discourse should be a pragmatic one. Specifically, he argues for the necessity of a model that explains how linguistic and contextual information are combined. And this is offered by pragmatic theories, especially RT.

The first to investigate temporal relations from a pragmatic perspective was Grice. He used them to introduce the distinction between *what is said* and *what is implicated*. Specifically

he treated temporal relations as defeasible implicatures. Carston (1988) argued that Grice's treatment of temporal and causal relations as implicatures was problematic. She followed Cohen (1971), who pointed out that what Grice called conversational implicatures was actually truth-conditional (under the scope of logical operators and connectives). Carston convincingly argued that temporal and causal interpretations are pragmatically determined aspect of what is said, therefore part of the explicature of the utterance. RT theorists make the distinction between the explicit meaning of an utterance (i.e. *what is said*) and the conventional (or "encoded") meaning of the linguistic expressions employed. Wilson and Sperber (1998) write that temporal and causal 'connotations' in examples (406)-(409) are not encoded in the meaning of the sentences uttered, and follow Carston's idea that they are pragmatically determined aspects of the explicit meaning of those utterances (i.e. explicatures).

The first experimental study of the identification and labelling by ordinary speakers of what is 'said' vs. what is 'implicated' was Gibbs and Moise (1997). In their paper, Gibbs and Moise designed their experiments to determine whether people distinguished what speakers say from what they implicate and if they viewed what is 'said' as being enriched pragmatically. They used five categories of sentences[99] and participants had to choose between a minimal vs. enriched interpretation. Example (421) illustrates the *temporal relation* type of sentence as well as the two possible interpretations (minimal or literal meaning and the pragmatically enriched meaning):

(421) 'The old king died of heart attack and a republic was declared'.
(422) <u>Minimal</u>: order of events unspecified
(423) <u>Enriched</u>: the old kind died and then a republic was declared

The experiments were designed in order to manipulate the type of sentence, the instructions and the context of the targeted sentence. In the first experiment, the instructions consisted in explaining the two types of interpretations of the sentence and no context was given. In the second experiment, the instructions were more detailed, including information about linguistic theories addressing the distinction between what is 'said' and what is 'implicated'. In the last two experiments, linguistic contexts were provided (a short story) in order to favour enriched interpretation (in the third experiment) as in example (424) and the minimal interpretation (in the fourth experiment) as in example (425), regarding *temporal relation* sentences.

(424) The professor was lecturing on the life of Jose Sebastian. He was a famous rebel in Spain who fought to overthrow the King. Many citizens wanted Sebastian to serve as their President. "Did Jose Sebastian ever became President?" one student asked. The professor replied, *The old king died of a heart attack before and a republic was declared.*

(425) Mike liked to take long bike rides each day. He also liked to sing as he rode because he has a terrific voice. Mike's roommate thought this was funny. He said to someone that *Mike likes to ride his bike and sing at the top of his lungs.*

---

[99] Cardinal (Jane has three children), possession (Robert broke a finger last night), scalar (Everyone went to Paris), time-distance (It will take us some time to get there) and temporal relations.

Gibbs and Moise's four experiments showed that speakers assume that enriched pragmatics plays a significant role in what is said: the enriched interpretation was preferred in the first 3 experiments but not in the last one where the context biased strongly for the minimal interpretation. Manipulation of instructions and training did not have any effect on the participants' judgements.

Three observations can be made concerning the *temporal relation* type of sentences: (a) temporal sequencing is an inference drawn contextually[100], (b) it is independent of the specific instructions that speakers received and (c) it can be blocked in a context biasing for the minimal interpretation, that is, the unspecified order. On the basis of their results, Gibbs and Moise argue that there might be two types of pragmatic processes, one that provides an interpretation for what speakers say and another one that provides an interpretation for what speakers implicate. They argue that this position can be explained by the principle of optimal relevance (Sperber and Wilson 1986) and they acknowledge the difficulty of testing it experimentally.

Moeschler (2000b) discusses the advantages of relevance theorists' explanation of the status of temporal and causal inferences. Firstly, the temporal interpretation corresponds to a pragmatic enrichment of the propositional form of the sentence and it contributes to its truth conditions. In example (426) from Wilson and Sperber (1998, 171), the disjunction is not redundant because each disjunct brings a genuine contribution to the truth-conditions of the utterance. This is based on the assumption that the events presented in each disjunct happened in a different order.

> (426)   It's always the same at parties: either I get drunk and no-one will talk to me or no-one will talk to me and I get drunk.

Secondly, relevance theorists' explanation focuses on *processing efforts* rather than on *cognitive effects*. Examples (419) and (420) with the PC produce two interpretations (either forward temporal inference or backward causal inference) and neither syntactic nor semantic structures indicate how the sentence should be interpreted. The interpretation is consistent with the communicative principle of relevance. This means that a temporal or a causal interpretation will be chosen depending on which manifest facts are more accessible to the hearer and based on the mutual cognitive environment.

Thirdly, forward temporal inference (called *temporal sequencing*) and backward causal inference (called *reverse-causal* interpretation) are not the only possible relations among eventualities. There are two other possible relations, namely *simultaneity* as in (427) and *indeterminacy* as in (428).

> (427)   Bill smiled. He smiled sadly. (Wilson and Sperber 1998)
>   'Bill souriait.IMP. Il souriait.IMP tristement.'
> (428)   Cette nuit-là, notre héros *but* la moitié d'une bouteille de whisky et *écrivit* une lettre à Lady Ann.

---

[100] In his Model of Directional Inferences (2000a, 2003), Moeschler makes the same prediction about temporal relations between eventualities. They have an inferential nature and are drawn based on contextual assumptions. They can be blocked (minimal interpretation) under certain specific linguistic and contextual conditions.

That night, our hero drink-PS half a bottle of whisky and write-PS a letter to Lady Anne
'That night, our hero consumed half a bottle of whisky and wrote a letter to Lady Anne.'

Moeschler (2000b) defines simultaneity and indeterminacy as it follows:

- Simultaneity: e1 covers (partially) e2 is a part of the eventuality denoted by e1 is included in the temporal interval defining e2
- Indeterminacy: the relation between e1 and e2 is undetermined if the determining the relation is not necessary for understating e1 and e2 or if determining the relation is not possible.

Fourthly, temporal sequencing does not seem to be central for temporal coherence in discourse. Causality plays an important role therefore the question concerning the relation between temporality and causality rises. In example (429), the only possible relations are forward causal and temporal relations while in (430) several relations are possible: forward temporal and causal, forward temporal and backward causal, backward temporal and causal.

(429) Socrate *but* un coup et *tomba* raide.
       Socrate drink.PS once and fall.PS stone
       'Socrate drank one and fell stone.'
(430) Marie *cria* et Pierre *partit*.
       Mary scream.PS and Peter leave.PS
       'Mary screamed and Peter left.'

These examples suggest that causal relations are a subset of temporal relations. Wilson and Sperber (1998) give an example where a causal relation occurs without a temporal relation as in (431).

(431) Susan is underage and can't drink.

Moeschler's proposal is that causal and temporal relations are two sets of relations that can have a Boolean junction. This means that for two eventualities e1 and e2, there can exist an intersection of causal and temporal relations for which [e1 causes e2] implicates [e1 precedes e2]. Two sentences can produce identical cognitive effects on the basis of different explicatures and implicated premises as in (432) and (433). In (432) temporal relation [e1 precedes e2] is part of the explicature while causal relation [e1 causes e2] is an implicated premise. In (433) causal relation [e1 causes e2] is part of the explicature while temporal relation [e1 precedes e2] is part of the implicated premise.

(432) Max *a laissé* tomber le verre (e1). Il s'*est cassé* (e2).
       Max dropp.PC the glass. It break.PC
       'Max dropped the glass. It broke.'
(433) Le verre s'*est cassé* (e2). Max l'*a laissé* tomber (e1).
       The glass break.PC. Max it dropp.PC
       'The glass broke. Max dropped it.'

It was stated that there are several types of possible relations among eventualities, and this

can be summarized in Figure 3-3. The model considers temporal and causal relations. As far as temporal relations are concerned, they can be or not forward temporal inferences (temporal sequencing). In the case where there is no temporal sequencing, there are two new possibilities: either there is or not a backward temporal inference. And finally, if there is no backward temporal inference, then the cases of temporal simultaneity or indetermination can be identified. Temporal sequencing can or not be accompanied by a forward causal relation, as in (429) and (431) respectively. Backward temporal inference can or not be accompanied by reverse causality, as in (434) and (435) respectively.

(434)   Max tomba. Jean l'avait poussé.
        Max fall.PS. John push.PQP him
        'Max fell. John had pushed him.'

(435)   Jean *prepara* son café. Il s'*était levé* sans entrain.
        John prepare.PS his coffee. He wake up.PQP without energy
        'Jean prepared his coffee. He woke up without energy.'

Figure 3-3 Possible relations among eventualities
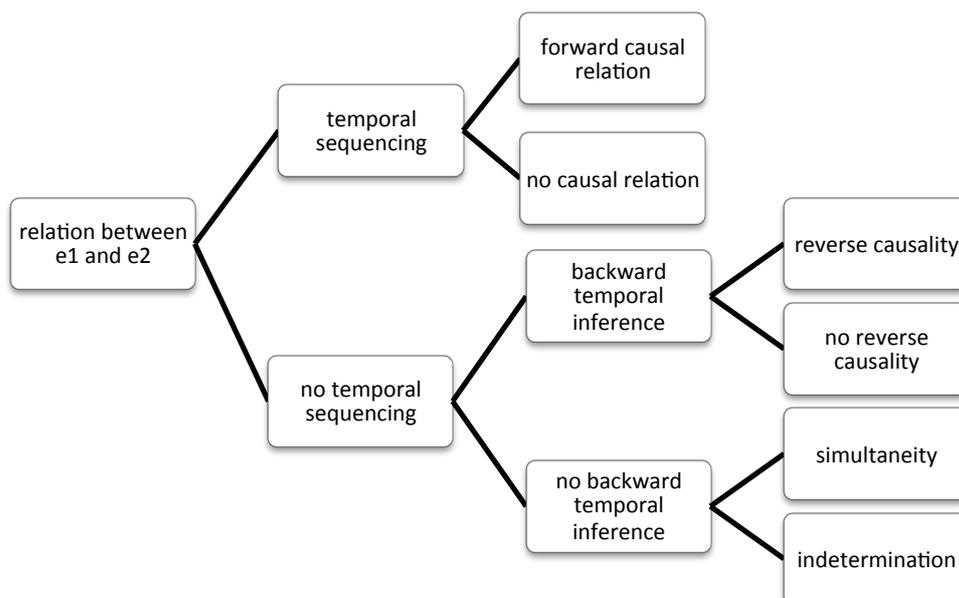
Moeschler's principle of temporal interpretation of the discourse is that during the comprehension process the hearer makes inferences about the temporal sequencing of eventualities, which are forward or backward temporal inferences. They correspond roughly to the discourse relations Narration and respectively Explanation in SDRT. These are not default inferences (in contrast to SDRT, where Narration is the default inference) but they are driven from linguistic expressions (encoding procedural and conceptual information) and non-linguistic information (contextual hypotheses and encyclopaedic knowledge).

Among procedural expressions, the most relevant ones for temporal interpretation at the discursive level are connectives and verbal tenses. For example, the conceptual relation holding between the verbs *pousser-tomber* ('push-fall') and the PC verbal tense, conveys a forward temporal and causal relation in (436) and a backward temporal and causal relation in (437). Examples (438) and (439) illustrate how the insertion of the connective changes the

direction of the temporal and causal relation: backward in the former (despite the forward direction conveyed by the conceptual relation) encoded by the connective *parce que* ('because') and forward in the latter encoded by the connective *et* ('and').

(436) Marie a poussé Jean. Il est tombé.
'Mary pushed John. He fell.'

(437) Jean est tombé. Marie l'a poussé.
'John fell. Mary pushed him.'

(438) Marie *a poussé* Jean <u>parce qu</u>'il *est tombé*.
'Mary pushed John because he fell.'

(439) Jean est tombé <u>et</u> Marie l'a poussé.
'John fell and Mary pushed him.'

Examples (440)-(448) illustrate the relation between verbal tense and connective. The PS in (440) and (442) conveys a forward temporal direction. Examples in (443) and (444) illustrate the compatibility of the PS with the connective *et*, which expresses explicitly the forward temporal relation. Example (445) expresses the incompatibility of the PS conveying a forward relation and the connective *parce que* that imposes a backward relation. This incompatibility disappears in (446), where the backward relation is maintained by the conceptual relation between the verbs. As seen in examples (436)-(439), the PC is not directional (i.e. it does not impose a temporal direction) and it is compatible with the direction imposed by the conceptual relation *pousser-tomber* ('push-fall') and the connectives *parce que* 'because' and *et* 'and'.

(440) Marie *poussa* Jean. Il *tomba*.

(441) 'Mary pushed John. He fell.'

(442) Jean *tomba*. Marie le *poussa*.
'John fell. Mary pushed him.'

(443) Marie *poussa* Jean <u>et</u> il *tomba*.
'Mary pushed John and he fell.'

(444) Jean *tomba* <u>et</u> Marie le *poussa*.
'John fell and Mary pushed him.'

(445) ?Marie *poussa* Jean <u>parce qu</u>'il *tomba*.
'Mary pushed John because and he fell.'

(446) Jean *tomba* <u>parce que</u> Marie le *poussa*.
'John fell because Mary pushed him.'

Regarding the case of the plus-que-parfait, the situation is opposite to that of the PS. It conveys a backward temporal relation as in (447), and this relation is expressed explicitly through the connective *parce que* in (448). Example (449) expresses the incompatibility of the plus-que-parfait conveying a backward relation and the connective *et* that imposes a forward relation.

(447) Marie poussa Jean. Il était tombé.
'Mary pushed John. He had fallen.'

(448) Marie *poussa* Jean <u>parce qu</u>'il *était tombé*.
'Mary pushed John because he had fallen.'

(449) ?Marie *poussa* Jean <u>et</u> il *était tombé.*
'Mary pushed John and he had fallen.'

The model developed by Moeschler (2000a, 2003) for temporal interpretation of discourse is called the Model of Directional Inferences (MDI). The basic assumption is that if linguistic and non-linguistic sources provide contradictory directional information, the conflict must be resolved in order to achieve the intended cognitive effects. The MDI postulates the following hierarchies for the various types of information that contribute to directional inferences (Moeschler 2000a, 7):

- Connectives >> tenses >> verbs
- Contextual assumptions >> connectives >> tenses >> verbs
- Contextual information >> linguistic information
- Contextual assumptions >> procedural information >> conceptual information

The first hierarchy considers the hypothesis that in case of mismatches the direction encoded by connectives prevails over the direction given by the verbal tense, which prevails over the direction given by the verbs (conceptual relation). The second and third hierarchies are based on the RT assumption that linguistic information is underdetermined and it is adjusted based on contextual assumptions. In case of mismatches, direction given by contextual assumptions prevails over temporal direction given by linguistic expressions. Finally, the fourth hierarchy considers the prevailing of procedural information (provided by connectives and verbal tenses) over conceptual information (provided by conceptual relations and situation types). Consequently, there is another assumption resulting from these hierarchies:

- verbs and verbal tenses bear weak directional features
- connectives and contextual assumptions bear strong directional features

Moeschler insists on the fact that the working hypotheses of the MDI should not be considered as fixed rules because they can be defeated (2003, 9). His idea is that the hearer's access to the intended interpretation is governed by the principles of *economy* (as defined in RT) and of *optimality*[101] (in RT *optimal relevance* is linked to the rapport between the hearer's minimal expectations and the presumption of relevance of an utterance to be treated). In his words (2003, 2):

> The combination of linguistic and non-linguistic information is directed by the general principle of optimality. This principle states that an optimal interpretation minimizes the conflict information: the less conflict you meet, the more optimal the interpretation you get.

In his discussion of the MDI, Saussure (2003) points to some limitations of Moeschler's model. The first regards the role of strong directional features. It seems that strong

---

[101] Optimality Theory (OT) (Prince and Smolensky 1993) is a generative grammar theory that explains the usage of linguistic expressions arises from the interaction between conflicting constraints. OT has three main components: GEN takes an input generates a list of possible outputs, CON provides the criteria, in the form of strictly ordered and violable constraints, used to decide between candidates, and EVAL chooses the optimal candidate based on the constraints, and this candidate is the output. Given two candidates A and B and a set of constraints, the candidate that incurs fewer violations of the highest-ranking constraint, is the best candidate and thus, the output. Optimality is defined as the rapport between the number of violations and the height of the constraint in the constraint's hierarchy.

directional features coming from connectives and contextual assumptions will always give the temporal interpretation of the discourse. It is not clear therefore which is the concrete role of weak directional features coming from verbal tenses and verbs for discourse interpretation. If for conceptual relations the MID suggests a reconversion, verbal tenses have ultimately no role for temporal interpretation of discourse. Both in cases of mismatch and of agreement, the result of the interpretation is given by the strong directional features. As far as directional features coming from conceptual relations are concerned, the MID predicts that they can be converted into strong features taking the form of contextual assumptions. For example, the conceptual relation push-fall is used as a causal inference *push causes fall* (initially used by Lascarides and Asher, 1993). In the MID, this conceptual relation produces the formulation of a contextual hypothesis *If Mary pushes John then John falls.* In this case, conceptual relations and contextual assumption should represent a unique feature.

The second limitation regards the identification of accessible contextual assumptions. Saussure points out three possible cases: (a) if contextual assumptions are built based on the presence of connectives (such as *parce que* or *et*), then the inference is triggered by linguistic expressions, therefore the directional features coming from contextual assumptions and connectives represent a unique feature; (b) if contextual assumptions are built based on conceptual rules then the directional features coming from contextual assumptions and verbs represent a unique feature; and (c) if contextual assumptions are built based on other contextual information, then this must be explained in the model.

The third limitation concerns ambiguous examples. This is the case of temporal indeterminacy, where no temporal direction can be determined, partial or total covering or relation whole-subparts, and cases where eventualities take place simultaneously (i.e. temporal simultaneity). Indeterminacy in (450)-(452) and simultaneity (453) are classical problematic examples (Kamp and Rohrer 1983; Reboul and Moeschler 1998; Saussure 2003). For these examples, the MID's explanation is that a contextual hypothesis coming from general world knowledge or conceptual rules cancels the temporal direction conveyed by other sources, such as verbs, tenses or connectives. Saussure argues that the MID does not provide the tools allowing non-directional temporal inferences, as for these examples with the PS.

(450) L'été de cette année là *vit* de nombreux changements dans la vie de nos héros. François *épousa* Adèle, Jean-Louis *partit* pour le Brésil et Paul *s'acheta* une maison à la champagne.
'The summer of that year saw several changes in our heroes' lives. François married Adele, Jean-Louis left to Brazil and Paul bought a house in the countryside.'

(451) Cette nuit-là, notre héros *but* une bouteille de whisky et *écrivit* une lettre à Lady Ann.
'That night, out hero drank a bottle of wisky and wrote a letter to Lady Ann.'

(452) Max *construisit* un château de cartes. Il *était* paisiblement à la maison.
'Max built a house of cards. He was peacefully at home.'

(453) Bianca *chanta* et Pierre l'*accompagna* au piano.
'Bianca sung and Igor accompanied her with the piano.'

Saussure (2003) proposes a slightly different model for processing temporal information in discourse. His main hypothesis is that the interpretation process is an algorithmic procedure. As far as temporal interpretation is concerned, verbal tenses play an important role in that

they set temporal reference of eventualities in relation to the moment of speech. In his words (2003, 179):

> La référence temporelle correspond au moment du temps, dans la conscience du destinataire, pour lequel les conditions de verité du process décrit sont vérifiées. Il est vraisemblable que l'esprit applique une stratégie aspectuelle pour se réprésenter les événements.[102]

For example, for processing the sentences in (454) and (455) from Saussure (2003, 179), the hearer does not determine a temporal interval lasting from a few seconds in the former to a few hours in the latter, but a punctual and bounded cognitive representation of the eventuality. This is mainly due to the assumption that the PS is a perfective verbal tense.

(454)   La bombe *explosa*.
        'The bomb exploded.'
(455)   Frédéric et Marie-Hélène *emplirent* la piscine.
        'Frédéric and Marie-Hélène filled the pool.'

As for sentence in (456) containing a telic eventuality (i.e. accomplishment), the hearer builds an unbounded cognitive representation due to the IMP. Saussure argues that the IMP imposes an imperfective reading of eventualities, despite their type (state, activity, accomplishment or achievement).

(456)   Luc *arriva* au stade. Augustin *courait* le 1500 mètres.
        Luc arrive.PS at the stadium. Augustin run.IMP the 1500 meters
        'Luc arrived at the stadium. Augustin was running the 1500 meters.'

At a higher level, temporal reference is an important factor for determining temporal sequencing of eventualities in the discourse. According to Saussure, temporal sequencing consists of three types of temporal relations: *positive* (i.e. forward sequencing), *negative* (i.e. backward) and *null* (i.e. simultaneity and indeterminacy). Temporal sequencing is the result of an algorithm, which consists of a *general* procedure and *specific* procedures. Procedural markers, such as verbal tenses and temporal connectives, trigger specific procedures taking place in the interpretation process. Moreover, conceptual relations (such as *push-fall*) and procedural markers impose constraints for determining the temporal sequencing of eventualities.

The general procedure suggested by Saussure (2003, 278-279) can be summarized as it follows:

- Input: the sentence is processed at the phonetic, syntactic and semantic levels. The central system dealing with pragmatic interpretation (as suggested in RT, see section 3.1.1) detects the verbal tense and a potential connective or a temporal adverbial.
- If a connective is detected, a specific procedure is triggered for determining a directional inference (as suggested in the MDI) called *Connective Directional Inference* (C)DI. If a temporal adverbial is detected, the system takes in consideration this

---

[102] 'Temporal reference corresponds to a moment in time when, for the hearer, the truth-conditions of the eventuality are verified. It is possible that the human brain applies an aspectual strategy for cognitively representing events.' (my translation)

restrictive temporal information and applies to the moment of reference R.

- The system captures the base semantics and the default instruction of the verbal tense detected:

  - If a temporal variable exists in the cognitive environment, the system extracts the directional instruction called *Verbal Tense Directional Inference* (VT)DI.
  - If there is no temporal variable in the cognitive environment, the system builds one based on the instruction given by the verbal tenses and potential temporal adverbials.
  - The system evaluates the presence of a conceptual relation, such as *push-fall* or *land-get off*. If this is the case, the system determines a directional inference called *Conceptual Relation Directional Inference* (CR)DI.

- The system tests for conflicting direction among (C)DI, (VT)DI and (CR)DI. If there is no conflict, the system retains the directional inference and passes at the instructions for ending the procedure (see below). In case of conflict:

  - If there is a connective, the system retains (C)DI and passes at the instructions for ending the procedure.
  - If the conflict is between (VT)DI and (CR)DI, the system evaluates the contextual relevance of the conceptual relevance.
  - In case of relevance (i.e. important contextual effects with respect to treatment cost), the system retains (CR)DI and passes at the instructions for ending the procedure.
  - In case of lack of relevance, the system searches for another conceptual relation. If there isn't any, the system fails interpreting the utterance.

- Instructions for ending the procedure: the system accesses again the secondary instructions given by the verbal tenses and builds a new temporal variable according to the DI obtained.

Saussure (2003) points out that this can be an acceptable model if these steps are understood as parallel tasks instead of sequential tasks.

A drawback in Saussure's model is the lack of distinction between Tense and Aspect, on which temporal reference depends. As pointed out in section 3.1.3.5, Saussure's model assumes that verbal tenses encode procedural information providing aspectual information. One of the problems of this approach where temporal and aspectual types of information are mingled is that it might lead to confusion regarding the semantic meaning of a verbal tense and its pragmatic uses. Secondly, it cannot be applied for a different language than that on which the model was developed, for example for languages where the distinction between Tense and Aspect is more relevant, such as other tense-prominent languages such as English where progressive aspect is morphologically marked, aspect-prominent languages and tenseless languages.

I would like to argue that both Moeschler's model (MDI) and Saussure's procedural model (PM) are potentially accurate models for French verbal tenses with respect to how the hearer processes temporal information at the discursive level. The two models have numerous similarities but also divergences. As far as similarities are concerned, I can suggest the following:

- They are both fine-grained models of semantic and pragmatic sources of temporal information.
- Both models make use of conceptual and procedural information provided by linguistic expressions, namely verbal tenses, connectives and temporal adverbials.
- None of the two proposals modelizes concretely grammatical and lexical aspect (an exception could be the suggestion that in the PM that a verbal tense such as the FR PC triggers a perfective representation of the process).
- In both models, the interpretation process is directed by the quest of relevance.
- Both models are theoretical models and lack consistent and objective empirical basis, such as corpus analyses and experimental validation of hypotheses.
- They are both monolingual models, and lack therefore cross-linguistic perspectives.

Among divergences, I can highlight the following:

- The PM assumes that a verbal tense provides a temporal direction by default whereas the MDI does not.
- The PM makes use of temporal relations whereas MDI makes use of both temporal and causal relations holding among eventualities.
- The two models suggest a similar method for resolving possible conflicting information: a hierarchy of features in MDI and a conditional procedure (i.e. of the type *if...then*) making use of the same hierarchy of features.

The model presented in this thesis offers the empirical and cross-linguistic testing of the theoretical assumptions behind MDI and PM. Moreover, it models temporal information coming from grammatical and lexical aspect. The model defended in this thesis is based on multifactorial analyses of data and addresses the issues of temporal coherence at the discursive and cognitive levels (see Chapter 7).

3.1.3.7 Account of Aspect and Aktionsart

In Chapter 2, I discussed the semantics of Aspect and Aktionsart. It was indicated that Aspect expresses information about the way in which the eventuality is presented, as perfective or imperfective. Aktionsart expresses inherent properties of the eventuality type, dividing eventualities into states, activities, accomplishments and achievements (Vendler 1957, 1967). These four aspectual classes can be described in terms of ontological features as telicity, durativity and dynamicity. In the literature, it has been argued that they are not inherent properties of the eventuality but of the verb phrase (i.e. verb and its arguments).

Previous research pointed out the role played by these two categories for temporal interpretation of a discourse (see sections 2.2.2 and 2.4.3). As far as temporal sequencing is concerned, aspectual theories (such as Dowty 1986) suggested that it depends on the lexical aspect of the eventuality. However, there are numerous counter examples that weaken the aspectual hypothesis. In a pragmatic framework, Saussure (2003, and previous research) argued that only Aspect and Tense play a role for determining temporal reference and temporal sequencing because they encode procedural instructions constraining the interpretative process. Moreover, he suggests that in case of conflict between an atelic eventuality and a perfective verbal tense (in other words, between Aktionsart and Aspect), the hearer builds a bounded conceptual representation of that atelic eventuality.

The question that arises at this point of the discussion regards the nature of the

information encoded by these two aspectual categories. Žegarac (1991) was the first to discuss the status of the information encoded by grammatical aspect in a relevance theoretic framework: conceptual or procedural information. As far as Aktionsart is concerned, Žegarac (1991, 44) points out that Vendler's time schematas (i.e. states, activities, achievements and accomplishments) are assumed to be universal and differently realized in individual languages (see also Smith 1986). Aspect, on the contrary to Aktionsart, is not related to inherent temporal properties of situation types. It expresses the speaker's viewpoint on the situation described.

Žegarac proposes a fine-grained analysis of both grammatical aspect (such as the oppositions *simple* vs. *progressive* in EN and *perfective* vs. *imperfective* in Slavic languages) and lexical aspect by looking at individual verbs in English and Serbian/Croatian. His contrastive analysis determines the following general conclusions: viewpoint aspect (i.e. Aspect) encodes procedural information constraining the explicit content of the utterance whereas situation aspect (i.e. Aktionsart) represents conceptual information contained in the lexical entries of verbs in the mental lexicon. Explicitly, he suggests with respect to the opposition *simple* vs. *progressive* in EN that the simple aspect is underdetermined for the sense of completion or entirety, which characterizes the perfective aspect in Slavic languages (p. 187). Sentences in (457) and (458) from Žegarac (1991, 187) provide evidence that the eventuality expressed by a SP may continue up to present and even beyond. They indicate therefore that the sense of completion with the SP is not determined by its encoded aspectual information but contextually.

(457) John ran for several hours this morning, and, for all I know, he may still be running.

(458) -How did Susan spend the morning? -She worked on Peter's paper all morning and she is still working on it.

The progressive on the other hand, encodes the instruction to instantiate (i.e. single event) the property denoted by a stative verbal predicate as in (459) and to present the eventuality expressed as being incomplete (460). Imperfective verbs in Serbian allow for two interpretations corresponding to either EN the progressive or the simple aspect, as in example (461) from Žegarac (1991, 184-185).

(459) He is being stupid to act like this.

(460) He was running when the tram stopped.

(461) Radi.
Work.IMPERF
'He/she works/is working.'

Imperfective verbs in Serbian retain the [-complete] feature in *when*-clauses whereas aspectually unmarked verbs can be either perfective or an imperfective interpretation, as illustrated by the contrast between (462) and (463) (Žegarac, 1991, 185). The sentence in (462) is understood as conveying the idea that the discussion took place after the analysing had finished whereas the sentence in (463) conveys that the answering took place while problems were talked about. He points out that example in (463) strongly supports the view that the imperfective aspect grammaticalizes the feature [-complete].

(462) Kad su *analizirali* problem, *raspravljali* su o mnogim pitanjima.
When they analyze.UNSPECIFIED_ASP the problem, discuss.IMPERF a lot of questions.

(463) Kad su govorili o tom problemu, odgovarali su na mnoga pitanja.
When they discuss.IMPERF about the problem, they answer.IMPERF a lot of questions

Hence, Žegarac's suggestion is that aspectual categories of English and Serbian are comparable categories, which can be explained in terms of the grammaticalization of *completion* and *instantiation*. The progressive of English and the imperfective of Serbian gramaticalize the lack of completion, in other words, the instruction to build an unfinished (in the sense of lack of completion) representation of the eventuality. The perfective aspect in Serbian encodes completion, whereas the simple aspect in EN is unspecified with respect to this feature. Furthermore, both the progressive and the perfective point indexically to a particular event instantiating the property denoted by the verbal predicated (i.e. Aktionsart) whereas the imperfective and the simple do not. Žegarac's cross-linguistic analysis illustrates that procedural information encoded by Aspect is both language independent and it presents language specificities.

Escandell-Vidal and Leonetti also (2011, 92) argue that Aspect encodes procedural information about how to construct the internal representation of the eventuality considered. They exemplify that the progressive marker in EN indicates that the event has to be viewed as an incomplete action in progress at a specific time.

In the following lines, I will discuss the status of Aktionsart with respect to the conceptual/procedural distinction. Žegarac (1991, 222) pointed out the different behaviour of states verbs and event verbs, and suggested that this difference may be captured in 'meaning postulates' or inference rules contained in the logical entries of the concepts denoted by these verbs. In other words, Aktionsart is of a conceptual nature and *duration* (from the durativity ontological features) is a primitive.

Moeschler (2002a, 2002b) suggested that lexical aspect encodes conceptual information and he gave several arguments. The first argument is linked to the fundamental assumptions of RT. RT is a representational theory stating that cognitive operations involve manipulating conceptual mental representations. These conceptual representations contain propositional content, i.e. information coming from nouns, verbs, adjectives, etc. On the other hand, functional categories encode procedural information about how to manipulate these conceptual representations. This initial parallel between lexical category/conceptual information and functional category/procedural information was refined based on empirical work, which brought evidence against a one-to-one correspondence[103]. Moreover, Moeschler et al. (2013) point out that Aktionsart has logical properties and it contributes to the propositional content of an utterance. Escandell-Vidal and Leonetti (2011, 92) suggest that durativity and dynamicity are formal linguistic traits involved in the description of situation classes in all natural languages.

In an utterance, the inherent temporal features of the eventuality combine with the instructions provided by Aspect. When they match, as in (464) where there is a dynamic telic

---

[103] For connectives, see Zufferey 2012; Wilson 2015; Blochowiak 2014, 2015a and Moeschler 2015a for theoretical accounts. For verbal tenses, Grisot and Moeschler 2013; and this thesis.

situation and a progressive Aspect, the hearer builds a mental representation of a dynamic event in progress. In (465) on the contrary, Aktionsart and Aspect do not match. Precisely, a progressive marker is applied to a stative predicate. The human brain processes these two types of information, and the hearer builds a mental representation of a dynamic situation in progress, i.e. John is behaving as a silly person in a particular situation.

(464) John is eating his sandwich.
(465) John is being silly.

Another example is the IMP in Romance languages. In Spanish for example, as Escandell-Vidal and Leonetti (2011, 93) note, the IMP encodes the instruction to view the eventuality as atelic or unbounded. Therefore, it combines most frequently with states and activities. When it combines with telic eventualities, there is an adjustment in the interpretation[104]. This can be expressed, for example as a habitual or ingressive reading of the sentence. Escandell-Vidal and Leonetti's proposal for this phenomenon is that procedural information encoded by Aspect is rigid and imposes a meaning adjustment for Aktionsart. This adjustment is inferential and takes place at the level of the propositional explicature.

In this thesis, I follow the hypotheses that Aktionsart represents conceptual information and that Aspect represents procedural information. Aktionsart was operationalized for the experiments carried out in terms of the actual realization of telicity, namely *boundedness*. As for Aspect, it represents the speaker's viewpoint on the eventuality and it is expressed morphologically in Slavic languages. A pilot experiment aiming at operationalizing grammatical aspect for the EN SP showed that it represents information inaccessible to consciousness pointing to its procedural nature (see Experiment 4).

3.2   On morphosyntactic theories

This section is concerned with the syntactic properties of tense and how they influence its meaning. Theories of the syntax of tense within the framework of generative grammar can be traced back to Noam Chomsky (1957, 1965, 1970, 1981, 1995), Stowell (1981, 2007), Pollock (1989), Belletti (1990), Zagona (1990, 1995), Guéron (1993, 2007, 2008), Giorgi and Pianesi (1997), among many others.

Chomsky's (1957, 1965) syntactic theory of the EN tense and aspectual auxiliary verb system was designed to generate grammatical EN sentences and their syntactic phrase structure representations. It was based on a set of phrase structure rules that included the categories sentence (S), noun phrase (NP), auxiliary phrase (Aux) and verb phrase (VP) among others. Each category was defined to have a given structure stipulated by a rule. The S category was formed of a ternary-branching structure that included NP, Aux and VP. Main verbs were included in the VP. Tense morphemes and auxiliary verbs were designed to the category Aux, where Tense has two values *past* and *present*. The phrase structure rules defined *Deep Structure representations* of sentences. Deep structures were converted into *Surface Structure representations* (actual linear order of words and affixes) by transformational rules. One

---

[104] This phenomenon has been investigated in semantics as *coercion* (for example, de Swart 1998, 2003, 2011).

of the rules, called *Affix Hopping* in EN for example, stipulated that the inflectional affixes for tense and agreement originated in Aux in the Deep Structure representation will occur in the Surface Structure representation either in Aux or in VP, depending on the auxiliary or main verb that immediately followed it. According to this model, the Tense position generates two tenses (past and present); all other tenses are generated in the Aux position (the future in the Modal position, the complex tenses originated as sequences of a tense affix and/or the modal *will* followed by an auxiliary+ affix combination, as in *has eaten* or *will have eaten*).

Chomsky's first theories assigned a flat structure to complex sentences including two or more auxiliaries. Constituency tests argued against a flat structure and in favour of hierarchical structures, where each auxiliary verb is the head of its own VP. Chomsky (1970) introduced the *X-bar* theory and the theory of *categorical distinctive features*. X-bar theory had two basic ideas: first, each phrase (VP, NP, AP, PP) inherits its syntactic category from its head (V, N, A, P), and, second, all phrases have the same internal structure and elements: an obligatory head, an optional specifier (Spec) and an optional complement. A lexical head can combine with one or more phrases that it selects as its complements to form a "non-maximal projection" of the head, called X′ (X-bar). X′ can combine with one or more constituents functioning as its specifier(s), to form a "maximal projection" X″ (X double bar), called XP.

Chomsky combined X-bar theory with a theory of categorical distinctive features. Lexical categories were defined in relation to distinctive values [±N] and [±V], for example the adverb lexical category is defined as [+V, -N]. This allowed more fine-grained types of XPs. Several subtypes of VPs can be thus defined based on the inflectional affixes *past, present, -en* and *−ing*. The idea is that these affixes stand for certain abstract categorical features applying on the verb that bears them. For example, a verb bearing the suffix *past* bears the features [+V, -N, +Tense, +Past]. The analysis of tenses and other inflectional affixes as distinctive features proved itself to be an important limit of the theory because of the lack of syntactic expression of the temporal semantics of tenses as discussed by Reichenbach, Comrie et others (see section 2.2.1).

Chomsky's (1981) Government-Binding theory of phrase structure introduced a new category Infl, consisting of abstract features for tense and subject agreement, which were realized morphologically as a single affix. Infl became the head of the clause based on the hypothesis of verb movement to I in order to combine with agreement and tense features. Pollock (1989) proposed that so-called Split-Infl Hypothesis, according to which the node I has to be split in two projections: one, called AGR, is for agreement features, and one, called T, is for temporal features. Pollock (1989) and Guéron (1993) proposed that the head of the clause is T, whereas for Belletti (1990) and Chomsky (1993) the head is AGR. It is not my purpose here to take sides, what interests this research is that Infl (T and AGR) is the head of the clause.

All these syntactic theories largely ignored the semantics of tenses. Referential approaches to tense semantics focused on its predicative and referential functions (Stowell 2007, 2012). The predicative function concerns the three temporal ordering predicates (before, after and simultaneous with) and their arguments consisting of pairs of coordinates. The referential function concerns temporal reference through the temporal coordinates S, R and E, function similar to that of pronouns (similarity identified by Partee 1973).

Zagona (1990) proposed a theory of the phrase structure of tenses that allowed the representation of the temporal ordering predicates (before, simultaneous, after) through the application of principles of binding (Chomsky 1981). Her fundamental hypothesis is that tenses are transitive predicates selecting "subject" and "object" predicates referring to times, specifically to S and E, respectively. She proposed a phrase representation similar to X-bar theory, where T[105] originates syntactically at the head of a functional category TP (maximal projection), where T selects VP as its complement, with an intermediate level that is T' (non-maximal projection). She further assumes that VP is the object argument of T referring to E and the Spec position is filled by the subject argument of T referring to S.

As Stowell (2012) points out, in Zagona's theory, it is not Tense itself that refers to times but rather it's arguments: the VP complement of T refers to E, and the Spec (subject) of TP refers to S. Her idea was that present tense selects a VP complement that behaves like an anaphor and must be bound to the Spec of TP, whereas past tense selects a VP complement that behaves like a definite referring expression that must not be bound to the Spec of TP. Stowell comments that the S and E must co-refer to the same time with the present tense, and must refer to different times with past tense. Zagona's theory stops at the distinction between past and present for tense, given that future has a modal status (occurs in Aux).

Stowell (1995, 2007) proposes a syntactic model that provides a more direct encoding of the temporal ordering semantics traditionally attributed to verbal tenses (see section 2.2.1). Since tense is a referential category, he suggests that it should have an internal structure analogous to other referential categories, such as Determiner Phrase (DP). Individual verbal tenses are analysed as predicates of temporal ordering, taking time-denoting expressions of their two arguments. Stowell proposes a ZP (Zeit-Phrase, zeit meaning *time* in German) that refers to a time much as the DP refers to an individual. T selects a ZP as its object argument (internal argument) and a null subject argument in the Spec position (external argument). A ZP with a VP complement denotes the time of the event or situation that the VP refers to, as shown the syntactic tree in Figure 3-4.

Figure 3-4 Syntactic tree of the Zeit-Phrase
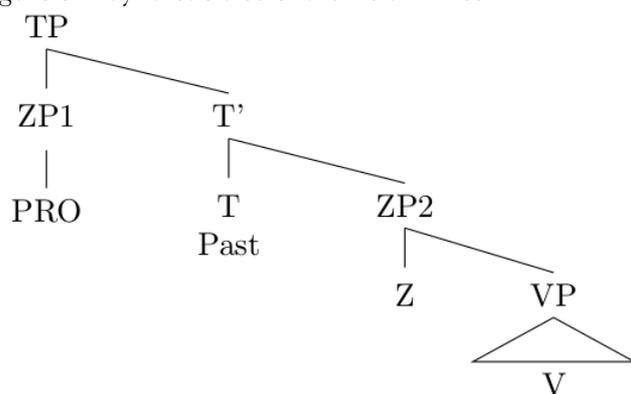
In simple sentences, such as *John left*, the ZP corresponds to E, while in complex sentences containing modals or auxiliaries, ZP corresponds more to what Klein (1994) called Topic

---

[105] Zagona uses the label F instead of the more familiar T for tense.

Time[106]. The external argument of the tense denotes a time that functions as the speech moment of the utterance. Stowell's theory assumes a conventional predicative semantics for tense with the meaning of "after[107]", "before" and "simultaneous (contained in)" as in examples (466)-(468). The tense itself imposes a relative temporal ordering on the reference time and the time of the event. For example, a past tense orders the speech time after the event time, (466) asserting that S is after the/a time at which Kim lived in London.

(466)   Kim lived in Paris.
(467)   Kim will live in Paris.
(468)   Kim lives in Paris.

Giorgi and Pianesi (1997) argued that relations between S/R and R/E represent different types of temporal categories called T1 and T2[108], respectively. T1 and T2 instantiate a relationship between two eventive variables: S and E. This relation is never realized directly being always mediated by R. They suggest the possible relations for T1 and T2 (1997, 27):

- T1: S<R future   T2: E<R perfect
  R<S past   R<E prospective
  S=R present   E=R neutral

The main assumption is that the various verbal tenses are the result of the composition of a relation of type T1 with a relation of type T2. For instance, the representation of the PRES is the result of the combination of S=R with E=R to yield S=R=S. The evidence for this framework comes from the interpretation of the future perfect (Comrie, 1985; Hornstein, 1990; Giorgi and Pianesi, 1997), as in (469).

---

[106] Klein (1994, 58) proposes temporal coordinates that are rather different than classical reichenbachian coordinates. He distinguishes between "finite" (FIN-time) and "non-finite" (INF-time) temporal components of an utterance. INF is a selective description of a possible situation and corresponds to the non-finite predicate of the utterance. INF can be specified by temporal adverbials such as *for two hours, two times*, etc. As for FIN, it depends on the type of sentence (declaratives, interrogatives, etc.). In an assertion, the content of FIN is the assertion made by the utterance. The *tense* used in that assertion imposes a temporal constraint on the assertion: it narrows down the assertion to some particular time. This particular time is called the topic time (TT): 'the time for which an assertion is made' (p. 58). The situation described in the assertion occurred at a time called the time of the situation (TSit) and it corresponds to INF-time. And finally, both TT and TSit are different than the time when the utterance was made, the time of utterance TU. In a sentence like *The light was on*, TSit is the time when the light was on, TT is the time for which such a claim was made and TU is the time when the assertion was uttered. Finally, the relation between FIN-time and INF-time corresponds to Aspect. Both tense and grammatical aspect can, in Klein's framework, be defined in terms of temporal relations, such as *before, after* and *simultaneous*. In Klein's words, 'they only differ in what is related to what' (p. 3).

[107] Each temporal relation is the contrary to Comrie's relations, in that the past tense predicate locates S (in the Spec) after E (the complement of the tense) in the hierarchical representation of the phrase structure.

[108] T1 and T2 are lexical categories assigning a T-role to their arguments, analogous to the notion of theta roles (the verb assigns theta roles to the subject and its internal arguments). The notion of T-role aims at capturing the observation that T must have a VP complement. In other words, it is a formal device allowing the identification of the event argument of the verb and the assignment of a temporal interpretation. For complex verbal tenses, T1 assigns a T-role to the auxiliary and T2 assigns its T-role to the verb. For simple verbal tenses, there is one TP and one AGR-P. The principle is that each event position can receive only one T-role and every T-role must be uniquely assigned to an event position (the so called T-criterion in Giorgi and Pianesi 1997, 29). According to the T-criterion, a verb cannot be temporalized twice (or more than twice), in other words a verb can have only one temporal morpheme.

(469)   John will have finished his manuscript by tomorrow.

The syntactic tree in Figure 3-5 illustrates Giorgi and Pianesi's theory (1997, 38). The two temporal projections, T1 and T2, lexicalize the tense relations, S/R and E/R respectively as indicated in the tree. The AGR nodes, AGR1 and AGR2 represent the morphological expression of person and number for AGR1 and of number and gender for AGR2[109].

Figure 3-5 Syntactic tree with agreement nodes and temporal nodes

```
AGR1-P
   |
AGR1      T1-P
           |
          T1      VP
          S/R      |
                  V      AGR2-P
                            |
                          AGR2      T2-P
                                     |
                                    T2      VP
                                    R/E      |
                                             V
```
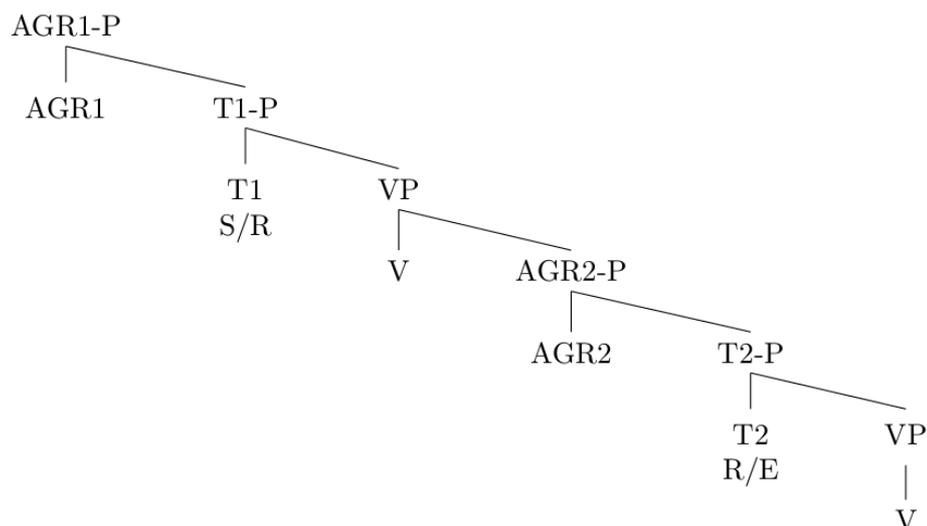
Sentence in (469) is ambiguous in that the exact temporal relation holding between the finishing of the manuscript and the S point is unknown. More specifically, John could have already finished the manuscript at the time in which the sentence is uttered (S); or has finished it exactly at S; or he will finish it in a close future lying between the time of the utterance and *tomorrow*. As Comrie and then Hornstein have pointed out, the revised Reichenbachian framework accounts better for the limited number of morphologized tenses, while original Reichenbachian framework predicts morphologically different realisations for each possible position of E. Comrie underlined that this is not the case. Giorgi and Pianesi (1997) hypothesize that the two relations T1 and T2 are realised by different morphemes, which have different features and morphosyntactic behaviours influencing the interpretations of Tense.

Following Chomsky (1993), Giorgi and Pianesi (1997) suggest that the syntactic structure of a phrase (simple sentence) in IT includes: at lest one V projection (VP), at least one T (TP) projection and at least one AGR projection (AGR-P) depending on the verbal tense, as shown in the syntactic tree in Figure 3-6. The structure of the PRES in (470) consists of a VP in the lowest part of the structure, which is the projection of the verb V *mangiare* (to eat) and an AGR-P for the agreement in person and in number. There is no T projection because the present tense is the unmarked tense and the IT present tense does not have a lexical T (a morpheme expressing present tense).

---

[109] Giorgi and Pianesi point out that their AGR1 and AGR2 resemble to Chomsky's Agreement subject and Agreement object. As for their T2 temporal relation, it is different Belleti's (1990) analysis that assigns an aspectual value to this node. Giorgi and Pianesi defend their model arguing that 'aspectual information is syncretically realised together with the temporal one and can be "scattered" when necessary' (1991, 39).

(470)   Mangio.
        Eat.PRES
        'I eat.'

Figure 3-6 Syntactic tree of the PRES

```
AGR-P
 |    \
AGR    VP
 -o    |
      (e)P
     mangi-
```
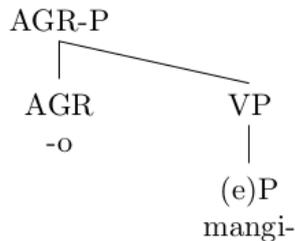
The structure of the IT Passato Composito (PC) in (471) consists of a VP in the lowest part of the structure, which is directly dominated by T2P- that is, by the head expressing the relation E and R; the T2P is followed by an AGR2-P that lexicalizes the agreement expressed on V. AGR-2 expresses the ending morpheme of the verb. Moreover, the structure contains a second VP, which the projection of the auxiliary and an AGR1-P that lexicalizes the agreement expressed on the auxiliary, as shown in the syntactic tree in Figure 3-7.

(471)   Ho mangiato.
        Eat.PC
        'I have eaten.'

Figure 3-7 Syntactic tree of the PC

```
AGR1-P
 |    \
-o     VP
       |   \
       V    AGR2-P
     Avere   |    \
           AGR2    T2-P
            -o      |    \
                   T2     VP
               (E_R)-at-   |
                           V
                         mangi-
```
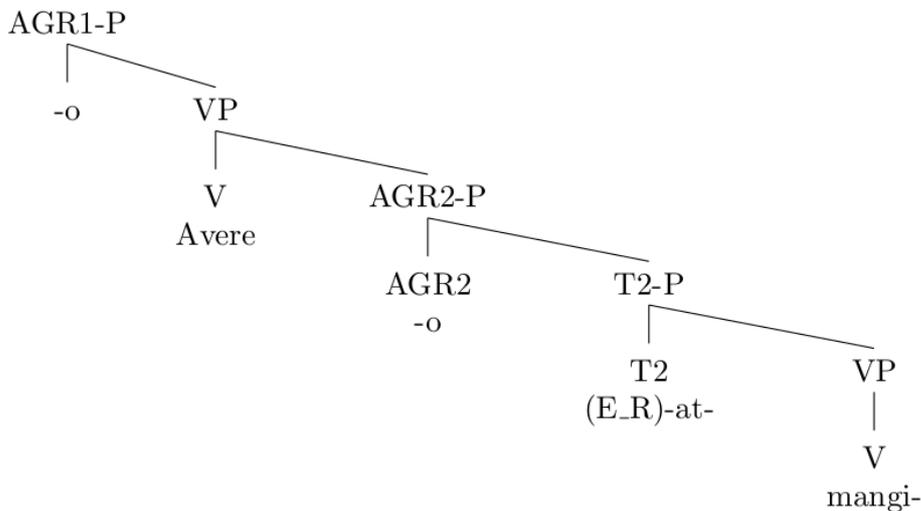
The structure of the IT past perfect from (472) and shown in Figure 3-8 is partly similar to that of the PC. For the past perfect, T1 (the head expressing the relation R/S) must assign its T-role and this is expressed on the auxiliary. Finally, in the highest position of the structure, there is the AGR1-P expressing agreement of person and number on the auxiliary.

(472)   Ebbi mangiato.

'I had eaten.'

Figure 3-8 Syntactic tree of the Past Perfect

```
AGR1-P
 ├── AGR1
 └── T1-P
      ├── (R_S)T1
      └── VP
           ├── V
           │   Avere
           └── AGR2-P
                ├── AGR2
                │    -o
                └── T2-P
                     ├── T2
                     │   (E_R)-at-
                     └── VP
                          └── V
                              mangi-
```
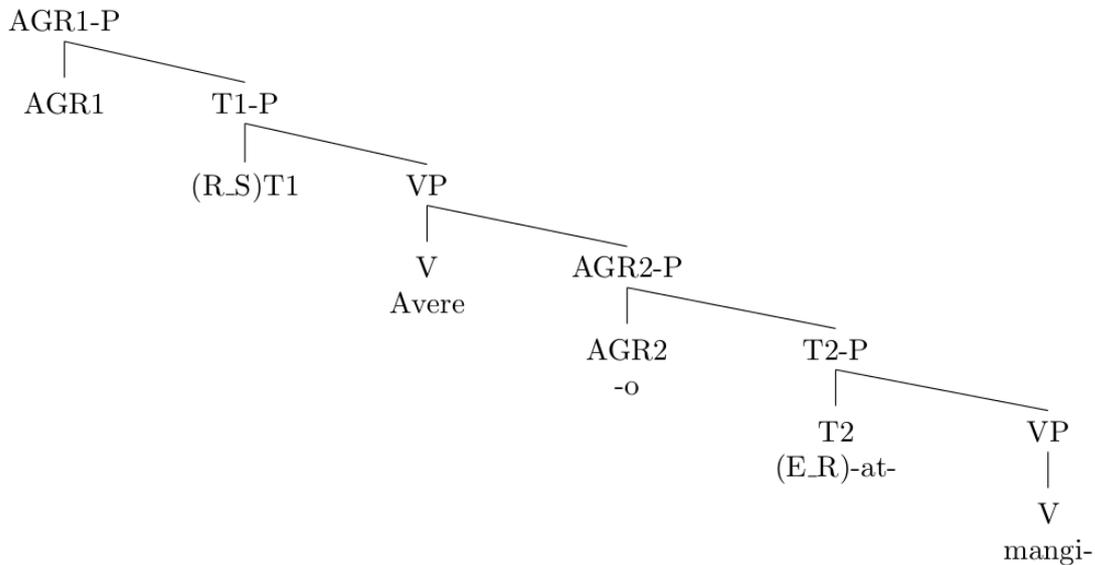
The structure of the IT Passato Semplice (the EN SP) in (473) consists of the VP in the lowest part of the structure, which is dominated by the T1P expressing the relation R/S. As far as the E/R relation is concerned, it is not expressed through a morpheme. Finally, the AGR-P expresses the agreement in number and person and the word boundaries, as shown in Figure 3-9.

(473) *Mangiai* una mela.
       Eat.PS an apple
       'I ate an apple.'

Figure 3-9 Syntactic tree of the PS

```
AGR-P
 ├── AGR
 │    -ai
 └── T1P
      ├── (R-S)T1
      │   mangi-
      └── VP
           └── V
               mangi-
```
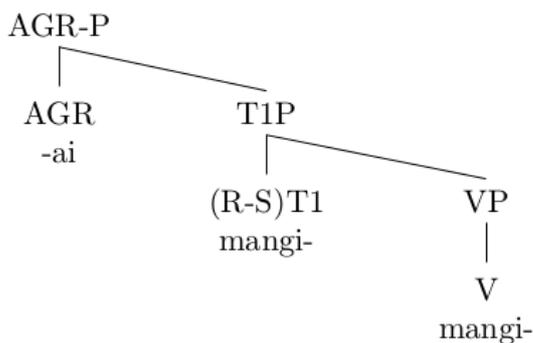
The SP is acceptable when the temporal argument of the predicate is not overtly specified and is interpreted relative to S in the absence of any other context, as in (473). This is not the case for the IMP, which must be interpreted, one the one hand, as a past with respect to S, and which requires that the temporal argument of the predicate be overtly specified, on the other hand. The sentence in (474) (Giorgi and Pianesi, 1997, 174) is odd if expressed without any context that would explicitly provide a temporal location from which the IMP can take its temporal reference.

(474) *Mangiavo* una mela.
       Eat.IMP an apple
       'I ate/was eating an apple.'

Giorgi and Pianesi argue that the EN verbal system presents important differences with respect to the IT one. Specifically, the verbal root is followed by only one suffix, which can be either the inflection for tense *–ed* or for agreement *–s* for the third person singular, as shown in (475)-(477). This is possible in other languages, like IT, FR, RO and German. Giorgi and Pianesi's proposition is that in EN, the features of AGR and of T belong to the same bundle projecting a single category called AGR/T. They suggest that the value of one feature determines the value of the other: the values of AGR features imply the value of T features. In other words, the values [±3rd person] of the AGR features imply the unmarked value of T features (i.e. [-past]), whereas the presence of the marked temporal value (i.e. [+past]) has no implication on the agreement values.

(475) He loves/loved.
(476) *He loveds.
(477) *He wills love.

To sum up, Giorgi and Pianesi suggest the following syntactic properties of verbal tenses that reflect their semantics:

- The IT verbal system has TP and AGR-P categories, whereas the EN verbal system has one category called AGT/T
- The T node consists of two projections, T1 and T2, which lexicalize the tense relations S/R and E/R.
- T1 and T2 assign T-roles to the verb arguments, which can or not be overtly specified depending on the verbal tense (not obligatory for SP but obligatory for the IMP)

In Chomsky's Minimalist Programme (Chomsky 1995, 2000; Cowper 2002, 2005), the inflection paradigm of finite verbs in English contains *uninterpretable* and *interpretable* features: (i) the uninterpretable agreement features (person, number), (ii) the interpretable features of [± perfective] Aspect[110], [± past] Tense and [± realis] Mood. Cowper (2002, 2005) proposes a *geometrical model,* which supposes a dependency structure, for the interpretable features of the Infl-P system. A coarse-grained version of her model is provided in Figure 3-10. In Cowper's view, features are monovalent, that is, the absence of a feature triggers a default interpretation of the node dominating the feature. In her model, default values are not overtly indicated.

---

[110] The perfective aspect is the marked values, in other words, a clause is interpreted as perfective unless it is overly marked as imperfective. Clauses are always marked for Tense in English, i.e. past tense must be overtly marked by one of the past-tense forms. Subjunctive or [-realis] does not have any distinct morphological forms in English but it can be expressed through past-tense forms.

Figure 3-10 Feature geometry for Infl
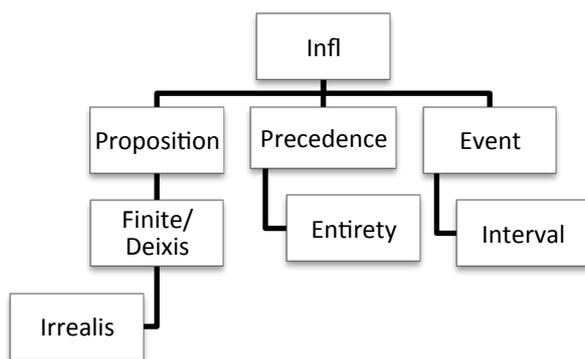
The features are explained as it follows (as provided by Cowper 2005). The *event* feature distinguishes eventualities that are events or states based on temporal properties, that is, states are interpreted as a singleton set of moments. The *interval* feature specifies that events are linked to an interval, that is, a nonsingleton set of moment, rather than to a single moment. Aspect is independent of the actual duration of the event (given through Aktionsart). This feature encodes imperfective Aspect. If *event* is absent, than the sentence is stative by default.

The *precedence* feature[111] establishes a marked relation (precedence rather than simultaneity or inclusion) between the clause and its temporal anchor, which can be established by a higher clause, by the discourse, as a whole or it can be the moment of speech S. If *precedence* is absent, the relation is taken to be one of simultaneity. Temporal anchoring of states is interpreted as the state holds at the temporal anchoring (for simultaneity) or the states holds at a moment prior to the temporal anchoring (for *precedence*). The *entirety* feature refers to moments within states and events. More precisely, with events, this feature requires the precedence relation to hold among all the moments associated with the event and the temporal anchor. With states, *entirety* requires all moments at which the state holds to precede the temporal anchor.

The *proposition* feature signals that the clause denotes a mental representation of a state or event, which has truth-functional values. The *finite* feature has purely syntactic content, it signals subject case and theta-role assignment on the verb. It indicates that the clause is a proposition. The *deixis* feature sets the temporal anchor of the clause to the deictic centre of the utterance/discourse, which is often the moment of speech. In Cowper (2005) there a second deixis, the *personal deixis*, which refers to *consciousness* as the set of propositions believed by the speaker at the moment of speech. The *irrealis* feature signals that the clause is compatible to the set of propositions denoted by the consciousness. The default case is therefore *realis*, indicating that the clause belongs to the set of propositions believed by the speaker at the moment of speech.

Moreover, this geometrical model includes entailment relations on the vertical axis, such as interval entails event, finite entails proposition. According to Cowper, these features and their entailment relations belong to the set of linguistic properties provided by Universal Grammar (Chomsky 2000). Languages differ in the way in which theses features are used. In other words, temporal interpretations of utterances can be linked to different features.

---

[111] It corresponds to both *predicative* and *referential* functions in Stowell's terminology.

Cowper argues, for example, that the *entirety* feature is not used in EN but it plays an important role in Spanish, whereas *interval* is not used in Spanish but it is used in EN. These two features are linked to the expression of viewpoint aspect in the two languages. Specifically, in EN, viewpoint aspect is directly encoded by the presence or absence of *interval*, a dependent of *event* (therefore Aktionsart). In Spanish, viewpoint aspect can be attributed to *entirety*, a feature of *precedence* (therefore Tense strictly speaking). Cowper's model in these terms recalls Giorgi and Pianesi's hypothesis (1997) that IT sentences have two functional heads (TP and ASP-P) whereas EN has only one AGT/T.

To sum up, in the morphosyntactic theories described in this section, Tense and Aspect belong to Infl, which is the functional head of a sentence. They are interpretable features, and are believed to be provided by the universal grammar. Languages differ in the way in which they make use of these features.

3.3 Human and automatic processing of temporal information

3.3.1 Human processing

If linguistic theories can make hypotheses about the relation between Tense and temporal information in discourse based on either built examples or naturally occurring examples from corpora, neurolinguistic account of Tense and its processing by the human brain is based on observation and experiments. Neurolinguistics is concerned with aphasic patients (brain damaged[112]) and with healthy speakers. Studying the dysfunction of language in aphasic patients reveals important aspects of the processing of language by healthy people. Neurolinguistic studies often build their working hypotheses and design their experiments based on existent linguistic theories. Neurolinguistics benefits thus of theoretical input while linguistics benefits of the validation or amendments of their theories. In this section, I will discuss the findings of several recent studies that investigated the way in which humans process Tense and it's relation to temporal information in discourse.

Agrammatic speech is characterized by omissions and substitutions of grammatical morphemes. It has been demonstrated across several languages that verb inflection (Tense and Aspect morphology) is problematic for agrammatic aphasic speakers (Friedmann and Grodzinsky 1997 for Hebrew; Bastiaanse et al. 2002 for Dutch; Burchert et al. 2005, Wenzlaff and Clahsen 2004, 2005 for German; Stavrakaki and Kouvava 2003 for Greek, among others).

In this section, I will discuss in more details these findings regarding the role and the interaction among Tense, Aspect and Aktionsart, as well as their implications for current

---

[112] Aphasia is a type of language disorder caused by the dysfunction of certain area of the brain. The dysfunction is due to brain damage caused most commonly by heart stroke or head injury. The area and extent of brain damage determines the type of the aphasia. Two main types of aphasia (which subsume several more specific types) are recognised: Broca's aphasia (also known as non-fluent or agrammatic aphasia) and Wernicke's aphasia (also known as fluent aphasia). Damage to Broca's area is associated with impairment of the ability to speak, their language becomes sporadic and agrammatic. Patients suffering from fluent aphasia, produce speech without any grammatical problem however they cannot convey the meaning (Wernicke's area being responsible for language comprehension) and thus they comprehension is severely deficient.

linguistic theories of temporal information in discourse. In morphosyntactic terms, the inflection paradigm of finite verbs in English contains *uninterpretable* and *interpretable* features (Chomsky 1995; Cowper 2005): (i) the uninterpretable agreement features (person, number), (ii) the interpretable features of [±perfective] Aspect, [±past] Tense and [±realis] Mood (see section 3.2.) Several proposals have been suggested in psycholinguistics and neurolinguistics with the purpose to characterize the impairments of agrammatic patients in the domain of finite verb flection. These proposals integrating theoretical linguistic assumptions, mainly syntactic ones, can be summarized as it follows.

The *Tree-Pruning Hypothesis* (Friedmann and Grodzinsky 1997; Friedmann 2008) is based on the syntactic hypothesis that there is a hierarchy of functional categories in which Tense and Agreement form separate functional categories and AgrP is lower than TP in clause structure (Pollock 1989). The tree-pruning hypothesis postulates a syntactic deficit according to which the TP and any other functional category above TP are omitted from agrammatic clause-structure representations yielding pruned phrase-structure trees with less hierarchical structures than the language of healthy people. According to this hypothesis, Tense is expected to be more impaired than agreement marking. However, other studies provide evidence that operations below the Tense and Agreement nodes are impaired in Greek, Dutch and English (Bastiaanse et al. 2003) and Aspect, which is below TP, is seriously impaired in Greek agrammatic speakers (Nanousi et al. 2006).

The *Impaired Interpretable Features* (Nanousi et al. 2006; Varlokosta et al. 2006) is related to the distinction between interpretable and uninterpretable features of functional categories (Chomosky 1995; Cowper 2005). This hypothesis states that functional "categories that carry interpretable features may cause more difficulties to non-fluent aphasic subjects" (Varlokosta et al. 2006, 742). By contrast, operations of agreement that involve uninterpretable features are expected to be unimpaired. However, other studies provide evidence that in Norwegian and Greek, Tense/Aspect production is impaired. Specifically, Norwegian patients did not produce any verb form referring to past (Simonsen and Lind 2002) and Greek patients had important problems with past Tense and perfective Aspect, both used to refer to past time (Stavrakaki and Kouvava 2003).

The *Tense Underspecification Hypothesis* (Wenzlaff and Clahsen 2004, 2005) states that the interpretable features of TP are underspecified, i.e. TP is not properly specified for either [+past] or [-past]. Consequently, there is a selective impairment for tense inflection. According to this hypothesis, Tense is expected to be more impaired than agreement and even more than mood [+realis]. Clahsen and Ali (2009) found that for EN, Tense marking is more impaired than both subject-verb agreement and Mood marking. Their findings bring evidence against the *Impaired Interpretable Features* hypothesis (Nanousi et al. 2006), which predicts better performances for agreement (tense and number being uninterpretable features) than for mood (an interpretable feature).

The *PAst Discourse Linking Hypothesis* (PADILIH) (Bastiaanse 2008; Bastiaanse et al. 2011) makes two predictions. Firstly, patients with agrammatic aphasia are selectively impaired in use of grammatical morphology associated with reference to the past, whereas, inflected form, which refer to the present and future are relatively spared. Secondly, this impairment is language independent and occurs in both production and comprehension. This hypothesis

integrates the theoretical distinction between *bound* and *discourse linking* reference[113]. Avrutin (2000, 2006) applied this principle of pronominal reference to Tense and suggested that Tense, being discourse linked, is difficult for agrammatic speakers. Zagona (2003) suggested that reference to present time should be considered as a kind of 'binding relation' based on the fact that temporal coordinates S, R and E are simultaneous. In the case of reference to past time, on the contrary, temporal coordinates do not coincide. Zagona argues in favour of a discourse linking relation between S, R and E regarding reference to past time. As far as reference to future is concerned, Zagona (2013) argued that it is a subclass of the present and therefore not discourse linked. Based on a series of experiments, Bastiaanse (2008) and Bastiaanse et al. (2011) observed that not only reference to past through Tense is discourse linked, but also through periphrastic verb forms ('has walked') as well. Their suggestion is reference to past time through verb inflection in general requires discourse linking, and is expected to be impaired in agrammatic speakers due to the more complex forms. It was shown (Faroqi-Shah and Dickey 2009) that reference to the past through verb inflection produces longer reaction times than verb forms referring to the present. Further evidence for the discourse linking nature of past-reference is brought through event-related brain potentials (ERP) and behavioural (reaction time and acceptability rating) data by Dragoy et al.'s study (2012). They designed a study that focused on processing of time reference violations in which verbal tenses do not match a time frame previously set by adverbial by healthy speakers. The stimuli classified in four conditions included sentences in Dutch such as in Table 3-3.

Table 3-3 Examples of sentences of four experimental conditions

Condition	Time reference in adverb	Time reference in verbal tense	Sentence
PrPr	Present	Present	The waiter who is now grinding the pepper doesn't get a tip.
PsPr	Past	Present	*The waiter who is just before grinding the pepper doesn't get a tip.
PsPs	Present	Past	The waiter who just before ground the pepper doesn't get a tip.
PrPs	Past	Past	*The waiter who now ground the pepper doesn't get a tip.

Dragoy et al.'s aimed at developing Baggio's (2008) findings regarding the link between temporal and pronominal reference. Specifically, Baggio's study proved that processing present time reference marked on the verb in a past time reference context is accompanied by the same ERP effects as processing locally bound pronouns. Consequently, Dragoy and colleagues designed a study that targeted the processing of past and present tense in incongruous contexts making the hypothesis that they rely on different neural processes. They investigated three types of measures: evoked brain responses (ERP), reaction times and

---

[113] According to Bastiaanse et al. (2011, 656-657) binding relations are established within a sentence, such as the relation between the subject and the finite verb, as in *he$_{3rd, sg}$ walks$_{3rd, sg}$*, or between a reflexive and its antecedent, as in *the boy$_i$ is washing himself$_i$*. Discourse linking relations are established beyond the boundaries of a sentence, as in *the boy$_i$ is washing him$_j$* where *him* refers to someone other than *the boy*. *Him* is not locally bound and its referent is identified through discourse linking.

acceptability judgments. Brain responses evoked by time reference violations were explored through several measures:

- P600 wave produced by the brain when it detects a morphosyntactic locally bound anomaly (usually 600 ms after the target word onset).
- N400 wave produced by the brain when it detects a lexical, semantic or conceptual anomaly (usually 400 ms after the target word onset).
- Left Anterior Negativity (LAN) wave produced by the brain when it processes a rule-governed compositional parsing of complex forms across linguistic domains, including both morphology and syntax (usually occurring during 300-500 ms after the target word onset).
- Numerous negative waves (different than N400) produced by the brain when it has difficulty to find a discourse-linked referent (for expressions such as ambiguous words, pronouns).

Turning now to the results of Dragoy et al.' study, the main findings can be summarized as follows. The analysis of the ERP data supports the idea that distinct neural areas, as signalled by different brain reaction patterns, process references to past and to present time. The processing of a past time context disrupted by a present tense verb produced a P600 response triggered by the targeted verb[114]. In contrast, the processing of present time context disrupted by a past time verb did not produce an immediate brain response. However, both past and present time reference produced sentence final negativity, which is a typical response to referential violations in general. Moreover, this ERP data is linked to behavioural data. Investigation of reaction times shows that present time reference violations by past tense verbs were detected later than past time reference violated by present tense verbs, which produced an immediate P600 response. Furthermore, the acceptability rating showed that relative clauses with an adverb referring to the present and a verb referring to past are considered to be less unacceptable than sentences with a past time context disrupted by a present tense verb. When a continuation of the relative clause is provided, participants find it easier to coerce into a meaningful sentence the combination present time adverbial/past tense verb than the combination past time adverbial/present tense verb. Dragoy and colleagues interpret the participants' willingness to wait for further contextual information before judging the relative as unacceptable as an indication for the discourse-linking view of past tense processing. They point out that participants notice the violation of the present/past context with past/present tense verb but respond to it in a qualitatively different manner. This response is showed by the negativity wave in ERP elicited by the end of the sentences.

Dragoy et al.'s study provided new evidence for the theoretical suggestion that time reference expressed by verbal inflections involves similar processing than pronominal reference (Partee 1973; Webber 1988) and that past-time and present-time reference involves different neural processes, dissociation observed both in healthy and aphasic

---

[114] Similar results were found by Steinhauer and Ullman (2002), who investigated only past time reference disrupted by a present time tense in sentences such as *Yesterday, I sail Diane's boat to Boston*. They found that tense disagreement elicited a LAN wave (300-500 ms after the verb onset) followed by a P600 wave. In a later study with a similar design, Newman et al. (2007) reported LAN and P600 effects occurring for disruptions with regular verbs and only a P600 for irregular verbs.

participants (Bastiaanse 2008; Bastiannse et al. 2009, 2011; Faroqi-Shah and Thompson 2007). Moreover, this study supports Zagona's (2003) suggestion that present tense processing requires establishing bound co-reference with the speech time (local binding, i.e. the present tense is deictic), whilst past tense processing requires establishing co-reference with another event time (discourse-linking, i.e. the past tense is anaphorical).

An important question that arises at this point of the discussion is whether these patterns about processing past and present time reference are directly linked to the processing of Tense or they are independent, therefore can be observed in tenseless languages. Qiu and Zhou (2012) and Bastiaanse et al. (2011) investigated this question among others. Qiu and Zhou (2012) designed a study having common features with Dragoy et al. (2012) and investigated brain responses to disagreements between a temporal context set by a temporal adverbial[115] (*jiangyao* 'to be going to' for future and *cengjing* 'in the past' for past time reference) or by the aspectual particle *guo* and temporal noun phrases, as in examples (479), (481) and (483).

(478)   *Next month* the United Nations V+***jiangyao/****will* dispatch a special investigation team.
(479)   **Last month* the United Nations V+***jiangyao/****will* dispatch a special investigation team.
(480)   *Last month* the United Nations V+***ceinging/***dispatch*ed* a special investigation team.
(481)   **Next month* the United Nations V+***ceinging/***dispatch*ed* a special investigation team.
(482)   *Last month* the United Nations V-***guo/***dispatch*ed* a special investigation team.
(483)   **Next month* the United Nations V-***guo/***dispatch*ed* a special investigation team.

Temporal marking in Chinese has to rely on either lexical semantics and discourse principles (in the case of temporal adverbials) or morphosyntactic processing (e.g. suffixation of verbs by the aspectual particle *–guo*). The authors found for Chinese similar patterns for time reference disruptions as those found for European tensed languages. Precisely, disagreements between NPs and temporal adverbials or the aspectual particle produced a P600 wave signalling the morphosyntactic violation and, an additional N400 wave only for the temporal adverbials due to their lexical nature. Moreover, a sustained negativity effect was found after the targeted words and the final words for all types of temporal markers, interpreted as the brain's attempt to correct errors and create a coherent representation of the sentence.

Bastiaanse et al. (2011) also argued that impairments regarding reference to past time occur not only for Tense morphology in tensed languages but also in tenseless languages, such as Chinese. Bastiaanse and colleagues designed a study where three different typological languages were compared (Chinese that expresses time reference through aspectual information, Turkish that has very complex verb inflection paradigms and English that has a combination of free and bound morphemes) and where reference to past, present and future time was tested. They used sentence production tasks and comprehension assessments and tested healthy and Broca aphasia patients. The healthy speakers from the control group all scored at ceiling (i.e. normal scores). Their findings regarding the production and comprehension of aphasic patients can be summarized as follows:

---

[115] According to Qiu and Zhou (2012), Chinese verbs can be combined with temporal adverbs and a small number of aspectual particles to establish temporal reference (as discussed in section 2.1.) Temporal adverbials and aspectual particles must agree with noun phases to provide temporal reference.

Firstly, in all languages, the agrammatic speakers were impaired in producing the grammatical forms for reference to past. English and Turkish speakers performed significantly worse for past than present and future reference. Chinese speakers performed well for sentences where that do not require a specific time reference but poorly for past, present and future reference. The authors assume that this is due to the fact that aspectual adverbs are not obligatory, contrary to English and Turkish verb inflection. A qualitative analysis of Chinese production shows that the aspectual adverb was most often omitted (the sentence remaining grammatical when a lexical adverb expressing the time frame also occurs) and when performed, the past (*le*) and future (*yao*) adverbs were substituted by the present marker *zai*.

Secondly, as far as production is concerned, there was no significant difference between the two tensed languages: speakers performed similarly regardless the complex verb inflection paradigm in Turkish and the use of periphrastic form in English. Finally, in all languages, the agrammatic patients were impaired for the comprehension of sentences containing reference to past. This is significantly worse than comprehension of sentences containing present time reference. Comprehension of future morphology is significantly worse than present but better than past time reference for all patients (though for the Turkish patients, the difference between past and future was only close to the significance threshold).

Bastiaanse and colleagues (2011) showed therefore that agrammatic speakers show performance patterns that are the same for all three languages: past time reference is more impaired than present reference, past reference is more or equally impaired than the future reference, and future reference is more impaired than the present reference. They suggest that this data should be interpreted at the morpho-semantic interface. Precisely, temporal information about the event moment relative to the moment of speech must be encoded (in production) and decoded (in comprehension) grammatically. In other words, temporal location through E related to S is encoded information both in tensed and tenseless languages. Other studies using aphasiological data indicate that not only is past time reference worse than present time reference, but also perfective aspect is more impaired than imperfective aspect in agrammatic aphasia (Nanousi et al. 2006; Stavrakaki and Kouvava 2003).

These studies suggest that irrespective of the category conveying temporal location of eventualities (Tense or/and Aspect) and the type of linguistic expressions (inflexions, auxiliaries, free or bounded morphemes), grammatical expressions for past and/or event completeness are impaired in aphasia, as Dragoy and Bastiaanse point out (2013, 114).

I will now consider the nature of Aspect deficits, as well as the relation between Tense and Aspect. Dragoy and Bastiaanse (2013) investigated for Russian aphasic patients the hypothesis suggested in Bastiaanse et al. (2011) that verb forms expressing reference to past time or conveying perfective semantics are more impaired than verb forms expressing reference to the non-past or conveying imperfective semantics, both for production and comprehension. Dragoy and Bastiaanse (2013) point out the strong relationship between

temporal reference through Tense, Aspect and Aktiobsart in Russian [116]. They note that Russian children strongly prefer to use perfectives to refer to past time, and imperfectives to refer to the present as suggested by Gagarina (2004).

Dragoy and Bastiaanse tested the following hypotheses: (i) past forms are more impaired than the present forms (according to PADILIH), (ii) the production of perfective verbs is more impaired than imperfective verbs production and (iii) due to time reference and aspect interaction in Russian, non-past time reference is advantageous only for imperfective verbs and past-time reference for perfective ones. They used sentence completion tasks and tested aphasic patients (both fluent/Wernicke aphasia and non-fluent/Broca aphasia). Their results showed a significant main effect of temporal reference (as predicted by PADILIH). There was no significant for Aspect uniquely (invalidating the second hypothesis) but there was a significant interaction of Tense and Aspect. Precisely, reference to non-past is better preserved than reference to past, but only for imperfective verbs. In contrast, for perfective verbs, reference to past is better preserved than reference to non-past (p. 120).

To sum up, Bastiaanse et al. (2011) and Bos et al. (2013) reported that veral tenses expressing reference to past or expressing perfective semantics are more impaired that verbal tenses expressing reference to non-past (present or future, one category according to Comrie 1985) or conveying imperfective semantics, both in comprehension and in production, for tensed and tenseless languages. Dragoy et al. (2012) and Qiu and Zhou (2012) found that processing past and present time reference involves different neuronal processes both for tensed and tenseless languages. Clahsen and Ali (2009) found that agrammatic patients are more impaired in Tense than in Mood (the subjunctive) or agreement marking. Dragoy and Bastiaanse (2013) provided evidence for an interaction between Tense and Aspect: imperfective aspect are better produced in the non-past, whereas production of perfective aspect is better preserved in the past time framework.

The question that arises at this point of discussion regards the way in which current semantic and pragmatic theories account for these neurolinguistic findings. Probably RT is the most appropriate framework to provide a plausible account. My assumption is that there would be two types of answers:

- Temporal reference through E related to S via R is procedural information encoded by Tense (see section 3.1.3.4)
- Temporal reference through E related to S is conceptual information (my proposal in Grisot and Moeschler 2014, and developed in this thesis in chapter 7).

I would argue that the current assumptions of the procedural account are insufficient for explaining these findings and raises several questions. Firstly, if temporal reference is procedural information encoded by Tense, then the status of temporal reference through Aspect, Aktionsart and other means in tenseless languages should be considered. More

---

[116] As for processing of lexical and grammatical aspect in general, Yap et al. (2009) demonstrated that the perfective sentences were processed more quickly and accurately with accomplishment verbs, whereas imperfective sentences were processed more quickly and accurately with activity verbs. Yap and colleagues interpreted these results arguing that accomplishments favour the construction of mental representations of perfective or completed situations because they have natural endpoints. The situation is parallel for activities that do not have natural endpoints and favour the construction of mental representations of ongoing or imperfective situations.

precisely, are temporal adverbs and aspectual particles in Chinese also procedural markers? How is it possible that in tenseless languages several types of linguistic expressions encode the same instruction as that encoded by Tense in tensed languages? Secondly, if tenses encoded only instructions about the temporal location of eventualities (therefore no specification of the type [+past] or [-past]), then there should be no significant difference in the processing of reference to past, present or future times. As the studies described in this section show, this is not the case.

In this thesis I argue in favour of the second account, more specifically, reference through E related to S is a conceptual type of information. Concepts are language independent and languages use different linguistic means to express them. Precisely, Tense is the dominant source in tensed languages, whereas Aspect and Aktionsart, temporal adverbs and other means (as discussed in section 2.1) are primary in tenseless languages. My suggestion is that it is reference to the concept of pastness that it is more impaired then reference to the concept of non-pastness (present and future). I will argue that past time verbal tenses in EN, FR, IT, RO among many other languages, encode an *ad hoc* concept of pastness, which is determined contextually through the contextual saturation of the temporal variables S and E (see section 7.2.1).

3.3.2 *Automatic processing*

Computational linguistics (CL), natural language processing (NLP) and machine translation (MT) are domains dealing with natural language from a different perspective than linguistic fields such as Semantics, Pragmatics and Syntax. Their general purpose is to make automatically what humans naturally do, that is produce and understand language. As for neurolinguistics and psycholinguistics, there is a strong relationship between linguistics and CL, NLP and MT. Specifically, the automatic treatment of language draws its models on linguistic theories describing qualitatively the functioning of human language, but also on high quantities of data and on frequent behaviours of linguistic expressions. Language models developed in CL, NLP and MT find patterns of linguistic expressions and semantic interdependencies, allowing researchers to generalize behaviours, such as, the parallel between temporal and pronominal reference suggested by Partee (1973) and further developed in CL framework by Webber (1988). Well-known works on discourse structure and lexical aspect, such as Dowty (1979; 1986), Moens and Steedman (1987; 1988), Steedman (1997), Moens (1987) among many others, were done in CL framework. The research presented in this thesis (and carried out in the COMTIS and MODERN projects) is original due to its aim of introducing pragmatic knowledge in language models used in NLP and MT for example, such as Grisot and Meyer (2014) and Meyer et al. (2013).

The literature on the processing of temporal reference has focused in the past few years on issues such as event ordering (relative to each other), time stamping (i.e temporal anchoring of a situation) and generation of words expressing temporal relations for individual languages, most often, for EN. I will present some of these studies in the NLP section. Other studies were interested in proposing cross-linguistic temporal reference mappings, most often closely linked to MT needs. I will present some of these studies in the MT section.

3.3.2.1   Natural Language Processing

In this section, I will describe three main previous studies in the NLP field related to temporal information. The first is a computational model of the semantics of Tense and Aspect (Passonneau 1988). The second is a model for processing and automatically annotating temporal information in discourse, namely the TimeML annotation scheme model proposed for EN by Pustejovsky and colleagues (2005) and adopted by Bittar (2010) for FR. The third is Li et al.'s (2001; 2004) model for processing Chinese.

Passonneau (1988) described a processing system called PUNDIT, which processes references to situations and the intervals over which they hold using an algorithm that integrates the analysis of verbal tenses (i.e. Tense) and aspectual information (i.e. Aspect and Aktionsart). The algorithm is developed for EN texts. Specifically, information from Tense and Aspect (perfective/perfect[117] or progressive), as well as temporal adverbials such as *before*, *after* and *when* are used for deriving three complementary pieces of information:

- Determine whether the situation is associated to realis or irrealis world. Different treatments are needed whether the situation refers to actual time or to potential time.
- Determine the internal temporal structure of the predicated situation, i.e. inherent temporal information of the VP, as one of three situation types: *state, process* and *transition event* (achievements and accomplishments in Vendler's terminology).
- Determine the temporal location of the actual situation with respect to the moment of speech/text production or to the times of other situations, with the help of Reichenbachian temporal coordinates E, R and S.

The internal temporal structure of a situation consists of one or more intervals. Each interval is characterized through two features, *kinesis* and *boundedness*. Kinesis pertains to the internal structure of an interval and can be *stative* or *active*. Stative kinesis signifies that "each subinterval is equivalent to any other subinterval with respect to the asserted situation" (Passonneau 1988, 47). Processes and transition events have active kinesis involving change from one subinterval to another. Boundedness regards the fact than an interval is bounded or not and constrains the manner in which the situations are located in time (i.e. temporal reference). The intervals associated with states are inherently unbounded, although they can become bounded by an appropriate temporal adverbial. Processes (activities in Vendler's terms) are generally unbounded and can become unspecified for boundedness if the verb is progressive. In (484), the clock time is interpreted at falling within the unbounded interval of *sounding*, but in (485), where the verb is not progressive, the clock time can be interpreted as falling at the inception of the process or roughly locating the entire process (Passonneau 1988, 47).

(484)   The alarm was sounding at 8am.
(485)   The alarm sounded at 8am.

These temporal pieces of information are assembled in a context-dependent compositional semantics framework. Passonneau points out the complexity of computing

---

[117] The model uses the term *perfect* for referring to the English Present Perfect and Past Perfect verbal forms. Perfect verbal forms (*relative* in Reichenbach's terms) present an R distinct than E.

temporal information coming from several sources, since the contribution of each distinct component can depend upon co-occurring elements. Her suggestion is a model of extracting temporal information based on separating temporal analysis into distinct tasks, each task targeting one type of temporal input. Each task provides input for the next stage of analysis, and this must be done as explicitly as possible to avoid conflicting with the subsequent processing. The algorithm for temporal analysis of an inflected verb contains three modules. The first module computes the actual time (realis) from temporal information coming from Aspect, Aktionsart and Tense. Only realis sentences are considered for further analysis. The second module derives the inherent temporal structure of the situation from two temporal parameters: lexical aspect and progressive aspect. The output of the second module, precisely an explicit representation of the situation's temporal structure and the event time, is sent to the third module, which derives temporal location of the situation from the last two parameters: perfect verbal form and tense. Temporal location is established with the help of Reichenbachian temporal coordinates. However, it diverges from Reichenbach primarily by distinguishing between the event time and the temporal structure of a situation (Passonneau 1988). Module three is illustrated in Table 3-4.

Table 3-4 Module 3: Temporal location

Parameter	Value	Rules
Perfect	Yes	E<R
	No	E=S
Tense	Past	R<S
	Present	R=S

The possible combinations of the values of all considered parameters are provided in Table 3-5. A situation is therefore located in time regarding the parameters of Aspect, Aktionsart and Tense and its interpretation depends on this temporal location. The Simple Present locates unbounded temporal structure coinciding with S, while processes and transition events do not refer to actual moment of speech of the utterance, as shown by the interpretation of (486). The SP locates the event time of any temporal structure prior to S. However, each temporal structure provides differences in interpretation regarding the surrounding of the event time. Perfect verbal forms provide supplementary information comparing to the simple forms, specifically about the relation between R and E.

(486) The pump operates.

Table 3-5 Possible combinations of temporal location of situations

Tense	Aspect	Stative	Process	Transition Event	Location
Present	Simple	unbounded	not actual time	not actual time	E=S=R
	Perfect	unbounded	unspecified	unspecified	E<R=S
	Progressive	unbounded	unbounded	unbounded	E=S=R
Past	Simple	unbounded	unspecified	bounded	E=R<S
	Perfect	unbounded	unspecified	bounded	E<R<S
	Progressive	bounded	unbounded	unbounded	E<R=S

To the best of my knowledge, Passonneau's account of temporal information in discourse is the first model that integrates semantic information coming from several linguistic sources.

Another semantic account of temporal information, called the Specification Language TimeML, was developed in the AQUAINT[118] programme. TimeML is a semantic annotation framework for temporal information in discourse and represents guidelines to trained humans who carry out the annotation (Pustejovsky, Ingria, et al. 2005; Pustejovsky, Knippen, et al. 2005)[119]. TimeML was designed to address four issues regarding temporal information:

- Temporal location of situations (identification and anchoring it in time)
- Ordering of situations with respect to one another (lexical and discourse ordering)
- Reasoning with contextually underspecified temporal expressions (such as *last week* or *two weeks before*)
- Reasoning about the persistence in time of situations

TimeML considers all temporal objects in a discourse, broadly grouped in *temporal expressions* (adverbials and connectives) and *events*. The class *events,* which includes *inflected verbs* and *event nominals*, is a generic term used for verbs describing various types of states and events. It makes reference to Reichenbach's (1947) description of verbal tenses, Vendler's (1957; 1967) aspectual classes, the distinction between *lexical* and *grammatical* aspect and Bach's (1986) notion of *eventualities*. The annotation language consists of a set of *basic tags* for expressing events, explicit temporal expressions and function words and a set of *links* between the annotated elements, which have different types, such as *temporal, subordination* and *aspectual*.

The tag <EVENT> is used to annotate both inflected verbs (predicative and non-predicative tenses) and events expressed by nouns. Verbal tenses are expressed in terms of a combination between *Tense* (with a choice among *present*, *past* and *future*) and *Aspect* (with a choice among *progressive, perfective, progressive-perfective* and *none*). Verbs are categorized in seven classes: *reporting, perception, aspectual, states, demanding an action, demanding a state* and *occurrences.* These classes are relevant for the type of relation (link) they require. The tag <TIMEX3> is used to mark up explicit temporal expressions referring to *day times, dates, durations, sets.* The tag <SIGNAL> is used to annotate function words, which indicate how temporal objects are to be related to each other. Signals are generally: temporal prepositions (*on, in, at, from, to, before, after, during*, etc.), temporal conjunctions (*before, after, while, when*, etc.) and special characters ("-" and "/", in temporal expressions denoting ranges, such as *September 4-6, April 1999/July1999*, etc.).

The tags <TLINK>, <SLINK> and <ALINK> serve to capture the different types of relations that exist between two events (in the broad sense used in this framework) and between an event and an explicit temporal expression. These links can have a *temporal* nature

---

(such as *before, after, includes, simultaneous, during, identity,* etc.), a *subordination* nature (such as *evidential, factive, counter-factive, conditional,* etc.) and an *aspectual* nature (such as *initiates, culminates, terminates, continues,* etc.).

The example in (487) is interpreted in TimeML as in (488) and it can be paraphrased in the following terms: the temporal adverb *today* is annotated with the tag TIMEX3, which expresses a date and has the identification tag t32; there is a temporal link with the value *before* between the event number 2 from the sentence and this adverbial shown by the TLINK tag at the end of the formal description. In the sentence two events are mentioned, the first one is expressed by the verb *learned* (which is described as a reporting verb, expressing past tense) and the second event is expressed by the verb *has taken* (which is described as an occurrence verb, expressing present tense and the perfective aspect). This kind of annotation carried out by trained humans allows the explicitation of temporal information that is implicit at the discourse level.

(487)   Finally, today we learned that the space agency has finally taken a giant leap forward.

(488)   <s>*Finally* <TIMEX3 tid="t32" type="DATE">*today*</TIMEX3>,*we* <EVENT eid="e1" class="REPORTING" stem="learn" aspect="NONE" tense="PAST" polarity="POS" pos="VERB"> *learned*</EVENT> *that the space agency has finally* <EVENT eid="e2" class="OCCURRENCE" stem="take" aspect="PERFECTIVE" tense="PRESENT" polarity="POS"pos="VERB">*taken*</EVENT> *a giant leap forward.* <TLINK eventInstanceID="e1" relatedToTime="t32" relType="BEFORE"/></s>

The metadata markup language TimeML is therefore a formal framework that integrates three types of semantic temporal information: (i) temporal anchoring of situations with respect to S and R, (ii) temporal ordering of situations relative to one another, both intrasententially and in discourse, (iii) the semantics for underspecified temporal expressions by integrating them in the overall interpretation of the discourse. Manually annotated corpora by humans with the TimeML language are useful tools for finer-grained analyses of temporal information. TimeML is a significant example of the efforts done by researchers to integrate temporal information coming for several sources and to explicitate the various types of relations existing among situations. However, as I have argued in section 3.1.3, processing temporal information cannot be done uniquely from linguistic or semantic sources.

Both Passonneau's model and TimeML are models developed for tensed languages, such as EN and FR. Li et al.'s (2001, 2004) developed a model for processing temporal reference in Chinese. They reported a computational model based on machine learning algorithms. The core model consists of a set of rules combined with a set of linguistic features for the purpose of temporal relation resolution. The linguistic features used are Chinese words that can function as temporal indicators: time words (e.g. *year, month)*, time position words (e.g. *a few days ago*), temporal adverbs (e.g. *lately, recently),* auxiliary words, and verbs, aspectual markers (e.g. *le, zhe* and *quo*), prepositions and special verbs among others. Temporal relations are described in terms of E, R and S (Reichenbachian coordinates). The TICS system (Temporal Information-extraction from Chinese Sources) receives as input financial texts, analyses each sentence one by one in order to extract temporal information, and represents each piece of information in a concept frame. All concept frames are linked based

on temporal relations holding among events. This model points to the necessity of NLP models to identify temporal relations holding among eventualities in order to have accurate results.

To sum up, in this section I have discussed three NLP studies targeting automatic processing of temporal information at the discursive level. It was shown that automatic systems make use of temporal information coming from various linguistic sources: verbal tenses, grammatical and lexical aspect, location of eventualities with respect to Reichenbachian coordinates E, R and S, temporal adverbials and other linguistic markers relevant especially in tenseless languages.

3.3.2.2   Machine Translation

In the MT field, two main types of automatic translation systems exist: *rule-based* and *statistical* systems. Rule-based systems were the first to be created in the 1970s, such as Systran (currently hybrid between rule-based and statistical system), then the German Verbmobil in the 1990s (for speech-to-speech translation) and the Swiss ITS-2 in 2008[120]. For these systems, a large set of lexical and/or syntactical rules had to be written by linguists and manually implemented. As pointed out by Meyer (2014), this costly procedure made it hard to adapt these systems to other language pairs, translations directions, or stylistical registers. The functioning of rule-based systems is designed at three levels. The first and bottom level consists of the translation is done word-by-word, with possible re-orderings of the words. The second and medium level, the system operates at the syntax level via transfer rules implemented on syntax trees from SL to TL. The third and most complex level is given by the building of an *interlingua*, which is 'completely language-independent semantic representation of the source text's meaning' used to directly generate the target text (Meyer 2014, 4). However, building the interlingua proved to be a very problematic task because of the difficulty to integrate world and domain knowledge.

Therefore, throughout the 1990s, most of the current research on MT focused on statistical systems. In SMT, where there is no rule-based processing, the goal for the system is to learn the correct translations of words, phrases and sentences from large corpora translated by humans, i.e. parallel corpora that nowadays exist in several languages such as EuroParl (Koehn 2005). Statistical machine translation systems make use of a statistical *language-model* and great quantities of training data. The language-model specifies the probability of the string of words[121] considered by the SMT system, as well as syntactic and lexical information of the source and target language. The correspondence between a source phrase (at the word level or chunks of words) and a target phrase is extracted from a parallel and aligned corpus. To build a *translation model*, the system uses the pairs of phrases and the

---

[120] To be found at http://www.systranet.com/translate, http://verbmobil.dfki.de/ww.html and respectively, http://latlapps.unige.ch/Translate.

[121] SMT systems use word or phrase alignment algorithms for aligning the words of a sentence in two languages, a SL and a TL. There are four types of alignment (Samardzic 2013): (i) *one-to-one* (when corresponding single words are identified, i.e. pairs of words), (ii) *one-to-null* (used to describe words that occur in one language but no correspondent can be found in the other language), (iii) *one-to-many* (when one word in a language corresponds to several words in the other language) and (iv) *many-to-many* (when no single word is an alignment unit). The first three types are called word-based alignments and the last one is called phrase-based alignment.

information from the language-model.

The functioning of an SMT system can be described in three phrases (as explained by Meyer 2014). The first is the *training* stage, in which the system learns the most likely correspondences and re-ordering the chunks of words from parallel corpora. The second is the *tunning* stage, in which the system trains on a much smaller text, ideally of the same register as the target text, in order to optimize the language pairs identified in the first stage. The third is the *decoding* or *testing* stage, in which a new text is handled to the system for translation. In this step the system tries to find the most likely phrase pairs and re-combines these hypotheses based on probability scores from the translation and the language model available. On of the most often used, freely accessible statistical MT system is Google translate.

Other attempts to improve the results of SMT systems were done. Two of them were to create hybrid systems using both linguistic rules and statistical methods (such as Systran, Reverso and Linguatec[122]) and to use additional knowledge within the SMT paradigm. For the latter, researchers proposed *factored translation models* (Koehn and Hoang 2007), which most often are used to add morphologic, semantic or pragmatic information. This information is provided to the system via annotation of the parallel data. The training data is enriched with the linguistic information wanted and it is automatically annotated by a *classifier*. A classifier is a tool that makes use of machine learning algorithms[123] based on, usually, human annotated data. A classifier is a machine-learning tool that will take data items and place them into one of the available classes. One type of classifiers is the *maximum entropy* (MaxEnt) classifier, which can be built with the Stanford Classifier Package (Manning and Klein 2003). The underlying principle of maximum entropy is that, when assigning a class if there is no external knowledge, one should prefer uniform distributions, thus assign uniformly the considered classes. Annotated data used for training these classifiers provide external knowledge and thus inform the automatic labelling technique where to be minimally non-uniform (i.e., where not to provide uniform distributions of the tags). Iterative runs of the classifier results in automatically labelled or annotated texts with the considered features. The classifier has a crucial role for an SMT because it produces automatically tags that increase the probability of a certain string of words in TL to be the correct translation. For this reason, much work has been done for the construction of the classifiers, such as the research carried out in the COMTIS and MODERN projects that focused on western-european tense-prominent languages.

Olsen et al. (2001) and Ye et al. (2006) aimed at improving the machine translation from Chinese to EN. Ye and colleagues (2007) were interested in machine translation from EN to Chinese. The different strategies used for encoding temporal information in EN and Chinese are challenging for the automatic translation of tense and aspect. Ye and colleagues point out that neither word-based alignment nor phrase-based alignment can capture the mapping between the tense markers in EN (morphemes) and the aspect markers of the corresponding

---

[122] Reverso is free online and it is available at http://www.reverso.net/text_translation.aspx?. Linguatec is private and it can be found at http://www.linguatec.net/.

[123] Samardzic (2013, 112) explains that the data which machine learning algorithms take as input are considered as *experience*. A computer programme "learns from experience" if its performance with respect to a task improves with experience, i.e. by dealing with the data.

Chinese verbs (lexemes).

Ye and colleagues (2006) built a classifier that generates tense marking in English. The classifier learns the mapping between English and Chinese from a set of features coming from a training set of data. Since the purpose of the SMT system is to translate into English, they used features of English for predicting tense marking. Their main argument is that NLP work must aim at building systems that follow the mechanisms of human brain in order to optimise their performances. In their words (2006, 50):

> The bottleneck in Artificial Intelligence is the unbalanced knowledge sources shared by human beings and a computer system. Only a subset of the knowledge sources used by human beings can be formalized, extracted and fed into a computer system.

The features based on the knowledge shared with human beings are called by Ye et al. *latent features*. Olsen and colleagues (2001) illustrated the value of latent features by showing how lexical aspect or the telicity of the VP improves the translation of temporal reference from Chinese to English. Ye and colleagues (2006) used several surface features (formal features) and two latent knowledge sources, precisely *telicity* as proposed by Olsen et al. (2001) and *event ordering* as implemented in the TimeML annotation scheme. The surface features used for generating tense markers in English are (2006, 50):

- Type of speech act.
- The syntactic structure in which the current verb is embedded.
- Occurrence of temporal adverbials and aspectual markers.
- Distance in number of characters between the current and the previous verb, and if the two verbs are in the same clause or not.

The two latent features are assumed to be used by human beings in tense resolution (though psycholinguistic and neurolinguistic studies have only recently begun to investigate them, as I have shown in section 3.3.1). Information about the lexical aspect is used in terms of *telicity* (i.e. the verb's ability to be bound in a certain time span) and *punctuality* (i.e. punctual verbs or achievements in Vendler's terms). The authors point out that a verb's telicity value is contextual-dependent. The second latent feature concerns temporal relations holding among eventualities. The authors defined temporal relations in terms of precedence, inclusion, overlapping and lack of temporal relation. They used therefore human-annotated data with these two latent features. The SMT system using the classifier trained on surface and latent features had significant better results than other systems, specifically than SMT systems without a trained classifier and SMT systems trained only on surface features. Ye and colleagues (2006) provided evidence that lexical aspect and temporal relations holding among eventualities are significant factors for predicting verbal tenses in a TL. In this thesis, specifically in sections 6.1.8 and 6.3, I suggest a model that uses several latent features, such as Aspect, Aktionsart, temporal and causal relations holding between eventualities (grouped under the [±narrativity] feature encoded by Tense) for predicting the verbal tense in several target languages. The advantage of the research presented in this thesis (see section 6.3 for of the application of the theoretical model of temporal reference to NLP and MT) compared to previous models for SMT, all features are captured automatically.

Ye and colleagues (2007) report the building of a classifier that generates aspectual

markers in Chinese, more precisely *le, zhe, guo* and NULL when none of the three occurs. Since the purpose of the SMT system is to translate into Chinese, the features used for predicting aspect marking correspond both to English and to Chinese. There were used five surface features and one latent feature (2007):

- Syntactic features that can influence the verb's tendency to take an aspectual marker.
- Positional features pointing to the fact the occurrence of a verb with another can influence the verb's tendency to take an aspectual marker.
- Signal lexeme features pointing to the fact that the aspectual markers considered present certain lexical occurrence patterns (for example with some auxiliary words and not with others).
- Phonological feature pointing to the fact that aspectual markers are incompatible with idioms that have four Chinese characters.
- English verbal tense feature pointing to the fact that verbal tenses play the same role as aspectual information in Chinese, more precisely expressing temporal reference.
- Lexical aspectual features pointing to the theoretical assumption that the inherent features of the VP play an important role for establishing temporal reference.

Verbal tense in EN and lexical aspect have been manually annotated. The classifier's performance was significantly better than a simple classifier, which always assigns the most frequent aspect marker (which is *le*). All features used for predicting aspectual markers in Chinese were significant but behaved differently for each of the three aspectual markers considered. For example, the lexical aspectual feature was significant only for the prediction of the aspectual marker *zhe*, whereas the EN verbal tense feature was significant for predicting the occurrence of *le* and NULL. These two studies involving the translation from and into Chinese, a tenseless language, point to the fact that dividing temporal information coming from Tense, Aspect and Aktionsart and using it as latent features is useful for improving the translation of a text with respect to temporal reference.

To sum up, in section 3.3, an account of human and computational processing of temporal information in discourse was given. Neurolinguistic studies indicated that the brain processes differently Tense, Aspect and Aktionsart when they make reference to past and when they make reference to non-past times. Computational models developed in NLP and MT pointed out that Tense, Aspect and Aktionsart, temporal connectives and adverbials, as well as their interrelations, are relevant linguistic information for having accurate NLP models and for improving the results of SMT systems.

In the COMTIS and MODERN projects framing this thesis, the research focused on finding which semantic and pragmatic features are candidates for improving the translation of verbal tenses, the human annotation of parallel corpora with these features, the building of the classifiers used to automatically annotate the data with the chosen semantic and pragmatic features and finally, the building and training of SMT systems on the automatically annotated data.

3.4   Conclusive remarks

This chapter gave an account of temporal reference and its ingredients that goes beyond purely semantic assumptions.

Firstly, I addressed cognitive pragmatic approaches and, in particular, Relevance Theory, which attached cognitive foundations to the language comprehension process. In RT language comprehension is a two-phase inferential process that includes *a decoding phase* and *a central inferential phase* in which a linguistically encoded logical form is contextually enriched and used to construct hypotheses about the speaker's intended meaning.

A central issue assumed in RT is that the linguistic expressions that a speaker utters underdetermine the content that she communicates not only at the level of implicatures but also the propositional contents she communicates explicitly (that is the explicature of the utterance). The hearer must therefore recover inferentially the speaker's intended meaning, at the level of explicatures and implicatures. In recent work, Moeschler argued in favour of the complexity of layers of meaning consisting of semantically determined meaning (presuppositions and entailments) and a set of pragmatically determined meanings (explicatures and implicatures). His proposal is that the different levels of meanings may be characterized in terms of explicitness, inferentiality, context-dependence, truth-functionality and speaker's commitment and behaviour with negation.

Another issue proposed in RT is that linguistic expressions encode conceptual and procedural (i.e. instructions for manipulating conceptual representations) information that contribute and, respectively, constrain, the interpretative process. As far as temporal reference and its ingredients are concerned, it is argued in the literature that Aktionsart encodes conceptual information whereas Tense and Aspect represent instructions to manipulate these conceptual representations. Moreover, it is currently suggested that temporal relations holding among eventualities should be treated as semantic discourse relations (in SDRT), as default interpretations associated to individual verbal tenses (as in Kamp and Rohrer 1983, Moeschler 2000b, Saussure 2003), as conversational implicatures (as suggested by Grice) or as pragmatically determined aspects of *what is said* (as suggested by Wilson and Sperber 1998). In more recent work, such as Grisot and Moeschler (2014) and more detailed in this thesis, it was suggested that temporal relations holding among eventualities represent procedural information encoded by Tense. Tense encodes the instruction to locate temporally eventualities with respect to one another, an instruction which leads to a positive or negative contextual value.

Secondly, syntactic theories have also pointed to the crucial role played by Tense and Aspect for the interpretation of a sentence. In the Chomskyan-based theories, it is generally accepted that in tensed languages the Infl Phrase (consisting of Tense and Aspect) is the functional head of the sentence. In tenseless languages, Aspect is the functional head of the sentence. Tense, Aspect and Mood are interpretable features (Chomsky 1995, 2000). In pragmatic terms, this means that their value must be worked out in the interpretative process. For example, in Romance languages and in EN, a clause is interpreted as perfective unless it is overly marked as imperfective, whereas in Slavic languages the situation is the opposite.

Thirdly, neurolinguistic studies pointed out to the fact interpretable features Tense, Aspect and Mood are particularly impaired in both Broca's agrammatic aphasia. Neurolinguistic evidence indicated that reference to past time is more impaired than reference to present time both in tensed and in tenseless languages. The same finding holds for temporal reference expressed by Tense, Aspect, Aktionsart or other means. Moreover,

investigations in both human and automatic processing of temporal information drew attention to the interrelations among Tense, Aspect and Aktionsart.

In conclusion, this chapter addressed recent developments in the study of temporal reference and its ingredients in pragmatics, syntax, neurolinguistics and automatic treatment of language fields by pointing to their advantages and their limits. The model of defended in this thesis is based on new empirical evidence with respect to the functioning of temporal reference. In what follows, chapter 4 is dedicated to the methodology followed in the empirical work, which will be described in chapters 5 and 6.

4 Contrastive studies, parallel corpora and linguistic experiments

4.1 Contrastive studies

Contrastive Linguistics, also known as Contrastive Analysis (CA), is defined by Johansson (2003, 31) as "the systematic comparison of two or more languages, with the aim of describing their similarities and differences". Current research in contrastive studies of typologically related or unrelated languages is interested in topics such as language universals and universal grammar, human and machine translation. When modern linguistic theories[124] flourished in the 20th century, a growing interest in methodological and theoretical components of contrastive studies appeared. If originally, all contrastive studies were pedagogically motivated and oriented, in the last thirty years, CA area has incorporated the distinction between *theoretical* and *applied* contrastive perspectives (Fisiak 1981, 2–9). According to Fisiak, theoretical and applied contrastive studies can be defined as it follows:

> Theoretical contrastive studies give an exhaustive account for differences and similarities between two or more languages, provide an adequate model for their comparison, determine how and which elements are comparable, thus defining such notions as congruence, equivalence, correspondence, etc…Applied contrastive studies are part of applied linguistics. Drawing on the findings of theoretical contrastive studies they provide a framework for the comparison of languages, selecting whatever information is necessary for a specific purpose, e.g. teaching, bilingual analysis, translating, etc.

If early contrastive studies were made in a more or less intuitive way, starting with the 20th century, contrastive studies are run on the basis of a well-established methodology. The three most known works that offer a methodology for contrastive studies are James (1980) and Krzeszowski (1990) and Chesterman (1998).

Mainly, the methodology used in contrastive studies consists of a first phase of monolingual description of the data (the phenomenon to be analysed), followed by the juxtaposition of the two or more monolingual descriptions and the analysis of the elements according to a *tertium comparationis* (TC) (James 1980; Krzeszowski 1990). Chesterman (1998), in contrast, proposes a different methodology, which focuses on formulating hypotheses about the data and testing them empirically.

The minimum prerequisite of CA is to discuss the considered phenomena in the two

---

[124] According to Krzeszowski (1990) the history of contrastive studies begins with Aelfric's *Grammatica,* that is a grammar of English and Latin, based on the implicit assumption that the knowledge of grammar of one language may facilitate the learning of another language. In the 17th century, another grammarian, John Hewes (1624) introduced the idea that the knowledge of the native grammar cannot only facilitate the learning of a foreign language but also interfere with it. The history of contrastive studies passes through other grammarians, such as Howel (1662), applied the idea of positive transfer of the native language grammar on a foreign language grammar and adjusted their grammars of English or of Latin to the needs of speakers of various native languages.

languages using the same descriptive model, in other words they should be cross-linguistically parallel (James 1980, 63). One of the possible methods to achieve parallel descriptions is to:

> … describe L1 and L2 data independently, using the models which yield the fullest descriptions of either language, and then translate these two descriptions into a form which is model-neutral. (p. 64).

This means to describe the phenomena considered in each language, respecting their language-dependent features, and then to formulate a language-independent and model-neutral description, taking into consideration the generalization of the two monolingual descriptions.

The second step of a contrastive analysis is the juxtaposition of the two descriptions and building of a TC. In James' words, the researcher decides what elements in one language can be compared with what elements in another language and on what basis (p. 64). If in early studies this decision was made based on a *bilingual competence* (Krzeszowski 1990, 36) of the linguists, in recent studies this decision has been taken according to certain well-established criteria. These criteria are to be set in accordance to a background of similarity. All comparisons involve the basic assumption that the objects to be compared have something in common, against which differences can be stated. This common platform of reference is the TC.

The setting of the TC is closely related to the principle of *comparability*, where comparability does not presuppose absolute identity, but merely a degree of shared similarity (James 1980, 168), established as follows:

> The first thing to do is to make sure that what we are comparing like with like: this means that two (or more) entities to be compared, while differing in some respect, must share certain attributes. This requirement is especially strong when we are *contrasting*, i.e. looking for differences, since it is only against a background of sameness that differences are significant.

The concept of *sameness* (used to define the TC as the background of sameness against which the differences are significant) is closely linked to that one of *equivalence* (James 1980, 175-178). The term *equivalence* itself is problematic because of the several types of equivalence (Krzeszowski 1990), such as *semantic* (two phrases are considered semantically equivalent if they have the same lexical and semantic form and if they refer to or express the same concepts in the real world) and *translation* equivalence (two phrases or sentences are considered to be translation equivalents if they share meanings in two languages but which are not formally constrained). According to Krzeszowski (1990, 17), some translation equivalents are often semantically non-equivalent. For example, the PresPerf in EN and the PC in FR (often used as translation equivalents) have a formal similarity but which is not matched by a complete semantic similarity. These two types of equivalence are often erroneously considered to be identical. Another type of equivalence that is important for contrastive studies is the *pragmatic* equivalence. The pragmatic equivalence between two phrases involves that the two sentences have the same function (both at the explicit and at the implicit level) in a discourse. Krzeszowski argued that the prerequisite for a good

translation is the pragmatic equivalence.

These issues raise the problem of what kind of equivalence should be included in the TC. According to James, in CA translation equivalents should be limited to those, which are both semantically and pragmatically equivalent. Besides the type of equivalence changes with the type of contrastive study being carried out (e.g. phonetic, phonological, syntactic) (Krzeszowski 1990, 21). In his words:

> …equivalence is the principle whereby *tertium comparationis* is established inasmuch as only such elements are equivalents for which some *tertium comparationis* can be found, and the extent to which a *tertium comparationis* can be found for a particular pair of items across languages determines the extent to which these elements are equivalent. Thus, equivalence and *tertium comparationis* are two sides of the same coin.

Building a consistent TC is crucial step for a reliable contrastive study because the choice of the TC constitutes a determining factor in establishing similarities and differences between the elements compared, and consequently, for the results of the comparison. Depending on the TC adopted, the elements to be compared can turn out to be similar or different (Krzeszowski 1990, 15).

Broadly speaking, in each area of comparison three situations may arise after the proper comparison step:

- An item from language A may be identical in some respects with an equivalent item in language B;
- An item from language A may be different in some respects with an equivalent item in language B;
- An item from language A has no equivalent in language B (Krzeszowski 1990, 37-38).

The items compared are identical only with respect to some selected property or properties they share that are mentioned in the TC.

A problem that may occur in studies using this CA methodology is that of *circularity* (Chesterman 1998, 52), defined as follows:

> If an item in one language has no equivalent in the other language, on what grounds are we comparing the two in the first place?

If the presumed equivalence is the justification for the choice of the items to be compared in the two languages, the same concept of equivalence should not be considered as the result of the contrastive analysis. One of the solutions considered by contrastive linguists and suggested by Chesterman is that meanings or the forms assumed initially as being equivalent turn out to be equivalent in some respects but not equivalent in others.

Chesterman (1998, 54-61) proposed an alternative methodology, called Contrastive Functional Analysis (CFA). This methodology does not include the notion of equivalence as identity and it does not place the result of the analysis before the initial assumptions. The main idea behind Chesterman's CFA is that the goal of cross-linguistic comparisons is to propose hypotheses that have to be empirically tested.

The CFA methodology consists of seven steps, which include notions such as *comparability criterion, similarity constraint, initial* and *revised hypothesis*. In what follows, I will depict briefly

Chesterman's methodology (Chesterman 1998, 54-61).

- *Primary data*: instances of languages behaviour (the phenomena studied) in different languages, both at the semantic and/or formal levels.
- *Comparability criterion and similarity constraint:* the trigger device of cross-linguistic comparison is the *perception of a similarity of some kind* of two elements from L1 and L2. This initial perception can be vague and unspecified but they are worth comparing. It is this initial perception and not an assumed equivalence that provides the initial comparability criterion. The similarity constraint consists in setting up an "upper bound to the range of phenomena considered being relevantly similar, for the purposes of a particular study" (Chesterman 1998, 56). The comparability criterion corresponds to what James and Krzeszowski refer to as the *tertium comparationis.*
- *Initial identity hypothesis*: it concerns the assumption of identity between the elements to be compared (similarly to the null hypothesis H_0 in experimental studies).
- *Hypothesis testing*: the initial hypothesis must be empirically tested. This means to establish (by using corpora, using one's intuition or bilingual competence and according to a theoretical framework) in what conditions the hypothesis holds and in what conditions it does not. The result of this step provides evidence in favour of the initial hypothesis (i.e. in what respects the elements compared are identical) and evidence for the adjusting the initial hypothesis (i.e. in what respects the elements compared are different).
- *Revised hypothesis* (if the identity hypothesis fails): the hypothesis is revised with respect to type of the relation between the elements compared, more precisely which are the common and the different features and which are the conditions in which the elements compared are identic or different.
- *Testing of the Revised hypothesis:* as the initial hypothesis, this revised hypothesis needs also to be tested aiming at having a more accurate hypothesis, that explains better the relation between the elements compared.

The evidence in favour of the initial hypothesis, more precisely, all cases and conditions in which the elements compared are identical, constitutes the TC:

> The *tertium comparationis* is thus what we aim to arrive at, after a rigorous analysis; it crystallizes whatever is (to some extent) common to X and Y. It is thus an explicit specification of the initial comparability criterion, but it is not identical with it, hence there is no circularity here. (Chesterman 1998, 60)

When the TC is the starting point of the cross-linguistic comparison, the analysis risks circularity. For Chesterman, the TC is the aim of the analysis and the essence of the results whereas the initial perceived similarity represents the starting point. This perceived similarity can become a TC only after the testing of the initial hypothesis.

Chesterman (1998, 60-61) points to the fundamental relation between corpus-based studies and hypothesis testing. More precisely, a hypothesis may arise both from mere intuition and from corpus studies. However, it should be tested rigorously (against a corpus, other speakers' intuitions, in a controlled experiment) in order to have a reliable validation.

In this thesis, I adopt the three steps proposed by James and Krzeszowski for their concreteness to investigate the usages and the translation of verbal tenses from EN into FR, IT and RO, as described in Chapter 5. In what concerns the methodology proposed by

Chesterman, the formulation and the testing of hypotheses about the elements compared is of capital importance for having a reliable study. The hypotheses formulated in this research with respect to temporal reference and its expression through Tense, Aspect and Aktionsart are expressed based on previous theoretical descriptions and on corpus work. They are tested in offline experiments with linguistic judgement task, as described in Chapter 6.

4.2 Parallel corpora

4.2.1 Corpus data

The corpus linguistics field has flourished in the last fifty years mainly due the growing interest of linguists in having objective, quantifiable and reproducible data and in using computers. For example, Kolaiti and Wilson (2014) carry out a corpus-based investigation of the unitary account from RT, and specifically lexical pragmatics, in which narrowing, approximation and metaphorical extension are explained within the same model. They argue that:

> Corpus-based evidence provides a valuable complement to more traditional methods of investigation, by helping to sharpen intuitions, develop and test hypotheses and reduce the possibility of intuitive data being mere artefacts of the linguist. (Kolaiti and Wilson 2014, 211)

They point to the fact that corpus studies are a valuable source of inspirations for theorists in RT who are mainly concerned with the mental processes that enable the hearer to infer the speaker's meaning. This is primarily due to the fact that corpus work forced 'us to consider examples that we might not have come up with ourselves, helping to sharpen and test our hypotheses, and raised new intriguing questions' (Kolaiti and Wilson 2014, 212).

Coming back to pure methodological issues about corpus work, in this section my aim is to describe the main features of corpora, the advantages and limits of using corpora as a source of data in pragmatics.

Defining a corpus can be an easy and a difficult task at the same time because of the numerous perspectives that can be underlined, such as the type of texts, the size, the purpose of creation, the way in which it can be analysed, etc. The well-known description of a corpus as being "a body of naturally occurring language" (McEnery et al. 2006, 4) is largely accepted in the corpus linguistics community, as well as in other domains that use corpora, such as empirical pragmatics, NLP, MT and translation studies (Baker 1993, 1995).

The main features of a corpus are that they have a *finite size* (which changed with time but which, in general, is pre-established so that the construction criteria like balancing could be applied), are a *representative sample* of the variety or varieties of the language analysed, and represent the *standard reference*. Corpora have been compiled for many different purposes and, thus, they have different kinds of design and nature of the texts included. Another definition of a corpus would be that it constitutes an empirical basis for identifying the elements and the patterns of the structure of a language to analyse variation for example, or it can be analysed distributionally to check how often and where a particular phonological, lexical, grammatical or pragmatic feature occurs.

Corpora have been used in linguistics before the development of computers, but it is the

computer use that gave an enormous boost to corpus linguistics around the early 1960s, by reducing the time of creation, use and analysis of a corpus and greatly increasing the size of databases. The definition of a corpus can thus be modified as it follows: "a corpus is a collection of texts in an electronic database" (Kennedy 1998, 3), and thus it has *machine-readable* form. This feature allows its compilation and analysis semi-automatically and automatically. As far as size is concerned, corpora become larger and larger and this is due to the possibility to be tagged, compiled and analysed automatically. The most important aspect to take into account when doing corpus work is to have an appropriate match of the research goal and the corpus type and size (Gries 2009).

Regarding the purpose of their creation, corpora are of many types, such as (see Kennedy 1998): (a) *general* (a text base for unspecified linguistic research) vs. *specialized* corpora (corpora which are designed for particular research programs); (b) *written* vs. *oral* corpora; (c) *sample-text* corpus (designed to be a representative sample of the total population of discourse) vs. *full-text* corpus (a total statistical population); (d) *synchronic* (an attempt to represent language or a text type at a particular time) vs. *diachronic* corpora (represents occurrences of a language over a period of time); (e) *monolingual* vs. *multilingual* corpora; (f) *comparable* (collection of original texts in two or several languages matched by criteria such as register, genre, domain, time of publication, etc.) vs. *translation* (collection of original texts written in a source language SL and their translation into one or several languages called target languages TL); (g) *unidirectional translation* (one translation direction) vs. *bidirectional translation* (two translation directions).

In the last twenty years, cross-linguistic studies have used more and more multilingual corpora, which helped in the same time the revitalization of research in this domain. Aijmer and Altenberg (1996, 12) indicate some of the benefits of corpus-based study in language comparison:

- They give new insights into the languages compared, informing the researcher on language-specific and language-universal information, as well as about typological and cultural differences;
- They illustrate differences between source texts (authentic texts) and their translations and between native and non-native texts;
- They can be used in numerous domains, such as language teaching, translation, typology, semantics, pragmatics, NLP and MT among others.

Aijmer and Altenberg note that there is a difference in use and interest between *comparable* corpora and *translation* corpora. Comparable corpora contain original texts in a certain language and have the particularity that the texts share in general, criteria such as the same stylistic genre, domain, purpose of creation, time of creation etc. Their main advantage is that they present natural language in use and have the property of *authenticity*. The most difficult problem with using these corpora is to know what to compare (e.g. relating forms which have similar meanings and pragmatic functions in the languages compares as suggested by Johansson 1998) and to what degree of insight the comparability is possible.

Translation corpora contain texts that are intended to express the same meanings and have the same discourse functions in the considered languages (Johansson 1998). Dyvik (1998) suggested that translations reveal semantic features of the source language. She argues that translation is a linguistic activity when the translator evaluates meaning relations

between expressions in an objective manner (in opposition with researches who aim at developing or testing a theory about meaning relations between expressions). Translations provide thus objective linguistic data. For this reason, translation equivalence has been considered the best basis for comparison, and it was used for a long time as the main principle for the construction of a *tertium comparationis* (Krzeszowski 1990) in CA. As a method used in linguistic research, corpus data have numerous advantages and a series of limitations. I will describe them more detailed in the following section.

4.2.2 *Corpus data: advantages and limitations*

A great advantage of corpus data is that it allows both *qualitative* and *quantitative* analyses. In qualitative analyses of data, all sentences are treated with equal attention and the results cannot be generalized being limited to the sample of language analysed. Besides, no attempts are made to assign frequencies to the linguistic features identified in the data. In quantitative analyses, frequencies are assigned to linguistic features identified in the data; features are classified, counted and summarized. A basic step in quantitative analysis of data is to classify sentences or items according to a certain schema, and the count how many items (called *tokens* or *occurrences*) are in each group of the classification schema (called *types*). The result of this his process is a *distribution* of the tokens in the corpus (McEnery & Wilson 1996).

Another advantage of working on corpora is that they represent an empirical basis for researchers' intuitions. Intuitions are the starting-point of any study but can be misleading and sometimes a few striking differences could lead to hazardous generalizations. Moreover, results of analyses of quantifiable data allow not only generalizations (through statistical significance tests) but also predictions through statistical analyses, such as correlations[125] or multiple regression models[126], which are often used for investigating such a complex phenomenon as language.

Corpus work is interesting when the researcher is concerned with a descriptive approach of the linguistic phenomenon considered, as well as the study of language in use, given the fact that the cotext is provided in the corpus. Corpora permit monolingual and cross-linguistic investigations. Furthermore, corpus work allows the researcher to uncover what is probable and typical, on the one hand, and what is unusual about the phenomenon considered, on the other hand.

Another advantage is that data from corpora can be annotated (enriched) with syntactic, semantic and pragmatic information, which allows more complex analyses of the corpus.

Annotation is the practice of adding interpretative linguistic information to a corpus (Leech 2005), and thus an enrichment of the original raw corpus. From this perspective, adding annotations to a corpus is providing additional value and thus increasing their utility

---

[125] Correlation is a monofactorial statistical method, which investigates the relation between one independent variable (the predictor) and one dependent variable (the phenomenon of interest). Correlation does not involve obligatorily causality between the two variables (they can be only associated) and can be used only when relationship is linear (Baayen 2008; Gries 2009).

[126] Multiple regressions are multifactorial statistical methods, which investigate the relation between several independent variables (predictors) and one dependent variable, as well as their interactions. The relation between independent variables and the dependent variable can be linear or non-linear. (cf. Gries 2009, Baayen 2008). See section 4.4.2.

(McEnery and Wilson 1996; Leech 2004). Firstly, annotated corpora are useful both for the researcher(s) who made the annotation and for other researchers, who can use them for their own purposes, modify or enlarge them. Secondly, annotated corpora allow both manual and automatic analysis and processing of the corpus and by assuring its multifunctional utilisation, the annotations themselves often revealing a whole range of uses which would not have been practicable unless the corpus had been annotated. Thirdly, annotated corpora allow an objective record of analysis open to future analysis, decisions being more objective and reproducible. Due to automatic analysis of the corpus, annotated corpora are often used for training of NLP tools, such as automatic classifiers (Meyer et al. 2013).

The most known and used types of annotation are *word-class (part-of-speech) tagging*[127], *lemmatization*[128]*, parsing*[129]*, semantic*[130] and *discourse tagging*. Discourse tagging is a domain that flourished in the past 20 years providing large annotated databases, such as the *Discourse Penn Treebank* (Prasard et al. 2004, 2008) for EN and LEXCONN (Roze et al. 2010) for FR connectives. Basically, discourse tagging consists of *discourse tags* for annotating items, which have an important role rather for discourse management than for the propositional content, such as *connectives*. Discourse tags were firstly used for annotating categories such as *apologies, greetings, responses* and *politeness* (McEnery & Wilson, 1996). Halliday and Hasan's study on *cohesion* in English (1976) was the most influential in this domain and it turned the attention towards the need to automatically mark the pronoun reference at the discourse level (*anaphoric annotation*). Cohesion is closely related to coherence in a text (i.e. *coherence relations* as approached in Rhetorical Structure Theory (Mann and Thompson 1989; Taboada and Mann 2006), Hobbs (1979, 1985), DRT (Kamp 1981), SDRT (Asher 1993, Asher and Lascarides 2003), Sanders and colleagues (1992, 1993); see also section 7.3) and it represents a great challenge for the machine translation domain.

One of the major issues with using translation corpora is related to its nature itself, the translation process and the bias that the influence of the source language can have on the target text (the so-called *translationese* in Gellerstam 1996). Secondly, Baker (1993, 1995) underlines that translated texts use *translation universals*, which are defined by Lefer (2009), quoting Laviosa (2002), as features of a translated language, existent independently of the source language, such as simplification, explicitation and normalization. Thirdly, translated texts are restricted to be compared only to their original texts and not to others.

---

[127] The most basic type of linguistic annotation is the word-class or part-of-speech tagging (also known as grammatical tagging or morphosyntactic annotation and shortly noted as POS tagging). The method consists of assigning to words their part of speech (e.g. adjective, verb, preposition) in order to increase the specificity of data retrieval from the corpus (McEnery and Wilson 1998). POS tagging represents the foundation of further forms of analysis such semantic tagging and discourse tagging.

[128] *Lemmatization* is an annotation tool that allows the distinction between words that are inflectionally or sometimes derivationally related to each other. These words are instances of a single word family, known a *lemma.*

[129] Parsing consists in the annotation of syntactic relations. It involves linguistic analysis according to some particular grammatical theory with the purpose to "identify and label the function of each word or group of words in a phrase or sentence" (Kennedy 1998, 231). A parsed corpus (called a *tree bank*) provides a labelled analysis for each sentence and it is useful to identify, for example, relative clauses in a text.

[130] Semantic tagging is mainly concerned with sense disambiguation. For example, the tagging of a polysemous word such as *once* (which can be an adverbial, a nound or a conjunction) needs sense disambiguation in order to make the POS tagging maximally accurate (Kennedy 1998).

Another methodological concern when working with translation corpora is that they need to be aligned (at sentence or phrase level) and be processed by parallel concordancers. As Lefer (2009) notes, alignment can be time-consuming because the automatic alignment requires manual control and correction for a complete accuracy of data. Most parallel concordancers, such *ParaConc* offer automatic pre-alignment tools. Two examples of very known and used parallel corpora are *Europarl*[131] and *Hansard*[132]. Europarl is a corpus extracted from the proceedings of the European Union Parliament. It includes versions in 23 European languages, and the version in 1996 contained 20 million words (Koehn 2005) . The Hansard corpus is a bilingual corpus (EN and FR) of the proceedings of the Canadian parliament.

Other difficulties are the insufficiency of multilingual corpora for less widespread languages or the predilection for 'form-based research' where there is an interest in a specific grammatical form (Granger 2003). These difficulties constrain researchers to carry out their research manually, including building their corpus themselves and annotating it if they are interested in other phenomena than a specific grammatical form, such as semantic or syntactic categories. Moreover, when the researcher is interested in infrequent phenomena[133], there will be insufficient occurrences in the corpus. Difficulties are also encountered when phenomena that are not lexically expressed such as world knowledge used in inferences as well as the cognitive basis of language are investigated. This is one reason why corpus data are more and more combined with other types of data, such as experimental data.

4.2.3 Translation spotting and cross-linguistic transfer of properties

Translation spotting

Translational spotting or *transpotting* is a technique that makes use of the translation of a specific word or linguistic expression in order to distinguish their meaning and disambiguate among their senses. This method was used for content words (Dyvik 1998; Noël 2003), but as well for discourse relations (Behrens and Fabricius-Hansen 2003) and connectives (Zufferey and Cartoni 2012; Cartoni et al. 2013). The term *translation spotting*, coined by Véronis and Langlais (2000), initially referred to the automatic extraction of a translation equivalent in a parallel corpus. Precisely, given a particular word or expression, translation spotting consists in detecting its translation in the target text, as shown in examples from Table 4-1 (Cartoni et al. 2013) for connectives and in Table 4-2 (Grisot and Moeschler 2014) for verbal tenses.

---

[131] http://www.statmt.org/europarl/

[132] http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC95T20

[133] For example, Grivaz (2012) who studied causality in certain pairs of verbs in a very large corpus and with human annotation experiments, found that less frequent pairs had a good causal correlation while very frequent pairs had a small causal correlation.

Table 4-1 Example of translation spotting for the connective *since*

	EN sentence	FR sentence	Transpot
1.	In this regard, the technology feasibility review is necessary, *since* the emission control devices to meet the ambitious NOx limits are still under development.	À cet égard, il est nécessaire de mener une étude de faisabilité, étant donné que les dispositifs de contrôle des émissions permettant d'atteindre les limites ambitieuses fixes pour les NOx sont toujours en cours de developpement.	étant donné que
2.	Will we speak with one voice when we go to events in the future *since* we now have our single currency about to be born?	Parlerons-nous d'une seule voix lorsque nous en arriverons aux événements futurs, puisqu'à présent notre monnaie unique est sur le point de voir le jour ?	puisque
3.	In East Timor an estimated one-third of the population has died *since* the Indonesian invasion of 1975.	Au Timor oriental, environs un tiers de l population est décédée depuis l'invasion indonésienne de 1975.	depuis
4.	It is two years *since* charges were laid.	Cela fait deux ans que les plaintes ont été déposées.	paraphrase

Table 4-1 is an example of the investigation of the usages of EN connective *since* carried out in translation corpora. The second column contains the translation of the original EN sentence into FR. The third column contains the linguistic expressions or types of linguistic expressions used in FR for translating the EN *since*, called *transpots*. The idea behind this analysis is that FR transpots provide information regarding the diverse contextual usages of the EN *since*. Table 4-2 is an example of the investigation of the usages of EN SP through looking at the verbal tenses used in FR for its translation.

Table 4-2 Example of translation spotting for the PresPerf verb tense

	EN sentence	FR sentence	Transpot
1.	If the Commission concludes that this objective *has been achieved* on a sustainable basis, the provisions of paragraph 3 shall cease to apply on 1 January 2001.	Si la Commission arrive à la conclusion que cet objectif *a été atteint* de façon durable, les dispositions du paragraphe 3 cessent de s'appliquer le 1er Janvier 2001.	PC
2.	The Member States take note of the fact that the Spanish Government *has embarked* upon the implementation of a policy of regional development designed in particular to stimulate economic growth in the less-developed regions and areas of Spain.	Les États membres prennent acte du fait que le gouvernement espagnol *est engagé* dans la mise en œuvre d'une politique de développement régional visant notamment à favoriser la croissance économique des régions et zones les moins développées de l' Espagne.	PRES
3.	Disputes between the European Central Bank, on the one hand, and its creditors, debtors or any other person, on the other, shall be decided by the competent national courts, save where jurisdiction *has been conferred* upon the Court of Justice of the European Union.	Les litiges entre la Banque centrale européenne , d' une part , et ses créanciers , débiteurs ou toute autre personne, d' autre part, sont tranchés par les tribunaux nationaux compétents, à moins que la Cour de justice de l' Union européenne *n'ait été déclarée* compétente.	Past subjunctive

Véronis and Langlais point out the difficulty of automatically spotting the words or sequence of words from the target language when there is no one-to-one correspondence between the source and the target language. Automatic spotting results have errors and NLP researchers' aim is to reduce as much as possible de number of errors. For this reason, other researchers (see Cartoni et al. 2013; Grisot and Moeschler 2014) performed the spotting

manually in order to get fully accurate data. Cartoni and colleagues agree that despite the fact that translations do not reproduce faithfully SL and have a number of inherent features (Baker 1993), they still can be used to shed light on SL. They suggest that theoretical insights developed based on analysis of parallel corpus should be validated through monolingual experiments.

The theoretical idea behind translation spotting is that similarities and differences in translation can reveal semantic features of the SL (Dyvik 1998; Noël 2003). Dyvik's idea is that the activity of translation is one of the very few cases where speakers evaluate meaning relations between expressions in an objective manner, without doing so as part of some kind of meta-linguistic, philosophical or theoretical reflection. From this perspective, he suggests to use translation corpora as a basis for semantic analyses. This method presupposes the existence of a *translational relation* between two languages. There are two aspects to be distinguished before determining a translational relation. The first is information regarding *parole*[134] and textual *token* items and the second is information about *langue* and *type* items. In the first case, translation choices are motivated only by reference to the particular text and its circumstances, whereas in the second case, translation choices are predictable and reflect translation correspondence relations between words and phrases seen as types rather then textual tokens. According to Dyvik, it is on this second aspect of language that a translational relation should be built. A translational relation consists of a series of properties or, more precisely, a series of senses shared partially by the linguistic expressions that stand in that translational relation. Translational relations can be identified through the translation spotting technique. Explicitly, in Cartoni and colleagues' study, the EN connective *since* is translated into FR through four linguistic expressions (three connectives and a paraphrase). Cartoni and colleagues showed in an experiment with a sentence completion task, that the four FR translation possibilities are clustered into two: a causal sense (for *étant donné que* and *puisque*) and a temporal sense (for *depuis* et *cela fait X que*). The translational relation of *since* and its transpots in FR consists therefore of two properties or senses shared partially by these linguistic expressions.

Translational relations reflect partial semantic equivalences between words and expressions in different languages. Therefore, they represent a concrete tool for developing cross-linguistic semantic representations. A semantic representation classifies together a set of linguistic expressions across languages that fall within the denotation of the representation (Dyvik 1998). Such cross-linguistically valid semantic representations are useful for improving the results of several NLP tasks, such as machine translations systems, multilingual dictionaries and concordances.

In this thesis, I use translation corpora built of original texts and their translations into three target languages. Transpotting was used in the analysis of a bilingual translation corpus (EN into FR and FR into EN) and of a multilingual parallel translations corpus (EN into FR, IT and RO. In monolingual analyses of the corpus written in the source language, the most frequent verbal tenses were identified. Consequently, in a cross-linguistic perspective, for each occurrence of the most frequent verbal tenses in SL, its translation in a TL was

---

[134] The well-known linguist Ferdinand de Saussure was the first to make the distinction between *parole* and *langue*, where the former refers to acts of language of individual persons and the latter refers to language as an abstract entity, proper to a linguistic community.

identified. The four most frequent forms used in TL were considered in a fine-grained analysis. Corpus analysis and its results are provided in section 5.1.2 for the corpus EN to FR, section 5.2.2 for the corpus FR to EN and section 5.3.2 for the multilingual corpus EN to FR, IT and RO.

Furthermore, corpus data is enriched with semantic and pragmatic information originating in offline experiments with linguistic judgement task (as discussed in sections 6.1 and 6.2). A multifactorial statistical model was fitted on this behavioural data enriched with information about Tense, Aspect, Aktionsart, as well as stylistical register, in order to predict the verbal tense used in TL (as discussed in section 6.1.8).

Cross-linguistic transfer of properties

Cross-linguistic transfer of properties is a novel technique that makes use of the notion of translational relation and its properties. My suggestion is that translation corpora permit cross-linguistic transfer of semantic and/or pragmatic information. Samardzic (2013) also made uses this novel methodology for investigating the translation equivalents of a range of EN light verb constructions into several languages. Unlike other European languages, Slavic languages encode morphologically Aspect. She applies the aspectual representation obtained in the EN-Serbian cross-linguistic setting to classify EN verbs into event duration classes.

In an experiment, native speakers of EN were asked to judge SP tokens with respect to Aspect and its two values *perfective* vs. *imperfective*. Judges found the task extremely difficult and they had a very low agreement rate. The cross-linguistic transfer of properties method was therefore used in order to have human annotated data (i.e. SP tokens) with aspectual information. Precisely, a native speaker translated the data into Serbian and she identified the contextual value of Aspect for each SP token. Based on the assumptions related to translation corpora, this aspectual information was transferred back to the initial EN source. This experiment and its results are discussed in section 6.1.5.

To sum up, translation corpora are particularly useful for investigations in semantics and pragmatics due to their inherent property of preserving in TL the meaning expressed in SL. However, it is more acceptable to speak about *similarity* than about *identity* of meaning. Consequently, similarities and differences in translation can reveal semantic features of the SL. Besides, corpora present another advantage for studying human communication, that is, they provide *natural* examples and the context in which the linguistic phenomenon of interest occurs. This feature is opposed to experimental data, which is artificial. However, experimental data represent an important complementary source of data because it targets the manipulation of variables. I will speak about experimentation and the complementary of corpus and experimental data in the next section.

4.3   Experimentation

Nowadays, researchers are beginning to require evidence that other people, besides the authors themselves, can understand and make judgments underlying the research reliability. Purely theoretical models are subjective and unique (in the sense of lacking of replicability). Research is judged according to whether or not the reader found the explanation plausible.

The main features of purely theoretical studies are: subjectivity, lack of replicability and lack of applicability to other linguistic domains such as NLP, CL, and MT. For this reason, experimental work became a privileged source of data in linguistic research.

4.3.1 Experimental data in linguistic research

In pragmatics, experimentation was shown to be extremely useful for studying issues from the semantics/pragmatics interface and testing theories concerning the psychological real competences that native speakers have regarding semantics and pragmatics (for example Katsos and Breheny 2008 for scalar implicatures and numerals).

Gilquin and Gries (2009) describe types of data used in linguistic research. They are presented in a descending order of naturalness of production and collection. Only corpora with written examples are produced for other aims than the specific purpose of linguistic research, and are thus the most natural kind.

- Corpora with written texts (e.g. newspapers, weblogs);
- Example collections;
- Corpora of recorded spoken language in societies/communities where note-taking/recording is not particularly spectacular/invasive
- Corpora with recorded spoken language from fieldwork in societies/communities where note-taking/recording is spectacular/invasive;
- Data from interviews (e.g. sociolinguistic interviews);
- Experimentation requiring subjects to do something with language they usually do anyway (e.g. sentence production as in answering questions in studies on priming or picture description in studies on information structure);
- Elicited data from fieldwork (e.g. response to "how do you say X in your language?")
- Experimentation requiring subjects to do something with language they usually do, on units they usually interact with (e.g. sentence sorting, measurements of reaction times in lexical decision tasks, word associations);
- Experimentation requiring subjects to do something with language they usually do not do (e.g. measurements of event-related potentials evoked by viewing pictures, eye-movement during reading idioms, acceptability/grammaticality/linguistic judgments) on units they usually do not interact with, involving production of linguistic output (e.g. phoneme monitoring, ultrasound tongue-position videos);

Two advantages of adopting experimentation as a type of methodology used in empirical and experimental pragmatics are (a) it makes possible systematic control of confounding variables, and (b) depending on the nature of the experiment, it permits the study of online processes (Gilquin and Gries 2009, 9). One difficulty with experimentation is the artificial setting experiments require that can influence the behaviour of the participants in this unnatural setting. If experimental pragmatics adopted the psycholinguistics methodology and the study of online processes (through EEG and eye-tracking techniques), empirical pragmatics focused mainly on offline experimentation, preserving the very essence of experimental studies: systematic manipulation of independent variables in order to determine their effect on dependent variables. Moreover, offline and online experimentation do not inform the researcher about the same type of information. Online experimentation targets the processing of languages itself whereas offline experimentation targets the result of

the processing of language.

Regarding the complementarity of corpus and experimental data, Gilquin and Gries argue that a corpus has a four-fold purpose in experimentation: (a) validator: the corpus serves as a validator of the experiment, (b) validatee: the corpus is validated by the experiment, (c) equal: corpus and experimental data are used on an equal footing and (d) stimulus composition: the corpus serves as a database for the items used in experiments. Corpus work deals with a larger range of phenomena that can be investigated compared to experimentation. Experiments, however, allow the study of phenomena that are infrequent in corpora. Moreover, corpus data is naturally occurring data whereas experimental data is artificial data. Finally, manipulation of variables is possible only through the experimental design. Corpora and experiments have advantages and limits that are complementary. For these reasons, nowadays linguists tend to use both of these empirical methods.

In this thesis two types of offline experimental designs were used: experiments with *linguistic judgement task* and *elicitation task*. This type of experimentation makes use of the intuitive and unconscious knowledge that native speakers have on their native language. It produces the so-called *behavioural* data, where only the output of the cognitive process is investigated. As mentioned earlier, the process itself may be investigated only in on-line experimentation.

Linguistic judgement task was used initially used for acceptability and grammaticality judgements but nowadays it is used for all types of linguistic information. The experiments carried out in this research are hybrid between typical experimentation (with manipulation of variables and built examples) and the intuitive analysis of corpus data. Explicitly, data from corpora (therefore, features as naturalness and original contexts were preserved) were submitted to judgement to native and naïve speakers. The judgements were made according to an objective set of guidelines established in advance. The results were evaluated through measuring the agreement between the two participants with the Kappa coefficient (Carletta 1996) (see section 4.3.2 for methodological issues regarding the usage of this coefficient in linguistic research). Consequently, the experimental data produced in this research represents *human annotated data*, which is used in NLP as training data for automatic classifiers. In other words, it provides the automatic tool different sorts of information (linguistic, contextual and world knowledge) that humans have and use in the language interpretation process.

Experiments with elicitation task are experiments where subjects are asked to fill in a sentence, for example by choosing among a series of choices or by proving themselves the material to be filled in. In this thesis, native speakers were asked to give the tensed form of a verb given at the infinitive form so that it would correspond to the surrounding cotext. The items used in this experiment were of two categories. The first category consists of items randomly selected from the corpus described in 5.2, which represent *naturally occurring* items judged in their *original contexts*. The second category consists of artificial sentences built for the purpose of the experiment. The results of this experiment were evaluated by counting the majority of answers for each item, since there were more than two participants. The number of concordant answers must be superior to agreement by chance, which is at 50% since there is a binary choice (i.e. past vs. non-past context). The hypotheses, the details of the design and the results of this experiment, as well as their theoretical implications, are provided in

section 6.1.7.

4.3.2 Other methodological issues

Since the mid 1990, computational linguists realized that pragmatic knowledge needed to be integrated into computational models of language. For developing language models, large amounts of human annotated data are required. This led to worries about the subjectivity of the judgments required to create annotated resources, much greater for semantic and pragmatic interpretations of language than for syntactic parsing for example.

It is largely acknowledged that human-annotated data can be used either to support an empirical claim or to develop and test language models, we have to be sure that such data are reliable. Authors in computational linguistics (CL) (e.g. Krippendorff 2004) argue that an annotated corpus with reliable data is similar to the results of a scientific experiment, and it is considered as being valid only if it is reproducible − if the same results can be replicated in an independent annotation experiment. There are the following requirements:

- An important quantity of data where samples are randomly chosen, where each category is equally represented;
- Data can be either naturalistic (from corpora) or artificial (built examples);
- Clearly expressed guidelines and written instructions for the judgment task fixed in advance and usable for other annotation experiments in order to allow replicability; The guidelines should include examples of all values of the categories;
- Participants: 2 or more native speakers who pass through a training phase and who work independently of each other − providing thus independent samples of annotated data;
- Evaluation: inter-annotator agreement rate with the Cohen's *kappa* coefficient (Cohen 1960; Carletta 1996) or other measure such as Krippendorff's *alpha* (Krippendorff 2004) or Scott's *pi* (Scott 1995) (see Artstein and Poesio 2008 for an extensive discussion of the methods currently used in CL for measuring inter-annotator agreement).

Data are reliable if annotators agreed on the categories to be assigned to each item. Artstein and Poesio (2008, 557) write that:

> If annotators produce systematically similar results, then we can infer that they have internalized a similar understanding of the annotation guidelines, and we can expect them to perform consistently under this understanding.

Reliability is a prerequisite for having valid data. There is though a problem with inter-annotator agreement rate, that is, agreement due to chance. There exist several methods to calculate inter-annotator agreement. The first is to calculate the *percentage agreement*, which is the ratio of observed agreements either between two judges either for the majority of opinions among several judges. If we consider the case of 2 judges, the amount of agreement we would expect to occur by chance (if annotators took a decision without taking into account the annotation guidelines) depends on the number of categories and if the categories are equally distributed or not, in other words if one category is much more common that the other.

It was argued that *percentage agreement* is biased in favour of small number of categories. Given two studies investigating the same phenomenon, the one that will use a smaller number of categories will have higher agreement rates just by chance. For example: for 2 equally distributed categories there are 50% of chances that when one judges makes a decision, the second judge will make the same decision (proportion based on the fact there only two choices; for 4 categories, there are 25% chances for the two judges to make the same judgment).

In this thesis, Cohen's *kappa* coefficient (from now on, the Kappa coefficient) was used to measure inter-judge agreement rate. The Kappa coefficient is a statistic measure that corrects for expected change agreement. It has values from 0 (i.e. there is no agreement other than that expected by chance) to 1 (i.e. there is total agreement). Another question concerning the Kappa coefficient is the threshold for reliable data. Generally, the threshold is set at 0.6-0.7 but it is not a fixed value and it depends on the type of research. Researchers proposed several possible explanations for low Kappa values. According to Bayerl and Paul (2011), the following factors have impact on agreement scores:

- Annotation domain;
- Number of categories in a coding scheme;
- Number of annotators;
- Whether annotators received training;
- The intensity of training;
- The annotation purpose;
- The method used for the calculation of percentage agreement.

Spooren and Degand (2010) point out that in linguistic studies, specifically in corpus-based analyses of discourse relations, low Kappa values occur very often. The question about the interpretation of these low Kappa values and the utility this data arises. They suggest three possible explanations for low Kappa values, as well as three methods to increase them in order to have reliable data (accepting the 0.8 threshold). First, linguistic ambiguity produces several interpretations. This means that a certain margin of disagreement should be allowed. Second, there is the case of coding errors. There can be two types: (a) errors regarding the initial working hypotheses, this means that the annotation guidelines (the coding scheme) do not capture entirely the considered phenomenon; and (b) there are individual strategies for each coder. Spooren and Degand (2010) suggest two methods to reduce coding errors. The strategies to improve Kappa value in order to have reliable annotated data-double coding are:

- Discussion of disagreements: individual strategies become cooperative strategies (e.g. Sanders and Spooren 2009 used double coding for their analysis of two connectives indicating causality in Dutch).
- One-coder-does-all (the method relies on subjective but systematic judgments).
- Application of linguistic tests with the purpose to render explicit the implicit information and thus reduce the amount of subjectivity for the decisions.

The offline experiments carried out in this thesis were designed to test theoretical assumptions currently existing in the literature with respect to temporal reference and its ingredients. Experiments were organised for each of the languages considered in this research: EN, FR, IT and RO. They made use of the methodology discussed in this chapter

and paid attention to these methodological issues. Explicitly, randomly selected corpus items were judged by the participants with respect to objective annotation guidelines established in advance. The Kappa coefficient correcting for chance agreement was used to evaluate the judgements made by the participants. In order to have higher Kappa values and therefore reliable data, disagreements were resolved using several techniques: discussion between the two judges, majority of answers when more than two judges participated at the experiment and application of linguistic tests.

Moreover, a cross-linguistic analysis was carried out on data consisting only of agreements. This analysis targeted possible correlations between usages of a linguistic expression in SL and its translation possibilities in TL.

Finally, *inter-judge agreement rate* is proposed as a measure for distinguishing between types of information encoded by linguistic expressions, namely, procedural and conceptual information (as proposed in Relevance Theory, see section 3.1.3). More precisely, in this thesis the Kappa coefficient is used but either Krippendorff's *alpha* or Scott's *pi* could be used. Explictly, the proposal is two-fold. On the one hand, judges have a poor ability to consciously evaluate procedural information and, consequently, the value of the coefficient used for measuring inter-annotator agreement is low. On the other hand, judges have a great ability to consciously evaluate conceptual information, and consequently, the value of the coefficient used is high. This proposal has been validated in very recent investigations of connectives (Cartoni et al. 2013) and of verbal tenses (Grisot and Moeschler 2014).

4.4   Statistical analyses

Linguists use more and more quantitative data because of the need to test empirically and quantitatively hypotheses about linguistic phenomena and predict their behaviour in larger amounts of data. According to Johnson (2009), the main goals of quantitative analyses of corpus data are (i) to describe the data in a systematic manner and (b) to generalize the conclusions from a representative set of observations to a larger universe of possible observations using hypothesis testing.

As Gries (2014) points out, linguistics is essentially "a distributional science" in the sense that linguists explore the distribution of linguistic elements on every level of linguistic analysis (i.e. phonology, morphology, syntax, semantics and pragmatics) in corpora. The linguistic expressions considered in a study are called *variables* or *factors*. Variables are of two types: *dependent* and *independent*. The dependent variable is the linguistic element of interest for the researcher. The independent variables are other linguistic elements that could be in relation with the dependent variable (also called predictors). Gries (2014) argues that predictors can be the individual variables, the levels of variables and the eventual interactions between the variables or their levels. The interaction of *n* variables is an unpredictable joint effect of the *n* variables on the dependent variable.

Most often linguistic data is *discrete* (also called *categorical* or *nominal*[135]). Linguistic elements

---

[135] The other types of variables are called *continuous* or *numeric*. They are of three types: *ordinal* (when three elements are ordered, it means that the first is less than the second, which is less than the third, such as rating scales), *interval* (elements that are measured on a scale and where there is no absolute 0 value, such as

are organised in categories. When two elements that are labelled the same way it means that they belong to the same category. Nominal variables do not have a meaningful order on a scale of any type.

Inferential statistics makes use of significance tests[136] for generalizing the observations based on a sample to the entire population. A significance test is based on two scientific hypotheses: H_1 and H_0. H_1 is called the *alternative hypothesis* and corresponds to researcher's hypothesis about the data. H_0 is called the *null hypothesis* and corresponds to states of affairs that would falsify the alternative hypothesis. An alternative hypothesis is non-directional (i.e. two-tailed) when no direction for the difference is provided or directional (i.e. one-tailed) when a direction for the difference is provided (more or less than 0). In the vast majority of cases, the null hypothesis states that the dependent variable is distributed randomly (or in accordance to some well known mathematically definable distribution) or it states that there is no difference between two or more groups, or that there is no relation between the independent variable(s) and the dependent variable and that the different or effect observed is due to chance or random variation. The result of a significance test allows the researcher to reject or not H_0, and therefore to accept or respectively, reject the alternative hypothesis. The threshold for this decision is the value of the probability of error p which must be is equal or less than 0.05 for rejecting the null hypothesis (which in generally means that there is a 95% confidence interval).

4.4.1 *Frequency tables*

Nominal data is most often analysed in frequency lists consisting of the frequencies of occurrence in corpora. A frequency list is summarized in a two- or more-dimensional frequency tables (also called *contingency tables* or *cross-tabulation tables*). Frequency tables provide information about variables' interrelations based on corpus or experimental data.

The basic idea of the analysis of frequency tables is to determine whether the frequencies observed in the corpus, and summarized in the cells of the frequency table, are distributed in a way that is significantly different from a random distribution (goodness of fit tests for frequencies) or from the distribution of another variable in independent samples (tests for differences/ independence).

Goodness of fit tests for one dependent nominal/categorical variable aim at testing the degree of fit between an observed and an expected distribution. Such questions are usually investigated with a Chi-Square test[137]. The null hypothesis postulates that the data are distributed randomly/evenly and that if there is a difference in the tested sample, it is due to

---

temperature measurement) and *ratio* (elements that are measured on a scale where there is an absolute 0 value, such as reaction times).

[136] The tests where the data is compared to a random or a mathematical distribution are called *goodness of fit* tests. The tests where two groups are compared or the relation between a dependent variable and one independent variable are called tests for differences/independence. These are monofactorial methods involving one dependent variable and one independent variable. Multifactorial methods involve one dependent variable and more than two independent variables, and include correlations, regressions and mixed models. These later methods aim at formulating a statistical *model* that best represents the data (see Gries 2009).

[137] Chi-square tests require that all observations are independent of each other and that at least 80% of the expected frequencies are larger than 5. If this is not the case, another test can be used called Fisher-Yates exact test.

random variation. The alternative hypothesis postulates that the frequency levels of the dependent variable are not identical. If the observed frequencies are significantly different than the expected ones (i.e. *p* is equal or less than 0.05), than the null hypothesis stating that the observed frequencies are due to random or even distribution[138] can be rejected. This means that the alternative hypothesis stating that the data is not distributed randomly can be accepted. The question that arises in this case is which are the independent variables that predict the observed distribution. This can be investigated through two- or more-dimensional frequency tables.

The distribution of one dependent variable can be compared to the distribution of one independent variable. This kind of investigation can be done with Chi-Square tests or with the Fisher Exact Probability test. The basic idea is to test whether the levels of the independent variable result in different frequencies of the dependent variable. In the case when the result of the test independence test is significant, the null hypothesis postulating that the levels of the dependent variable do not vary according to the levels of the independent variable can be rejected, and therefore, the alternative hypothesis can be accepted. Moreover, the contribution of each of the cells of the table to the overall result can be investigated through inspecting the so-called *Pearson residuals*. If the value of the Pearson residuals in a cell is positive/negative, then the observed frequency in that cell is greater/less than the expected frequency in that cell. Secondly, the larger the Pearson residual deviates from 0, the stronger that effect. The size of the effect is calculated with the *phi* coefficient or with *Cramer's V* coefficient. The values of the coefficients range from 0 ('no effect') to 1 ('perfect correlation'). The correlation coefficient is not affected by sample size.

The contribution to the overall significative Chi-Square of every cell (levels of the dependent and independent variable) can be visualized through an association plot, consisting of a box for the correlation of each level of the dependent variable with each level of the independent variable (as, for example, in section 6.1.5 indicating the correlation between Aspect and the verbal tense used in TL).

These types of analyses use monofactorial methods and investigate the relation between the dependent variable and maximally one independent variable. But very often, there is more than one factor that influences the dependent variable. This type of investigation uses multifactorial methods.

4.4.2 Multifactorial methods

Multifactorial methods are used to explore variation due to multiple factors and this exploration involves formulating a *statistical model*. A statistical model is *a statistical description of the structure of the data that provides the best possible characterization of the data* without violating Occam's razor[139] by including more parameters than necessary or by assuming more complex relations between variables than necessary (Gries 2009, 238).

The standard stepwise procedure used in multifactorial methods consists of:

---

[138] In case of binomial nominal variables, the expected distribution is 50% for each level (just as tossing a fair coin many times).

[139] The most fundamental principle of scientific reasoning based on frequency data is the *principle of parsimony* known as Occam's razor. It prohibits the inclusion of unnecessary explanatory notions into an analysis.

- Determining of a maximum model, i.e. the model that includes all predictors and their interactions;
- Iteratively deleting the least relevant predictors (starting with the highest-order interactions) and fitting a new model;
- Arriving at the minimal adequate model, which contains only predictors that either significant themselves or participate in higher-ordered interactions.

For categorical or nominal variables, methods such as binary logistic regression or Poisson regressions can be performed. Binary logistic regressions are used for dependent variables with two levels whereas Poisson regressions are used for dependent variables with more than two levels. Another multifactorial method serving at predicting the values of the dependent variable is given by the *generalized mixed models* (Baayen 2008, chapter 7). The main difference between regression analyses and mixed models regards the possibility to add random factors in the mixed model besides fixed factors. Random effects usually refer to items (for example words) or subjects (the participants to the experiment). Since the items/subjects are chosen randomly from the population, we would want the model containing fixed factors to predict correctly also when applied to new items/subjects. In other words, random factors account for the variation in the data that is due to subjects and items chosen for the experiment. Fixed factors are the factors hypothesized to explain most of the variation observed in the data.

In this thesis, statistical analyses were carried out with R programme[140]. A mixed model was built on the bilingual data, resulting after Experiments 3 to 5. The results and their discussion are provided in section 6.1.8.

4.5   Conclusive remarks

This chapter addressed the methodology used in this research: systematic contrastive analysis as proposed in the CA field, corpus and experimental work.

Firstly, CA methodology points to an element that is crucial for investigating contrastively linguistic phenomena in two or more languages, namely the *tertium comparationis* (TC). Literally meaning 'third term of comparison', it represents the background of sameness against which the differences are significant (James 1980). The information contained in the TC is required to be cross-linguistically valid so that the linguistic phenomenon studied could be positioned with respect to the PC. Moreover, in Chesterman's framework, the goal of cross-linguistic comparisons is to propose hypotheses that have to be empirically tested and validated. In this thesis, the theoretical model proposed in Chapter 7 plays the role of the TC for contrasting verbal tenses in EN, FR, IT and RO, which could be applied to other languages.

Secondly, the methodology developed in Corpus Linguistics became a valuable source of evidence in pragmatic research. Two essential advantages of corpus data are *naturalness* and *the usage of a linguistic expression in its original cotext.* The preservation of naturalness and of the original cotext is opposed to the artificial data used in the experimental design strictly speaking. Moreover, translation corpora are particularly interesting for cross-linguistic

---

[140] Documentation and free downloading at http://www.r-project.org/.

investigations due to the fact that the very essence of the translation process is to provide in a TL the same meaning as the in the SL. Therefore translation corpora illustrate both linguistic similarities and linguistic differences. Based on this property of translation corpora, a novel method was proposed called *cross-linguistic transfer of properties*, which allows the researcher to formulate hypotheses about linguistic features of a TL based on the linguistic features observed in the SL. Translation corpora were investigating using the translation spotting technique, which was successfully used in previous studies.

Thirdly, offline experimentation with linguistic judgement task was used in order to test native speakers' intuitions (used in relation to Chomsky's notion of *competence*) with respect to the features of their first language. The participants' judgements were evaluated using Cohen's Kappa, a statistical measuring inter-annotator agreement rate and taking into consideration agreements due to chance. It was pointed out that experimentation and corpus work being complementary methods, are both necessary in order to test empirically theoretical hypotheses and build new models. Experimental data was investigated quantitatively through descriptive and inferential statistical methods.

In conclusion, this chapter described the general methodology used in this research and pointed to specific methodological issues. Chapters 5 and 6 are dedicated to the empirical work carried out in this thesis.

5 Analysis of translation corpora

For the specific needs of this research, parallel (also called *translation*) corpora consisting of texts of four registers have been assembled. The qualitative and quantitative analyses of the corpora were carried out in two steps: (a) a first monolingual step in order to see which verbal tenses occur in the corpus and to calculate their frequency in the source language (SL) and (b) a second bilingual step in order to identify the verbal tenses used as translation possibilities in the target language (TL) for a certain tense in SL, as well as to calculate their frequency. Analysis of frequency of tenses in SL provided information about verbal tenses that are potential problematic candidates for machine translation systems. The assumption of this procedure is that frequent erroneously translated verbal tenses decrease the quality of the translated text more than infrequent incorrectly translated verbal forms. Hence, in this chapter, I will provide the description and the results of the analysis of corpora.

The purpose of this chapter is to describe the corpus work, which is the first layer of the empirical work presented in this thesis. This research is partly based on parallel or translation corpora, consisting of texts written in EN and their translations into three target languages. Three corpora have been built. The first and the second corpora are bilingual, consisting of texts written in EN and their translation into FR, and respectively, texts written in FR and their translation into EN. The third corpus is a multilingual one, consisting of texts written in EN and their translations into FR, IT and RO. All texts have been randomly selected and appertain to four stylistic registers: literature[141], journalistic, legislation and EuroParl (Koehn 2005)[142]. The literature register consists of texts coming from several novels, written in EN and translated into FR, and respectively, written in FR and translated into EN. The EN-FR-IT-RO corpus consists of randomly selected passages from "Alice in the Wonderland" by Lewis Carol and their translations into the target languages. The journalistic register consists of texts coming from several journals, which have an on-line version. For the multilingual corpus, all texts were randomly selected from the Press Europ website[143] and aligned manually. Legislation register consists of the European Union law texts collected into a multilingual corpus called JRC-ACQUIS parallel corpus and in the EuConst Corpus[144]. The EuroParl register consists of the transcription of parliamentary debates. The language used in the EuroParl corpus is spoken but transcribed, presenting therefore features of both spoken and written language.

---

[141] A detailed presentation of the texts from the four registers is available in the Appendix section. The corpora are in part aligned at the sentence level by other researchers and are available for download online and in part created during the COMTIS research project (typed and aligned manually by the author).

[142] The EuroParl corpus is a collection of the proceedings of the European Parliament during 1996-2011. It is available online at http://www.statmt.org/europarl/.

[143] The translation of the journalistic articles into all the languages spoken in the European Union is available online at http://www.voxeurop.eu.

[144] The JRC-ACQUIS corpus was collected by the Language Technology team of the European Commission's *Joint Research Centre* (JRC) in the context of the workshop *Exploiting parallel corpora in up to 20 languages*, held in Arona, Italy, on 26 and 27 September 2005. The EuConst corpus is a parallel corpus collected from the European Constitution (Tiedemann 2009).

The purpose of the monolingual analysis is to identify frequent and less frequent verbal tenses whereas the purpose of the cross-linguistic analysis is to identify *translation divergences*, i.e. each verbal tense that is consistently translated into TL through more than one verbal tense. In sections 5.1, 5.2 and 5.3 I will describe the corpora and provide the results of the corpus analysis.

5.1 Bilingual Corpus: EN-FR

5.1.1 Monolingual analysis

The EN-FR bilingual corpus consists of texts in EN and their translations into FR belonging to four different stylistic genres, having the following proportions: literature 15%, journalistic 16%, legislation 38% and EuroParl 31%. The corpus contains 1670 occurrences of predicative verbal tenses, occurring in a total of 725 sentences[145]. A total of 1281 predicative verbal tenses have been considered[146], which represent 77% of the verbal tenses occurring in the corpus, as shown in Table 5-1. The rest of 23% of verbal tenses have not been considered[147] in the analysis.

Table 5-1 Verbal tenses per register in the EN-FR bilingual corpus

Register	No. of sentences	No. of verbal tenses	No. of verbal tenses considered	% of verbal tenses considered	% of verbal tenses not considered
Literature	118	255	232	14%	1%
Journalistic	155	275	228	14%	3%
EuroParl	136	512	403	24%	7%
Legislation	316	628	418	25%	13%
Total	725	1670	1281	77%	23%

Figure 5-1 illustrates the frequency of verbal tenses analysed[148] in the EN-FR bilingual corpus, where the most frequent tenses are the Spres (32%), the SP (25%) and the PresPerf (13%), whereas the past and present progressive, as well as the past perfect verbal forms are much less frequent. This figure shows the inequality of occurrence of verbal tenses in a corpus containing a total of 1670 predicative verbal forms pertaining to different stylistic registers. One possible explanation for the higher frequency of certain verbal tenses is that some tensed forms are highly context-dependent, their interpretation depending on the various contextual hypotheses.

---

[145] I use the word sentence to refer to a chunk of text, consisting of one or several complex clauses. Since verbal tense is a referential category (as argued in section 2.2.1) and its meaning is underdetermined (as argued in section 3.1.3), contextual and contextual information is needed for determining its meaning. Therefore, the segmentation was done manually in order to decide the size of the text chunks relevant for determining the meaning of a verbal tense.

[146] The considered tenses are several tenses from the indicative mood: the simple present and past tenses, the present and the past perfect, the present and the past continuous.

[147] Non-analysed tenses are other tenses from the indicative mood (present and past perfect continuous, all future tenses), EN verbal tenses with conditional and subjunctive readings and modal verbs.

[148] Legend: SP = Simple Past, PresPerf = Present Perfect, PresCont = Present Continuous, Spres = Simple Present, PastCont = Past Continuous, PP = Past Perfect and Non-analysed

Figure 5-1 Frequency of EN verbal tenses in the EN-FR bilingual corpus
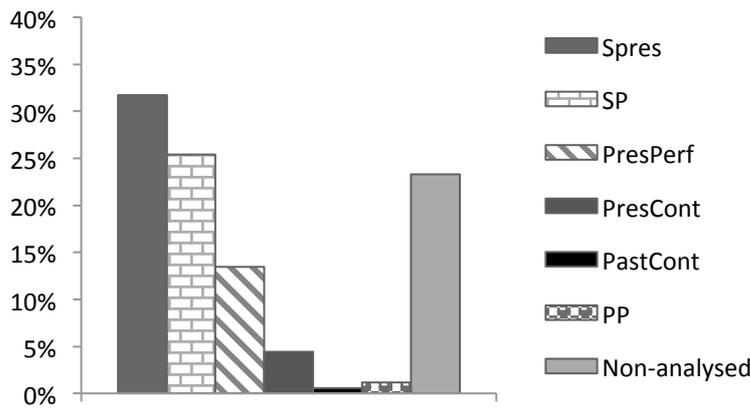
Figure 5-2 presents the frequency of three most frequent verbal tenses in each register. It shows that the SP is the preferred tense (representing 74% of the predicative tenses used) in the literature register, whereas the Spres and PresPerf are much less frequent. The distribution of these verbal tenses is more balanced than in the journalistic and EuroParl registers, where the SP occurs in 45% of the cases in the former and 25% of the cases in the latter. Finally, in legislation the Spres is the most frequent tense (58%), followed by the PresPerf and the SP.

Theses distributions correspond to the expected ones. Firstly, the SP is preferred in narratives instructing the addressee to order temporally eventualities with respect to one another. Secondly, the legislation register is a prospective and deontic register and the Spres is an appropriate verbal tense for expressing these interpretations (similarly to the FR Présent, see section 2.5.1.2). The journalistic and EuroParl registers consist of mixed types of texts (small narratives, comments, descriptions, etc.). In the journalistic register, the SP is used in 45% of the cases, the Spres in 34% and the PresPerf in 14% of the cases. In the EuroParl register, the Spres is used in 43% of the cases, the SP in 25% and the PresPerf in 22% of the cases.

Figure 5-2 Frequency of EN tenses per register (column distribution)
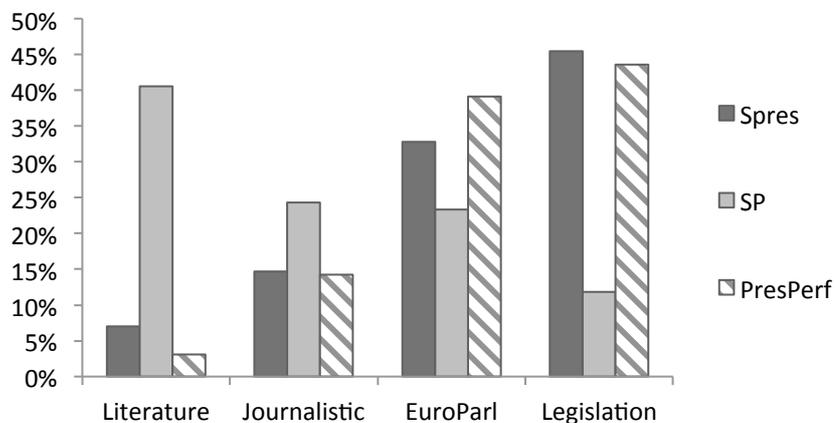
Figure 5-3 shows the distribution of all SP occurrences per register. 41% of SP occurrences come from the literature register, and the rest of 59% is shared between

journalistic and EuroParl (24% and 23%) and finally, the remaining 12% occur in the legislation register.

Figure 5-3 The distribution of the SP per register (row distribution)
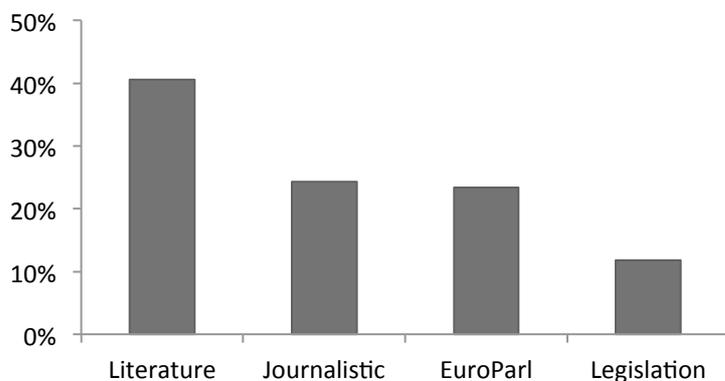
To sum up, the monolingual analysis of this corpus revealed that the most frequent verbal tenses are the Spres, the SP and the PresPerf. In section 5.1.2, I will provide the results of the cross-linguistic analysis, which will show which verbal tenses have consistently more than one translation possibility in FR (i.e. are ambiguous for MT systems).

5.1.2 Cross-linguistic analysis

The cross-linguistic analysis was performed using the *translation spotting* method (see section 4.2.3 for a methodological presentation) in order to identify *translation divergences*. It is considered a translation divergence a verbal tense for which in TL there are at least two translation possibilities which are much more frequent than all the other translation possibilities. The analysis revealed two translation divergences among the verbal tenses considered, precisely the SP and the PresPerf. The results from Table 5-2 indicate that each of the first four verbal tenses is consistently translated into FR through one frequent verbal form. Explicitly, the Spres is most often translated through the PRES verbal tense[149], the PP is most often translated though the PQP verbal tense[150], the PastCont is most often translated through the IMP verbal tense[151] and the PresCont is most often translated through the PRES verbal tense[152].

The PresPerf is one of the two translation divergences identified. The PC is most often used whereas the PRES is much less used, though it is more frequent than any other

---

[149] The Others category consists of very infrequent cases, such as 0 translation (5%), present participle, past participle and modal verbs (2% for each form), conditional, future, IMP, PC, PS, infinitive and noun (1% for each form), infinitive and PS (0.2% for each form), forming a total of 19%.

[150] The Others category consists of IMP (3 occurrences representing 16%), PC (2 occurrences representing 11%), subjunctive, participle and anterior past (1 occurrence representing 5% for each form), forming a total of 42%.

[151] The Others category consists of PQP, noun and *était en train de* (1% for each form representing 11%), for a total of 33%.

[152] The Others category consists of IMP (4 occurrences representing 5%), 0 translation (3 occurrences representing 4%), modal verbs (2 occurrences representing 3%), future and PC (1 occurrence representing 1% for each form), forming a total of 15%.

forms[153]. Finally, the SP is the most important translation divergence. It is translated into FR though four tenses. The first three are past time tenses (PC, IMP and PS) and the fourth one is the present tense (i.e. the PRES)[154].

Table 5-2 Translation possibilities EN-FR

EN	Spres	PP	PastCont	PresCont	PresPerf	SP
FR	PRES 81%	PQP 58%	IMP 67%	PRES 85%	PC 68%	PC 33% IMP 29%
					PRES 13%	PS 18% PRES 5%
	Others 19%	Others 42%	Others 33%	Others 15%	Others 19%	Others 15%

My hypothesis for explaining this linguistic variation of the forms used in FR for the translation of each of the verbal tenses considered is that the most frequent translated tenses share the semantics and pragmatics with the source verbal tense and are predictable forms. Less frequent forms (included in the Others category), on the contrary, are context-dependent, i.e. they depend on specific type of text, its purpose, the translator's personal choice, etc., and are unpredictable forms. Following Dyvik (1998), I suggest that predictable forms are about *langue* and *type* items, whereas unpredictable forms are about *parole* and *token* items. As far as this thesis is concerned, I will deal only with the predictable forms of the SP translation divergence.

The SP translation divergence is interpreted as following: the SP has several usages that correspond to several FR tenses used as its translation possibilities. The FR tenses used to render the semantic and pragmatic meaning of the SP are: the IMP, the PC, the PS and the PRES. The PC is used most frequently in the EuroParl and the journalistic registers whereas the PS is used most frequently in the literature register and the PRES used in 9% of the cases in the legislation register in order to create a certain effect in deontic contexts, as shown in Figure 5-4. This distribution shows that register is not a good predictor of the verbal tense used in TL: in the literature genre the SP is translated through an IMP in 44% of the cases and through a PS in 40%.

---

[153] The Others category consists of past participle (11 occurrences representing 5%), subjunctive (7 occurrences representing 3%), noun (8 occurrences representing 4%), 0 translation (4 occurrences representing 2%), IMP, *venir de*, past infinitive, anterior future, PQP (2 occurrences representing 1% for each form), participle and past conditional (1 occurrence representing 0.5% for each form), forming a total of 19%.

[154] The Others category consists of 0 translation (14 occurrences representing 3%), past participle, PQP, subjunctive (10 occurrences representing 2% for each form), conditional, past infinitive, noun and present participle (4 occurrences representing 1% for each form), past conditional, infinitive (2 occurrences representing 0.5% for each form) and *venir de* (1 occurrence representing 0.2%), forming a total of 15%.

Figure 5-4 Translation possibilities of the EN SP into FR (column distribution)
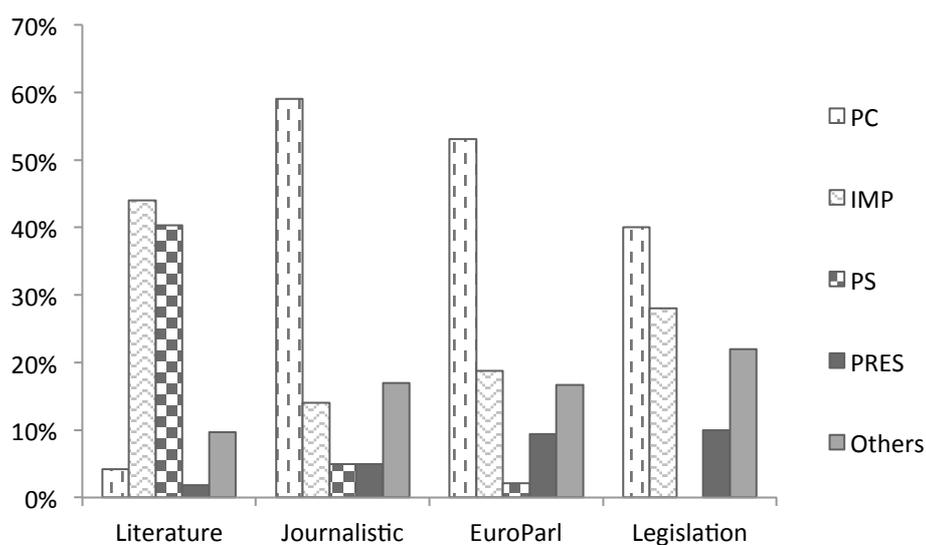
Examples (489)-(491) depict the translation divergence of the English SP: in (489) the SP is translated by the French IMP, in (490) by the PC, in (491) by the PS and in (492) by the PRES.

(489) EN/SP: The atmosphere *had* more to do with the negative aspects of a great European project and vision than a positive promotion of what is deep and good about the European dream, and that is a disappointing feature of Nice. (EuroParl Corpus)
FR/IMP: 'L'ambiance *avait* plus à voir avec les aspects négatifs d'un grand projet et d' une grande vision pour l' Europe qu'avec une promotion positive de ce que le rêve européen a de profond et de positif, et c'est là un aspect décevant de Nice.'

(490) EN/SP: I welcome the consultation process and can assure colleagues that in my Member State the authorities *took care* to carry out a broad and meaningful consultation. (EuroParl Corpus)
FR/PC: 'Je me félicite du processus de consultation et je peux assurer mes collègues que les autorités de mon pays *ont pris soin* de mener une consultation vaste et significative.'

(491) EN/SP: Cyril had very little affection for him, and was only too glad to spend most of his holidays with us in Scotland. They never really *got on* together at all. (Literature Corpus)
FR/PS: 'Cyril avait fort peu d'affection pour lui, et n'était que trop heureux de passer l'essentiel de ses vacances avec nous en Ecosse. Ils ne s'*entendirent* jamais véritablement.' (Literature Corpus)

(492) EN/SP: Something else they *had in common* was that they either conflicted with existing legal instruments or duplicated them. (EuroParl Corpus)
FR/PRES: Ces initiatives *ont* également *en commun* que tantôt, elles sont en contradiction avec les instruments juridiques existants, tantôt, elles les dupliquent.

Corpus analysis revealed that there is a mismatch between theoretical descriptions of verb tenses and actual usages in corpora. Certain verb tenses that are predicted by theoretical literature as being ambiguous for translation purposes, such as the English PastCont or the PastPerf, are infrequent in the corpus described in this section. Others, such as the English

Spres and SP, are ambiguous and frequent, representing therefore an important translation divergence.

Regarding the theoretical description of the SP in terms of the Reichenbachian coordinates S, R and E (see section 2.2.1), the SP shares the same configuration only with the PS (E=R<S). Even if the IMP has the same configuration as the PS, Reichenbach (1947) underlines that the two verbal tenses are different: the first is extended (i.e. progressive) and the latter non-extensive. Moreover, the PC, which is the most frequent verbal tense used to translate the SP, has a different temporal configuration than the SP, precisely E<R=S. Finally, the fourth tense used for translating the SP is the PRES, which is described as E=R=S. There are two questions that arise at this point of the discussion. The first regards the relation between the source and the target languages, specifically, what do the verbal tenses used in TL reveal about the verbal tense used in SL? The second question regards the factors that explain and predict this cross-linguistic variation. Several candidate features are tested experimentally in Chapter 6, where section 6.1.8 provides a multifactorial analysis of the data. Finally, an answer to these two questions is given in Chapter 7.

5.2 Bilingual Corpus: FR-EN

5.2.1 Monolingual analysis

The corpus consists of texts written in FR and their translations into EN belonging to four different genres, having the following proportions: literature 24%, journalistic 25%, legislation 21% and EuroParl 31%. The corpus contains 1283 occurrences of predicative verbal tenses, occurring in a total of 603 sentences. A total of 1031 predicative verb tenses have been considered[155], which represents 80% of the verb tenses occurring in the corpus, as shown in Table 5-3. The rest of 20% of verbal tenses have not been considered[156] in the analysis.

Table 5-3 Verbal tenses per register in the FR-EN bilingual corpus

Register	No. of sentences	No. of verbal tenses	No. of verbal tenses considered	% of verbal tenses considered	% of verbal tenses not considered
Literature	162	305	275	21%	2%
Journalistic	172	320	220	17%	8%
EuroParl	180	392	332	26%	5%
Legislation	89	266	204	16%	5%
Total	603	1283	1031	80%	20%

Figure 5-5 illustrates the frequency of verbal tenses[157] in the corpus, where the most frequent ones are the PRES (37%), the PC (19%) and the IMP (14%), whereas the PS and PQP are much less frequent (9% for the former and 3% for the latter).

---

[155] The considered tenses are several tenses from the indicative mood, such as IMP, PS, PC, PRES and PQP.

[156] Non-analysed tenses are other tenses from the indicative mood, other moods and modal verbs.

[157] Legend: PRES= Présent, PC=Passé Composé, IMP=Imparfait, PS= Passé Simple, PQP= Plus-que-parfait and Non-analysed

Figure 5-5 Frequency of FR verbal tenses in the FR-EN bilingual corpus
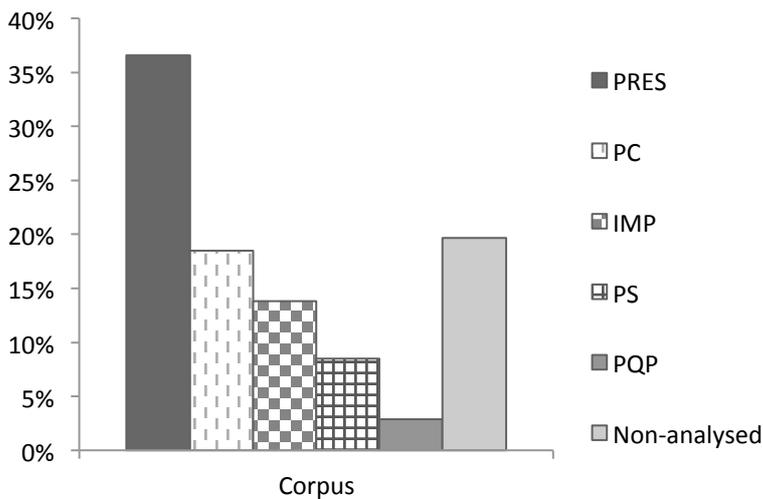
Figure 5-6 presents the frequency of the analysed verbal tenses in each register. The PRES is the preferred tense (29% in the former and 26 in the latter) in the journalistic and legislation registers, whereas the PC, IMP and PS are much less frequent (expect the PC in legislation used in 14% of the cases). The distribution of these verbal tenses is more equilibrated in the literature and EuroParl registers.

Figure 5-6 Frequency of FR verbal tenses per register (row distribution)
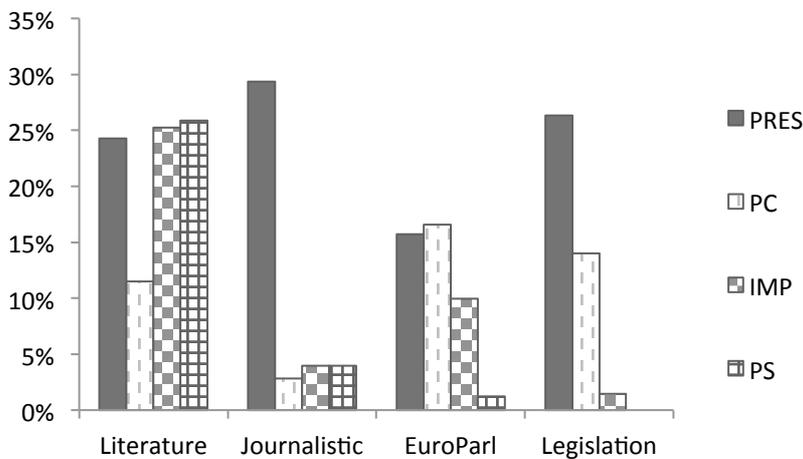
Figure 5-7 presents the distribution of each FR verbal tense considered in the four registers. The PRES verbal tense occurs most often in the journalistic (33%). 27% of the PRES tokens analysed occur in legislation and 24% in EuroParl. Finally, 16% of the tokens come from the literature register. These distribution show that the PRES verbal tense is not specialized for any stylistic register. The IMP is highly used in literature (44% of the IMP tokens) and in EuroParl (41% of the IMP tokens). The PS and PC also seem to be stylistically specialized. Precisely, most of the PC tokens occur in the EuroParl and legislation registers (51% and 27 % respectively) whereas 72% of the PS tokens occur in the literature register.

Figure 5-7 Distribution of FR verbal tenses in all registers (column distribution)
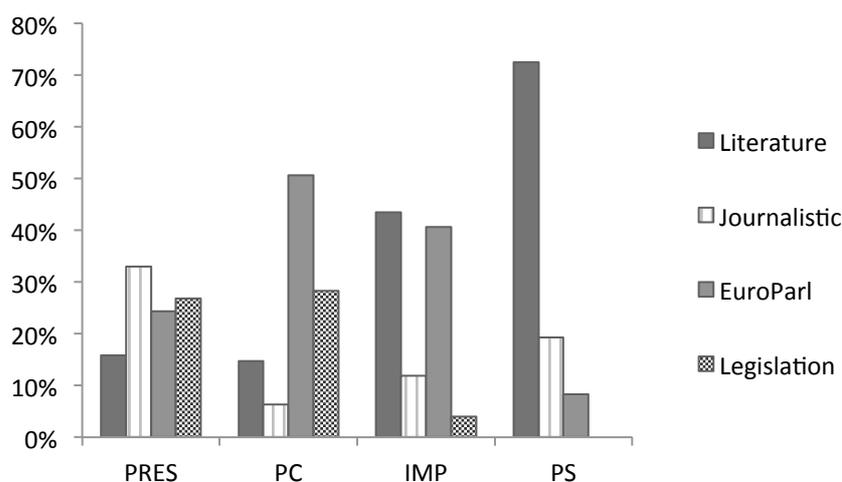
These results refute the predictions made in the literature, especially in classical discourse analysis field, with respect to using verbal tenses exclusively in one or another stylistic register or type of discourse. For example, Weinrich (1973) makes the prediction that the FR PS is used only in texts falling under the *monde raconté* 'stories' (i.e. literature) opposed to texts coming under the *monde commenté* 'commentaries' (i.e. journalistic, legislation, parliamentary disccusions among others), where other past time verbal tenses, such as the PC, are used[158]. Figure 5-7 indicates the PS is not exclusively used in the literary register but also in journalistic and EuroParl registers.

To sum up, the monolingual analysis of this corpus revealed that the most frequent verbal tenses are the PRES, the PC and the IMP. In section 5.2.2, I will provide the results of the cross-linguistic analysis, which will show which verbal tenses have consistently more than one translation possibilities in EN.

5.2.2 Cross-linguistic analysis

The cross-linguistic analysis of the parallel corpora performed through the translation spotting method, revealed two translation divergences among the verbal tenses considered, precisely the PC and the PQP. The results from Table 5-4 indicate that each of the first three verbal tenses considered (i.e. IMP, PS and PRES) is consistently translated into EN through one verbal form (i.e. the most frequent translation possibility into the TL). Explicitly, the IMP is most often translated through the SP verbal tense[159], the PRES is most often translated through the Spres verbal tense[160] and the PS is most often translated

---

[158] For a critical discussion of discursive and textual theories regarding FR verbal tenses, see Saussure (2003).

[159] The Others category consists PP (12 occurrences representing 7%), *would* (7 occurrences representing 4%), PresPerf, gerund (3 occurrences representing 2% for each form), Spres, PastCont, 0 translation, infinitive, past perfect continuous (2 occurrences representing 1% for each form), forming a total of 18%.

[160] The Others category consists of future (62 occurrences representing 13% and occurring exclusively in the legislation register), SP, PresCont, 0 translation (23 occurrences representing 5% for each form), PresPerf (20 occurrences representing 4%), modal verbs (13 occurrences representing 3%), gerund (11 occurrences representing 2%), infinitive, past participle (4 occurrences representing 1% for each form), Pcont (2 occurrences representing 0.4%) and *would* (1 occurrence representing 0.2%), forming a total of 39%.

through the SP verbal tense[161]. The PC is one of the two translation divergences identified in the FR into EN direction. Precisely, the SP and the PresPerf are the most frequent translation possibilities compared to the other forms[162]. Finally, the PQP is the second translation divergence. It is translated into EN though three tenses: SP, PP and PresPerf[163].

Table 5-4 Translation possibilities FR-EN

FR	IMP	PS	PRES	PC	PQP
				SP 47.6%	SP 52.6%
	SP 82%	SP 93%	Spres 61%	PresPerf 42.9%	PP 28.9%
EN					PresPerf 10.5%
	Others 18%	Others 7%	Others 39%	Others 9.5%	Others 7.8%

Examples (493) and (494) illustrate the translation divergence of the French PC: in the former example the PC is translated through the SP while in the latter through the PresPerf verbal tense.

(493) FR/PC: Une chance à laquelle, comme l'*a dit* notre collègue Böge, nous devons maintenant donner une forme concrète. (EuroParl Corpus)
EN/SP: 'An opportunity that must be given concrete shape, as the honourable Member Böge *said*.'

(494) FR/PC: J'*ai volé* un peu partout dans le monde. Et la géographie, c'est exact, m'a beaucoup servi. (Literature Corpus)
EN/PresPerf: 'I *have flown* a little over all parts of the world; and it is true that geography has been very useful to me.'

Examples (495)-(497) illustrate the translation divergence of the French PQP: in the first example the PQP is translated through the SP, in the second through the PP and in the third through the PresPerf.

(495) FR/PQP: Dans les années 1570, le sang des protestants massacrés *avait* littéralement *ruisselé* dans les rues de Paris, et le conflit qui s'en était suivi avait déchiré le pays pendant des générations. (Journalistic Corpus)
EN/SP: In the 1570s, Paris literally *flowed* with the blood of slaughtered Protestants, and the ensuing conflict tore the country apart for generations.

(496) FR/PQP: Le père du jeune Fergusson, un brave capitaine de la marine anglaise, *avait associé* son fils, dès son plus jeune âge, aux dangers et aux aventures de sa profession. (Literature Corpus)

---

[161] The Others category consists of PP (4 occurrences representing 4%), modal verbs (2 occurrences representing 2%) and gerund (1 occurrence representing 1%).

[162] The Others category consists of Spres (7 occurrences representing 3%), PP, 0 translation (5 occurrences representing 2% for each form), PresCont, modal verbs (2 occurrences representing 1% for each form) and gerund (1 occurrence representing 0.4%), forming a total of 9.5%.

[163] The Others category consists of past participle, gerund and Spres (3% for each form), forming a total of 7.8%.

EN/PP: 'Ferguson's father, a brave and worthy captain in the English Navy, *had associated* his son with him, from the young man's earliest years, in the perils and adventures of his profession.'

(497) FR/PQP: De plus, ce n'est pas la première fois que j' interviens dans un parlement - y compris celui -ci - et jamais personne ne m' *avait accusé* de faire de la flibusterie, bien au contraire. (EuroParl Corpus)

EN/PresPerf: 'Furthermore, this is not the first time I have spoken in a parliament - this not even the first time I have spoken in this one - and nobody *has* ever *accused* me of filibustering.'

As for the EN-FR direction of translation, corpus works reveals mismatches between theoretical descriptions of verbal tenses and their actual usage in human communication. The PC and PQP represent cases where the theoretical description with the help of Reichenbachian temporal coordinates seems to need improvements. Precisely, the FR PC is described as having the same temporal configuration as the PresPerf (i.e. E<R=S). In other words, the PC and the PresPerf are expected to be in a perfect translation relation, i.e. to share the same semantic and pragmatic content. The corpus work described in this section brings evidence against this association and questions the classical configuration suggested for the PC. A linguistic theory about the meaning of the FR PC should explain cases where the PC is translated through an SP and cases where it is translated through a PresPerf.

Another interesting case is the PQP, which, is considered to have the same temporal configuration as the EN PP (i.e. E<R<S). However, corpus work reveals that the PP is only one of the three verbal tenses used for the translated of the PQP into EN (in 29% of the cases). As shown in Table 5-4, the SP is used in 58% of the cases and the PresPerf is used in 11% of the cases. Similarly to the case of the PC, theoretical semantics and pragmatics need to provide an explanation for the PQP translation divergence.

To sum up, sections 5.1 and 5.2 provided quantitative and qualitative analyses of verbal tenses and their usage in SL and their translation possibilities into a TL. Two directions of translation have been considered, EN into FR and FR into EN. Cross-linguistic analyses have indicated most problematic translation divergences in the two directions of translation. In this thesis, one translation divergence is systematically investigated, more precisely, the translation of the SP into a TL. In order to increase the empirical basis of this research, two other Romance languages were added: the IT and the RO. The results of multilingual corpus analysis are provided in the following section. As for the other translation divergences identified, they need to be explored in further research.

5.3   Multilingual Corpus

The multilingual corpus consists of texts written in EN and their translations into FR, IT and RO. This kind of corpus is called *parallel translations* corpus (Granger 2003). The main advantage of parallel translations corpora is that one can identify language-independent patterns, i.e. systematic choices of translators in TLs when dealing with the same form in SL. The multilingual corpus described in this section was built with the aim of identifying language-independent patterns for the translation of the EN SP verbal tense. In section 5.3.1, I will describe how data was collected, and in section 5.3.2 I will provide the results of

the corpus analysis per target language.

5.3.1 Data collection

The multilingual corpus was created with the specific purpose of analysing the translation of the EN SP into three target languages. The chosen languages belong to the same family of languages, i.e. Romance languages. Within the family, however, they belong to different groups. As noted by Hall (1964), RO belongs to the eastern group whereas IT and FR belong to the Italo-Western group, which is divided into Western Romance (Portuguese, Spanish, Catalan, Occitan and FR) and the Proto-Italian (IT). This choice of language allows controlling for cross-linguistic variance that is due to structural differences among languages.

With the aim of guaranteeing comparability with the bilingual corpus (EN-FR, described in section 5.1), the multilingual corpus consists of texts belonging to the same stylistic registers: literature, EuroParl, legislation and journalistic[164]. The occurrences of the SP were randomly selected from the texts written in EN. They were aligned with their translations into FR, IT and RO. Texts from all languages and all stylistic registers are parallel translations, except for the EN-RO data pertaining to the EuroParl register. Since Romania joined the European Union later than France and Italy, the RO data in EuroParl is available only after 2004. Therefore, the data EN-FR-IT consists of parallel translations and the EN-RO data is a separate file. Table 5-5 provides the percentage per type of register of the number of SP occurrences in the source texts. 513 occurrences of the SP and their translation into three TLs (a total of 1281 sentences in the four languages) were analysed.

Table 5-5 Description of the multilingual corpus

	Literature	EuroParl	Legislation	Journalistic
EN- FR/IT	38%	19%	25%	18%
EN-RO	39%	16%	26%	18%

5.3.2 Analysis and Results

The corpus was analysed from a cross-linguistic perspective through the translation spotting method. The results from Table 5-6 indicate that all three target languages make most frequently use of the same verbal forms. Specifically, the FR data from the multilingual corpus is comparable to the FR data[165] from the bilingual corpus, described in section 5.2.2. The IT data shows that the PC[166] is in 33% of the cases, followed by the PS in 22% of the

---

[164] A detailed presentation of the texts used for data collection is provided in the Appendix section.

[165] The FR Others category consists of noun (12 occurrences representing 3%), 0 translation, past participle, PQP and subjunctive (10 occurrences representing 2% for each form), gerund, infinitive, rephrase (6 occurrences representing 1% for each form), conditional (1 occurrence representing 0.2%) forming a total of 16%.

[166] For the sake of simplicity, I will use the same name for the same verbal tense occurring in all three TLs. Specifically, I use PC for the FR Passé Composé, the IT Passato Prossimo and the RO Perfectul Compus; PS for the FR Passé Simple, the IT Passato Remoto and the RO Perfectul Simplu; IMP for the FR

cases, the IMP in 18% of the cases, by the PRES in 5% of the cases and, finally, by several other linguistic forms included in the Others[167] category. In RO, the PC is by far the most frequent verbal tense used (49%), followed by the PS (18%), the IMP (15%), the PRES (5%) and by other linguistic forms included in the Others[168] category.

Table 5-6 Translation possibilities of the SP into FR, IT and RO in the multilingual corpus

	FR	IT	RO
PC	37%	33%	49%
IMP	24%	18%	15%
PS	16%	22%	18%
PRES	8%	5%	5%
Others	16%	21%	13%

Table 5-7 provides the frequency of each of the verbal tenses considered in each register for each TL. It can be seen that for all three languages, for each register, verbal tenses have similar distributions. Explicitly, in the literature register, the most frequent verbal tenses are the PS and the IMP. In EuroParl, legislation and journalistic, it is the PC that is most frequently used whereas the PS is almost inexistent. This distribution could be interpreted as a register specialization on the PS showing the complementarity of the two verbal tenses expressing past time. In each register and for all three languages, the IMP is the second most frequent tense. Based on these data and on theoretical considerations (see section 2.5.1), I suggest reducing the SP translation divergence to a three branch divergence: PS/PC, IMP and PRES.

Imparfait, the IT Imperfetto and the RO Imperfect; PRES for the FR Présent, the IT Presente and the RO Prezent.

[167] The IT Others category consists of past participle (17 occurrences representing 4%), noun, 0 translation, PQP, subjunctive, rephrase (12 occurrences representing 3% for each form), gerund, infinitive (3 occurrences representing 1% for each form) and conditional (1 occurrence representing 0.7%), forming a total of 21%.

[168] The RO Others category consists of noun (11 occurrences representing 3%), 0 translation, past participle (9 occurrences representing 2% for each form), PQP, subjunctive, gerund and conditional (3 occurrences representing 1% for each form), infinitive and future (3 occurrences representing 0.5% for each form), forming a total of 13%.

Table 5-7 Frequency of verbal tenses in FR, IT and RO per register

	Verbal tense	Literature	EuroParl	Legislation	Journalistic
FR	PS	**40%**	0%	0%	1%
	IMP	*35%*	*17%*	*14%*	*22%*
	PC	10%	**45%**	**63%**	**49%**
	PRES	1%	17%	12%	9%
IT	PS	**55%**	0%	0%	9%
	IMP	*28%*	*12%*	*14%*	*10%*
	PC	1%	**40%**	**64%**	**52%**
	PRES	0%	10%	7%	8%
RO	PS	**45%**	0%	0%	0%
	IMP	*29%*	*7%*	*5%*	*9%*
	PC	13%	**76%**	**75%**	**66%**
	PRES	1%	3%	11%	8%

This interpretation is also shown in Figure 5-8, Figure 5-9 and Figure 5-10 respectively for each TL. From these figures, one can see that in FR 99% of the PS occurrences are in the literature register and the remaining of 1% in the journalistic register. In IT 93% of the PS occurrences are in the literature register and the remaining of 7% in the journalistic register. In RO, all occurrences of PS belong to the literature register. As for the PC, it has the lowest frequencies in the literature register in all three languages, with a minimum of 1% in IT. Regarding IMP, most of the occurrences are in the literature register in all three languages. EuroParl, legislation and journalistic register also make use of the IMP with a frequency of 12% in FR, 10% in IT and 7% in IT. Finally, the lowest frequencies of the PRES verbal tense are in the literature register whereas the highest frequencies are in EuroParl in FR (40%), EuroParl and legislation in IT (36% for each language) and in legislation in RO (57%).

Figure 5-8 Frequency of FR verbal tenses (row distribution)
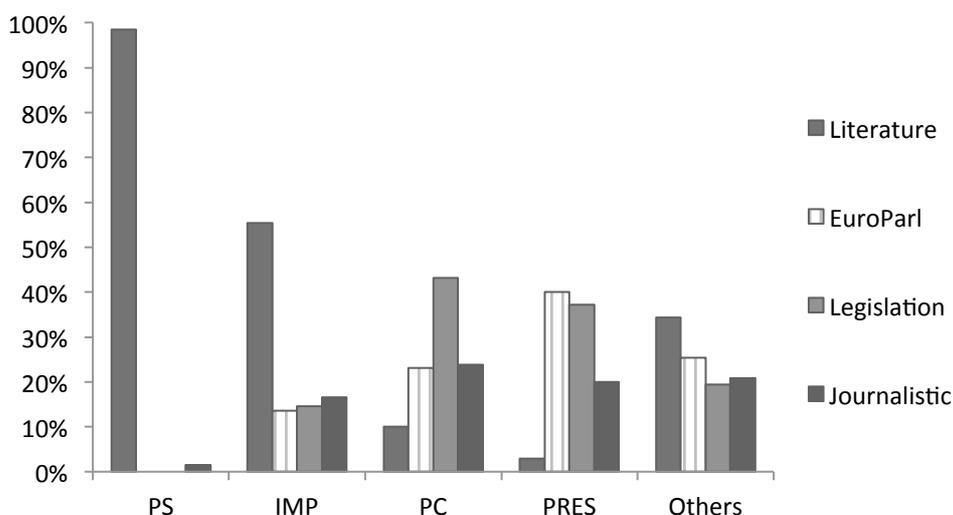
Figure 5-9 Frequency of IT verbal tenses (row distribution)
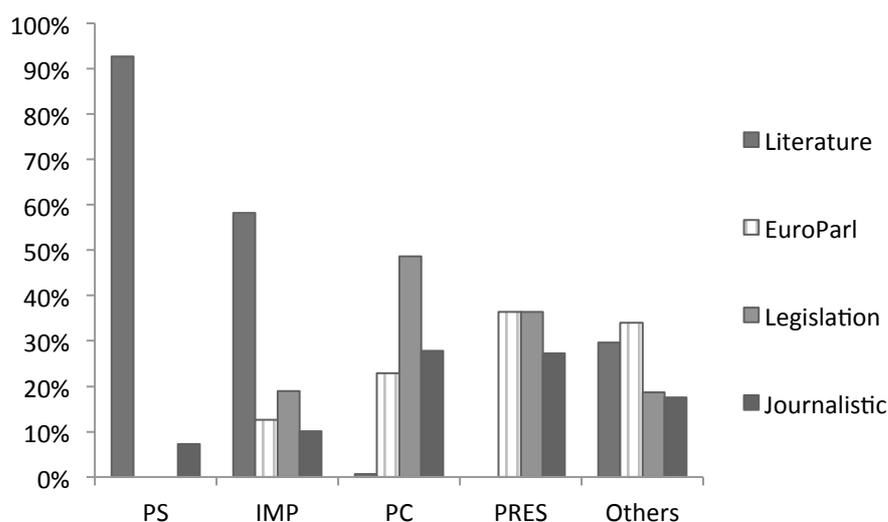
Figure 5-10 Frequency of RO verbal tenses (row distribution)
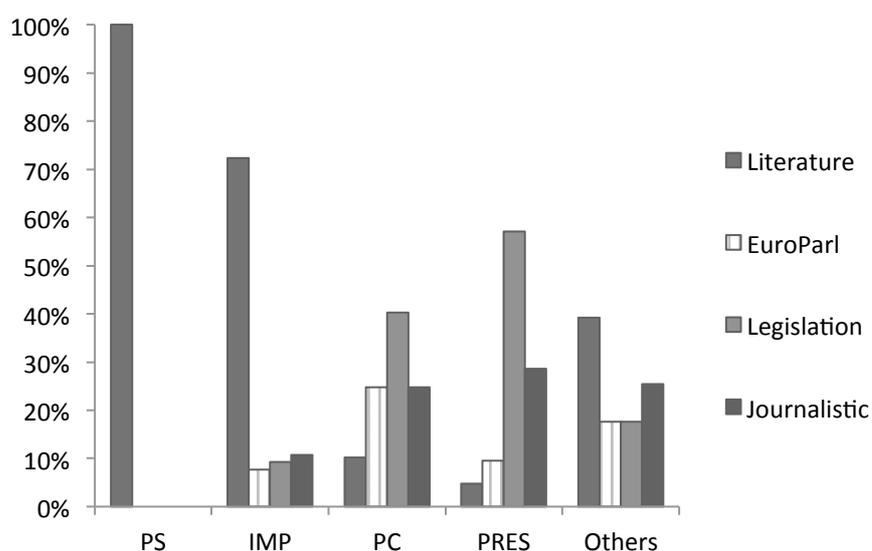
To sum up, the translation divergence of the SP identified in the bilingual corpus EN-FR is confirmed in the multilingual corpus. The SP is most frequently translated into FR, IT and RO through a PS in the literature register and through a PC in the other three registers. Similarly, the PRES is used almost exclusively in the EuroParl, legislation and journalistic registers in all three languages. Finally, the IMP verbal tense is used in all four registers for translating an SP.

5.4 Conclusive remarks

In this chapter, an account of the corpus work carried out in this research was given. Bilingual and multilingual parallel corpora were built and analysed according to the methodology described in Chapter 4. The results of the corpus work described in sections 5.1, 5.2 and 5.3 can be summarized as follows:

- The EN-FR parallel corpus revealed two main translation divergences: PresPerf and SP. These two verbal tenses are both frequent in the corpus and ambiguous i.e. each of them is systematically translated into TL through at least two verbal forms.

- The FR-EN parallel corpus revealed two main translation divergences: PC and PQP. The PC is both frequent and ambiguous whereas the PQP is ambiguous but much less frequent.

- The parallel translations corpus confirmed the SP translation divergence identified in the bilingual corpus. The data on IT and RO brought further evidence for including the PC and PS in a unique category and reduced, therefore, the initial four-branch divergence to a three-branch divergence.

The SP translation divergence was chosen for further investigations aiming at proposing a multilingual theoretical and empirical disambiguation model. In this thesis, the term *disambiguation* does not imply that the SP is polysemous. On the contrary, as argued in section 3.1.3, Tense is an underdetermined linguistic category which must be contextually worked out. Consequently, a verbal tense does not have several *meanings* but several contextual *usages*. The notion of *disambiguation model* refers therefore to disambiguation among the various usages of the SP. The basic idea is that the SP has several usages and each of these usages may be translated into a certain TL through a different verbal tense. Of course, the question that arises at this point of the discussion regards the nature and the source of these usages. I will deal with this topic in section 6.

6 Offline experiments with linguistic judgment task

In section 4.2.3, I introduced methodological matters regarding two techniques that use the properties of translation corpora: translation spotting and cross-linguistic transfer of properties. Both methods make use of the assumption that translation is a source of perceived similarities across languages. The main consequence is that translation corpora can be used to investigate semantic and pragmatic properties of an SL and one or several TLs. The former method consists of identifying the verbal tenses used in TL when translating a source verbal tense and making hypotheses about the properties included in the translation relation. The latter method consists of the cross-linguistic transfer of properties from L1 to L2. However, these methods allow only the formulation of hypotheses about possible semantic and pragmatic properties of the linguistic phenomena considered. These features must be validated monolingually in offline experiments with linguistic judgment task. In order to have reliable data, in the offline experiments the inter-annotator agreement is evaluated with the Kappa measure, which corrects for chance agreement (see section 4.3 for methodological issues). In this section, I will describe the experimental work for the monolingual validation of features. Section 6.1 is dedicated to bilingual data and section 6.2 to multilingual data. The annotation guidelines used for each of the experiments carried out are provided in the Appendix section.

6.1   Bilingual Data

The analysis of translation corpus through the translation spotting method has shown that the EN SP translation paradigm consists of four verbal tenses in FR, three verbal tenses expressing past time (PC, PS and IMP) and a verbal tense expressing present time (PRES). FR verbal tenses expressing pas time received abundant attention in the literature (as discussed in Chapters 2 and 3) and were described regarding three features provided below. Table 6-1 provides the summary of the classical description of FR verbal tenses expressing past time in terms of these four features (see section 2.5.1.2 for detailed description).

- Temporal sequencing: in the interpretation process, the hearer makes directional inferences regarding the temporal relations holding among eventualities based on the information provided by the verbal tense. This information was operationalized as the [±narrativity] feature.
- Aktionsart: in the interpretation process, the hearer makes inferences about the inherent temporal structure of an eventuality. This information was operationalized as the [±boundedness] feature.
- Aspect: in the interpretation process, the hearer makes inferences about the completion/entirety vs. ongoing status of an eventuality. This information was operationalized as the [±perfectivity] feature.

Table 6-1 FR verbal tenses and their semantic and pragmatic properties

Verbal tense/Feature	Temporal sequencing	Aktionsart	Aspect
PS	Temporal progression	No restriction	Perfective
PC	Neutral	No restriction	Perfective
IMP	Lack of temporal progression	No restriction	Imperfective

The experimental work presented in this section consists of seven experiments, described as follows:

- Test the theoretical link between verbal tenses and the [±narrativity] feature

  - Experiment 1: FR PC, PS and IMP with the [±narrativity] feature
  - Experiment 2: FR IMP with the [±narrativity] feature
  - Experiment 3: EN SP with the [±narrativity] feature

- Test the theoretical description of EN SP regarding aspectual information

  - Experiment 4: EN SP with [±boundedness] feature
  - Experiment 5: EN SP with [±perfectivity] feature

- Test the configuration of Reichenbachian temporal coordinates

  - Experiment 6: EN SP and PresPerf with Reichenbachian coordinates
  - Experiment 7: FR verbal tenses expressing past and non-past time with Reichenbachian coordinates

One way of evaluating human annotation is to calculate the inter-annotator agreement with the help of Cohen's Kappa coefficient (Carletta 1996; as discussed in section 4.3.2). One issue that influences corpus annotation by annotators is the subjectivity of judgments, which can be quite substantial for semantic and pragmatic annotations (Artstein and Poesio 2008). It can be tested whether different judges produced consistently similar results, so that one can infer that the annotators have understood the guidelines and that there was no agreement just by chance. The Kappa statistic factors out agreement by chance and measures the effective agreement by two or more raters. The Kappa coefficient is 1 if there is a total agreement among the annotators and 0 if there is no agreement other than the one expected to occur by chance. In what following sections, I will describe the procedure and the results for each of the seven experiments.

6.1.1 Experiment 1: French PC, PS, IMP and the [±narrativity] feature

Experiment 1 was designed to test the existent theoretical assumptions regarding the link between verbal tense and temporal sequencing and for this purpose the [±narrativity] feature was suggested. The [±narrativity] feature is a pragmatic feature referring to information about temporal relations holding among eventualities. It is a binary feature with the *narrative* and *non-narrative* values. Narrative usages point to the case when the two eventualities are temporally linked (both forward and backward inferences). Non-narratives usages point to the case when eventualities are either not temporally linked or they occur simultaneously. In (498), the first three eventualities expressed through an SP have a

narrative usage whereas the last and the fourth are used non-narratively.

(498) John screamed [e2]. His leg was broken [e3]. Mary pushed him [e1]. She felt betrayed [e4].

As shown in Table 6-1, the link between FR verbal tenses and temporal sequencing was suggested in the literature. The assumption that that a verbal tense has narrative and non-narrative usages was made. Experiment 1 was designed to test the following three hypotheses:

- The PS has more often narrative usages than non-narrative usages.
- The PC has more often narrative usages than non-narrative usages.
- The IMP has more often non-narrative usages than narrative usages.

Participants

There were 76 FR native speakers, who were first year students at Faculty of Letters from University of Geneva. Their participation at the experiment was organized during a Linguistics class. It was however benevolent and anonymous.

Procedure and material

The materials used consisted of 300 items[169] randomly chosen from FR part of the parallel corpus organized in 19 sets. Each participant received a set of 15 items. The data contained 127 occurrences of IMP described by the literature as most frequently non-narrative, 173 occurrences of PS/PC (101 PS and 72 PC) described as most frequently narrative.

The annotation guidelines included two tasks. The first task was to read and understand the definitions of narrative and non-narrative usages, as follows:

- The eventualities are temporally linked. This means that e1 happened before e2. The relation may be explicitly expressed in the sentence or it may be implicit (you understand it based on the context).
- The eventualities are not temporally linked. This means that e1 and e2 happened either in the same time (simultaneously) or they are not temporally linked (opposite to case 1 above).

Each definition was accompanied by two explained examples, as given in (499) where V1 (PS) and V2 (PS) have a narrative usage while V3 (IMP) has a non-narrative usage. The second task was to read each item and decide if the verb highlighted has a narrative or a non-narrative usage.

(499) On raconte qu'un Anglais *vint* [V1] un jour à Genève avec l'intention de visiter le lac; on le *fit* [V2] monter dans l'une de ces vieilles voitures où l'on *s'asseyait* [V3] de côté comme dans les omnibus.

---

[169] An item consists of a sentence where the tense form of interest occurs (PS, PC or IMP for Experiment 1 and SP for Experiment 2) and another sentence, either the preceding or following one. This choice is explained by the need to have sufficient co(n)text for a pragmatic decision.

It is said that un Englishman come.PS one day to Geneva with the intention to visit the lake; he send up.PS in one of these old cars when you sit.IMP on the sides as in a local train

Participants were trained on 6 items followed by a collective discussion. The evaluation was done manually according to an evaluation scheme defined as follows. The results were evaluated by counting the majority of answers for each item, since there were more than two judges. The number of concordant answers must be superior to agreement by chance, which is at 50% since there is a binary choice (i.e. narrative vs. non-narrative usage). When the situation was equal (2 over 2 judges), than the item was evaluated as *inconclusive*. Inconclusive items were disregarded from further analysis. Moreover, labels given by participants were compared to a baseline established on the basis of theoretical descriptions of the considered verbal tenses (provided in section 2.5.1.2).

Results

Table 6-2 provides the results of this annotation experiment, where 221 tokens of IMP, PC and PS were considered. Among the 300 items annotated by four judges 79 received equal answers (no majority), hence they were inconclusive. These items were not considered in the analysis. In the clean data of 221, judges agreed with the theoretical reference in 182 items (82% of the data) with a Kappa value measuring inter-annotator agreement of 0.63. This value signals reliable results.

The table shows that the narrative feature was identified for 86% of the annotated tokens according to the theoretical predictions (i.e. PS and PC together, 110 items labelled as narrative out of 128 existent in the corpus) and the non-narrative feature for 77% of the cases (IMP, 72 items labelled as non-narrative out of 93 existent in the corpus). Chi-square test performed on this result shows that the correlation between the annotator's judgment and the theoretical reference is statistically significant (Chisq 86.96, df=1, p<.0001).

Table 6-2 Narrativity for PS/PC and IMP: Majority of judges and Reference

		Majority of judges		
		Narrative	Non-narrative	Total
Reference	PC/PS	110	18	128
	IMP	21	72	93
	Total	131	90	221

More precisely, as shown in Table 6-3, judges clearly recognized a primary narrative usage for the PS (92%) but did not make the same clear judgment for the PC (narrative in 77% of the cases) or the expected non-narrative primary usage of the IMP (77.5%).

Table 6-3 Judgment for individual verbal tenses

Verbal tense/Narrativity	Narrative	Non-narrative
PS	92%	8%
PC	77%	23%
IMP	22.5%	77.5%

This leads to about 23% of non-expected usages, namely non-narrative usages for the PC and 22.5% narrative usages for the IMP. This result opened a path for further finer-grained investigation, namely an annotation experiment of the IMP with the [± narrativity] feature.

6.1.2 Experiment 2: the French IMP and the [±narrativity] feature

Experiment 2 was designed to test the existent theoretical assumption regarding the two usages of the IMP: non-narrative and narrative. The narrative usage of the IMP, known as the breaking IMP ("imparfait de rupture"), is characterized by the presence of a subjectivity marker or a temporal adverbial or connective that encodes an immediate transition towards a resulting state. This information is inferential and it directs discourse computation towards temporal sequencing. Theoretical assumption about the information expressed by narrative and non-narratives IMPs is summarized in Table 6-4. Explicitly, both narrative and non-narrative IMPs express reference to past time and are viewed as continuous eventualities. Non-narrative IMP does express temporal sequencing and is not viewed as being completed, whereas narrative IMP expresses temporal sequencing and it is views as being completed (the final boundary is expressed by a temporal adverbial or it expresses a punctual eventuality). The former is illustrated in example (500) and the later in (501).

Table 6-4 IMP and its semantic and pragmatic properties

Usage of IMP/Feature	Temporal location	Continuity	Temporal sequencing	Completion
Non-narrative	Past time	Continuous	Lack of temporal progression	Not completed
Narrative	Past time	Continuous	Temporal progression	Completed

(500) Il y a une heure Max *boudait* dans son coin, et ça n'est pas près de changer.
An hour ago Max sulk.IMP in a corner, and this will not change very soon.

(501) Elle a fini par fuguer à Kaboul, où elle a été recueillie par une femme généreuse. Quelques mois plus tard, elle *épousait* un jeune cousin de sa bienfaitrice dont elle était tombée amoureuse.
She finally run.PC to Kaboul, where receive.PC.Passive by a kind woman. A few months later, she marry.IMP a younger cousin of her benefactor of whom she fell in love.PQP.

Experiment 2 aimed at validating that these two usages of the IMP are identifiable by native speakers in corpus data.

Participants

There were 2 FR native speakers, who were students at Faculty of Letters from University of Geneva. They were paid for their participation at the experiment.

Procedure and material

The material consisted of a total of 230 items containing IMP tokens. 120 items were randomly selected from the FR part of the parallel corpus, where FR was the SL. 110 IMP

were translations of SP items into FR, where FR was the TL. The two annotators received annotation guidelines, consisting of the definition and examples for each type of usages. They were trained on 6 items followed by a group discussion. Evaluation was done by calculating the inter-annotator agreement rate using the Kappa coefficient.

Results

The results are presented in Table 6-5. Out of 230 annotated tokens, judges agreed on the annotation of 179 tokens (77%) representing a Kappa value of 0.24. This very low Kappa value is explained by the fact that the two categories (narrative and non-narrative) are not equally distributed, and therefore the non-narrative category is the default case. The judges were not aware that there is a default case and they assigned the categories by judging the sentences according to the annotation guidelines. If only the 176 cases of agreement are considered in the analysis, the IMP was categorized in 90% of the cases as being non-narrative and in 10% of the cases as being narrative.

Table 6-5 Narrativity for IMP: Judge 1 and Judge 2

| | | Judge 2 | | Total |
		Narrative	Non-narrative	
Judge 1	Narrative	17	35	52
	Non-narrative	19	159	178
	Total	36	194	230

The annotation results have also been analysed regarding the original language. For the 120 tokens of IMP written in FR as SL, judges agreed on 90 items (75%). Among the agreement cases, the IMP was labelled as being non-narrative in 84% of the cases and narrative in 16%. As for the 110 tokens of IMP written in EN as SL, judges agreed on 86 items (78% of the cases). Among the agreement cases, the IMP was labelled as being non-narrative in 97% of the cases and narrative in 3%. The results of this experiment show that categorization of the IMP in terms of narrative and non-narrative usages presents different patterns regarding the SL. However, using a Fisher exact test, the difference of categorization between the two source languages is shown not to be statistically significant (p=0.09).

6.1.3 Experiment 3: the English SP and the [±narrativity] feature

According to the cross-linguistic transfer of properties method, there is equivalence between the SP and the FR tenses used for its translation with respect to the [±narrativity] feature. The assumption is that the SP has narrative and non-narrative usages. Experiment 3 was designed to test the following hypotheses:

- Narrative usages of SP are translated more often through a PS/PC than through an IMP.
- Non-narrative usages of the SP are translated more often through an IMP than through a PS/PC.

There two participants, EN native speakers, who were studying Linguistics at the Bachelor level at the Faculty of Letters from University of Geneva. They were originating from the United Kingdom. Their participation at the experiment was paid.

Procedure and material

The material used consisted of 458 items containing 458 SP tokens randomly chosen from the EN part of the parallel corpus. As in the first two experiments, judges received annotation guidelines and went through a training phase. The first task from the annotation guidelines was to read and understand the instructions containing definitions of narrative and non-narrative usages (as given in Experiment 1). They also included two examples for each usage, as given in the examples (502) and (503). The second task was to read each item and decide if the verb highlighted has a narrative or a non-narrative usage. Participants were trained on 10 items followed by a discussion where each annotator had to "think loud" his/her answers.

In the first example below, there are two events, i.e. 'the marriage that happened' and 'the wealth which was added'. The second event is presented in relation to the first (first he got married and then he added to his wealth), which is why the SP verbs happened and added are in narrative usage. In the second example, there are three states (was a single man, lived and had a companion) that describe the owner of the estate. States are not temporally ordered, which is why this example illustrates the non-narrative usage of the SP.

(502) By his own marriage, likewise, which happened soon afterwards, he *added* to his wealth. (Literature Corpus)

(503) The late owner of this estate was a single man, who *lived* to a very advanced age, and who for many years of his life, had a constant companion and housekeeper in his sister. (Literature Corpus)

Evaluation of inter-annotator agreement rate was done with the Kappa coefficient. As regards cross-linguistic evaluation, the judged items were compared to a reference baseline containing the verbal tenses used for the translation of the SP into FR from the FR part of the parallel corpus.

Results

Results are provided in Table 6-6. Judges agreed on 325 items (71%) and disagreed on 133 items (29%). The value of Kappa coefficient was 0.42. This value is above chance, but not high enough to point to entirely reliable linguistic decisions. Among the 113 items of disagreement, 19 items were signalled as having insufficient context for a pragmatic decision. They were disregarded from further analysis.

Table 6-6 Narrativity for SP: Judge 1 and Judge 2

| | | Judge 2 | | Total |
		Narrative	Non-narrative	
Judge 1	Narrative	180	83	263
	Non-narrative	50	145	195
	Total	230	228	458

Error analysis showed that the main source of errors was the length of the temporal interval between two eventualities, which was perceived differently by the two annotators. This lead to ambiguity between temporal sequence or simultaneity, each of them corresponding to narrative, respectively, non-narrative usage, as in example (504) where the eventualities "qualify" and "enable" were perceived as being simultaneous by one judge and successive by the other.

(504) Elinor, this eldest daughter, whose advice was so effectual, possessed a strength of understanding, and coolness of judgment, which qualified her, though only nineteen, to be the counsellor of her mother, and *enabled* her frequently to counteract, to the advantage of them all, that eagerness of mind in Mrs. Dashwood which must generally have led to imprudence. (Literature Corpus)

A possible explanation is the fact that personal world knowledge is used to infer temporal information, such as the length of the temporal interval between two eventualities, i.e. information that allows the judge to decide whether the eventualities are temporally ordered or not. Cases where the length of the temporal interval between two eventualities was very reduced were ambiguous for the judges, so each of them decided differently whether it was long enough for temporal sequencing or too short, so that the simultaneity meaning was preferred.

Disagreements (114 items) were resolved in a second round of the annotation experiment, where the narrativity feature was identified with a new linguistic test that was explained to two new participants[170](as suggested by Spooren and Degand 2010, see section 4.3.2 for a discussion). Judges were asked to insert a connective, such as *and* and *and then* when possible, in order to make explicit the 'meaning' of the excerpt, namely the temporal relation existent between the two eventualities considered. The connective *because* (for a causal relation) has also been proposed by annotators under the [+narrative] label showing that causal relations should also be considered. The inter-annotator agreement rate in this second phase of the experiment was corresponds to a Kappa of 0.91, signalling very strong and reliable agreement.

In the data containing agreements, the SP was judged as having narrative usages in 59% of the cases and non-narrative usages in 49% of the cases. This finding suggests that the SP is not specialized for one or the other possible values of the [±narrativity] feature. The cross-linguistic application of these findings consists of the observation of a pattern in the parallel corpus. The data containing agreements from both annotation rounds (435 items) was

---

[170] The new participants were the author and a research peer, who was not aware of the purpose of the research. They were not native speakers but were fluent in spoken and written EN, and used EN as professional language.

investigated and analyzed against the reference baseline, defined based on the parallel corpus. The two alternative hypotheses are:

- Non-narrative SP is translated more often with an IMP.
- Narrative SP is translated more often with a PC/PS.

The results are provided in Table 6-7. They show that the narrative usages of the SP correspond to narrative usages in the FR part of the corpus (translation by a PC or PS) and the non-narrative usages of the SP correspond to the non-narrative usages in the FR text (translation with an IMP) in 338 items (78%). Using a Chi-Square significance test, this correspondence is shown to be statistically significant (Chisq 124.26, df=1, p<.001). This correlation is shown in Figure 6-1. The effect is intermediately strong with a Phi-coefficient of 0.52. The remaining of 22%, for which annotators agreed on the narrativity label but which are not consistent with the verbal tense used in FR, point to narrative usages of the IMP and to non-narrative usages of the PC.

Table 6-7 Narrativity for SP: Judges vs. Baseline

| | | Baseline | | Total |
		PC/PS	IMP	
Judges	Narrative	208	49	257
	Non-narrative	48	130	178
	Total	256	179	435

Figure 6-1 Correlation Narrativity and Target tense
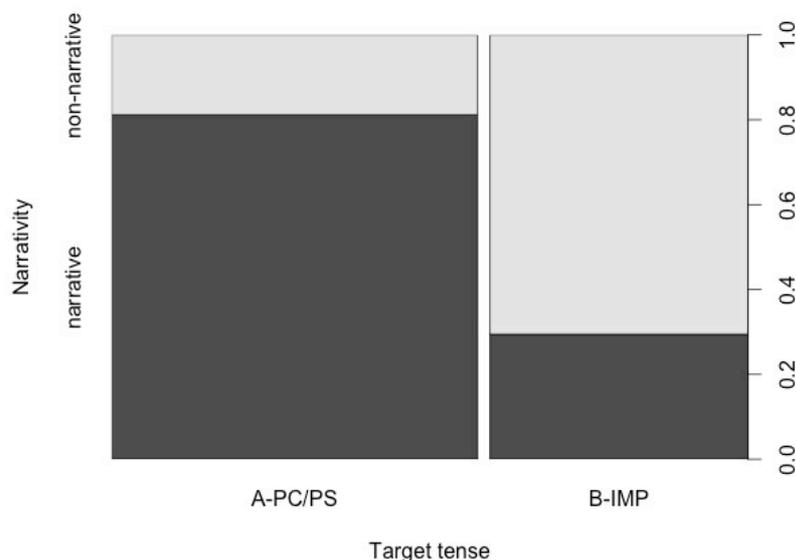
The association plot in Figure 6-2 shows the contribution to the overall significative Chi-Square of every cell (levels of the dependent and independent variable). In this plot, the area of the box is proportional to the difference in observed and expected frequencies. The black rectangles above the dashed line, indicating observed frequencies exceeding expected frequencies, correspond narrative usage of the SP positively correlated to the PC/PS value of the Target tense dependent variable and to the non-narrative usage of the SP positively correlated to the IMP value of the dependent variable. The grey rectangles below the dashed line, indicating observed frequencies smaller than expected frequencies, correspond to the lack of correlation between non-narrative usage of the SP and the PC/PS and narrative

usage of the SP viewpoint and the IMP value of the dependent variable.

Figure 6-2 Association plot for Narrativity and Target tense: Residuals
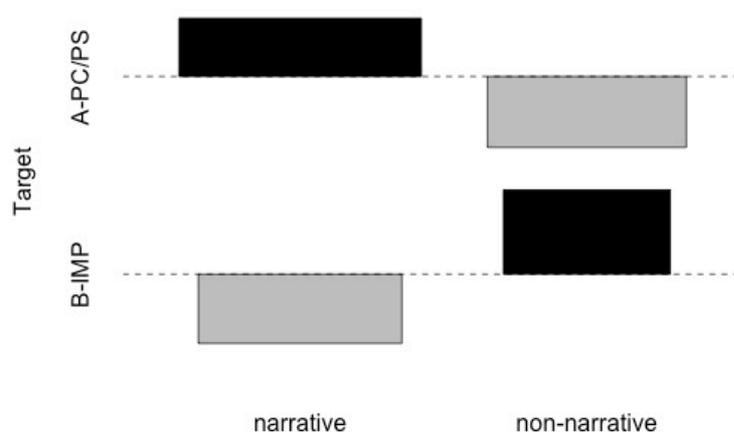
The experiment described in this section showed that native speakers of EN have a poor ability to consciously evaluate temporal interpretations triggered by Tense, operationalized as the [±narrativity] feature. The difficulty to make conscious evaluations of this type of information provides strong empirical evidence for the procedural nature of this feature, which is described as not easily accessible to consciousness. When speakers do not have conscious access to the instructions encoded by linguistic items, this information can be uncovered through other means. Precisely, participants were asked to propose a connective that would render explicit the implicit temporal relation (such as *and then*) or the implicit lack of temporal relation (such as *and at the same time*). The results showed that explicitating the implicit relation is an easier task for speaker than consciously evaluating these temporal relations. This represents strong empirical evidence for the procedural nature of this feature.

The results of Experiments 1-3 indicated that the narrativity feature is identifiable both for FR and EN verbal tenses after the second phase, when the judges inserted temporal connectives in order to render explicit the implicit temporal relation existing among the eventualities expressed. From a cross-linguistic perspective, the narrative usage of the SP is translated with PC or PS (themselves having a narrative usage), while an IMP is used to translate the non-narrative usage of the SP. Moreover, when investigated in translation corpora, narrative usages of the SP also point to narrative usages of the IMP (known as the *historical/breaking/narrative* IMP). Hence, it is conceivable to hypothesize a cross-linguistic character of this feature.

6.1.4 Experiment 4: the English SP and the [±boundedness] feature

Experiment 4 was designed to test if speakers can categorize the inherent aspectual properties of VPs, i.e. the [±boundedness] feature, and to investigate the relation between the type of eventuality and the verbal tense used. Eventualities are theoretically distinguished between *bounded* (generally, achievements and accomplishments) and *unbounded* (generally, states and activities) (see section 2.4.2 for a theoretical discussion). Dowty (1986) suggested the link between eventuality type, temporal progression and verbal tense. Precisely, he

argued that bounded eventualities trigger temporal progression as in examples (505) and (506) whereas unbounded eventualities express lack of temporal progression as in examples (507) and (508) (see also section 2.4.3).

> (505)   John entered in the president's office. The president walked toward him.
> (506)   John entered the president's office. The president stood up.
> (507)   John entered the president's office. The president sat behind a huge desk.
> (508)   John entered the president's office. The clock on the wall ticked loudly.

From a bilingual perspective, Kozlowska (1998) argued that in FR there is temporal progression with bounded eventualities expressed with the PS as in (509) and (510) and there is no temporal progression with unbounded eventualities expressed with the IMP as in (511) and (512), where examples (509)-(512) are the FR translation of examples (505)-(508).

> (509)   Jean entra dans le bureau du président. Le président *s'avança* vers lui.
> (510)   Jean entra dans le bureau du président. Le président *se leva*.
> (511)   Jean entra dans le bureau du président. Le président *était* assis derrière un énorme bureau.
> (512)   Jean entra dans le bureau du président. L'horloge murale *marchait* bruyamment.

This experiment tested the following hypotheses issued from the literature:
- Bounded eventualities expressed through an SP are more frequently translated into FR through a PC/PS
- Unbounded eventualities expressed through an SP are more frequently translated into FR through an IMP

Participants

A previous pilot experiment with the same feature showed that judging lexical aspect required a certain level of theoretical knowledge and that training was not sufficient to improve their results. In order to have reliable data annotated with the [± boundedness] feature, two research peers were asked to participate at this experiment. They were not native speakers but were fluent in spoken and written EN, and used EN as professional language[171]. They were not paid for their participation at the experiment.

Procedure and material

The material used is the clean data resulted from Experiment 3, that is 435 items containing SP tokens. Participants received annotation guidelines, consisting of the two tasks. The first task was to read and understand the descriptions of bounded and unbounded eventualities according to their behaviour with the linguistic tests (see section 2.4.1) provided in Table 6-8. The second task was to read each item and decide whether the verb highlighted expressed a bounded or unbounded situation. Evaluation of inter-annotator agreement rate was done with the Kappa coefficient. As regards cross-linguistic evaluation,

---

[171] For more accurate results, in further research, this experiment could be carried out again, with native speakers.

the labelled items were compared to a reference baseline containing the tenses used for the translation of the SP into FR from the FR part of the parallel corpus.

Table 6-8 Linguistic tests for the [± boundedness] feature

Test	Bounded eventualities	Unbounded eventualities
in/for adverbials	*in* adverbials	*for* adverbials
homogeneity	-	+
entailment with progressive	-	+

Results

Results are provided in Table 6-9. Judges agreed on the label for 401 items (92%) and disagreed on 34 items (8%). The agreement rate corresponds to a Kappa value of 0.84. All 34 disagreements were resolved in a second phase consisting of a discussion between the two judges.

Table 6-9 Boundedness for SP: Judge 1 and Judge 2

		Judge 2		Total
		Bounded	Unbounded	
Judge 1	Bounded	210	8	218
	Unbounded	26	191	217
	Total	236	199	435

In a cross-linguistic perspective, the data containing agreements from both annotation rounds (435 items) was investigated and analysed against the reference, defined based on the parallel corpus. The results are provided in Table 6-10. They show that bounded eventualities expressed with a SP correspond to a translation through a PC or PS and unbounded eventualities expressed with a SP correspond to a translation with an IMP for 360 items (82%). Using a Chi-Square test, this correspondence is shown to be statistically significant (Chisq 182.62, df=1, p<.001). This correlation is shown in Figure 6-3. The effect is intermediately strong with a Phi-coefficient of 0.661.

Table 6-10 Boundedness for SP: Judges and Reference

|		Judges		Total
		Bounded	Unbounded	
Reference	PC/PS	208	28	236
	IMP	47	152	199
	Total	255	180	435

Figure 6-3 Correlation Boundedness and Target tense
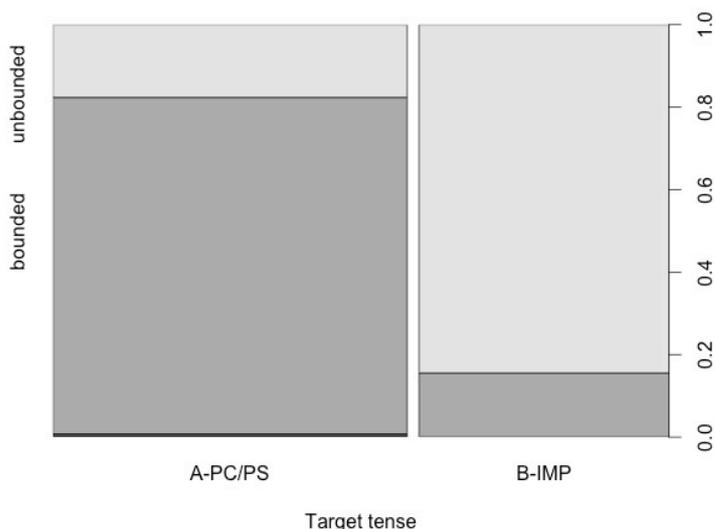
The association plot in Figure 6-4 shows the contribution to the overall significative chi-square of every cell. The black rectangles above the dashed line, indicating observed frequencies exceeding expected frequencies, correspond to the bounded type of situations positively correlated to the PC/PS value of the Target tense dependent variable and to the unbounded type positively correlated to the IMP value of the dependent variable. The grey rectangles below the dashed line, indicating observed frequencies smaller than expected frequencies, correspond to the lack of correlation between unbounded situations and the PC/PS and bounded situations and the IMP.

Figure 6-4 Association plot for Boundedness and Target tense: Residuals
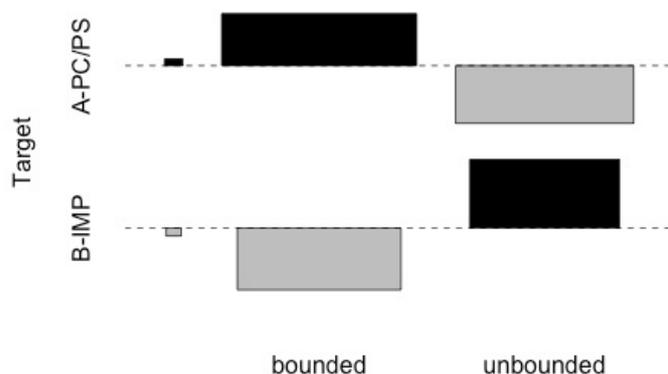
To sum up, this experiment showed that the SP is compatible with both bounded and unbounded eventualities and that this is observable on natural data. In this experiment, the two judges had a very high agreement rate. According to Sperber and Wilson's (1993) description of the cognitive foundations of the conceptual/procedural distinction, the information dealt with in this experiment is conceptual. From a cross-linguistic point of view, unbounded situations are most frequently correlated with the IMP whereas bounded situations with the PC/PS in the target language. This correlation is statistically significative. Therefore, one could expect that the [± boundedness] feature is a significant factor for predicting the verbal tense used in TL. This will be investigated in a multifactorial analysis (see section 6.1.7).

6.1.5 Experiment 5: the English SP and the [±perfectivity] feature

Experiment 5 was designed to test if speakers are able to categorize the aspectual viewpoint on eventualities, i.e. the [±perfectivity] feature. Eventualities can be presented with a perfective or an imperfective point of view (see section 2.3.2 for a theoretical discussion). The imperfective aspect restraints temporal progression by presenting the situation as ongoing or by setting a focus on an internal phase as in (513). The perfective aspect favours temporal expression by presenting the situation as a completed whole (Comrie 1976; Dowty 1986) as in (514).

(513) John entered the president's office. The president was writing a letter.
(514) John entered the president's office. The president wrote a letter.

From a bilingual perspective, the FR PS and PC are described as expressing the perfective aspect whereas the IMP is associated to the imperfective aspect in its non-narrative usages. However, the IMP has also narrative usages that present the situation as a completed whole (similarly to the perfective aspect), more precisely in its narrative usages. Experiment 7 was therefore designed to test the following hypotheses:

- Perfective usages of the SP are more often translated into FR though a PC/PS
- Imperfective usages of the SP are more often translated into FR through an IMP

The [±perfectivity] feature was first tested in a pilot experiment that I will describe below. The results of the pilot experiment not being reliable (value of Kappa coefficient below the threshold of acceptability), another technique was used the identification of grammatical aspect, i.e. translation of the EN data into Serbian, identification of the two aspectual categories and their transfer to the EN initial source data. I will describe this procedure after the description of the pilot experiment.

Participants

The participants of the pilot experiment were two EN native speakers, originated from the United Kingdom. They were the same participants from Experiment 3, in which SP tokens were annotated with the [±narrativity] feature. Their participation at the experiment was paid.

Procedure and material

The material used consisted of 62 items containing SP tokens chosen randomly from the data annotated in Experiment 3, precisely from the 22% of the cases for which judges' label did not correspond to the verbal tenses used in TL in the translation corpus. The participants received annotation guidelines, consisting of the definition of the perfective and imperfective viewpoints, as well as two examples for each category. The training was done on 13 items and it was followed by a collective discussion, where each judge had to think loud his/her decisions.

The two judges agreed on the label for 41 items (66%) and disagreed on 21 items (33%). The agreement rate corresponds to a Kappa value of 0.32. Disagreements were not resolved after the discussion between the two judges. The results of this pilot experiment show that the data annotated with the [±perfectivity] feature is not reliable. In order to have reliable data annotated with this feature another method was used.

Translation and cross-linguistic transfer of properties

A native speaker translated the data consisting of 435 items containing SP tokens into Serbian. The translator was a linguistics student from University of Geneva, native speaker of Serbian. The participation at the experiment was paid. Grammatical aspect was identified in Serbian (see section 2.1 for a discussion of the verbal system in Serbian) for each item and transferred on the EN initial source based on the cross-linguistic transfer of properties method. The SP was labelled as perfective for 204 items (47%) and as imperfective[172] for 231items (53%).

The results of the contrastive analysis between the value of Aspect and the verbal tense used in FR are provided in Table 6-11. They show that the perfective viewpoints expressed with a SP correspond to a translation through a PC or PS and imperfective viewpoint expressed with a SP correspond to a translation with an IMP for 339 items (78%). Using a Chi-Square test for independence, this correspondence is shown to be statistically significant (Chisq 132.86, df=1, p<.0001). This correlation is graphically illustrated in Figure 6-5. The effect is intermediately strong with a Phi-coefficient of 0.557.

Table 6-11 Perfectivity for the SP: Judgement through translation and Baseline

| | | Judgment through translation | | Row Total |
		Perfective	Imperfective	
Baseline	PC/PS	144	36	180
	IMP	60	195	255
Column Total		204	231	435

---

[172] For 7 items, the translator was free to choose between perfective and imperfective, both aspects being possible. The verbs which occurred in these sentences are *to promise, to spend, to reproach, to organize, to despise, to stay* and *to try.* All these verbs express atelic situations.

Figure 6-5 Correlation Perfectivity and Target tense
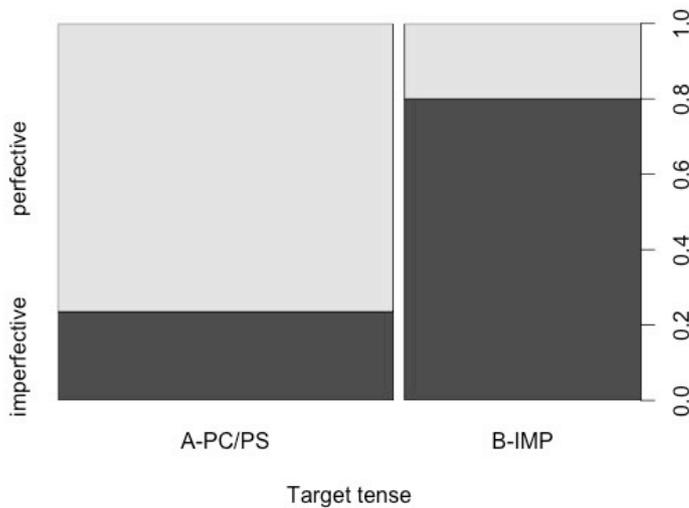
The association plot in Figure 6-6 shows the contribution to the overall significative Chi-Square of every cell. The black rectangles above the dashed line, indicating observed frequencies exceeding expected frequencies, correspond to the perfective viewpoint positively correlated to the PC/PS and to the imperfective viewpoint positively correlated to the IMP. The grey rectangles below the dashed line, indicating observed frequencies smaller than expected frequencies, corresponding to the lack of correlation between the imperfective viewpoint and the PC/PS and the perfective viewpoint and the IMP.
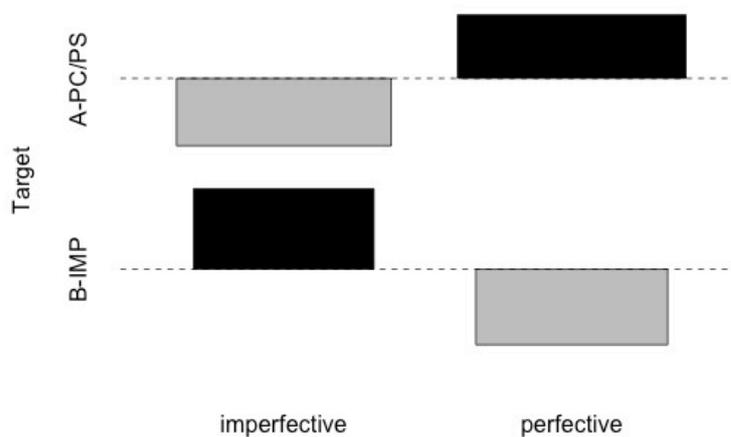
Figure 6-6 Association plot for Perfectivity and Target tense: Residuals

The experiment described in this chapter showed in the first place that native speakers of EN have a poor ability to consciously evaluate the meaning of Aspect, operationalized in this research as the [±perfectivity] feature. The difficulty to make conscious evaluations of the type of viewpoint provides strong empirical evidence for the procedural nature of this feature, which is described as not easily accessible to consciousness. When speakers do not have conscious access to the instructions encoded by linguistic items, this information can be found elsewhere. Based on parallel corpora, the novel cross-linguistic transfer of properties

technique was used in order to reveal for EN verbs procedural information that is expressed morphologically in Slavic languages.

Secondly, translation data annotated with the [±perfectivity] feature was analysed cross-linguistically. The results pointed to the strong correlation between perfective usages of the SP and the FR PC/PC and imperfective usages of the SP and the FR IMP. Another finding is the existence of less frequent cases, such as imperfective usages of the SP and the FR PC/PC and perfective usages of the SP and the FR IMP.

6.1.6 Experiment 6: the English SP, PresPerf and Reichenbachian coordinates

According to Reichenbach (1947) the EN SP and PresPerf have different configurations. The former express a past eventuality dislocated from the present time that is the moment of speech (E=R<S) as in example (515) whereas the latter expressed a past eventuality linked to the present (i.e. the past eventuality is completed but it has present relevance as in example (516) or the hearer can infer that the eventuality is not completed as in example (517) (see section 2.2.1 for a detailed theoretical discussion).

The theoretical assumption behind this experiment is that the difference in meaning between the SP and the PresPerf is not a difference in terms of conceptual information (both express E<S) but a difference in terms of procedural and aspectual information. Specifically, the SP is a preterit form (it locates of an eventuality prior to S) as in (515), whereas the PresPerfect is a perfectal form (it presents an eventuality as completed and having current relevance), as shown in examples (516) and (517).

(515)   I lived in Paris for two years and then in London for five years.
(516)   I have lived in Paris for two years. I can explain you how to get to the Eiffel tower.
(517)   I have lived in Paris for two years and I hope to remain here for the rest of my life.

Participants

There two participants, EN native speakers, who were studying Linguistics at the Bachelor level and were originate from the United Kingdom. Their participation at the experiment was paid.

Procedure and material

The material used consists of 30 items containing 15 SP tokens and 15 PresPerf tokens, which were randomly chosen from the EN corpus. Participants received annotation guidelines, consisting of the description of the two categories and two examples for each category chosen from the corpus. The description of the two categories was done in simple words such as *past situation* vs. *present result of a past situation* instead of the corresponding technical terms E=R<S vs. E<R=S. This experiment involves dealing with both conceptual and procedural information encoded by the SP and PresPerf. Precisely, for both verbal tenses the hearer is brought to build an ad-hoc concept E<S. However, these two verbal tenses differ with respect to the position of R: R=E for the former and R=S for the latter.

Judges were asked to read each item and judge the highlighted verb according to the annotation guidelines. The two judges participated to Experiment 3, and because they were

acquainted to the procedure, there was no supplementary training. Evaluation of the inter-annotator agreement was done with the Kappa coefficient.

Results

The results are provided in Table 6-12. Judges agreed on the label for all 30 items (100%). This agreement rate corresponds to a value of the Kappa coefficient of 1.

Table 6-12 Reichenbachian coordinates for SP and PresPerf: Judge 1 and Judge 2

		Judge 2		
		Past situation	Present result of a past situation	Total
Judge 1	Past situation	15	0	15
	Present result of a past situation	0	15	15
	Total	15	15	30

There are two possible explanations for the Kappa value of 1 signalling perfect agreement between the two judges. The first is that the PresPerf in British EN has remained a true perfect expressing that the eventuality denoted by the verb is already completed at the time of reference, which is S for PresPerf. This aspectual information might facilitate the hearer's access to the relation R=S. In other words, judging procedural information could be rendered easy to access through conscious thinking in cases when the hearer has supplementary sources of information. Similarly, in Experiment 3 on the [±narrativity] feature and the SP, disagreements were resolved by asking the judges to propose a connective that would render explicit the implicit temporal relation holding among eventualities. Judges agreed on the connective to be inserted at a very high rate. Connectives themselves encoding procedural information, one could have expected a low rate of agreement. My suggestion is therefore that supplementary sources of information facilitate the hearer's task in the interpretation process.

The second explanation, which is in the same time a drawback of this experiment, is the reduced number of items. A more important number of items might have resulted into a lower rate of agreement than 1 but high enough to have reliable data. Future research should take this aspect into consideration.

6.1.7 Experiment 7: French verbal tenses and Reichenbachian coordinates

Experiment 7 was designed to test whether speakers are able to categorize the configuration of two reichenbachian coordinates, E with respect to S. The main assumption behind this experiment is that verbal tenses encode a pro-concept TIME, which is semantically incomplete. In the interpretation process, the addressee builds an ad hoc concept E<S (i.e. past) or E≥S (i.e. non-past) based on contextual information. This offline experiment with elicitation task was designed to test this precise theoretical assumption.

Participants

The participants were FR native speakers, Bachelor students at the Faculty of Letters of

the University of Geneva and of the University of Neuchâtel. Their participation at the experiment was voluntary and benevolent. They did not receive training before participating at the experiment.

Procedure and material

The participants were asked to give the tensed form of a verb provided at the infinitive so that it would correspond to the surrounding cotext. The items used in this experiment were of two categories. The first category consists of 90 items randomly selected from the corpus (as described in section 5.2), which represent *naturally occurring* items judged in their *original contexts*. The second category consists of 36 artificial sentences built for the purpose of the experiment. Each item comprised a first sentence that sets the context and a second sentence that contains the targeted verb, as shown in examples (518) and (519).

(518) Le jeune soldat mis en cause a agi contre les orders de ses supérieurs, il (être) aujourd'hui incarcéré et en attente d'être jugé pour meurtre. (Literature register)
'The suspected young soldier behaved against his superior's orders, he (to be) today imprisoned and waiting to be judged for murder.'

(519) Marie a pris du poids. Avant de casser sa jambe, Marie (courir) tous les soirs pendant une heure. (Built example)
'Mary gained weight. Before breaking her leg, Mary (to run) every evening for an hour.'

The first sentence had the role to set a past or a non-past time context. Participants had the task to provide the tensed form of the verb in the second sentence in accordance with the context set by the first sentence. All experimental items were distributed in sets of 15 items for the corpus data and 18 items for the artificial sentences. Participants received annotation guidelines, in which they were asked to read each item and provide the tensed form of the verb in the infinitive. Each participant received either corpus (natural) or built (artificial) data.

The results of this experiment were evaluated by counting the majority of answers for each item, since there were more than two participants. The number of concordant answers must be superior to agreement by chance, which is at 50% since there is a binary choice (i.e. past vs. non-past context). When the situation was equal (3 over 6 judges), than the item was evaluated as *inconclusive*. Inconclusive items were disregarded from further analysis. Finally, for an item, if less than 50% of the judges (maximum 2 over 6 judges) made the same judgment than it was considered as a disagreement. Due to the reduced number of participants who saw each item, that is 6 per item, the evaluation was done manually. Moreover, labels given by participants were compared to a baseline established based on the translation corpus for the natural data and as defined by the experimenter when the data was built for the artificial data.

Results

A total of 126 items were evaluated according to the evaluation scheme described above. The judged agreed on their label for 119 items (94.4%) and disagreed on 3 items (2.4%). 4 items were evaluated as inconclusive (3.2%). Disagreements and inconclusive items were

disregarded from further analysis. Table 6-13 provides the results of the comparison between the label provided by the judges and the reference baseline (from translation corpus) for all data. The correspondence between the judges' label and reference of 111 items (93.3%) corresponds to a Kappa of 0.86.

Table 6-13 Judges vs. Reference baseline for past/non-past distinction in all data

| | | Judgment | | Row Total |
		Past	Non-past	
Baseline	Past	57	2	59
	Non-past	6	54	60
Column Total		63	56	119

Regarding the two types of data, natural vs. artificial, all the 3 disagreements and the 4 inconclusive items were natural data. Therefore, judged agreed on the label provided to all artificial items. When compared against the reference baseline, there is a one-to-one correspondence between the judges' labels and the baseline. This corresponds to a Kappa of 1.

As for the natural data, judges agreed on 83 items (92%). Among the agreements, the items were judged as expressing reference to non-past in 45 items (54.2%) and reference to past in 38 items (45.8%). Table 6-14 provides the results of the comparison between the label provided by the judges and the reference baseline for natural data only. The correspondence between the judges' label and reference of 75 items (90%) corresponds to a Kappa of 0.80.

Table 6-14 Judges vs. Reference baseline for past/non-past distinction in natural data

| | | Judgment | | Row Total |
		Past	Non-past	
Baseline	Past	39	2	41
	Non-past	6	36	42
Column Total		45	38	83

Discussion of results

This experiment aimed at testing whether speakers are able to categorize the configuration of two reichenbachian coordinates, E with respect to S. The hypothesis defended in this research is that the relation between these two coordinates is of a conceptual nature, and the ad hoc concept is built contextually. According to the qualitative features proposed by Wilson and Sperber (1993) for conceptual and procedural information, it was argued that judging conceptual information results in high Kappa values. This experiment provided evidence the conceptual information encoded by verbal tenses, that is past vs. non-past, is determined contextually and that the agreement among the participants produced high Kappa values: 1 for artificial data, 0.80 for natural data and 0.86 for all the data.

The difference of results with respect to natural vs. artificial data is that natural data used in this research is much more complex and more difficult to understand than artificial data built for the purposes of the experiment. This is partly due to the type of data, which originates in parliamentary discussions, legislation, journalistic and literature stylistic registers. The two types of data are exemplified in examples (520) for the natural data, where

in the baseline reference to past time was expressed through a PS and (521) for the artificial data, where reference to past time was expressed through an IMP.

<blockquote>
(520) De son côté, l'Eglise catholique avait organisé, en 1986, la Rencontre nationale ecclésiale cubaine (ENEC), qui - tout en rappelant que Cuba est une nation chrétienne - (prendre acte) de la société cubaine telle qu'elle était et non telle que l'Eglise l'aurait souhaitée. (Journalistic register)
'On its side, Catholic church had organized, in 1986, the Cuban National Eclesiastic Meeting, which - recall that Cuba is a Christian nation – (take cognizance of) the Cuban society as it was and not as the Church would have wished it.'

(521) Après son accident, Marie était très triste. Elle ne pouvait plus faire ce qui la rendait si heureuse. Marie (jouer) du piano. (Built example)
'After her accident, Mary was very sad. She could not do anymore what used to drive her so happy. Mary (play) the piano.'
</blockquote>

At this point of the discussion, I think it is necessary to point out that using the Kappa coefficient is a first attempt to have quantitative measures for conceptual and procedural information. I think that it is necessary in the future to create online experimental designs that could bring more light on this distinction. Section 6.1.8 is dedicated to multifactorial analyses of data resulted from the experiments described in sections 6.1.1 to 6.1.5.

6.1.8 Multifactorial analyses

The experimental work presented in sections 6.1.1 to 6.1.5. consisted of two types of investigations. Firstly, the experiments were designed for testing experimentally the validity of several theoretical assumptions: (i) the link between verbal tenses and temporal progression in FR and EN and (ii) the link between verbal tenses and aspectual information that is grammatical and lexical aspect. Secondly, the experiments were designed for testing the validity of the cross-linguistic transfer of properties method based on translation corpora.

The results of the experiments showed that EN SP, on the one hand, and FR PC/PS and IMP, on the other, are correlated regarding three type of encoded information: narrativity feature (i.e. temporal and causal relations), Aspect and Aktionsart. Explicitly, the SP is used both for bounded and unbounded situations presenting them from a perfective or an imperfective viewpoint having narrative or non-narrative interpretations. Cross-linguistic analysis of translation corpora revealed that different combinations of these features correspond to translations into FR either through an IMP or a PC/PS.

In order to investigate the relationships among the [± narrativity], [± boundedness] and [±perfectivity] features for predicting the verbal tenses used in a TL, multifactorial statistical analyses were performed. The theoretical principles lying behind multifactorial analyses were introduced in section 4.4.2. In this section, I will provide the results of the multifactorial analyses performed with R statistical program and their interpretation.

The data used in multifactorial analyses consists of 435 items containing SP tokens annotated for which the following information is known:

    i.   the verbal tense used in TL

    ii.   the verb in SL at the infinitive

    iii.   the stylistic register

iv. the value for each item in SL of the [±narrativity], [±boundedness] and [±perfectivity] features

The dependent variable is a binary categorical variable, i.e. the verbal tense used in TL, more precisely 255 occurrences of PC/PS and 180 occurrences of IMP. The independent variables were classified as fixed predictors, that is the [±narrativity], [±boundedness] and [±perfectivity] features, and random predictors, that is the verb and the stylistic register. The three fixed predictors are correlated as shown by the two-by-two figures below. The correlation Perfectivity and Boundeness is statistically significant (Chisq 224.57, df = 2, p<0.05) corresponding to 0.469 Cramer's V value. Perfectivity and Narrativity is statistically significant (Chisq 95.71, df = 1, p<0.05) corresponding to 0.469 Cramer's V value. Finally, Narrativity and Boundedness is statistically significant (Chisq 147.28, df = 2, p<0.05) corresponding to 0.582 Cramer's V value.

Figure 6-7 Correlation Perfectivity and Boundedness
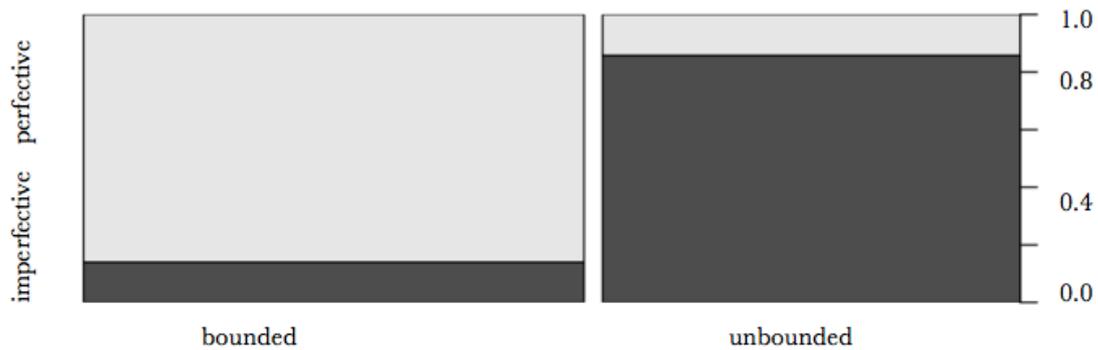
Figure 6-8 Correlation Perfectivity and Narrativity
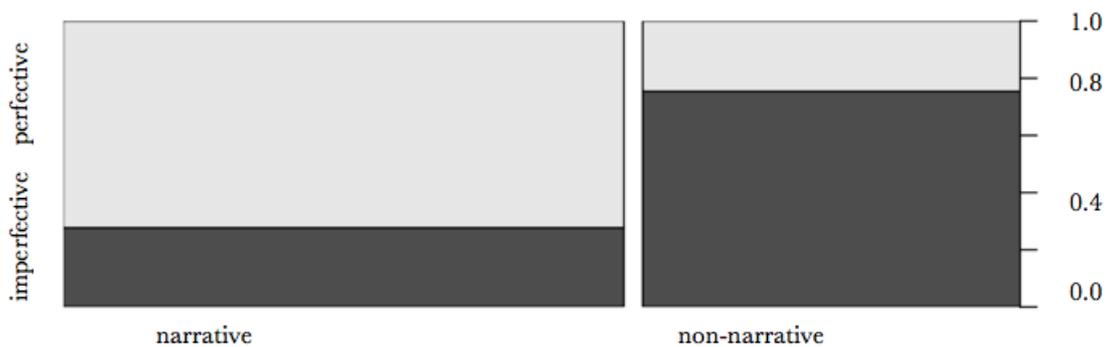
Figure 6-9 Correlation Narrativity and Boundedness
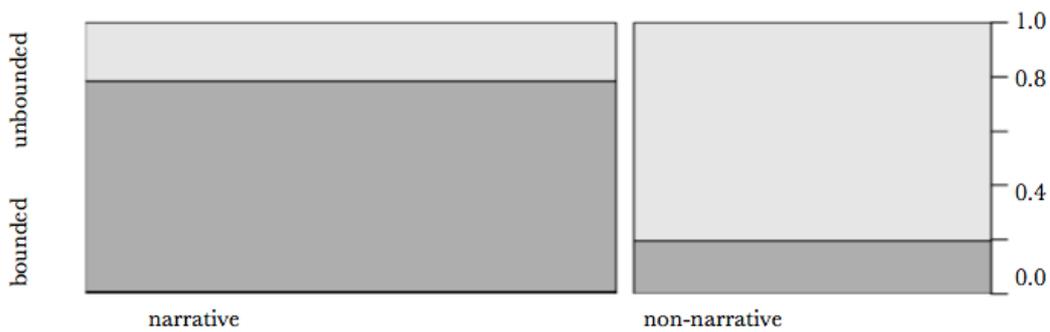
Figure 6-10 presents the distribution of the data regarding the three fixed predictors established. It can be seen that there are two main tendencies and that all combinations are possible for the SP. The first main tendency is perfective view associated for bounded situations in narrative contexts whereas the second is imperfective viewpoint associated to unbounded situations in non-narratives contexts.

Figure 6-10 Mosaic plot of the data with three fixed predictors: narrativity, perfectivity and boundedness
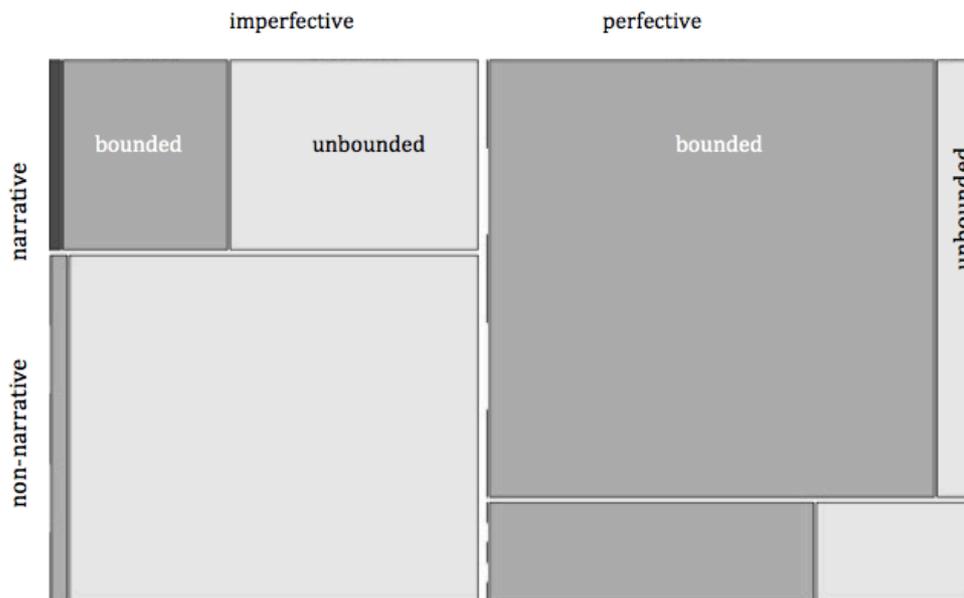
The order of the predictors for finding the best model (i.e. balance between high within-dataset accuracy and high predictive accuracy for new data) was calculated with the Step function. An ANOVA performed on the results of the Step function are provided in Table 6-15. It can be seen that there are four significative predictors, one significative interaction (signalled by the ':' colon symbol) between Aktionsart and narrativity and one almost significative interaction (between Aktionsart and Aspect).

Table 6-15 Order of predictors and their *p* value

Predictor	Df	Chi-Square *p*
Boundedness	2	<.0001
Narrativity	1	<.0001
Perfectivity	1	0.001
Boundedness:Narrativity	1	0.03
Boundedness:Perfectivity	1	0.08

Following the standard stepwise procedure aiming at respecting Occam's razor, a maximal model was built, i.e. the model that includes all fixed and random predictors and their interactions. Secondly, other models were built by iteratively deleting the least relevant predictor. Finally, an ANOVA was performed on all models and the significant model with the highest number of degrees of freedom was kept. The model best fitting the data is the model that considers the three fixed predictors and the interaction between boundedness and narrativity, as well as one random predictor, the verb. Table 6-16 provides the results of the best fitting model and shows that narrativity and perfectivity, as well as the interaction between lexical aspect and narrativity, are statistically significant factors for predicting the verbal tense used in TL.

Table 6-16 Results of the mixed model[173]

Fixed factors	P value
Boundedness	0.968
Narrativity	<.0001 ***
Perfectivity	0.004 **
Boundedness:Narrativity	0.04 *

Moreover, perfective viewpoint is negatively correlated with the IMP verbal tense whereas narrative usages of the SP are positively correlated with the IMP. Moreover, bounded situations in non-narrative contexts are also negatively correlated with the IMP. This interaction is seen in Figure 6-11. This model's predictive force in new data is of 0.83.

---

[173] The number of * signals the level of significance: *** highly significant, ** very significant, * significant.

Figure 6-11 Interaction Boundedness*Narrativity
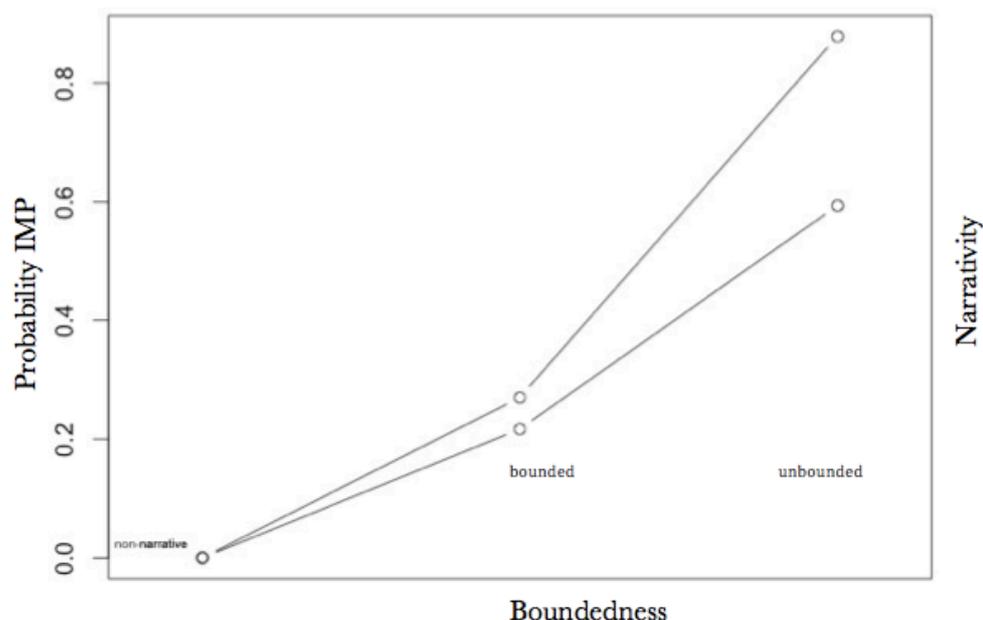
The results of the multifactorial analyses described in this section point to the cross-linguistic correlations between contextual usages of a verbal tense in a SL and to the corresponding verbal tenses used in a TL. A mixed model fitting the data indicated that there are three significative factors for predicting the verbal tense in TL. Specifically, the FR IMP can be predicted based on the procedural feature encoded by Tense, operationalized as the [±narrativity] feature, the procedural feature encoded by Aspect, operationalized as the [±perfectivity] feature and, thirdly, the interrelation between the procedural feature [±narrativity] that constrains the interpretation of conceptual information encoded by Aktionsart, operationalized as the [±boundedness] feature. My suggestion is that humans treat temporal information coming from these three sources in a coherent manner. Specifically, this linguistic data points to *discourse temporal coherence* established at the level of the discourse. With respect to the addressee's cognitive faculties involved in the interpretation process, my suggestion is that the human brain treats this temporal interpretation in a coherent manner, and therefore, one can speak about *cognitive temporal coherence* (see section 7.3 for a more detailed discussion).

The predictive force in new data of the model of 0.83 illustrates that when dealing with human language there is a part of variability, which cannot be predicted nor modelized[174]. This part may be explained by the speaker's personal choices, as well as the translator's personal choices. As for the variability that could be predicted, some specifications could be made. Four fixed factors and two random factors (i.e. stylistic register and the verb itself) have been considered in this mixed model. Other factors that might be studied are the conceptual difference between past and non-past, the speaker's subjective viewpoint (see section 7.2.2.2) and the usage of English progressive. The first was not included in this model

---

[174] I would like to thank Andrei Popescu-Beliş (IDIAP Research Institute, Martigny, Switzerland) for pointing out at the defence of this thesis (2015) the expectation to have a deterministic linguistic model and for suggesting that there might be a part of variability due to speaker's personal preference regarding, for example, the choice between a PC or a PS.

because all the verbal tenses from the target language are past time verbal tenses whereas the second could not, at least for the moment, be operationalized in order to have reliable experimental data. Finally, the third factor should be considered in future research since, according to the literature, it is closely linked to a translation through the IMP. At this point of the discussion, I would like to note that previous researches in NLP and MT pointed out that too fine-grained, as well as too general features, do not provide good results for the automatic treatment of the language. This model includes medium-coarse grained features which proved to be useful for the task.

In the remaining of this chapter, I will address the [±narrativity] feature and its empirical validation for IT and RO, as well as the replication of the findings for FR (section 6.2). Finally, I will describe NLP and MT experiments carried out on the human annotated bilingual data described above in this chapter (see section 6.3).

6.2 Multilingual Data

6.2.1 Motivation

In section 6.1, I described the motivation of the experiments carried out with bilingual EN and FR data. Principally, these experiments aimed at testing empirically numerous theoretical assumptions existent in the literature with respect to FR verbal tenses expressing past time, grammatical and lexical aspect and their interactions for determining temporal reference at the discursive level. For the bilingual data, the following features were tested:

- The [±narrativity] feature, which operationalizes the addressee's hypotheses with respect to temporal and causal relations holding among eventualities;
- The [±boundedness] feature, which operationalizes the addressee's hypotheses about the inherent temporal structure of an eventuality;
- The [±perfectivity] feature, which operationalizes the addressee's hypotheses about the completion/entirety of an eventuality;

The results of the experiments showed that the first two features are identifiable by native speakers both for EN and FR verbal tenses. These two features are however processed differently by hearers. Firstly, the two judges found challenging to deal directly with the [±narrativity] feature (Kappa values around 0.42 in the first round). If a temporal connective that explicitates this information is provided, judges have better results when judging the feature (Kappa values around 0.95). This result was interpreted as pointing to the procedural nature of the [±narrativity] feature, which is expected to be accessed with difficulty through conscious thinking. However, if it is accessed via a connective, the task is easier and the agreement rate is very high.

Secondly, the two judges dealing with the [±boundedness] feature had very good results in terms of inter-annotator agreement rate (Kappa value of 0.84 in the fist round and of 1 in the second round). This result was interpreted as pointing to the conceptual nature of the [±boundedness] feature, which is expected to be easily accessed through conscious thinking.

The results of the experiment dealing with the [±perfectivity] feature showed that native speakers of EN were not able to identify it when judging tokens of the EN preterit (Kappa

value of 0.32). This result was interpreted as pointing to the procedural nature of the [±perfectivity] feature, which is expected to be accessed with difficulty through conscious thinking. In order to have annotated data with this feature, the cross-linguistic transfer of properties method was used.

Given these results for bilingual data, the [±narrativity] feature was tested experimentally in other Romance languages, namely IT and RO. In order to have comparable data, a sample of 85 items containing SP tokens in EN and a PS, PC or IMP in TL were randomly chosen from the parallel translations corpus (i.e. corpus where containing texts originally written in EN were translated into FR, IT and RO by professional translators; see section 5.3). This data was used in the experiments with linguistic judgement task described below.

6.2.2 Experiment 8: Italian PC, PS, IMP and the [±narrativity] feature

Experiment 8 was designed to test the [±narrativity] feature in IT. According to theoretical assumptions, the following hypotheses were formulated:

- The PS has more often narrative usages than non-narrative usages
- The PC has more often narrative usages than non-narrative usages
- The IMP has more often non-narrative usages than narrative usages

Participants

There were two participants, IT native speakers originating from the southern part of Italy (Naples). Their participation at the experiment was voluntary and benevolent.

Procedure and material

84 items containing 37 PC, 27 PS and 21 IMP were randomly chosen from the IT part of the multilingual translations. These items were originally written in EN and the targeted IT verbal tense corresponds to a SP in the source language. Judges received annotation guidelines and went through a training phase. The first task from the annotation guidelines was to read and understand the instructions containing definitions of narrative and non-narrative usages (as given in Experiment 1 for FR and 3 for EN in section 6.1). They also included two examples for each usage, as given in the (522)-(524), where (522) is an example of non-narrative usage whereas (523) and (524) are examples of narrative usages.

(522) V'erano porte tutt'intorno alla sala, ma *erano* [IMP] tutte serrate. (Literature Corpus)
'There were doors all 'round the hall, but they *were* all locked.'

(523) Ma, risalito dopo pranzo con tale proposito, appena varcata la soglia, *scorsi* [PS] lì dentro una ragazza che, inginocchiata davanti al fuoco e circondata da scope esecchi di carbone. (Literature Corpus)
'On coming up from dinner, however, on mounting the stairs with this lazy intention, and stepping into the room, I saw a servant-girl on her knees surrounded by brushes and coal-scuttles'.

(524) Malgrado le misure di controllo adottate dalle autorita delle isole Faroe, nel 2004 *sono stati segnalati* [PC] alla Commissione nuovi focolai della malattia. (EuroParl Corpus)

> 'Despite the control measures undertaken by the Faroe Islands, further outbreaks of ISA occurred and were notified by that State to the Commission in 2004.'

The second task was to read each item and decide if the verb highlighted has a narrative or a non-narrative usage. Participants were trained on 6 items followed by a discussion.

Results

Judges agreed on 64 items (76%) and disagreed on 21 items (33%). The value of the Kappa coefficient was 0.41. Disagreements were discussed in a second round of the experiment. The final results are provided in Table 6-17. Judges agreed on 76 items (89%), which represents a Kappa value of 0.74. This value signals reliable data.

Table 6-17 Narrativity for IT verbal tenses: Judge 1 vs. Judge 2

| | | Judge 2 | | Total |
		Narrative	Non-narrative	
Judge 1	Narrative	55	4	59
	Non-narrative	5	21	26
	Total	60	25	85

As far as the analysis of individual verbal tenses is concerned, the data containing only agreements was considered (76 items). Specifically, 16 IMP were judged as non-narrative (84%), 30 PC were judged as narrative (88%) and 22 PS were judged as narrative (96%).

Table 6-18 Narrativity for IT PS, PC and IMP

	Narrative	Non-narrative	Total
IMP	3	16	19
PC	30	4	34
PS	22	1	23
Total	55	21	76

The results of this experiment indicate that the [±narrativity] feature is identifiable by native speakers with reliable Kappa values. This information receives narrative values most frequent for the PS and the PC, and non-narrative values for the IMP. Similarly to EN and FR speakers, IT speakers have poor abilities to consciously evaluate temporal relations triggered by verbal tenses. They have betters results when they are asked to insert connectives, which express explicitly the same implicit content. These findings are solid empirical basis for arguing that the [±narrativity] feature is procedural and that it is a cross-linguistically valid feature.

6.2.3 Experiment 9: Romanian PC, PS, IMP and the [±narrativity] feature

Experiment 9 was designed to test the [±narrativity] feature in RO. According to theoretical assumptions, the following hypotheses were formulated:
- The PS has more often narrative usages than non-narrative usages
- The PC has more often narrative usages than non-narrative usages
- The IMP has more often non-narrative usages than narrative usage

There were two participants, RO native speakers. One of the judges is a research peer and the other is a Bachelor student from University of Geneva, Faculty of Letters. Their participation at the experiment was benevolent.

Procedure

85 items containing 50 PC, 14 PS and 21 IMP were randomly chosen from the RO part of the multilingual translations corpus. These items were originally written in EN and the targeted RO verbal tense corresponds to a SP in the source language. Judges received annotation guidelines and went through a training phase. The first task from the annotation guidelines was to read and understand the instructions containing definitions of narrative and non-narrative usages (as given in Experiment 1 on IT data, see section 6.2.2). They also included two examples for each usage, as given in (525)-(527), where (525) is an example of non-narrative usage whereas (526) and (527) are examples of narrative usages.

(525) Erau uşi de jur împrejurul holului dar toate *erau* [IMP] încuiate. (Literature Corpus)
'There were doors all 'round the hall, but they were all locked.'

(526) Aşa că, întorcându-mă de la masă, urcai scările cu intenţia de a-mi petrece după-amiaza lenevind. Când să intru în odaia mea, *văzui* [PS] o tânără servitoare, îngenuncheată lângă sobă, înconjurată de perii şi găleţi cu cărbuni. (Literature Corpus)
'On coming up from dinner, however, on mounting the stairs with this lazy intention, and stepping into the room, I saw a servant-girl on her knees surrounded by brushes and coal-scuttles'.

(527) Cu toate că autorităţile din insulele Feroe au pus în aplicare măsuri de combatere au apărut alte focare de AIS, care *au fost notificate* [PC]Comisiei de această ţară în 2004. (EuroParl Corpus)
'Despite the control measures undertaken by the Faroe Islands, further outbreaks of ISA occurred and were notified by that State to the Commission in 2004.'

The second task was to read each item and decide if the verb highlighted has a narrative or a non-narrative usage. Participants were trained on 6 items followed by a discussion.

Results

The results are provided in Table 6-19. Judges agreed on 64 items (75%) and disagreed on 21 items (25%). The value of Kappa coefficient was 0.42[175].

---

[175] This experiment was carried out in two rounds. 42 items were judged in the first round and 43 items in the second round. Due to the unfortunate lack of availability of the two judges, only the data from the first round was judges a second time for resolving the disagreements. For the first 42 items, the Kappa value improved from 0.23 (agreement in 62% of the cases) to 0.75 (agreement in 88% of the cases). The results provided in Table 6-19 represent the data obtained after the second round with the first 42 items and the unique round with the other 43 items. The low Kappa value of the entire data is due to the fact that disagreements in 43 items were not resolved.

Table 6-19 Narrativity for RO verbal tenses: Judge 1 vs. Judge 2

		Judge 2		Total
		Narrative	Non-narrative	
Judge 1	Narrative	47	0	47
	Non-narrative	21	17	38
	Total	68	17	85

As far as the analysis of individual verbal tenses is concerned, the data containing only agreements was considered (64 items). Specifically, the IMP was judged as non-narrative in 10 cases (71%), the PC was judged as narrative in 30 cases (83%) and the PS was judged as narrative in 13 cases (93%).

Table 6-20 Narrativity for RO PS, PC and IMP

	Narrative	Non-narrative	Total
IMP	4	10	14
PC	30	6	36
PS	13	1	14
Total	47	17	64

This experiment shows that, similarly to IT, that the [±narrativity] feature is identifiable by RO native speakers with reliable Kappa values. This information receives most frequently narrative values for the PS, PC and non-narrative values for the IMP. Moreover, native RO speakers have poor abilities to consciously evaluate temporal relations triggered by verbal tenses. They have betters results when they are asked to insert connectives, which express explicitly the same implicit content. These findings confirm the proposition that the [± narrativity] feature is procedural and that it is a cross-linguistically valid feature.

6.2.4 Experiment 10: French PC, PS, IMP and the [±narrativity] feature

Experiment 10 was carried out to replicate the Experiment 1 from section 6.1 and to test the [±narrativity] feature on FR data originating from the multilingual parallel translations corpus. The same hypotheses were formulated as in Experiment 1 from section 6.1:

- The PS has more often narrative usages than non-narrative usages
- The PC has more often narrative usages than non-narrative usages
- The IMP has more often non-narrative usages than narrative usage

Participants

There were 49 participants, FR native speakers, who were first year students of Linguistics at University of Geneva[176]. Their participation at the experiment was organized during a Linguistics class. It was however benevolent and anonymous.

Procedure and material

84 items containing 18 PS, 36 PC and 30 IMP were randomly chosen from the FR part

---

[176] There was a lapse of time of two years between the first experiment on FR (described in section 6.1.1 and this experiment. There were therefore different participants.

of the multilingual translations corpus. The data was organised in 7 sets of 12 items. Several participants judged each set, more precisely, 4 sets were judged by 6 participants and 3 sets by 7 participants[177]. The evaluation of each item was done through calculating the majority of judgments[178]. Judges received annotation guidelines and went through a training phase. Annotation guidelines consisted of the definition of non-narrative as in (528) and of narrative usages as in (529) and (530).

(528)   Il y avait plusieurs portes autour de la salle, mais elles *étaient* [IMP] toutes fermées à clé. (Literature Corpus)

         'There were doors all 'round the hall, but they were all locked.'

(529)   [...] Je gravis donc l'escalier dans cette intention paresseuse; mais, en entrant dans la pièce, je *vis* [PS] une servante à genoux, entourée de brosses et de seaux à charbon. (Literature Corpus)

         'On coming up from dinner, however, on mounting the stairs with this lazy intention, and stepping into the room, I saw a servant-girl on her knees surrounded by brushes and coal-scuttles'.

(530)   Malgré les mesures de lutte mises en œuvre par les autorités des îles Féroé, d'autres foyers d'AIS sont apparus et *ont été notifiés* [PC] par ce pays à la Commission en 2004. (EuroParl Corpus)

         'Despite the control measures undertaken by the Faroe Islands, further outbreaks of ISA occurred and were notified by that State to the Commission in 2004.'

The second task was to read each item and decide if the verb highlighted has a narrative or a non-narrative usage. Participants were trained on 6 items followed by a discussion.

Results

One item was not judged for unknown reasons. It was therefore deleted from the corpus. Among the 83 items annotated by 6 judges, 10 items received equal answers (no majority). To avoid disregarding them and reducing the data, for 8 of these items the judgment coming from the 7th judge was considered in order to have a majority of answers. These other 2 inconclusive items were not considered in further analysis. In the clean data of 81 items, the label given by the judges correlates with the verbal tense used in the corpus for 59 items (73% of the data).

Table 6-22 shows that the narrative feature was identified for 69% of the annotated tokens according to the theoretical predictions (39 items labelled as narrative out of 52 existent in the corpus) and the non-narrative feature for 73% of the cases (20 items labelled as non-narrative out of 29 existent in the corpus of 81 items). This agreement rate

---

[177] This was due to the number of students present in the linguistics class the day when the annotation experiment was carried out, i.e. 45 students.

[178] Similarly to Experiment 1 and 7, the evaluation procedure is the following. The number of concordant answers must be superior to agreement by chance, which is at 50% since there is a binary choice (i.e. past vs. non-past context). When the situation was equal (3 over 6 judges), than the item was evaluated as inconclusive. Inconclusive items were disregarded from further analysis. Finally, for an item, if less than 50% of the judges (maximum 2 over 6 judges) made the same judgment than it was considered as a disagreement. Due to the reduced number of participants who saw each item, that is 6 per item, the evaluation was done manually.

corresponds to a Kappa value of 0.42.

Table 6-21: Majority of judges and Reference

		Majority of judges		
		Narrative	Non-narrative	Total
Reference	Narrative	39	13	52
	Non-narrative	9	20	29
	Total	48	33	81

More precisely, judges clearly recognized a primary narrative usage for the PS (94%) but did not make the same clear judgment for the PC (65%) or for the IMP (69%) as being primarily non-narrative, as shown in Table 6-2.

Table 6-22 Narrativity for PS, PC and IMP: Majority of judges and Reference

		Majority of judges		
		Narrative	Non-narrative	Total
Reference	IMP	9	20	29
	PC	22	12	34
	PS	17	1	18
	Total	48	33	81

These results replicate Experiment 1 carried on bilingual data.

6.3   Application to NLP and MT

Human annotated texts serve as training data for machine-learning tools. One of the purposes of this research was to improve the results of a statistical machine translation (SMT) system in what concerns the translation of verbal tenses. Current SMT systems have difficulties in choosing the correct verb tense translations because these depend on a wider-range context than SMT systems consider. SMT systems aiming at modelling intersentential relations, such as temporal information conveyed by Tense, Aspect and Aktionsart, require large amounts of annotated corpora with semantic and pragmatic information to be used in the training phase of the statistical system.

In order to have large amounts of annotated data, one can either produce it manually or automatically. Unfortunately, manual annotation of large amounts of data is time consuming and very expensive. For these reasons, manual annotation is usually made on smaller amounts of data. As for the automatic annotation, one can choose to use existing automatic tools dealing with temporal information in the discourse such as the TimeML markup language or to build a *classifier* (as discussed in section 3.3.2.1). A classifier trains on the small amount of annotated data and learns through machine-learning algorithms the annotation scheme. The classifier is used thereafter to annotate large amounts of data, necessary to the SMT system.

One issue that is worth pointing out at this point of the discussion regarding the type of data regards the trade-off between using small quantity of accurate data (generally human annotated or human post-edited) on the one hand, and using a large quantity of imperfect data on the other hand. Large quantities of imperfect data can be used in what is called *on-*

line and *unsupervised learning* (i.e. the system learns all the patterns emerging from the data) and they are very useful in binary classifications for unambiguous cases. However, for ambiguous (and also underspecified) cases, which are difficult to classify, using large quantities of imperfect data remain limited in their usefulness. For these cases, in general, human intervention is required in order to have an accurate judgement. Hence, small accurate quantities of data are necessary especially for the classification of difficult cases and are used in what is called *supervised* learning. The choice of one of the two types of data depends on the task and they are used as complementary methods.

In the COMTIS and MODERN projects two classifiers were built in order to annotate automatically data with labels learnt from human annotated texts[179]. The first classifier annotates automatically texts with the [±narrativity] feature (work published in Grisot and Meyer 2014; Meyer 2014). The second classifier deals with the [±boundedness] feature (work published in Loáiciga and Grisot 2015). Human annotation experiments with these two features were described in section 6.1. I will describe the automatic annotation experiments in section 6.3.1. Several SMT systems were built which were trained on the data annotated by the two classifiers. The results of the MT experiments provided in section 6.3.2 showed that SMT systems that are aware of the linguistic information provided through annotation experiments (i.e. information about temporal ordering of eventualities and about lexical aspect) translate more accurately verbal tenses and have a better lexical choice (Meyer et al. 2013; Loáiciga and Grisot 2015).

6.3.1 Automatic annotation experiments

The data used in the automatic annotation experiments consists of 435 items of EN SP, initially used in the human judgement experiments described in section 6.1.3 with the [±narrativity] feature and in section 6.1.4 with the [±boundedness] feature. A classifier was built for each of these features and trained on the human annotated data. For each classifier, a series of surface features were considered.

Annotation of the [±narrativity] feature

The training data contained 257 narrative and 178 non-narrative EN SP items (a total of 435). The performance of the classifier was tested on a smaller and earlier manually annotated sub-portion of the corpus with the same stylistic genre distribution, consisting of 118 items of EN SP: 75 instances of narrative and 43 of non-narrative. Surface features were obtained from syntactic parsing and temporal analysis of the text with the TimeML markup language. The surface features used were the following:

- Neighbouring verb word forms.
- The position of the verbal tense in the sentence.
- The POS tags of all the words in the sentence.
- The syntactic tree structure of the sentence.
- Temporal markers (such as *while, since, weeks/day after or before, subsequently, repeatedly* and

---

[179] The NLP and MT work described in this section was done by my colleagues Thomas Meyer, Sharid Loáiciga and Andrei-Popescu Belis, in the COMTIS project.

the like) from a hand-made list of 66 temporal discourse markers.

- Type of temporal markers (from TimeML) such as temporal *simultaneity* or *sequencing* for temporal markers, *infinite, participle, future* for the class of verbal tense and *perfective* or *imperfective* for the grammatical aspect).

With these features, a MaxEnt classifier performs at 0.72 F1 score[180]. Out of the 118 test instances, the classifier correctly annotates 90 items corresponding to 76.3%. Moreover, the Kappa value for the agreement between the classifier and the reference is 0.46. The results of the classifier are reliable enough to automatically label the SP verbal tenses in the EN side of a large parallel corpus so that it can be used for training machine translation systems.

Annotation of the [±boundedness] feature

The training data contained 236 bounded and 199 unbounded instances (a total of 435). As for the testing of the performance of the classifier, in the absence of a previously annotated corpus, a *cross-validation* approach was used. The data was automatically split in several equal sub-parts (10 in this case, therefore a 10-fold cross-validation). The classifier trains iteratively on 9 parts and its performance is tested on the 10th part. At the end the classifier's performance is calculated as the average of the results it had for each of the 10 folds.

For this experiment, several features resulting from the human judgement experiments described in section 6.1 and from human editing of the data were used. There were two classes of features: syntactic and temporal features. The syntactic features are the following ones:

- SP token: SP forms were manually identified in order to eliminate parsing errors, such as when past participle forms were tagged as an SP or, for passive forms, only the auxiliary was labelled.
- Infinitive form of the verb.
- Grammatical aspect originating from the translation of the corpus into Serbian, see Experiment 5 in section 6.1.5.
- FR verbal tense identified in the FR part of the translation corpus.

The temporal features are the following:

- The same temporal markers as those used in the annotation experiment with the [±narrativity] feature.
- The [±narrativity] feature itself.

With these features, the classifier performs at 0.89 F1 score for the bounded class and 0.86 for the unbounded class. These scores signal very good performances of the classifier. The most informative features in a descending order are: grammatical aspect, verbal tense used in FR, narrativity and the infinitive form of the verb in EN. Among these features, grammatical aspect and narrativity (as well as boundedness with respect to its interaction

---

[180] The metrics used in computational linguistics for evaluating classification results are: *accuracy* (percentage of correctly classified instances), *precision* (percentage of correctly classified instances among correctly identified ones) and *recall* scores (percentage of correctly classified instances over all instances) (Meyer 2014, 50). Precision and recall correspond to Type I and Type II errors in statistics. Precision and recall are used (their harmonic mean) to determine the F1 score, which ranges from 0 (worst score) to 1 (best score).

with narrativity) turned out to be significant also in the mixed model fitted in order to predict the verbal tense used in TL (see section 6.1.8).

In a second experiment, certain surface features were generated automatically, such as the SP token, the infinitive form, the position in the sentence, the POS tags of the verbs and the POS tags of their arguments (the verb phrase)[181]. Three features originating from the human annotated data were not used in this experiment: narrativity, grammatical aspect and FR verbal tense.

Since human annotated data is costly and time-consuming, this second experiment aimed at testing whether the classifier has reliable results if it is trained only on automatically extracted surface features, which might have errors. With these features, the classifier performs at 0.82 F1 score for the bounded class and 0.78 for the unbounded class. Compared to the first experiment, these scores still represent reliable results and they show that both the bounded and the unbounded category is more difficult to predict solely based on automatically generated features. The difference of approximately 8% for each category between the results of the two experiments is shown to be statistically significant, $p<0.05$, according to a paired t-test. This result can be interpreted in terms of the quality of human annotated data compared to automatically generated data, which contains a percentage of errors.

6.3.2 Machine translation experiments

The classifiers presented above were built with the purpose of annotating automatically large amounts of data, necessary for the training of SMT systems. In what follows, I will describe machine translation experiments performed with SMT systems aware of the [±narrativity] feature (Meyer et al. 2013) and the [±boundedness] feature (Loáiciga and Grisot 2015).

MT experiments with the [±narrativity] feature

One main question that arose at this point of the research was how to provide the linguistic information conveyed in the labels given by the classifier to a SMT system. Two methods were tested[182]:

- Concatenation of the label with the SP verb form considered as a new word to be translated, as in example (532) containing an input sentence for the SMT system in which the concatenation is shown by the '-' symbol.
- Use factored translation models, which allow for any linguistic annotation to be considered as an additional feature, next to the basic features of the phrase-based models as in example (533) containing an input sentence for the SMT system, in which the factorisation is shown by the '|' symbol.

To evaluate the gain that the [±narrativity] feature brings to the quality of the translation made by an SMT system, three systems were built using a 5-gram language model. The first

---

[181] For technical details regarding the exact description on syntactic and temporal features, see Loáiciga and Grisot (2015).

[182] For technical details regarding the building of the SMT systems, data using for training, tuning and testing, as well as evaluation metrics, see Meyer et al. (2013) and Meyer (2014, chapter 7).

one, called baseline system, is a statistical system trained on plain text input, without verbal labels, as in (531). The second, called tagged system, is a statistical system using a phrase-based translation model and trained on plain text input containing narrativity labels concatenated on the verb, as in (532). The third one, called factored system, is a statistical system using a factored translation model and trained on texts, where each SP has a narrativity label whereas all the other words have a |Null label, as in (533) where the Null labels were omitted for readability.

(531) Baseline SMT: On Wednesday the čssd declared the approval of the next year's budget to be a success. The people's party was also satisfied.

(532) Tagged SMT: On Wednesday the čssd declared-*Narrative* the approval of the next year's budget to be a success. The people's party was-*Non-narrative* also satisfied.

(533) Factored SMT: On Wednesday the čssd declared|*Narrative* the approval of the next year's budget to be a success. The people's party was|*Non-narrative* also satisfied.

For the labelling of the data with the [±narrativity] feature, the classifier described in section 6.3.1 was used to annotate the EuroParl corpus (Koehn 2005) containing 321,577 sentences with 66,143 instances of SP identified by the POS tagger used. The classifier labelled 30,452 narrative SPs and 35,691 non-narrative SPs.

The results of the three SMT systems were evaluated using two measures: BLEU and TER[183]. Table 6-23 provides the results of the evaluation of the SMT systems in terms of BLEU and TER scores. The factored model improves performance over the baseline by +0.2 BLEU and -0.2 TER (since smaller scores represent better translation), and these differences are shown to be statistically significant, $p<0.05$ according to a t-test (signalled by the * in the table). Meyer et al. (2013) explain that the lower scores of the tagged model may be due to the sparsity of the data, i.e. verbal forms were altered by concatenation with the narrativity label. As for the small improvement of the factored model, it can be explained by the fact that narrativity feature improved the translation of the verbal tense solely and that the translation of the other words in the sentence is unchanged compared to the baseline.

Table 6-23 Evaluation of SMT systems aware of temporal information

Translation model	BLEU	TER
Baseline	21.4	61.9
Tagged	21.3	61.8
Factored	21.6*	61.7*

A human evaluation of the performance of baseline and factored systems was also

---

[183] The BLEU score (Bilingual Evaluation Understudy; Papineni et al. 2002) counts the overlap in terms of matching number of words and n-grams between the candidate translation and one or more reference translations. The more matches there are for 4-, 3-, 2- and 1-grams in a candidate translation compared to its reference, the higher the BLEU score. The values of the score range from 0 to 100, which is reached for identical translations. Most often, the existent SMT systems have scores between 11-33 BLEU points. BLEU is appreciated to be the metric that approaches most of human judgments of translation quality, especially when averaged over a large quantity of text. TER (Translation Error Rate; (Snover et al. 2006) computes the number of edits (called edit-distance) required to transform a candidate translation into one of its reference. The smaller the edit-distance is, the lower the score and therefore, the better the translation is (Meyer 2014, 50).

performed on the 207 first instances of SP. Bilingual evaluators (EN and FR) scored the translation by looking at the source sentence and its reference translation from the parallel corpus. The scoring targeted the following criteria: the correctness of the narrativity label, the improvement of the lexical choice, of the choice of verbal tense and of the VP compared to the baseline system. Human evaluation revealed that the narrativity feature helped the factored system to generate more accurate FR verbal tenses in 10% of the cases and to have better lexical choices for verbs in 3.4% of the cases. For example, the input EN sentence in (534) was translated as in (535) by the baseline system and as in (536) by the factored system. The SP *looked* is translated by the baseline system as *considérés* (infelicitous lexical choice, past participle form, wrong number agreement) whereas by the factored system it is translated as *semblait* (better lexical choice, IMP verbal tense, correct agreement in number).

(534)   Tawa hallae looked|*Non-narrative* like many other carnivorous dinosaurs.

(535)   Tawa hallae *considérés* comme de nombreuses autres carnivores dinosaures.

(536)   Tawa hallae *semblait* comme de nombreux autres carnivores dinosaures.

Another issue identified through the human evaluation process concerns the cases where the factored model performed worse than the baseline system. Some of these cases are due to errors in the POS tagging used to find the SP instances to label. For example, for passive forms of the verb, only the auxiliary was tagged and this introduced noise and errors in the automatic annotation process. Moreover, the factored translation model seems to operate at the local level despite the pragmatic nature of the [±narrativity] feature. Meyer et al. (2013) suggest that for widening the context captured by the translation model, a possibility would be to label the entire VP in hierarchical or tree-based syntactical models. Overall, the factored system compared to the baseline system produces better translations of the SP VP in 9% of the cases.

The improvement of translation presented here is important because it points out that adding pragmatic knowledge regarding temporal relations holding among eventualities is useful. This value is dependent on the classifier's performance, which produces reliable but not perfect results (70% correctly labelled SPs). This aspect should be considered in future research.

MT experiments with the [±boundedness] feature

Another series of experiments performed targeted the [±boundedness] feature. Since the factored model proved to be the most accurate one, in these experiments only two systems were built: a baseline and a factored system using a 5-gram language model. The systems were trained on the EuroParl corpus, containing 321,577 sentences with 73,081 instances of SP identified by the POS tagger. The corpus was automatically annotated with the [±boundedness] feature by the classifier described in section 6.3.1. For example, an input sentence, as in example (537), is labelled as follows: the verb receives a *unbounded* label (which is one of the two values of the boundedness feature) whereas all other words from the sentence receive a Null label (Loáiciga and Grisot 2015).

(537)   Max ran for an hour.

As in the case of the narrativity classifier, a factor does not determine entirely the translation of a verb, i.e. there is no exact correspondence between a label and a verbal tense in FR. For instance, a *bounded* label does not necessarily lead to a translation into FR through a PC. Instead, factors are considered when estimating the translation probabilities computed over the entire parallel corpus.

The performances of the two translation systems were evaluated with the BLEU measure. The results provided in Table 6-24 indicate that the factored system using the lexical aspect labels obtained an increase of 0.12 points, reflecting a small improvement in the overall quality of translation. This score may be further analyzed in the following way. The output of lexical aspect aware translation system was compared to the output of the baseline SMT on the translation of the same 300 sentences. For each sentence, a BLEU score was computed. The analysis of the 300 BLEU scores showed that in 35% of the sentences, the BLEU scores of the aspect-aware system are highest than the scores of the baseline system. For the remaining of the 65% of the sentences, the scores of the two systems are equal.

Table 6-24 Evaluation of SMT systems aware of lexical aspect

Translation model	BLEU
Baseline	20.55
Factored	20.67

A human evaluation of the performances of the two systems and of the performance of the classifier was performed on 156 randomly selected instances of SP. The classifier correctly identified the SP instances in 90% of the cases and annotated them correctly as *bounded* or *unbounded* situations in 74% of the cases. For the factored system, compared to the baseline system, human evaluation indicated a better translation of SP instances into FR in 23% of the cases, a similar translation in 76% of the cases and an inferior translation in only 5% of the cases. The cases of similar translation can be explained by the fact that the baseline system had itself a good performance since it provides PC labels by default and that the distribution of the verbal tenses used in the translation into the TL is highly skewed in favour of the PC. Therefore the improved cases correspond to cases where an IMP was used in the reference and the aspect-aware system translated correctly an SP through an IMP.

For example, the input EN sentence in (538) was translated as in (539) by the baseline system, as in (540) by the factored system aware of lexical aspect and as in (541) by a professional translator, the reference translation coming from the parallel corpus.

(538) The vice-chairman of the ODS, Petr Nečas *said* that the concept of an interim government supported by the ČSSD, ODS, and Green Party, *was* evidently no longer working.

(539) Le vice-président, de l'ODS Petr Nečas, *dit* que le concept d'un gouvernement intérimaire soutenu par les ČSSD, ODS, et parti vert, *a* apparemment aucune *fonctionne* plus.

(540) Le vice-président, de l'ODS Petr Nečas, *a déclaré* que le concept d'un gouvernement intérimaire soutenu par les ČSSD, ODS et aux verts, *était* manifestement, de ne plus travailler.

(541) Le porte-parole de l'ODS Petr Nečas *a déclaré* que l'idée d'un cabinet administratif soutenu par le ČSSD, l'ODS et le Parti des verts ne *fonctionnait* manifestement plus.

The first SP instance, *said*, was labelled by the classifier as *bounded* and the second SP instance, *was*, as *unbounded*. Both verbal tenses were translated with a PRES tense by the baseline system. The factored model, instead, produced the same verbal tenses as the reference, PC for the former SP and IMP for the latter SP.

Overall, the factored system compared to the baseline produces better translations. An improvement can also be observed if the two factored systems (i.e. one aware of temporal information and the other one aware of lexical aspect) are compared. The aspect-aware SMT system produced better translation than the narrativity-aware SMT system. This is mainly due to the higher performance of the classifier producing boundedness labels than the classifier producing narrativity labels. The second reason is a better identification of correct instances of SP. This was due to using the POS tagger improved with a series of rules (for more technical details, see Loáiciga and Grisot 2015).

To conclude, I would like to point to the importance of the granularity of the linguistic features. In order to be useable, linguistic features must be medium-coarse grained. In other words, too fine-grained features are either are not sufficient for explaining the variation in the data or they are not implementable. For example, the mixed statistical model fitted on the manually annotated corpus of 435 sentences (see section 6.1.8), shows that the FR verbal tense in TL is significantly determined by the interaction between the narrativity status and the lexical aspect of EN verbs. This theoretical insight is unfortunately very difficult to model in NLP and to apply for SMT. This is an important issue to be investigated in further research.

The medium-coarse grained features proposed, i.e. [±narrativity] and [±boundedness], are not only successfully implementable but also their implementation in NLP and application to MT produced significant improvements of the results of the automatic systems. Thus, these ameliorations represent an empirical indirect but solid validation of the theoretical model proposed.

6.4 Conclusive remarks

This chapter described the empirical work carried out for investigating temporal reference and the main linguistic sources contributing to its expression and processing in the discourse.

Corpus work revealed the absence of one-to-one cross-linguistic correspondences of verbal tenses. One of the identified divergences is the EN SP translation divergence. The analyses of bilingual and multilingual parallel corpora showed that the four verbal tenses most frequently used in FR, IT and RO for translating the SP are three verbal tenses expressing past time (PC, PS and IMP) and the present tense (PRES). The question that arose regarded the linguistic and non-linguistic factors that explain this cross-linguistic variation.

Experimental work revealed that when interpreting a text, hearers take into consideration temporal information originating from several sources and treat them as a coherent whole.

According to the existing rich literature, several possible factors were defined and tested in the experimental work, i.e. off-line experiments with linguistic judgement task, namely:

- Temporal location of eventualities with respect to the moment of speech S;
- Temporal relations existing among eventualities, which can either be implicit or be expressed explicitly through temporal connectives, were operationalized as the [±narrativity] feature;
- Aspectual viewpoint on the eventuality operationalized as the [±perfectivity] feature;
- Temporal information inherent to the eventuality (i.e. lexical aspect of the VP, for which temporal adverbials have a significant contribution) operationalized as the [±boundedness] feature;

Several conclusions were drawn from the analysis of results of the experimental work, as described briefly below and developed in Chapter 7.

Firstly, temporal location of eventualities with respect to S is conceptual information encoded by Tense. Verbal tenses can therefore be classified as locating E<S (i.e. pastness) or E≥S (non-pastness).

Secondly, temporal relations existing among eventualities is procedural information encoded by Tense, which is the functional head of the sentence. It's a binary feature and its value is determined pragmatically in every context based on linguistic and non-linguistic factors.

Thirdly, the speaker's viewpoint on the eventuality is procedural information encoded by Aspect. It is a binary feature and its value is either morphologically expressed or determined pragmatically.

Finally, temporal information inherent to the VP is conceptual information and its value is computed contextually based on the lexical aspect of the verb and other factors that influence it, such as temporal adverbials, countable/uncountable NPs and grammatical aspect.

A mixed model fitted on the data showed that information about temporal relations holding among eventualities, aspectual viewpoint and lexical aspect are components of the temporal interpretation of a coherent discourse. The value of each component can be predicted based on the values of the others. Specifically, these three factors were used to predict the verbal tense used in the TL side of a translation corpus. Finally, the data was used for improving the results of SMT systems.

In the following chapter, I will propose a theoretical reanalysis of temporal reference and of its ingredients based on the results of the experimental work carried out in this research. The main hypothesis is that temporal reference is expressed through Tense, Aspect and Aktionsart. As for the role played by temporal connectives and adverbials, they have been considered indirectly in this study. Precisely, temporal information given by connectives such as *then, before, in the same time* was modelled in the narrativity feature in the sense that judges were asked to propose a connective that would render explicit the implicit temporal relation existing among the eventualities expressed. If eventualities were not temporally related or they occurred simultaneously, then no connective would be inserted and respectively, the *at the same time* discourse marker would be provided. The temporal information given by temporal adverbials such as *yesterday, in 1978* and *two weeks ago* was partly modelled in the boundedness feature since they can change the aspectual class of the VP.

7 Temporal reference in discourse: a reanalysis

7.1 Introduction

Chapter 2 described semantic approaches to temporal reference and its ingredients in discourse. Referential approaches to Tense claimed that it plays a central role for establishing temporal reference at the discursive level. However, in these approaches the generic notion of *verbal tense* was used to refer also to the contributions to temporal reference offered by Aspect. The various interrelations between verbal tenses and aspectual classes have been pointed out by formal discourse semantic theories, such as Kamp and Reyle (1993), Kamp and Rohrer (1983), Dowty (1986), ter Meulen (1995) regarding EN and FR and Smith (2005, 2006) with application to Mandarin Chinese. Their different accounts of temporal reference make use of principles, such as for example Dowty's Temporal Discourse Interpretation Principle or Smith's Bounded Event Constraint, as well as numerous semantic rules. SDRT framework (Lascarides and Asher 1993; and Asher and Lascarides 2003) accounted for temporal reference in terms of coherence relations and reasoning using different types of inferences.

Chapter 3 assessed the question of temporal reference in a cognitive pragmatic framework, precisely RT (Sperber and Wilson 1986/1995; Wilson and Sperber 2004, 2012), as well as in morphosyntax, neurolinguistics and automatic treatment of language. In RT, the ingredients of temporal reference encode procedural and conceptual information that constrain, and respectively, contribute to the propositional, the explicit and the implicit contents of an utterance. Unfortunately, the generic notion of *verbal tense* continued to be used also in these pragmatic approaches to temporal reference having as result a blurred image and models that cannot be used cross-linguistically. Morphosyntactic models of temporal reference (Chomsky 1981, 1993, 1995, 2000; Pollock 1989; Guéron 1993; Belletti 1990; Giorgi and Pianesi 1997; Stowell 2007, 2012; Zagona 1990; Cowper 2002, 2005; Demirdache and Uribe-Etxebarria 2007) agree that Tense and Aspect represent the functional head of the sentence (despite the dissimilarity with respect to the competition between Tense and Aspect for the highest position). The main assumption is Tense and Aspect are interpretable features and are provided by the Universal Grammar. Languages differ in the way in which they make use of these features. Neurolinguistic studies found that humans process differently reference to past time and reference to present time, and that ingredients of temporal reference are impaired in language disorders such as aphasia among others[184], both in tensed and in tenseless languages.

In this chapter, I aim at proposing a theoretical cross-linguistically valid reanalysis of temporal reference, which is empirically and experimentally supported and which might provide an explanation for neurolinguistic findings (as discussed in section 3.3.1).

The issue of temporal reference (i.e. the location of eventualities in time) in a pragmatic

---

[184] See for example Fiorin (2010), who investigated the processing of Aspect, pronouns and quantification by Italian dyslexic children.

approach was addressed in a systematic manner by Moeschler et al. (1998)[185] with application to FR verbal tenses. Moeschler (1998, 159) notes with respect to the notion of eventuality and reference to eventualities that:

> Un événement se caractérise par sa nature (c'est un événement de tel ou tel type, *courir, manger, pleuvoir*, etc.), par ses participants (agent ou patient), par ses circonstances spatio-temporelles (il s'est produit à un moment et dans un lieu donné) et par ses relations à d'autres éventualités, événements ou états. En d'autres termes, un événement est la projection complète, saturée, d'une référence temporelle virtuelle (temps verbal) sur une référence lexicale virtuelle (prédicat), combinée aux références actuelles des arguments de la phrase. [186]

In other words, determining the temporal reference of an eventuality requires therefore virtual and actual temporal and lexical references, where:

- Virtual temporal reference is provided by Tense through temporal coordinates E,R and S.
- Actual temporal reference is provided by the contextual saturation of temporal coordinates E, R and S.
- Virtual lexical reference is provided by the aspectual class of the verb (i.e. state, activity, accomplishment and achievement).
- Actual lexical reference is provided by the predicate and the arguments of the VP.

My assumption is that this picture represents only one piece of the puzzle. My suggestion is that the global interpretation of temporal reference at the discursive level is determined, on the one hand, by the linguistic means existent in a language, and on the other hand, by their ad-hoc inferential contextual saturation.

Figure 7-1 provides a possible model of the functioning of temporal reference in discourse[187].

---

[185] Developing previous work by Reboul (1994) and Moeschler (1994, 1996), where the distinction between *actual* and *virtual* reference (introduced by Milner 1982 for referential expressions) was used (see section 3.1.3.6).

[186] 'An event is characterized by its nature (it is an eventuality of a type or another, *run, eat, rain*, etc.), by its participants (agent or patient), by its spatial and temporal circumstances (it takes place at a certain moment and in a certain place) and by its relations to other eventualities (events or states). In other words, an eventuality is a complete and saturated projection of a virtual temporal reference (i.e. a verbal tense) on a lexical virtual reference (a predicate) combined with actual references of the arguments of the VP.' (my translation)

[187] A very interesting account of EN Tense morphemes is suggested by Chilton (2005, 2007, 2013) in the framework of Deictic Space Theory (DST). The DST postulated the existing of a universal conceptual space consisting of three dimensions: discourse referents on the *d*-axis, conceived time on the *t*-axis and epistemic modality on the *m*-axis. The three dimensions converge at *S*, which stands for the speaker or the conceptualizer, which is also the zero point of now-here-real. These four elements represent a *frame of reference*. There is a base coordinate system called *R* standing for the reality conceptualized by the speaker. The conceptual space contains of every linguistic expression several frames of reference, treated as copies of R, which are *virtual realities* anchored at different deictic points in S's R (2014, 237). The relative distances one each axis represents the speaker's *cognitive distance*. Around S there is a *peripersonal space* standing for the speaker's peripersonal *time* consisting the memory of recent past and anticipation and planning for immediate future. I found Chilton's geometrical modelization convincing in that it provides a visual support of the speaker's conceptualization of time. The three-dimensional deictic space consisting of a base frame with the speaker's here-now-real in the centre and the other frames of reference nested in the base system helps us visualize the complexity human's conceptualization of time. Chilton's model integrates fundamental existent assumptions that have been made about tense: (i) Reichenbach's (1947) temporal

Figure 7-1 Encoded temporal information
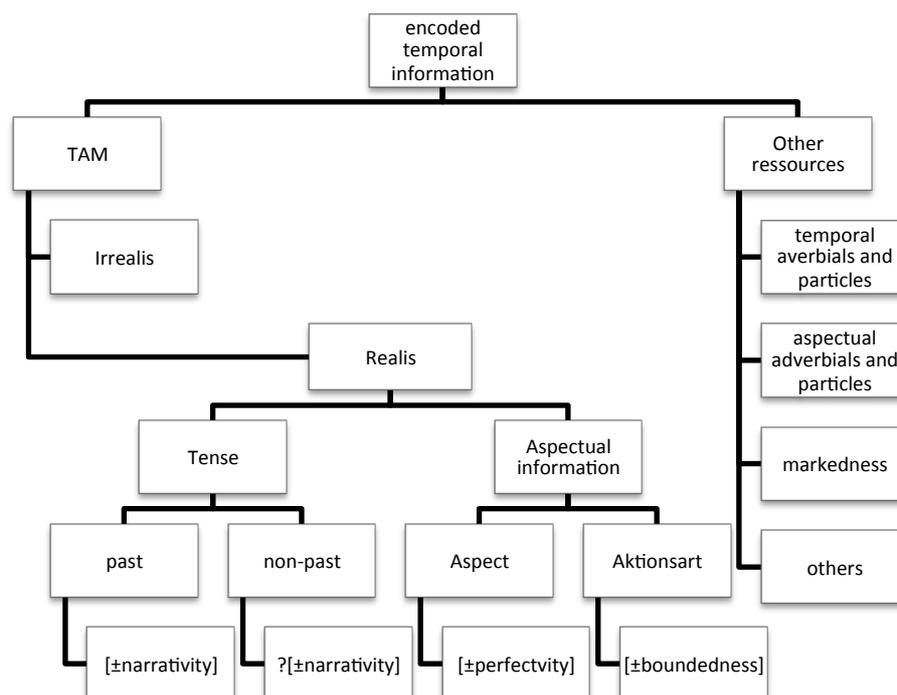
A first distinction is proposed between linguistic means typically used in tensed and tenseless languages. Tensed languages, as the languages under study in this research, make use of TAM markers, namely *tense, aspect* and *mood*. In morphosyntactic terms, these are *interpretable* features (Chomsky 1995; Cowper 2005): [± past] Tense, [± perfective] Aspect and [± realis] Mood, where the past, imperfective and irrealis are the unmarked forms (i.e. sentences are interpreted as perfective, non-past and realis in the absence of overt markers of the other category). Sentences in tensed languages have Infl (i.e. from Inflection, following Chomsky 1957, 1967) as functional head (consisting of Tense and Agreement features). My hypothesis is that the hearer follows several parallel steps in the interpretation process (as suggested in RT, see section 3.1.3.1). With respect to the temporal structure of a discourse, the hearer makes hypotheses about:

- Location in Realis or Irrealis of the eventuality or series of eventualities.
- If the Realis pathway is chosen, the hearer makes hypotheses about the contextual values of Tense and Aspect
- As far as Tense is concerned, the hearer makes a hypothesis about location with respect to S: past (E<S) or non-past (E≥S).
- If the past time path is selected, a second hypothesis is made about temporal location of an eventuality with respect to another eventuality, operationalized as the [± narrativity] feature.

coordinates, (ii) Damourette and Pichon' (1911-1940) observation that languages is used to communicate the speaker's psychological attitudes, (iii) Jaszczolt's (2009) account of the relation between tense and modality in the Default Semantic framework, (iv) Saussure's perspectival interpretation of tenses (2013). Moreover the distinction between the base system and the frames of reference, which are copies of the base system, allows the modelization of 'default' semantics and the meanings triggered by supplementary lexical material or pragmatic factors.

- As far as Aspect is concerned, the hearer makes hypotheses about the possible contextual values of grammatical and lexical aspects.
- As far as Aktionsart is concerned, the hearer makes hypotheses about the actual realization of (a)telicity.

In tenseless languages, as Mandarin Chinese (see section 2.1), the speaker and the hearer make use of other means to express temporal reference than the Tense branch. The Aspect branch and the Other resources branch are much more developed than in tensed languages. Sentences in Chinese have an aspectual functional head ASP, which can be perfective or imperfective. Additionally, other resources are used to determine the contextual value of the ASP head.

The empirical research described in this work dealt with two branches: *tense* expressing past time reference and the *aspect* branch applied to verbal tenses expressing past time reference. More precisely, it focused on three verbal tenses: the simple past, the compound past and the imperfective. Due to the specific applied purpose of this research regarding machine translation, there are other verbal tenses expressing past time reference that were not considered such as the pluperfect and EN past progressive. As for the branch *tense* expressing non-past time reference, the PRES was described in three target languages. However, this verbal tense was not considered in the experiments as far as the procedural [±narrativity] feature is concerned. The theoretical model proposed in Figure 7-1 is therefore validated for the branches investigated empirically in this research.

In section 7.2 I discuss theoretical assumptions regarding Tense and Aspect and the conceptual/procedural distinction proposed in RT. Section 7.3 is dedicated to temporal coherence at the cognitive and discursive levels. Section 7.4 deals with monolingual descriptions of the verbal tenses considered in this research and section 7.5 concludes this chapter.

7.2 Temporal reference and its ingredients: a reanalysis

In what follows, I will propose a reanalysis of Tense, Aspect and Aktionsart in the light of the empirical findings described in Chapters 5 and 6. Chapter 5 was dedicated to the corpus work carried out on bilingual and multilingual translation corpora. Analysis of translation corpora indicated that verbal tenses have dissimilar frequencies in SL, as well as dissimilar translation paradigms and translation relations (i.e. as discussed in 4.2.3, a translation relation consists of a series of properties and senses shared partially by the linguistic expressions that stand in that translation relation). Precisely, the SP is both a frequent verbal tense and a problematic one because of its rich translation paradigm, which consists of four verbal tenses in TL (i.e. PC, PS, IMP and PRES in the three Romance languages considered) (as discussed in section 5.1.2). Another example is the case of the FR PC, which is a frequent verbal tense with a binary translation paradigm (i.e. SP and PresPerf in EN) (as discussed in section 5.2.2.).

Chapter 6 addressed the offline experimental work carried out in order to test a series of properties included in the translation relation shared by the EN SP and the PC, PS, IMP and PRES in FR, IT and RO. These properties have been defined according to the literature on verbal tenses and operationalized as the following features: [±narrativity],

[±boundedness], [±perfectivity] and the conceptual past/non-past distinction. Experiments carried out confirmed that the SP, PS, PC and IMP can be described in terms of the instructions encoded by the categories of Tense, precisely the [±narrativity] feature, Aspect precisely the [±perfectivity] feature and by Aktionsart, precisely the [±boundedness] feature. Moreover, the category of Tense encodes the pro-concept TIME that is contextually specified with one of the following two ad hoc concepts, past and non-past.

Cross-linguistically speaking, all verbal tenses investigated in this research encode these instructions. However, each language presents dissimilar associations between one verbal tense and the values of the three temporal and aspectual categories: Tense, Aspect and Aktionsart, as it will be accounted in section 7.4. The following sections are organised as follows: section 7.2.1 deals with conceptual and procedural information with respect to Tense and the types of meaning layers that they trigger, and in section 7.2.2 with types of meaning layers linked to Aspect and Aktionsart.

7.2.1 Tense

As far as Tense is considered, neurolinguistic studies brought evidence that humans process differently sentences referring to past and non-past time (see section 3.3.1.) Two observations have been made. Firstly, reference to past time is impaired in aphasia whereas reference to non-past (present and future) is relatively spared. This finding holds for both tensed and tenseless languages. Moreover, healthy speakers have longer reaction times when processing sentences with reference to past time than sentences with non-past time reference.

Secondly, there are different brain reactions for sentences containing time reference violations, where time reference can be provided by temporal adverbials and verbal tenses. Sentences containing a disagreement between a past time adverbial and a present time verbal tense produce a P600 wave triggered by the verb, longer reaction times and reduced acceptability ratings than the opposite configuration (i.e. present time adverbial and a past time verbal tense). In both cases the brain produces LAN waves at the end of the sentence signalling that it detected the disagreements and is trying to resolve it in a meaningful way. Participants find easier a meaningful interpretation for the present time adverbial/a past time verbal tense condition when a larger context is provided.

The question that arises is how is the difference between the processing of past and non-past time reference accounted for by the current theoretical models regarding temporal reference in discourse. In the literature, two main trends are opposed as far as the nature of the encoded content is concerned:

- Tense encodes procedural information
- Tense encodes both procedural and conceptual information

According to the first trend, verbal tenses encode rigid procedural information that help the hearer reconstruct the intended representation of eventualities (Nicolle 1997, 1998; Wilson and Sperber 1993; Moeschler 1994, 1998; Moeschler et al. 1998; Aménos-Pons 2011, Saussure 2003, 2011). Saussure (2003) proposes algorithms to follow based on these instructions in order to grasp the intended meaning of a verbal tense at the discourse level. Nicolle (1998, 4) argues that tense markers impose constraints on the determination of temporal reference and thus they "may be characterized as exponents of procedural

encoding, constraining the inferential processing of conceptual representations of situations and events". Concerning the status of the temporal coordinates, Saussure and Morency (2012) argue that tenses encode instructions on how the eventuality is to be represented by the hearer through the positions of temporal coordinates. They consider thus that temporal location with the help of S, R and E is of a procedural nature. I want to argue that location through temporal coordinates does not constrain the inferential processing but contribute to the propositional content of the utterance. I will however make a difference among Reichenbachian coordinates: E and S are crucial for all verbal tenses and contribute to building an ad hoc concept whereas R is more likely to be linked to the procedural information encoded by Tense. I suggest a *conceptualist* view of Tense. The hearer makes use of pragmatic inferences in order to recover the speaker's meaning with respect to temporal location of eventualities. This takes place on two levels: ad hoc narrowing of the pro-concept TIME through contextual saturation of two Reichenbachian coordinates and relating eventualities with respect to one another, that is, the [±narrativity] feature making use of the R coordinate.

7.2.1.1   Reichenbachian coordinates: E and S

Wilson and Sperber (1993, 151) argue that conceptually encoded information contributes either to explicatures (to the proposition expressed and to high-level explicatures) or to implicatures whereas procedurally encoded information represents constraints either on explicatures (to the proposition expressed and to high-level explicatures) or on implicatures (as discussed in section 3.1.3.3). They argue for the idea that during interpretation process, the hearer builds conceptual representations and uses encoded procedures for manipulating them. A conceptual representation differs from other types of representations in that it has logical properties and truth-conditional properties. The sentence in (542) that has the logical form (543) and the propositional form (544). They argue that the logic form recovered through decoding and the propositional form recovered by a combination of decoding and inference are conceptual representations.

(542)   Peter told Mary that he was tired.
(543)   x told y at t_i that z was tired at t_i.
(544)   Peter Brown told Mary Green at 3.00 pm on June 23 1992 that Peter Brown was tired at 3.00 pm on June 23 1992.

The hypothesis defended here is that the configuration of temporal coordinates S and E is conceptual information acting like *pro-concepts* (Wilson 2011, Sperber and Wilson 1998). Pro-concepts are semantically incomplete, they are conveyed in a given utterance and have to be contextually worked out through a pragmatic enrichment process similar to lexical-pragmatic processes. The pro-concept TIME can be specified through narrowing in the ad hoc concepts of pastness (i.e. $E<S$) and of non-pastness[188] ($E≥S$). Tense encodes this base

---

[188] Similarly to lexical pragmatics, where, for example, the pro-concept OPEN may be specified to numerous ad hoc concepts (e.g. open a can, open a door, open a bank account, open a file, etc.), one could image that the pro-concept TIME can be narrowed to express more specific categories of temporal remoteness (such as in Bantu languages, cf. Comrie 1985), omnitemporality (E holds *before*, *at* and *after* S) and atemporality. A

semantic and conceptual information and it is contextually worked out. Due to repeated and constant activation of the same ad hoc concept, certain verbal tenses became specialized for activating the concept of pastness whereas others for activating the concept non-pastness (a similar account of interjections is given by Padilla Cruz 2009, as discussed in section 3.1.3.4). For example, the classically described 'verbal tenses expressing past time', such as the PC, PS, IMP and PQP, have undergone this specialization for the ad hoc concept of pastness. However, this does not prohibit a verbal tense to make reference to another time or to no time at all, if contextual information directs it.

This temporal information is not defeasible, i.e. cannot be cancelled. Let's consider Wilson and Sperber's example (1993, 157) given in (542) and the propositional form given in (544). I add to this propositional form the information that eventualities of *saying* and of *being tired* took place before the moment when the sentence was uttered. The extended propositional form would be something like the one given in (545). This temporal information cannot be cancelled or contradicted, as shown by the incompatibility with the adverb *now* or *tomorrow* in (546) and (547), as well as the compatibility with the adverb *yesterday* in (548).

(545) Peter Brown told Mary Green at 3.00 pm on June 23 1992 (a moment before the present moment/in the past) that Peter Brown was tired at 3.00 pm on June 23 1992 (a moment before the present moment/in the past).

(546) *Peter Brown told Mary Green at 3.00 pm on June 23 1992 which is now (a moment contemporary with the moment of speech)/ tomorrow (a moment which is after the moment of speech) that Peter Brown was tired at 3.00 pm on June 23 1992 which is now/tomorrow.

(547) *Now/tomorrow Peter told Mary that he was tired.

(548) Yesterday, Peter told Mary that he was tired.

The contextual values and the relation between S and E (i.e. E<S for past and E≥S for non-past) are pragmatically determined. As suggested above, the pro-concept TIME is specified through narrowing to ad hoc concept based on contextual linguistic and non-linguistic information. For example, the FR PC allows reference both to past time and to future time. In (549), the PC, which is specialized for expressing the ad hoc concept of pastness, locates the eventuality of *finishing* prior to S. In (550) on the contrary, the hearer builds an ad-hoc concept of non-pastness making use of linguistic information, precisely the temporal adverb *tomorrow,* and therefore it expresses reference to future time (i.e. E>S). Since the building of the ad hoc concept and the computation of the instructional content, operationalized as the [±narrativity] feature, are simultaneous processes, the hearer can readjust his initial hypotheses during the interpretative process.

(549) J'*ai fini* mon livre.
    I finish.PC my book
    'I *finished* my book.'

(550) Demain, j'*ai fini* mon article.
    Tomorrow I finish.PC my article

---

future study investigating this matter empirically (corpus-based study) and experimentally would be very interesting.

'Tomorrow, I *will have finished* my paper.'

If we consider example (551) and imagine two different contexts, the distance on the time line between E and S, even if S=E for present tenses is contextually adjusted based on world knowledge. In a first context, a husband is upstairs and his wife is downstairs in their house, he calls her and she answers (551). In the second context, the wife has an hour ride from work to home, he calls her to see when she comes back home and she answers (551). The distance between E and S is between immediately and 2-3 minutes in the first context and a few minutes and one hour (or even more) in the second context.

(551)   J'arrive!
        I arrive.PRES
        'I am coming!'

Experiment 7 was designed to test the nature of the relation E vs. S. According to the qualitative features proposed by Wilson and Sperber (1993) for conceptual and procedural information, it was argued that judging conceptual information results in high Kappa values. Experiment 7 provided evidence that the conceptual information encoded by verbal tenses, i.e. past vs. non-past, is determined contextually and that the agreement among the participants produced high Kappa values: 1 for artificial data, 0.80 for natural data and 0.86 for all the data.

Moreover, I suggested that translating conceptual information triggers little cross-linguistic variation whereas translating procedural information is source of important variation. This quantitative feature makes use of Moeschler's et al. (2012) suggestion that conceptual information is easily translatable whereas procedural information is translatable with difficulty. This suggestion is linked to the fact that conceptual information represents concepts that are constituents of the language of thought and therefore, language-independent. Based on this observation, it is expected that translating conceptual information triggers small variability in the target language(s) whereas translating procedural information triggers high variability. In Grisot and Costagliola (2014) and in this thesis (see section 5.3 for analysis of parallel translations corpora), it was shown that for translating the SP into three Romance languages, verbal tenses expressing past time are used in more than 70% of the cases (precisely, 73% in FR, 72% in IT and 83% in RO) whereas the PRES is used in less than 8% of the cases (precisely, 8% in FR, 5% in IT and in RO).

Regarding procedural information encoded by verbal tenses, it helps the hearer to access the right contextual hypotheses conforming to the communicative and cognitive principles of relevance to get the intended cognitive effects (Wilson and Sperber 1998). Carston (1998) points out that under normal conditions discourse material is presupposed to be relevant and, when information is not explicitly given, it is filled in. The linguistic content of utterances is hence enriched in the interpretive process. In this case, the basic temporal location of the eventuality (E/S) is enriched via procedural information. In (552), Binnick (2009) following Grice[189] (1989) argues that the material in brackets is implicit. Sentence in

---

[189] Binnick's example is a typical example for conversational implicatures (in Grice's terms, 1989) that follow the maxim "Be orderly". Carston (1998, 2002) and Sperber and Wilson (1986/1995) treat this content as

(552) is an example of temporal ordering, and thus the procedural feature [±narrativity] of the SP is active.

(552) He took off his boots and [*then*] got into bed.

To sum up, my suggestion is that Tense encodes a very broad pro-concept TIME. Each verbal tense in a language is constantly used in activation to reference to past or non-past (distinction recognised also in neurolinguistics, see section 3.3.1) and it becomes, therefore, specialized for these ad hoc concepts. Another case is when a verbal tense does not have a temporal interpretation. My hypothesis is that procedural information encoded by Mood (i.e. realis vs. irrealis) constraints the building of the ad hoc concept. In all cases, the hearer is brought to make inferences regarding the ad hoc conceptual meaning of a verbal tense and this is constrained by procedural types of information encoded by Tense and Aspect. Section 7.2.1.2 is dedicated to the procedural information encoded by Tense.

7.2.1.2 [±Narrativity] and Reichenbachian R

Moeschler et al. (2012) propose a theoretical model permitting a cross-linguistic description of the usages of verbal tenses. It contains three hierarchical features: [±narrative], [±subjective] and [±explicit] as shown in Figure 7-2.

Figure 7-2 Types of usages of verbal tenses
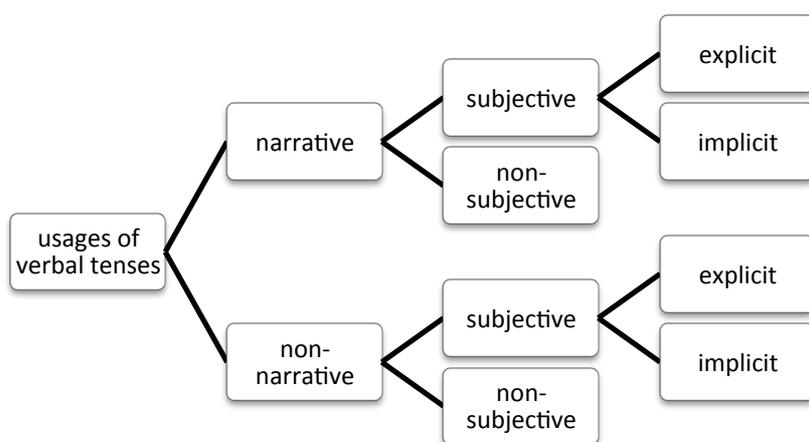
This model was initially proposed[190] to describe the use of the historical present and the Free Indirect Style in FR, EN and Japanese. It was thereafter developed in relation to a general inferential model for temporal information in discourse, i.e. the MDI model (Moeschler 2000a, 2003) (see section 3.1.3.6) in order to describe semantic and pragmatic usages of verbal tenses. Experimental work carried out in this thesis and in Grisot (in preparation) allowed me to refine this model by partly validating and partly challenging the theoretical assumptions behind this model. I will focus on the [± narrative] feature in this

---

pragmatically determined aspects of what is said, thus an explicature. See Blochowiak (2014a, 2014b) for a presuppositional account of temporal and causal connotations of 'and'.

[190] Research carried out by Jacques Moeschler (University of Geneva), Anne Reboul (L2C2, ISC Lyon) and Izumi Tahara (Meiji Gakuin University, Tokyo) (2010-2012).

section and on [± subjective] feature in section 7.2.2.2.

The notion of *narrativity* has already been used in the literature by numerous scholars and more importantly, in various frameworks. For example, Labov and Waletzky (1967) argued that two sentences, which are interpreted as being temporally successive, form a narrative text. In DRT (see section 2.2.2), Kamp and Rohrer (1983) argued that certain verbal tenses, such as the FR PS, impose a narrative (i.e. temporal progression) interpretation of the discourse where it occurs. In SDRT (see section 2.2.2), discourse segments can be linked through discourse relations, such as *narration*, which is the default coherence relation. Narration involves sentences where textual order matches temporal order of eventualities in the real world. Finally, Smith (2003) uses the notion *narrative discourse mode*[191], defined based on aspectual criteria, specifically the type of eventualities expressed (events and states), and interpretation semantic principles (as discussed in section 2.4.3). All these usages of the *narrativity* notion have in common the temporal progression interpretation of the discourse. However, they propose dissimilar explanations of how this interpretation is carried out. In this thesis, the *narrativity* notion shares only partly with other frameworks the idea of temporal progression interpretation. Moreover, a different explanation with respect to how hearers arrive at this interpretation is suggested. In this research, the [±narrativity] feature is meant to model forward and backward temporal inferences triggered by verbal tenses, as well as simultaneous temporal relations. In other words, verbal tenses encode procedural information instructing the hearer to determine the contextual value, i.e. positive or negative, of the feature.

In Moeschler et al. (2012) and Grisot and Moeschler (2014) four arguments were given in favour of the procedural nature of this feature (see also section 3.1.3.3 for a theoretical discussion of procedural information). Firstly, the [±narrativity] feature is information that constraints the inferential phase of constructing explicatures. It does not contribute, but constraints the construction of the propositional content of utterance (Wilson and Sperber 1998, Binnick 2009, Escandell-Vidal and Leonetti 2011). Secondly, temporal sequencing is a discourse property: it needs at least two eventualities for the [±narrativity] feature to be active. Procedural content gives information about how to manipulate conceptual representations, corresponding to more than one discourse entity. If a tense has a narrative usage, it means that as soon as its reference time is set, it is used to construct the temporal reference of the next event, and thus time advances. Binnick (2009) pointed out the role of verbal tenses for discourse coherence as temporal anaphors (discourse interpretation depends on the identification of their antecedents). In example (553), the SP of the verb *take* (specifically *took*) is bound by that of the verb *go* (specifically *went*). Time advances in a narrative sequence because the R point of one eventuality is located just after the preceding one.

---

[191] The *narrative* discourse mode is a type of *temporal* discourse modes (next to report and description), which are contrasted to atemporal discourse modes (informative, argument-commentary). The narrative mode makes use of two types of discourse entities: states and events. Smith (2003) and Dowty (1982, 1986) propose two principles that are involved for interpreting verbal tenses in the narrative mode. Firstly, if a sentence expresses a bounded event, the reference moment R increases and the verbal tense expresses continuity. Secondly, if the eventuality expressed is not a bounded event (therefore a state), then R does not change and the verbal tense is used anaphorically.

(553) John went home early. He took the subway.

Thirdly, temporal sequencing can hardly be paraphrased (as with synonyms for conceptual representations), but it can be rendered explicit with the help of temporal connectives, such as *and, then, afterwards, because*. And fourthly, the [±narrativity] feature is information inaccessible to consciousness resulting in low agreement rates among annotators.

Grisot and Moeschler (2014) speak about a *discursive* model[192]. If [±narrativity] feature is positive, then a procedure of temporal ordering calculus is set on. A verbal tense has a narrative usage (i.e. there is temporal progression from one eventuality e_1 to another eventuality e_2, therefore $R_1 \rightarrow R_2$) as in (554) or a non-narrative usage (i.e. there is no temporal progression from one eventuality e_1 to another eventuality e_2, therefore e_2 has the same R_1) as in (555).

(554) Erkshine *rose* from his seat, and going over to a tall inlaid cabinet, that stood between the two windows, unlocked it, and *came back* to where I was sitting, carrying a small panel picture set in an old and somewhat tarnished Elizabethan frame. (Literature Corpus)

(555) It was enough for her that he *appeared* to be amiable, that he *loved* her daughter, and that Elinor *returned* the partiality. (Literature Corpus)

The identification of reference time R is either linguistically triggered (through verbal tense form or a temporal adverb, for example) or pragmatically inferred by the hearer based on contextual and world knowledge. This procedure of temporal ordering calculus is not a default procedure, as Asher and Lascarides (2003) state, but it is triggered by the activation of the [±narrativity] procedural feature. Generally speaking, I would like to suggest that verbal tenses do not encode one of the two possible values of this feature by default, as it is assumed for example by Saussure (2003). He suggested that the FR PS encode the narrative value by default whereas the IMP is not specified for this instruction, which means that the [±narrativity] procedural feature is not applicable for the IMP. According to the model developed in this thesis, the category of Tense encodes this feature, and consequently, all verbal tenses encode it. Each verbal tense may be more frequently associated with one or the other possible values; however, it does not encode that specific value.

I think that this problem can be investigated and answered only in experiments targeting the on-line processing of temporal information. My prediction, regarding the hypothesis that a verbal tense encodes by default one of the two values, is that longer reaction times should be found when participants process a sentence where a verbal tense, the PS for example, has a non-narrative usage than a sentence where it has the default narrative usage. On the contrary, regarding the hypothesis defended in this thesis, my prediction is that similar reaction times should be found when participants process a sentence where the PS has a non-narrative usage and a sentence where the PS has a narrative usage. I suggest the same prediction for the PC and the IMP. This issue will be addressed in further experimental research.

Grisot and Moeschler's model is determined by the requirement to disambiguate usages

---

[192] Kamp and Rohrer (1983) also argued for their discourse semantics model that the meaning of a verbal tense could be established only at the discursive level.

of the SP and to improve its translation into FR. Consider example (556) with an isolated SP and example (557) containing the target sentence and its cotext. With respect to its translation into a TL, the isolated token is ambiguous. In (557), the second sentence introduces another eventuality and the two eventualities are temporally and causally related. According to the model, the SP has a narrative usage and it is translated into FR by a PS/PC as in (558) and (559). In (560) on the other hand, the second sentence introduces an eventuality that takes place simultaneously. More specifically, the R period of the first SP includes the R moment of the second eventuality. According to the model, the SP has a non-narrative usage and it is into FR by an IMP, as in (561).

(556)   John *slept*.
(557)   John *slept*. He got rest.
(558)   Jean *a dormi*. Il s'est reposé.
        John sleep.PC. He get rest.PC.
(559)   Jean *dormit*. Il se reposa.
        John sleep.PS. He get rest.PS.
(560)   John slept. He had a dream.
(561)   Jean *dormait*. Il fit un rêve.
        John sleep.IMP. He have.PS a dream.

Further research was carried out in order to test empirically the theoretical assumptions suggested in Moeschler et al. (2013). The experiments involving bilingual data research were presented in Grisot and Moeschler (2014). In this thesis, sections 6.1.1, 6.1.2 and 6.1.3 for bilingual data, as well as 6.2.2, 6.2.3 and 6.2.4 for multilingual data, describe the experiments carried out for FR, EN, IT and RO verbal tense. The experiments carried out on multilingual data confirm Grisot and Moeschler's model and validate it for two additional Romance languages, IT and RO.

In this thesis, I use the Kappa coefficient to measure the agreement among judges that occurs beyond chance. The suggestion I made is that the Kappa coefficient can be used for evaluating procedural and conceptual information. Specifically, judging conceptual information results in high agreement rates, outcome due to the high accessibility to consciousness of conceptual information (according to the qualitative features proposed by Wilson and Sperber, 1993). On the contrary, judging procedural information results in low agreement rates, outcome due to the low accessibility to consciousness of rules and instructions to manipulate conceptual representations. However, judging procedural information could be rendered easy to access through conscious thinking when the hearer has supplementary sources of information (as suggested in Experiment 6 testing the SP and the PresPerf).

Moreover, an important variability was identified with respect to the usage of specific verbal tenses expressing past time, which can be discriminated based on procedural information encoded by Tense and Aspect. Precisely, for translating the SP in FR the PC, IMP and PS had the following frequencies of usage: 34%, 23% and respectively 16%. Similar values were found for IT (33%, 17% and respectively 22%) and RO (49%, 15% and 18%).

The experimental work described in Chapter 6 regarding the [±narrativity] feature

elicited the following findings:

Firstly, native speakers recognize narrative and non-narrative usages of FR, EN, IT and RO verbal tenses expressing past time. The judges agree with each other and this agreement is not due to chance.

- The low Kappa values of the first rounds of the experiments for each language points to the difficulty hearers/readers have in the interpretation process to conceptualize the language rules they have and make decisions about their functioning.
- The high Kappa values from the second rounds of the experiments for each language points to the procedural nature of the feature, given that one of the characteristics is the possibility to render explicit the instructions encoded with the help of discourse markers.

Secondly, each of the considered languages presents a language-specific behaviour for the [±narrativity] procedural feature. My hypothesis is that this is linked to the aoristicization process (Squartini and Bertinetto 2000, see section 2.5.1) undergone by the PC form. Specifically, the PC suffers a change from a pure perfect (as it remained in Spanish and Portuguese) to an aorist (the value of PS)[193]. The Romance languages considered in this research, FR, IT and RO, find themselves at a different point in the aoristicization process. Precisely, the RO PC is more advanced than the IT PC, which, it its turn is more advanced than the FR PC. The [±narrativity] feature is meant to capture the instruction to temporally relate (i.e. temporal and causal sequencing vs. temporal simultaneity) one eventuality with respect to another. The IMP most often performs non-narrative values in Romance languages.

Corpus analysis showed that the PC is more frequently used in RO than in IT and in FR. Moreover, experimental work indicated that it is perceived and judged by native speakers more often as narrative in RO, than in IT and in FR. As for the PS, corpus analysis showed that its frequency of usage decreases in the three languages considered: less frequently used in RO than in IT and FR. Experimental work did not show significant differences in judgment among the three languages.

EN language presents a different pattern mainly because the PresPerf did not develop aorist functions, as the PC in Romance languages did. Moreover, the SP carries out narrative and non-narrative usages with comparable percentages (60% narrative and 40% non-narrative as shown in section 6.1.3. Unfortunately, the EN past progressive form was not considered in the analysis. This is mainly due to its infrequency in the corpus (only 1%, see section 5.1.1).

These empirical findings show that the [±narrativity] procedural feature is a language-independent feature that has a language-specific behaviour. The results of the annotation experiments of the data used in this research are summarized in [±Narrativity] feature and its cross-linguistic realization by each verbal tense considered

---

[193] In future work, the [±narrativity] feature could be tested for Spanish and Portuguese PS, PC and IMP. My prediction is that it will produce a very different pattern for the PC. Precisely, it might be judged more frequent as non-narrative than as narrative due to the fact that it does not undergo the aoristic drift taken by the PC in FR, IT and RO.

Table 7-1 [±Narrativity] feature and its cross-linguistic realization by each verbal tense considered

Language	VT	Narrative	Non-narrative
EN	SP	59%	41%
	PS	**92%**	8%
	PC	**77%**	23%
FR	IMP	16%	**84%**
	PS	**96%**	4%
	PC	**88%**	12%
IT	IMP	16%	**84%**
	PS	**93%**	7%
	PC	**83%**	17%
RO	IMP	19%	**81%**

In other words, there is a cross-linguistic variation among the individual verbal tenses that encode this instruction and its contextual values. My prediction is that, for example, a narrative usage of the SP can be translated into a TL trough a narrative usage of a verbal tense, be it PS, PC, IMP (i.e. the so-called narrative IMP) or even PRES (i.e. the so-called historical PRES), as shown in examples (562)-(565), where the former is the original text in EN and the others are its translation[194] into FR, IT and RO respectively. In these texts, the original SP form having a narrative usage is translated through a narrative IMP in FR, and through a narrative PS into IT and RO.

(562) Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do. Once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," *thought* Alice, "without pictures or conversations?"

(563) Alice commençait à se sentir très lasse de rester assise à côté de sa soeur, sur le talus, et de n'avoir rien à faire: une fois ou deux, elle avait jeté un coup d'oeil sur le livre que lisait sa soeur; mais il ne contenait ni images ni dialogues: « Et, *pensait* Alice, à quoi peut bien servir un livre où il n'y a ni images ni dialogues? »

(564) Alice cominciava a sentirsi mortalmente stanca di sedere sul poggio, accanto a sua sorella, senza far nulla: una o due volte aveva gittato lo sguardo sul libro che leggeva sua sorella, ma non c'erano imagini nè dialoghi, "e a che serve un libro," *pensò* Alice, "senza imagini e dialoghi?"

(565) Alice începuse să se simtă foarte obosită stând pe bancă lângă sora ei şi neavând nimic de făcut: o dată sau de două ori trase cu ochiul la cartea pe care sora ei o citea dar nu avea poze sau dialoguri „Şi care e rostul unei cărţi" *se gândi* Alice, „fără poze sau dialoguri?".

I made the hypothesis that other factors, such as Aspect and Aktionsart influence the choice of the verbal tense in TL, as in examples (566)-(569), the former is the original text in EN

---

[194] The examples come from parallel corpus (see section 5.3) consisting of texts translated by professional human translators.

and the others are its translation into FR, IT and RO respectively from the JRC corpus. Experimental work with respect to Aspect and Aktionsart showed that the perfective aspect and bounded type of situations are significantly correlated with the PC/PS verbal tense, whereas imperfective aspect and unbound type of situations is correlated with the IMP (see section 6.1.7 for the results of the multifactorial analysis, and section 7.3 for a theoretical interpretations).

(566) The field experiment [...] It was accompanied by measurements at four fixed stations, with 15 mobile units, with an aircraft and balloons and *included* model calculations on the basis of a detailed emission inventory.

(567) L'expérience sur le terrain [...] Elle a été accompagnée par des mesures dans quatre stations fixes et avec 15 unités mobiles, un avion et des ballons et *prévoyait* des calculs par modèle sur la base d'un inventaire détaillé des émissions.

(568) L'esperimento sul campo [...] è stato accompagnato da misurazioni in quattro stazioni fisse, con l'ausilio di quindici unite mobili, un aereo e palloni aerostatici, e *ha incluso* calcoli di modello sulla base di un inventario dettagliato delle emissioni.

(569) Experimentul de teren […] a fost însoţit de măsurători la patru staţii fixe, cu 15 unităţi mobile, cu un avion şi baloane şi *a inclus* calcule conform unui model bazat pe un inventar detaliat al emisiilor.

In (567), the FR translator made use of the verb *prévoir* 'to foresee, to anticipate, to envisage' which is an unbounded eventuality, and chose the IMP. In IT and RO, the translators made use of the same verb as in EN *to include*, which is a bounded eventuality, and chose the PC. As far value of the [±narrativity] procedural feature is concerned, in these texts the SP and the verbal tenses used in TL receive the non-narrative value (i.e. the eventualities *accompany* and *include* are temporally simultaneous). This value is carried out by the IMP in FR and by the PC in IT and RO.

7.2.2 *Aspectual information*

7.2.2.1 Aspect and Aktionsart

The distinction existent in the literature between grammatical aspect and lexical aspect was adopted in this thesis and I used the notation Aspect for the former and Aktionsart for the latter. In RT, it is assumed that Aspect encodes procedural information constraining the interpretative process by imposing the speaker's viewpoint on the eventuality (as discussed in section 3.1.3.7). To be more precise, perfective aspect constrains the hearer to build a completed representation of the eventuality denoted by the verb, in other words a single whole with highlighted boundaries. Žegarac (1991) suggested that the perfective aspect points indexically to a particular event instantiating the property denoted by the verbal predicate. For example, in the sentence in (570), the PresPerf conveys that the eventuality of having breakfast is completed and it makes reference to a particular instance of having breakfast, in principle at some relatively proximate time in the past. The analysis is similar for (571) except that the eventuality took place at some time farther in the past. The difference of the meanings of the two utterances with respect to the lapse of time between E and S follows from the communicative principle of relevance.

(570) I have had breakfast.

(571) I have been to Tibet.

The imperfective aspect constrains the hearer to build an unfinished representation of the eventuality, in other words to focus on the internal structure of the situation or on a moment other than the initial or the final boundaries. For example, in the sentence in (572), the progressive instructs the hearer to build an unfinished representation of the raining event and makes reference to a particular event instantiating the property denoted by the verb. On the contrary, the SP in (573) located the eventuality of raining at some time in the past without making reference to a particular instance of raining (Žegarac 1991, 155).

(572) It was raining.

(573) It rained.

Hence, Aspect encodes procedural information which constraints the explicit content of an utterance. More specifically, Aspect imposes constraints on Aktionsart: these conceptual representations are viewed from the speaker's point of view as being completed or not. Aspect presents the features proposed by Wilson and Sperber (1993) for procedural information: inaccessible to consiousness and not available through conscious thought in languages where they are not expressed morphologically. Besides, they are translatable with difficulty as shown by the lack of one-to-one correspondence between EN and Serbian or between EN and FR for example.

The experimental work described in section 6.1.5 confirmed these theoretical assumptions. Two judges were asked to evaluate SP items with respect to perfective vs. imperfective viewpoint and they agreed in 63% of the cases, which corresponds to a Kappa of 0.32. This Kappa value is beyond chance value but it is, however, below the threshold of reliable data (around 0.6). This result shows the difficulty judges have in deciding on the type of viewpoint from which the eventuality was expressed. This result points to the procedural nature of the [±perfectivity] feature. As far as the interpretation process is concerned, my suggestion is that hearers assign through an inferential procedure a contextual value of the [±perfectivity] feature and this takes place at the level of the explicature. In other words, the [±perfectivity] feature represents procedural information constraining the formulation of the explicature of the utterance. Due to necessity to have reliable annotated data with this feature for training an automatic classifier, in this research another method was used: the cross-linguistic transfer of properties based on translation corpora (see section 4.2.3 for details).

Considering Aktionsart, it is assumed that it represents conceptual information that undergoes the constraints imposed by procedural information. Aktionsart has logical properties and it contributes to the propositional content of an utterance (Moeschler et al. 2013). Scholars have identified the ontological distinctive features of aspectual classes cross-linguistically pointing to their language-independent character. Moreover, Aktionsart presents the qualitative features proposed by Wilson and Sperber (1993) for conceptual information: speakers have access with easiness to lexical aspect, it can be reflected on through conscious thought and represents easily graspable concepts (as shown in Experiment 4).

The experimental work described in section 6.1.4 confirmed these theoretical assumptions. Two judges were asked to evaluate SP items with respect to one distinctive ontological feature of Aktionsart, that is boundedness. Scholars showed that Aktionsart is sensitive to both Tense and Aspect, and therefore it was operationalized as the [±boundedness] feature. Judges were asked to evaluate SP items with respect to bounded vs. unbounded situations and they agreed in 92% of the cases, which corresponds to a Kappa of 0.84. The disagreements were resolved in a second phase. This Kappa value is beyond chance value and also beyond of the threshold of reliable data. This result signals the easiness judges had to decide on the type of eventuality using three linguistic tests. This results points to the conceptual nature of the [±boundedness] feature, which contributes to the explicatures of the utterance and has truth-conditional value. As far as the interpretation process is concerned, my suggestion is that hearers assign through an inferential procedure a contextual value of the [±boundedness] feature.

As pointed out by Žegarac (1991, 59) with respect to aspectual categories, and largely accepted nowadays in the relevance theoretic community:

> Rather than fully determining the propositional content of the utterance, the natural language code provides evidence on the basis of which the hearer arrives at the speaker's intended interpretation.

Hence, Aspect encodes procedural information representing instructions of the natural language expressions in context. They constrain the conceptual content of the proposition expressed by the utterance, specifically Aktionsart, and therefore they constrain the construction of the explicature.

Moreover, numerous scholars talked about the close relationship between Aspect and Aktionsart (as discussed in section 2.4.2) but this relationship, as far as I am aware of, was not quantified before this thesis (as discussed in section 6.1.8).

7.2.2.2   Subjectivity

Subjectivity is a notion used in connection with the linguistic expression of the way the speaker or other subject of consciousness[195] is involved in or looking at the situation(s) denoted by the utterance (Žegarac 1991, 78). Moeschler et al. (2013) use the [±subjectivity] feature to model the notion of *perspective* or *point of view* from which the events are narrated. Verbal tenses are sensitive to this focalization (Binnick 2009). A narration may be non-focalized [-subjective] or it may adopt the perspective of either an internal or an external focalizer [+subjective] (also Binnick 2009; Fleischmann 1990). Initially, it was assumed that

---

[195] The discussion of *subjectivity* and *point of view* (*perspective*) has been carried out mainly in the field of literary structuralism (Genette 1972; Fleischman 1990) and Cognitive Grammar (see Langacker 1991, 1999, 2002, 2006). Scholars suggested that subjectivity is a complex pragmatic phenomenon associated with usages of verbal tenses (Banfield, 1982/1995; Reboul 1992; Schlenker 2004; Reboul 2012; Sanders and Redeker 1996) pronouns, imperfective aspect (Fleischmann 1995; Boogaart 1999), modality and evidentiality (Nuyts 2001), discourse relations and connectives (Mann and Thompson 1986; Verhagen 1995; Sanders 1992, 1993; Sanders et al. 2012; Stukker and Sanders 2012) and connectives (Pander Maat and Sanders 2000; Pander Maat and Degand 2001; Pit 2003, 2007; Sanders and Spooren 2009; Stukker et al. 2009). Subjectivity has also been identified as an important factor in the processing of discourse, both for humans and for machines (Wiebe 1990, 1994; Chen 2008).

subjectivity is expressed linguistically through the first person personal pronoun and the PRES verbal tense (Benveniste 1966). Banfield (1982) challenged this idea by showing that subjectivity can be identified for other tenses (the IMP) and pronouns (third person), specifically for their usage in the free indirect discourse (containing represented speech and thought). Scholars (Reboul 1992; Vuillaume 1990; Tahara 2000) extended subjectivity to include the second person pronoun and argued that subjectivity can occur in narratives that are not free indirect discourse, such as in (574). Verbs in italics are in the preterit form (PS) and they express in this fragment the advancement of time seen from Emma's point of view (she was terrified and exhausted).

(574) L'idée qu'elle venait d'échapper à la mort faillit la faire s'évanouir de terreur; elle ferma les yeux; puis elle tressaillit au contacts d'une main sur sa manche: c'était Félicité. – Monsieur vous attend, madame, la soupe est servie. Et il *fallut* descendre! Il fallut se mettre à table! Elle *essaya* de manger. Les morceaux l'étouffaient. (Flauber, *Madame Bovary*, cited by Tahara, 2000)
'The idea that she had just escaped death failed to make her pass out of terror; she closed her eyes; then she shuddered in contact with a hand on her sleeve: it was Félicité. – Sir (Charles) is waiting for you madam; the soup is served. And she *had* to go downstairs! She *had* to sit to the table! She *tried* to eat. The bites of food suffocated her.'

Banfield argued that the subjective or objective reading of the sentence is conveyed by tense and deixis. In other words, the PS conveys the objective narration of the seeing of the moon event and the IMP conveys the subjective experiencing interpretation.

(575) Elle *vit* la lune.
'She saw the moon.'
(576) Elle *voyait* la lune (maintenant)
'She saw the moon (was seeing the moon now).'

Grisot (in preparation) aimed at testing experimentally the link made in the literature between subjectivity and verbal tenses. One of the first problems encountered was the lack of a generally accepted definition and of a list of sources of subjectivity. Smith (2003) is the only study who distinguished among most plausible sources, classified them and used them in a formalized model in DRT framework. She distinguished the following classes: *communication, contents of mind, evaluation and evidentials,* and *perception and perspective*. The last category consists of perception and perspectival cues[196]. Perception may be direct (a verb of seeing or hearing

---

[196] The *communication* category includes sources of subjectivity such as discourse structure (direct, indirect speech and reported speech with their respective corresponding deictics such as person and tense), verbs of communication (*say, ask, request, declare, confess, advise, insist, claim, shout, read, sing, remark, note, announce,* and many others) and performative verbs (*command, swear, promise, baptize*). The *contents of mind* category consists of expressions of mental states such as thoughts, beliefs and attitudes. The third category, *evaluation and evidentials*, is a rich category containing adverbs expressing speaker's attitude (*frankly, honestly*), adverbs expressing speaker's commitment (*clearly, likely, certainly*), adjectives (*peculiar, surprising, certain*) and modals in their epistemic reading (*must, may, can*), other verbs (such as *seem, appear, suggest*). As far as evaluation in concerned, there are evaluative discourse markers (*yet, anyway, still, but, after all*), gradable and non-gradable terms (see below Chen's fine-grained classification of linguistic cues).

introduces a complement which expresses the object of the perception), contextual (such as cases where the first sentence presents the perceiver and the second sentence conveys the object of the perception) and inferred (cases where the situation implies a perception due to world knowledge). Perspectival cues are related to the assignation of reference (temporal and spatial deictics), reflexive and possessive pronouns and imperfective viewpoint.

Grisot (in preparation) followed Smith (2003) and described subjectivity by the correspondent linguistic cues using her fourth category, namely perception (direct, indirect and inferred) and perspectival cues related to the assignation of reference (temporal and spatial deictics), reflexive and possessive pronouns and the imperfective viewpoint. The speaker's ability to judge this feature was tested experimentally for EN and FR verbal tenses. Subjectivity was hence defined as *the speaker's psychological perspective and perceptions included into the description of a situation.* With respect to this definition, sentences can be *subjective* or *not subjective* (objective). A sentence is *subjective* when the description of a situation or a series of situations is *centred* on the speaker's psychological perspective. A sentence is *not subjective* when the speaker merely reports a situation or a series of situations that are related in the world.

The experimental design is similar to experiments described in Chapter 6 of this thesis. The first experiment was carried out on 99 items containing occurrences of the SP. The two judges received annotation guidelines containing the definition of subjectivity and a list of its sources as explained above, and went to a training phases on 10 items. The two judges agreed on the label for 54 items (55%) and disagreed on 45 items (45%). For this data, the Kappa has a value of 0.0045, which is close to agreement due to chance. Among the items that received identic labels, the sources of subjectivity are most often linked to the type of perception. In very few cases, the judges based their decision on a perspectival cue, such as temporal or spatial deictics, reflexive and possessive pronouns and imperfective viewpoint. This data annotated with the subjectivity feature shows that subjectivity is a feature hard to pin down for native speakers of EN.

The second experiment was carried out on FR data, containing occurrences of the PS, PC and IMP. In total there were 80 items. Participants received annotation guidelines containing the definition of subjectivity and its sources. They went through a training phase on 10 items. The two judges agreed on the label for 58 items (73%) and disagreed on 21 items (27%), which represent a Kappa value of 0.31. When the explanations provided by the two judges were investigated, an unexpected pattern emerged. Among the 58 cases of agreements, in 47 cases the judgement was based on the grammatical aspect of the considered verb (81%). Specifically, the imperfective viewpoint was correlated to subjective usages (and therefore, by exclusion, the perfective viewpoint was correlated to non-subjective usages). In 19% of the cases, the judgement was based on the cues linked to the type of perception and perspectival cues other than grammatical aspect.

Hence, Grisot (in preparation) found that the [±subjectivity] feature is not linked to Tense. She suggested two possibilities: either [±subjectivity] is procedural feature linked to Aspect constraining the construction of implicatures and not truth-functional or it is an implicature carried trough general inference. These two assumptions must be tested experimentally in further work. Žegarac (1991, Chapter 3) argues against the possibility that

the subjective point of view is conveyed by a particular aspectual category (i.e. Aspect) and that it is encoded information, in his words 'pertaining to linguistic competence' (p. 78).

There are also scholars who do not completely reject his possibility. Trnavac (2006) shows in a corpus-based analysis that modal meanings are expressed with both imperfective and perfectives. However, subjective modal meanings most often correlate with imperfective aspect. Boogaart (2007) suggests, in application to Romance and Germanic languages, that modal and subjective readings of the imperfective are related because they both represent specific instantiations of the underlying anaphoric semantics of the imperfective. Specifically, imperfective aspect imposes the constraint that the eventuality is simultaneous to some previously established point of reference R (i.e. E=R). The reference point R can be instantiated as a subjective point of perspective or a point of evaluation for the truth-conditional content of the clause. Boogaart suggests that when R is instantiated as a subjective point of perspective, the imperfective has a subjective modal interpretation. For example, in (577) the R of the IMP is given by the event presented in the preceding discourse which is a PS in this case. The R of the PS is a point of perspective in which John notices the room was dark. It can be identified by means of pragmatic inferencing: it was probably dark before, at and after the moment in which John entered the room.

(577) Jean *entra* dans la chambre. Il *faisait* noir comme dans un four.
       John entered-PS the room. It was-IMP pitch dark.
       'John entered the room. It was pitch dark.'

Trnavac (2006) argues that one can find evidence that supports the hypothesis about an imperfective aspect-subjectivity correlation, evidence that goes against this hypothesis and evidence that is neutral to this hypothesis.

As far as this thesis is concerned, I suggest that both possibilities should be taken into consideration and tested experimentally.

7.2.3 Layers of temporal meaning

I adopt Moeschler's view about a complex layered meaning, consisting of non-inferential and inferential layers, and I propose a description of temporal meaning expressed by Tense, Aspect and Aktionsart as follows. From a morphosyntactic point of view, Tense and Aspect are, together with Mood, interpretable features belonging to Infl, which is the functional head of the clause (Chomsky 1981, 1995, 2000; Pollock 1989; Cowper 2003, 2005). Languages differ in the way in which they make use of the features. For example, it seems that Romance languages have two separate projections of Infl, T-P (i.e. Tense phrase) and ASP-P, whereas EN has only one, as suggested by Giorgi and Pianesi (1997) for IT and EN, and Cowper (2005) for Spanish and EN.

From a semantic and pragmatic point of view, Tense and Aspect and Aktionsart encode procedural and conceptual information, which guide the interpretation process either through contributing or through constraining to the content expressed. Conceptual information most often represents a pro-concept TIME that must be adjusted contextually in the form of an ad hoc concept. As for procedural information, it operates at two levels: syntactic computation and pragmatic interpretation. The layers of temporal meaning are

summarized in Table 7-2.

Table 7-2 Layers of temporal meaning

RT (2004)	Conceptual/ procedural information	Temporal reference	Inferential status	Truth- functionality status
Explicature	Conceptual (contribution)	E/S; Aktionsart	Inferential	Truth-functional
	Procedural 1 (constraining)	Narrative vs. non-narrative (via R) Perfective vs. imperfective		
Implicature	Procedural 2 (constraining)	?[197]Subjective vs. not-subjective	Inferential	Non truth-functional
Implicature	General inference	?Subjective vs. not-subjective	Inferential	Non truth-functional

Firstly, inflectional morphemes (temporal and aspectual) contribute to the content of the proposition expressed:

- Temporal location of an eventuality with respect to S must be specified, and this is done at the level of the explicature. It represents inferential and truth-functional content.
- The type of eventuality or Aktionsart with respect to its actual realization is inferentially processed at the level of the explicature: the [±boundedness] feature.

Secondly, the presence of inflectional morphemes (temporal and aspectual) in a sentence constrains the interpretative process:

- Tense instructs the hearer to order temporally and causally eventualities. The result of this inference is an explicature and it is truth-functional content.
- Aspect instructs the hearer to identify the speaker's viewpoint of the eventuality expressed. The result of this inference is an explicature and it is truth-functional content.

Finally, as far as the subjectivity feature is concerned, two hypotheses are proposed (as indicated in Grisot, in preparation). The first is that the presence of inflectional aspectual morphemes in a sentence constrains the interpretative process: it instructs the hearer to identify the speaker's subjective perspective on the eventuality. The result of this inference is an implicature and it is not truth-functional. The second is that subjectivity arises due to a general inferential process and it is not triggered by procedural or conceptual encoded information.

In this section, I proposed a layered representation of temporal meaning established based on conceptual information contributing to truth-conditional content of an utterance and on procedural information constraining the formulation of the explicature and implicatures associated to an utterance. Table 7-2 also indicates a concrete example of the

---

[197] The question mark signals that the suggestion is a hypothesis, which requires experimental investigation.

possibility of having two types of procedural information, as suggested by Wilson and Sperber (1993) (see Figure 3-2 in section 3.1.3.3). The first type represents constrains on explicatures and is truth-conditional whereas the second type embodies constraints on implicatures and is not truth-conditional.

Additionally, I argued in favour of a holistic interpretation of temporal information coming from various sources and I suggest that temporal coherence takes place both at the discursive and cognitive levels. The following section is dedicated to this matter.

7.3 Temporal coherence

The essential feature that makes a piece of a text a discourse is the coherent succession of sentences forming a whole and referring to the same entities (nominal or eventualities). However, example (578) is not coherent even if "he" can refer to "John" (Hobbs 1979, 67). Hobbs argues that there is an expectation of coherence, which is deeper than the notion of a discourse just being "about" some set of entities:

(578)   John took a train from Paris to Istanbul. He likes spinach.

Coherence was principally characterized in terms of *coherence relations.* Coherence relations have been investigated from three points of view: theoretical linguistics, computational linguistics and psycholinguistics. Theoretical linguistics described the factors that contribute to discourse coherence and has sought to categorize the different types of coherence relations that have the role to connect clauses and sentences. Halliday and Hasan (1976) proposed the terms *cohesion* and *cohesive ties* for the linguistic devices used to build coherence between sentences[198]. They define *cohesion* as a semantic concept and a property of a text that occurs when the interpretation of some elements in the discourse is dependent on that of another. They identify *grammatical* and *lexical cohesion* and each type has its own cohesive ties and methods. Grammatical cohesion involves methods such as *substitution*, *anaphora* and *ellipsis.* Anaphora is exemplified in (579) where the pronoun *she* has Mary as antecedent and *done* sends back to *send a picture of the children*. Ellipsis is exemplified in (580) where the elliptical verbal group is the second one. Grammatical cohesion consists of grammatical cohesive ties such as pronouns, discourse connectives and verbal tenses whereas lexical cohesion occurs through repetition of a word, for example as in example (581), where the repetition of the word *apple* has a cohesive effect.

(579)   Mary promised to send a picture of the children, but she hasn't done.
(580)   Are you dieting? I have been for some time.
(581)   Wash and core six apples. Put the apples into a fireproof dish.

As far as relations that link clauses are concerned, there is an enormous variability both for their names and types. The term of "coherence relations" is due to Hobbs (1979), who investigated them from a computational linguistics perspective. As Kehler (2004) suggests, all proposals are based on data analysis but they do not pursue the goal of descriptive accuracy

---

[198] For a critical discussion on cohesion, see Reboul and Moeschler (1998).

to the same extent. He points out that "an explanatory theory of coherence requires a set of externally driven principles to motivate and ultimately constrain the relation set." This is the direction taken by Sanders and colleagues (Sanders et al. 1992, 1993; Sanders and Noordman 2000; Sanders 1997, 2005) by proposing a theory in which psychological plausibility is the primary motivating factor:

> Understanding a discourse means constructing a mental coherent representation of that discourse by the hearer. […] An adequate account of the relations establishing coherence has to be psychologically plausible, because coherence relations are ultimately cognitive relations. (Sanders et al. 1992, 1)

Sanders and colleagues (1992) argued that coherence relations point to coherence in the cognitive representation of discourse and they see coherence relations as *cognitive entities* (Hobbs 1979; Mann and Thompson 1986; Sanders et al. 1992, 1993). Sanders (2005) argues that *causality* and *subjectivity* are two fundamental cognitive and discursive principles.

From the perspective of human comprehension, temporal reference in discourse is established based on three components: Tense, Aktionsart and Aspect. Temporal coordinates S, R and E combine with the predicate's Aktionsart and contribute to the explicature of the utterance, whereas procedural information encoded by Tense and Aspect constrain the formulation of contextual hypotheses and implicated conclusions. The hearer makes use of all three components in order to recover the speaker's meaning, that is, her overtly intended content.

From a different perspective, I would like to argue that the human brain tends to treat these different sources of temporal information in a coherent whole. I suggest the notions of *cognitive temporal coherence* and *discourse temporal coherence*. This idea is based on Sanders et al.'s cognitive approach of coherence relations[199] (also Hobbs 1979, 1985; Mann and Thompson 1986). In a series a works (Sanders et al. 1992, 1993; Sanders 2005 among others), Sanders and his colleagues argue that 'coherence relations are ultimately cognitive relations' (Sanders et al. 1992, 1, 3). For them, 'understanding a discourse may be regarded as the construction of a mental representation of the discourse' by the hearer/reader (Sanders et al. 1992, 1).

Sanders and Noordman (2000) show through experimental work that coherence relations (causal vs. additive vs. contrastive), as well as their linguistic marking (explicit vs. implicit) affect text processing. Precisely, they found that causal relations and explicit marking resulted in faster processing and are better recalled[200]. This could mean that discourse relations and their explicit marking are important for processing and for the building of mental representations of the content of the discourse. Previously, Sanders (1997) argued that readers make a match between the relational meaning of the connective and the

---

[199] Sanders et al. (1992) propose a classification of coherence relations according to the *relational criterion*, that is, 'a criterion concerning the meaning of two or more discourse segments that cannot be described in terms of the meaning of the segments in isolation' (p. 25). According to this criterion, discourse relations are *causal* and *additive*. Each type consists of a series of more fine-grained types, defined with respect to the source of coherence (semantic vs. pragmatic), the order (basic vs. non basic) and polarity (positive vs. negative). Temporal relations are in this classification a type of additive relations.

[200] Participants were asked to tell what they remembered from the short narrative they received in the experiment. Their output was evaluated with respect to the quantity and the quality of the information remembered.

meaning of the content of the segment. This finding is consistent with the assumptions from RT with respect to procedural and conceptual information: procedural information represents instructions to manipulate mental representations consisting of conceptual information. Sanders and Noordman (2000) also found a difference regarding the type of discourse relations. Causal relations are more strongly connecting than additive relations, therefore language users have a preference for causal over additive relations. Sanders and Noordman's interpretation is that 'the reader will arrive at an additive relation only if no causal relation can be established'.

In this thesis, I am interested in temporal relations and some of the linguistic markers used to express them[201]. My suggestion was that temporal reference takes place at two levels. The first is location of eventualities with respect to S. The second is locating eventualities with respect to each other. These two types of location can be either explicitly marked through discourse markers such as *before, after, and then, simultaneously, in the same time* and the kind, or implicitly expressed. In both cases, the temporal relation must be determined pragmatically through inference by the speaker. My prediction is that the text is processed faster when discourse connectives encoding procedural information are explicit. However, procedural information encoded by Tense and Aspect is sufficient to guarantee the recovery of the speaker's intended meaning.

While processing an utterance/a series of utterances, the hearer has at his disposal conceptual and procedural information about temporal knowledge. He treats it in a coherent way for recovering the speaker's meaning. Temporal coherence takes place, therefore at two levels:

- At the cognitive level, when we speak about the hearer's processing of the utterance.
- At the discursive level, when we speak about discourse relations and their linguistic marking.

A multifactorial analysis of the data gathered in this research revealed that the different sources of temporal information in discourse are significantly correlated and have significant interactions (as discussed in section 6.1.8). Precisely, the occurrence a verbal tense can be predicted based on the contextual values of the [±narrativity], [±boundedness] and [±perfectivity] features. For example, for EN SP all combinations of features are possible. Nevertheless, two main tendencies were observed. The first main tendency is perfective viewpoint associated with bounded situations in narrative contexts whereas the second is imperfective viewpoint associated with unbounded situations in non-narratives contexts. As for the FR PC/PS and IMP, the best predictive model provides two statistically significant factors and one interaction: procedural types of information encoded by Tense and by Aspect, as well as the interaction between Aktionsart and procedural information encoded by Tense.

From this data, it was seen that all these sources of temporal information occur in a coherent discourse. Eventualities are temporally (or/and causally) related and this relation is determined by the contextual value of Tense, Aspect and Aktionsart. Experiment 3 from section 6.1.3 also showed that judges (i.e. readers/hearers) agreed more often on their

---

[201] I did not deal directly with temporal connectives and temporal adverbials. However, this matter should be considered in further research.

judgment of items (i.e. pieces of discourse) when they were asked to insert a connective that would render explicit the implicit temporal relation. Furthermore, one can make the hypothesis that human brain treats sources of temporal information in a coherent manner. Cognitive temporal coherence is assumed to mentally represent the perceived temporal discursive coherence.

Experiment 3 also revealed that speakers make a connection between temporal and *causal* relations among eventualities. Precisely, the two judges were asked to propose a connective that would render explicit the existent implicit temporal relations. They proposed the causal connectives *because* and *as a consequence*. In Sanders' cognitive approach of coherence, it is argued that *causality* and *subjectivity* are two fundamental cognitive and discursive principles. The question that arises at this point of the discussion is about the status of *temporality* with respects to *causality* and *subjectivity*. Before suggesting a possible answer to this question, I would like to come back to the findings of Grisot (in preparation) with respect to subjectivity in EN and FR. For both languages, the hypothesized link between subjectivity and Tense was not validated. Additionally, judges pointed to Aspect as the best 'cue' of subjectivity. However, this link has to be validated experimentally (see for example Trnavac 2006 for an exploratory corpus-based and contrastive investigation in this direction). On the other hand, subjectivity has vastly been attested for connectives in Dutch and French (for example Knott and Sanders 1998; Sanders and Noordman 2000; Pit 2003, 2006; Pander Maat and Degand 2001; Pander Maat and Sanders 2000, 2001) and it is assumed to be a fundamental cognitive principle.

I would like to suggest several possibilities of accounting for these findings. Firstly, regarding temporality and causality, one could argue that (i) temporality is dependent on the processing of causality or (ii) temporality is an independent discourse and cognitive principle. Sanders et al. (1992) however, decided not to propose temporality as a basic categorizing principle for two reasons. Their first reason is that temporal meaning is too dependent on the referential content of the segments and temporality cannot be ignored by language users whereas causality can. Their second reason is that it is not a categorizing principle as productive as causality and additively. Since Sanders and colleagues reject temporality as a categorizing principle for discourse relations, they don't count it as a fundamental cognitive and discursive principle. My suggestion is that temporal information is a crucial component of discourse comprehension and that temporality has cognitive relevance. Regarding the productivity of temporality for categorizing discourse relations, I cannot argue for or against Sanders et al.'s position because in this research I did not investigate this particular issue.

Coming back to the possible relations between temporality and causality, if temporality has independent cognitive relevance, then we expect to find a similar pattern as for causality and the linguistic markers associated with it. Precisely, conceptual contents of Tense and of Aktionsart contribute to building the mental representation whereas procedural meaning of Tense and Aspect provide instructions about manipulating these conceptual representations. Moreover, temporal connectives and temporal adverbials also play a role for expressing temporal reference, as indicated in section 0 with respect to the relation between temporal adverbials and Aktionsart and in section 3.1.3.6. with respect to temporal connective and their explicitation of otherwise implicit temporal relations. One prediction is therefore that when the contextual values of Tense and Aspect (narrative vs. non-narrative, perfective vs.

imperfective) are explicitly expressed through temporal connectives and adverbials, the text is processed faster than in the opposite condition, when they are implicit. On the contrary, if temporality were dependent on causality, not only temporal but also causal connectives would have a facilitating role for processing a text with respect to temporal information. Experiment 3 seems to provide evidence in favour of the second possibility. In my opinion, these two possibilities have to be tested experimentally with a carefully built design using both offline and online data in further research.

Secondly, regarding temporality and subjectivity, I think the situation is even more complex. One of the problems concerning subjectivity is the multitude of definitions and approaches existing in the literature. I think that until the moment when scholars will define subjective in a similar, if not identic way, there is no comparison possible among the findings. Grisot (in preparation) reduced subjectivity to Smith's (2003) *perception* and *perspective* category and used notion of *subject of consciousness* (Pander Maat and Sanders 2001; Pit 2003) to describe subjective and non-subjective sentences. Grisot (in preparation) found that subjectivity is not directly linked to Tense and suggested that it could be linked to Aspect but this was not, at least as far as I know, validated experimentally. Both Tense and Aspect are sources of temporality in discourse, the assumption, is that there is a link between temporality and subjectivity. However, I think that further research is needed in order to investigate (i) what is the exact nature of link between temporality and subjectivity, (ii) if temporality is an independent cognitive principles next to causality and subjectivity and finally (iii) if temporality is dependent on causality.

7.4   Verbal tenses monolingually

The model defended in this thesis, recalling the procedural pragmatic approach of FR verbal tenses (Moeschler et al. 1998, Moeschler 2000a,b, 2002; Saussure 2003), assumes that verbal tenses underdetermine the speaker's communicated content. The hearer must therefore recover inferentially the speaker's intended meaning with respect to temporal reference, which is defined broadly. The hypothesis formulated in this research is that Tense, Aspect and Aktionsart are parameters considered by the hearer in the interpretative process and that the human mind tends to treat these parameters in a coherent manner.

Based on the theoretical model described in section 7.2 several predictions can be made for individual verbal tenses in EN, FR, IT and RO. They all share the following features and this represents the common *tertium comparationis* required for rendering possible their contrastive analysis:

- Their meaning is underdetermined and must be contextually worked out
- They encode conceptual and procedural information: E/S and [±narrativity]
- The same form expresses both Tense and Aspect: [±perfectivity]
- They apply to all types eventualities: [±boundedness]

The cross-linguistic investigation carried out in this thesis showed not only that these parameters are operationalized differently in each language, but also that they receive values that change from a context to another. Table 7-1 provided in section 7.2.1.2 and repeated below as Table 7-3 points out that the situation in EN is quite different than in Romance

languages[202]. A Fisher Exact Probability Test shows that the difference between the SP and each of the verbal tenses used in a target language is statistically significant ($p<0.05$). One of the reasons for this is that the PC in the Romance languages investigated entered in the aoristicization process whereas the EN PresPerf remained a perfect, with resultative and non-narrative usages. Consequently, there is a competition between the PS and the PC for operationalizing narrative contexts only in Romance languages. Moreover, the IMP in Romance is not specialized for non-narrative usages and has only a partial correspondence with the imperfective aspect. An empirical and experimental comparison between the EN progressive and the IMP in Romance is required in order to have an accurate understanding. This should be carried out in future research.

Table 7-3 [±Narrativity] feature in English and Romance

Language	VT	Narrative	Non-narrative
EN	SP	59%	41%
	PS	**92%**	8%
FR	PC	**77%**	23%
	IMP	16%	**84%**
	PS	**96%**	4%
IT	PC	**88%**	12%
	IMP	16%	**84%**
	PS	**93%**	7%
RO	PC	**83%**	17%
	IMP	19%	**81%**

Within the Romance area, these numbers indicate that IT and RO are more advanced in the aoristicization process than FR: 88% for the IT PC and 83% for the RO PC compared to 77% for the FR PC. The difference between FR and each of the other two Romance languages is shown to be statistically significant through a Fisher Exact Probability test applied to a 2x2 contingency table ($p=0.03$). The difference between IT and RO is not statistically significant. As far as the IMP is concerned, Table 7-3 indicates that the RO IMP has more often narrative usages than the IT and FR IMP, however this difference is not statistically significant.

As for Aspect and Aktionsart, empirical work carried out in this thesis revealed differences between EN and FR[203]. With respect to the [±perfectivity] feature, the difference between EN and FR is statistically significant both for the PC/PS and for the IMP (Fisher Exact Probability test with $p<0.05$). As for the [±boundedness] feature, only the difference between the SP and the IMP is statistically significant (Fisher Exact Probability test with $p<0.05$). As far as this thesis is concerned, no experiments were carried out with aspectual

---

[202] The values written in bold signal the cases of high frequency of the association of a verbal tense and one of the value of each feature.

[203] The total of values for each verbal tense should be considered per feature : [±perfectivity] and [±boundedness].

information for IT and RO verbal tenses.

Table 7-4 [±Perfectivity] and [±Boundedness] in English and French

Language		Perfective	Imperfective	Bounded	Unbounded
EN	SP	46.9%	53.1%	48.3%	43.9%
FR	PC/PS	**33.1%**	8.3%	**47.8%**	34.9%
	IMP	11.2%	**44.8%**	10.8%	**41.4%**

In what follows, I will propose a description of the usages of each of the verbal tenses considered in the four languages under study. This section is mainly intended as a parallel section to the classical description of these verbal tenses provided in section 2.5.1. The reanalysis suggested in these pages is characterized by the separation of the categories of Tense, Aspect and Aktionsart and their meanings. My assumption is that an accurate description of the usages of a verbal tense may be done only if purely temporal (i.e. triggered by Tense) interpretations are separated by aspectual interpretations (i.e. triggered by Aspect and Aktionsart). Moreover, describing the usages of these verbal tenses in terms of procedural information triggered by Tense, Aspect and Aktionsart, as well as ad-hoc concepts associated to Tense, is a model that can account for all their descriptive and interpretative uses, as they were depicted in section 2.5.1. I will give concrete examples in what follows regarding each verbal tense analysed.

7.4.1 English

The EN SP is expressed through the *–ed* morpheme and through irregular forms of certain verbs. It's meaning is underdetermined and must be contextually worked out. The SP participates at determining temporal reference in a broad sense at two levels: location of eventualities with respect to S via conceptual information and location of eventualities with respect to one another via procedural information. I would like to argue that all the usages and features identified and described in grammars of English (as temporal and non-temporal usages, see section 2.5.1.1.2) can be accounted for based on the cross-linguistic valid parameters suggested in this thesis. In the following lines, I will address each of them.

Conceptual information

The SP encodes conceptual information in the form of a pro-concept TIME, which can be operationalized as the localization of E with respect to S. In its temporal usages, the SP locates eventualities in the past, i.e. a moment prior to S. As I have argued in section 7.2.1.1, Reichenbachian coordinates are variables that have to be saturated contextually. Their contextual saturation takes place, in the interpretation process, at the level of the explicature of the utterance. This process is inferential and the content is truth-conditional. The contextual value of this conceptual information must be compatible with temporal information, coming from other sources, such as temporal adverbials for example as shown

in examples (582)-(585). The ungrammaticality of sentences in (584) and (585) shows that the requirement that E and R must be different than and previous to S.

(582) Yesterday, he *walked* to school.
(583) Today, he *walked* to school.
(584) *Now, he *walked* to school.
(585) *Tomorrow, he *walked* to school.

With respect to the competition between the SP and the PresPerf, the difference between the two verbal tenses lies in the position of R (i.e. R=E and, respectively, R=S). My assumption is that the speaker's choice between the two possible forms depends on her intention to focus on the eventuality that is located in the past time or on the current resultative state/current relevance of a past eventuality. Experiment 6 showed that native speakers[204] identify this difference with easiness and without errors (inter-annotator agreement rate measured with the Kappa coefficient is 1. My hypothesis is that the PresPerf in British EN has remained a true perfect expressing that the eventuality dented by the verb is already completed at the time of reference, which is S for PresPerf. The PresPerf expresses therefore a strong temporal link with the present time (S). This aspectual information might facilitate the hearer's access to the relation R=S. In other words, judging procedural information could be rendered easy to access through conscious thinking when there the hearer has supplementary sources of information.

As pointed out in grammars of EN, the SP has non-temporal interpretations, such as modal, conditional and, respectively, politeness usages as in (586)-(588) from Huddleston and Pullum (2006) and Aarts (2011). The model defended in this thesis accounts for these usages as well, though it is not its primary aim.

(586) I wish they *lived* nearby.
(587) If you *left* now, you would miss the rush-hour traffic.
(588) I *wanted* to ask you a little about the Exorcist.

I argued that among the TAM markers, Mood is situated higher in the hierarchy than Tense and Aspect. The two possible values of Mood are realis and irrealis. My suggestion is that the hearer can arrive at temporal interpretations of the SP only when Mood has the realis value. On the contrary, modal, conditional and politeness interpretations are possible only when Mood has the irrealis value.

Procedural information

The SP encodes procedural information, which is operationalized in this thesis as the [±narrativity] feature. This procedural information instructs the hearer to locate

---

[204] As pointed out in section 6.1.6 (Experiment 6), the participants of this experiment where originate from United Kingdom. It is acknowledged that the difference between the SP and PresPerf is kept in British English and lost in American English. One could make the hypothesis that in American English, the PresPerf passed through the aoristicization process and lost it's perfectal aspectual interpretation that is correlated with non-narrative usages of Tense.

eventualities with respect to one another and to verify for temporal relations. Examples (589) and (590) illustrate the fact that the interpretation of the SP in the first sentence is dependent on the following sentence. Indeed, in (589) the two eventualities are temporally and causally related, and the SP in the first sentence presents an atelic eventuality (to sleep) perfectively and bounded. In (590) the two eventualities are simultaneous from a temporal point of view and are not causally related. Moreover, the SP in the first sentence presents an atelic eventuality (to sleep) imperfectively and bounded. My suggestion is that in (589) the [±narrativity] feature encoded by the SP has the narrative value whereas in (590) it has the non-narrative value. Furthermore, the SP from the first sentence in (589) and (590) can be judged as ambiguous with respect to its contribution to the overall temporal (and causal) interpretation of the discourse.

(589) John *slept*. He got rest.
(590) John *slept*. He had a dream.

I argued this information is encoded by Tense (in contrast with other theoretical assumptions stipulating that temporal relations should be treated as general implicatures, see section 3.1.3.5) using empirical and syntactical arguments. Firstly, annotation experiments from Chapter 6 revealed that the [±narrativity] feature displays the characteristics suggested for procedural information by Sperber and Wilson (1993) and Moeschler et al. (2013): it is inaccessible to consciousness and it is not available through conscious thought but it can be rendered explicit through discourse markers, such as connectives. Judges dealing directly with this information in four languages (EN, FR, IT and RO) agreed on their judgment with agreement rates (measured with the Kappa coefficient) that are beyond chance agreement. Disagreements were resolved through the explicitation of this implicit procedural information, namely, insertion of connectives such as *and (then), because* for the narrative value and *and (in the same)* for the non-narrative value. Example in (180) given in section 2.5.1.1.2 and repeated below as (591) illustrates both the narrative and non-narrative usages of the SP. The eventualities *grabbed* and *twisted* are temporally successive, temporal relation that can be rendered explicit through the pragmatic connectives *then* or *and then*. The SP has in these cases a narrative usage. As for eventualities *let go* and *was* are temporally simultaneous and this temporal relation is signalled by the temporal adverbial *when*. The SP has in these cases a non-narrative usage.

(591) I *grabbed* his arm and I *twisted* it up behind his back and when I *let go* his arm there *was* a knife on the table and he just *picked* it up and *let* me have it and I *started* bleeding like a pig.

Secondly, I suggested that the [±narrativity] procedural feature is encoded by Tense contained in the Inflection category (Infl-P), which is the functional head of the sentence (see section 3.2). Escandell-Vidal and Leonetti (2011) advocate that procedural information can operate at two levels: that is syntactic computation and that of interpretation. Their assumption is that there are instructions that operate primarily at the syntactic level and are not 'visible' at the interpretative level. Conversely, there are instructions which, in addition to their syntactic role, are crucial for the interpretative process. These correspond to what

has been called procedural information in RT (2011, 84). Following Escandell-Vidal and Leonetti's idea, I want to suggest that the [±narrativity] procedural feature operates both at the syntactic and the interpretation levels. At this level of the research, this suggestion is a hypothesis that has to be tested experimentally. I will submit some predictions for experimental testing in section 8.3.

Aspect

Grammars of EN, such as Huddleston and Pullum (2006), point out that the SP expresses both the perfective and imperfective viewpoints. This idea was confirmed in Experiment 5. The data containing 435 SP tokens was translated into Serbian, a Slavic language where grammatical aspect is marked morphologically, and the contextual value of Aspect for each SP token was transferred back in EN based on the properties of translation corpora. The SP was labelled as perfective for 204 items (47%) as in (592) and as imperfective for 231 items (53%) as in (593).

> (592) I welcome the consultation process and can assure colleagues that in my Member State the authorities *took care* to carry out a broad and meaningful consultation. (EuroParl Corpus)
> 'Podržavam savetodavni proces i mogu da uverim kolege da *su se* nadređeni u mojoj državi *pobrinuli* da realizuju jasan i značajan dogovor.'

> (593) He certainly was wonderfully handsome. People who *did* not *like* him, philistines and colege tutors, and young men reading for the Church, used to say that he was merely pretty. (Literature Corpus)
> 'Bio je izrazito lep, prelep. Ljudi koji ga *nisu voleli*, filistri i učitelji i crkveni mladići bi govorili da je bio samo lep.'

For 7 items, the translator was free to choose between perfective and imperfective, both aspects being possible. The verbs that occurred in these sentences are *to promise, to spend, to reproach, to organize, to despise, to stay* and *to try*. All these verbs express atelic situations, as in example (594).

> (594) Musharraf and his political allies *tried* to adjust to this new reality, but their patience ran out when the Supreme Court took up petitions against Musharraf's decision to run for president. (Journalistic Corpus)
> 'Mušaraf i njegovi politički saveznici *pokušali su/pokušavali su* da dodaju ovome novu stvarnost, ali su postali nestrpljivi kada je Vrhovni sud prihvatio peticije protiv Mušarafove odluke da se kandiduje za predsednika.'

Žegarac (1991) described the SP as being associated to the simple aspect, one member of the simple/progressive opposition (as discussed in 0). Knowing that progressive aspect appertains to imperfective aspect (as suggested by Comrie 1976), the question that arises is whether simple aspect appertains to perfective aspect. The results of Experiment 5 seem to suggest that this question should receive a positive answer. This issue should however further investigated regarding the simple/progressive opposition in EN and the perfective/imperfective aspectual opposition in Slavic languages.

Grammars of EN point out that the SP is compatible with all aspectual classes. Classically, Aktionsart is classified in EN with respect to the criterion of compatibility with the progressive and it is most often referred to as states and events (containing activities, accomplishments and achievements). In this thesis, information about Aktionsart was operationalized in terms of *boundedness*, which represents the actual and contextual realization of telic and atelic eventualities. In Experiment 4, eventualities expressed through the SP were judged as being bounded or unbounded based on a series of linguistic tests (as discussed in section 2.4.1). The results showed that the SP is compatible with both bounded and unbounded eventualities. In this experiment, the two judges had a very high agreement rate. According to Sperber and Wilson's (1993) description of the cognitive foundations of the conceptual/procedural distinction, the information dealt with in Experiment 4 is conceptual. The main assumption defended in these lines is that Aktionsart, and more specifically, (a)telicity is underdetermined and it must be contextually worked out. The actual realization of (a)telicity is done in terms of (un)boundedness. Due to its conceptual nature, this aspectual information contributes the to the propositional content of the utterance and it is worked out at the level of explicatures. The SP, a preterit form, is compatible with both values of this parameter, bounded and unbounded. Moreover, there is a cross-linguistic correlation between EN and FR: SP expressing a bounded eventuality is more often translated into FR through a PC/PS form as in (595), and SP expressing an unbounded eventuality is more often translated into FR through an IMP form as in (596).

(595) Where Bulgaria and Romania are concerned, we *decided* to support these two countries in their efforts to achieve their aim of accession in 2007. (EuroParl Corpus)
'En ce qui concerne la Bulgarie et la Roumanie, nous *avons décidé* de soutenir ces deux pays dans leurs efforts pour atteindre l' objectif de l' adhésion en 2007.'

(596) At that time, Arafat was still regarded as a terrorist, he was barred from obtaining a visa to go to the United Nations headquarters in America and there *was* no general consensus regarding the diplomatic strategies to be adopted. (EuroParl Corpus)
'C'était l ' époque où M. Arafat était encore considéré comme terroriste, il n ' avait pas de visa pour se rendre en Amérique au siège des Nations unies et il n'y *avait* aucun consensus quant aux stratégies politiques à adopter.'

In this section, I have argued that the EN SP is a verbal tense encoding conceptual and procedural information. Conceptual information contributing to the propositional form of the utterance is inferentially worked out at the level of explicatures. Procedural information constrains the interpretative process through instructing the hearer to locate eventualities with respect to one another. As for its relation to aspectual information, the SP combines felicitously both with Aspect and Aktionsart.

7.4.2 French

In this thesis I investigated four FR verbal tenses: the preterit or aorist PS, the perfect PC, the IMP and the PRES. Traditionally, the first three tenses express reference to past time (for the IMP, only in its temporal interpretations). Numerous approaches aimed at

explaining the difference among them: classical, aspectual, anaphoric, textual and pragmatic approaches (see section 2.5.1.2). Among the pragmatic approaches, procedural pragmatics initiated by Saussure (2000) argued that these verbal tenses have descriptive and interpretative usages, the latter being triggered by the combination of semantic and pragmatic temporal procedures with contextual assumptions. The PS, the PC and the IMP encode instructions that guide the interpretative process. The main assumption is therefore that verbal tenses are underdetermined and that their meaning is determined inferentially based on the instructions encoded by Tense and Aspect for each of these tenses. Regarding the PRES, it is generally accounted as expressing reference to present time (E=S) but also to past time in its historical usage. In this thesis, PRES is accounted for from a theoretical point of view, and principally, with respect to its opposition to the PC, the PS and the IMP, established on the conceptual information E=S vs. E<S.

The model defended in this thesis assumes, on the one hand, that Tense encodes both conceptual and procedural information and, on the other hand, that it combines with Aspect and Aktionsart.

Conceptual information

The PS, PC and IMP encode conceptual information in the form of a pro-concept TIME, which can be operationalized as the localization of E with respect to S. These three verbal tenses share the same conceptual meaning most frequently expressed as the ad-hoc concept E<S (i.e. *pastness*). Similarly to the EN SP, the hearer builds contextually an ad-hoc concept, which specifies the temporal location of an eventuality with respect to S. All three coordinates E, S and R are variables saturated contextually based on linguistic and non-linguistic knowledge. As far as R is concerned, it accounts for the instruction encoded by Tense to locate eventualities with respect to one another (i.e. the [± narrativity] feature).

The PS, PC and IMP share conceptual information monolingually speaking, but also cross-linguistically (i.e. with the EN SP, as well as with the PS, PC and IMP in IT and RO). The analysis of translation corpus described in section 5.3 indicated that there is little cross-linguistic variation for the conceptual content of the EN SP, namely reference to past time. Specifically, past time tenses are used a TL in more than 72% of the cases while the PRES is used only in 5% of the cases. At this level of the content, the PS, PC and IMP are interchangeable. In actual usage, procedural information as well as computability with Aspect and Aktionsart provide supplementary information and reduce the number and type of cases when the PS, PC and IMP are interchangeable.

In section 2.5.1.2.4, I briefly described the FR PRES and its usages. The main assumption is that the PRES is opposed to the PS, the PC, and the IMP with respect to their conceptual information. More specifically, if the former tenses most frequently instantiate an ad hoc concept E<S, the PRES instantiates most frequently an ad hoc concept E=S. The results of Experiment 7, which tested whether native speakers provide the correct verbal tense in a given context, indicated that there is no ambiguity for participants to provide a verbal form expressing reference to past or present time. Specifically, Experiment 7 provided evidence the conceptual information encoded by verbal tenses, that is past vs. non-past, is determined contextually and that the agreement among the participants produced high

Kappa values: 1 for artificial data, 0.80 for natural data and 0.86 for all the data.

Considering that the meaning of a verbal tense is worked out with respect to its conceptual and procedural information, there are cases where the PRES is interchangeable with the PS, the PC and the IMP, namely, in their narrative usage. This usage of the PRES is called *present historique* (PH) 'historical present'. In this circumstance, the hearer builds for the PRES an ad hoc concept E<S based on contextual information. The question that arises at this point of the discussion regards what allows the shift from E=S to E<S, therefore from PRES to PH. The literature suggested that the shift is linked to the notion of *subjectivity* and the Free Indirect Discourse (Benveniste 1966; Banfield 1982; Schlenker 2004; Moeschler 2014; cf. Reboul et al. 2015 for a critical investigation of the previous proposals). Moeschler (2014) argued that subjectivity is a pragmatic feature of natural language and that the PH triggers two pragmatic effects: temporal sequencing [+narrative] and subjectivity [+subjective]. As for its semantics, the PH may be described through a configuration of Reichenbachian temporal coordinates: E, R and S. There are two possibilities that permit reference to past time. The first is E=R<S, which also corresponds to the PS or E<R=S which also corresponds to the PC. Moeschler's suggestion is to dissociate the tripartite configuration in three pairs of relations: E&R, R&S and the inferred relation E&S. For the PH, the situation is the following (2014, 7):

> Dans le PH, si E est cotemporel à R (E=R), la seule contrainte de R est qu'il soit distinct de S (R≠S). […] Ce qui est encodé linguistiquement dans le PH est la relation entre E et R, à savoir E=R. La disjonction R≠S est inférée pragmatiquement sur la base des traits pragmatiques [±narratif] et [±subjectif].[205]

In other words, a context allowing narrative and subjective pragmatic features permits the shift from inferring E=S with PRES to inferring E<S via R≠S with the PH. This description explains the lack of interchangeability between the PH and the other three FR verbal tenses expressing reference to past time. Firstly, the PH is not interchangeable with the PS, with it shares the [±narrativity] feature, because it is compatible with a subjective point of perspective. Secondly, the PH is not interchangeable with the PC, because it requires the disjunction R≠S. Finally, the PH is not interchangeable with the IMP, which has been described as a subjective verbal tense (as discussed in section 2.5.1.2.2) because it combines [±narrativity] and [±subjectivity] features.

As far as this thesis in concerned, my suggestion is that ad hoc concept of pastness (E<S) is constructed contextually based on positive values of [±narrativity] and [±subjectivity] features, as well as other cues as temporal adverbials. The [±narrativity] feature represents procedural information encoded by Tense validated experimentally for the PS, PC and IMP. Future research should investigate how the [±narrativity] feature behaves with the PRES and its usages, as the PH among the others.

Procedural information

---

[205] 'For the PH, if E is temporally simultaneous to R (E=R), the only constraint of R is that it must be different than S (R≠S). […] What is linguistically encoded in the PH is the relation between E and R, precisely E=R. The disjunction R≠S is inferred pragmatically on the basis of the pragmatic features [±narrative] and [±subjective].' (my translation)

Procedural information encoded by the PS, PC and IMP is operationalized in this thesis as the [±narrativity] feature. Experiment 1 and Experiment 2 investigated the behaviour of these three verbal tenses with respect to the [±narrativity] feature. In the literature it was observed that the PS and the PC are more often used in narrative discourses whereas the IMP is used in non-narrative discourses where it expresses background information. These observations received different types of explanations, as described in section 2.5.1.2. One of these explanations was suggested in the procedural pragmatics framework (Saussure 2003), according to which all verbal tenses encode uniquely procedural information. In this framework, it is assumed that the PS encodes by default the instruction for temporal progression, that the IMP instructs the hearer to build an unsaturated P variable interior to the event (which will be saturated contextually either as R or as a moment of consciousness C) and, finally, that the PC has a base value where it locates the eventuality prior to S (E<S) and two contextual values distinguished by the position of R: R=E in its *anteriority* usage and R=S in its *resultative* usage.

According to the model defended in this thesis, I want to suggest that the PS, PC and IMP encode procedural information operationalized as the [±narrativity] feature, that is, they instruct the hearer to determine if the eventualities expressed are temporally and/or causally related. A positive value of this feature points to a narrative usage of the verbal tense considered whereas a negative value of this feature points to a non-narrative usage of the verbal tense. This hypothesis was tested in Experiment 1. The results of this experiment showed that judges clearly recognized a primary narrative usage for the PS (92%) but did not make the same clear judgment for the PC (77%) or the expected non-narrative primary usage of the IMP (77.5%). This result opened a path for further finer-grained research, namely an annotation experiment on IMP with the [±narrativity] feature, which was done in Experiment 2. In this experiment, the IMP was categorized as non-narrative in 90% of the cases and as narrative in 10% of the cases.

In the light of these results, I would like to make a few suggestions. My first suggestion is that the [±narrativity] feature accounts for Harris' (1982) and Squartini and Bertinetto's (2000) hypothesis about the aoristicization process undergone by the PC in Romance languages (except Portuguese and Spanish). Their suggestion is that the PC undergoes a change from a true perfect towards an aorist and that this scalar process is visible for the PC in FR, IT and RO. My assumption is that the perfect aspect (such as the EN PresPerf and the PC in Portuguese and Spanish) is correlated with the non-narrative value of the [±narrativity] procedural feature whereas the aorist (such as the PS in FR, IT and RO) is correlated with the narrative value of this feature. If this were true, the PresPerf and Spanish PC would be judged in an annotation experiment as having more frequently non-narrative usages than narratives ones. On the other hand, the FR, IT and RO PC would have more frequently narrative usages than non-narratives ones. This former prediction was confirmed in Experiment 1 and Experiment 10 for the FR PC, in Experiment 8 for the IT PC and in Experiment 9 for the RO PC.

The experiments on FR, IT and RO confirmed the scalar orientation of these languages in the aoristicization process. Specifically, the FR PC was judged as narrative in an average of 71% of the cases, the IT PC in 88% of the cases and the RO PC in 83% of the cases. The difference between FR and the other two Romance languages is statistically significant.

However, the difference between IT and RO is not statistically significant. These results raise two issues with respect to Squartini and Bertinetto's aoristicization scale.

- They suggest that IT precedes FR (i.e. standard FR and IT) in the aoristicization process, underlying in the same time that there is a significative regional difference (north vs. centre vs. south). The results of experiments carried out in this research suggest that FR precedes IT. In other words, the FR PC precedes the IT PC in the aoristic drift. This result might indicate that the IT PC continues to evolve in the aoristic drift differently than FR. As I will point out in section 7.4.3, the PC was judged as narrative by speakers of IT from the Sourthern part of Italy in 86% of the cases. Consequently, it undergoes its perfectal functions (i.e. non-narrative) in only 14% of the cases.

- Squartini and Bertinetto suggest that IT and FR precede RO. According to their scale, a higher percentage of narrative usages is expected for the RO PC than for the FR and IT PC. The results of experiments carried out in this research confirm the relation between FR and RO. As for the relation between IT and RO, the observed difference between the two languages is not statistically significant (83% in RO vs. 88% in IT).

My second suggestion is linked to the PC and its description in the literature, according to which the PC has a base value where it locates the eventuality prior to S (E<S) and two contextual values distinguished by the position of R: R=E in its *anteriority* usage as in (597) and R=S in its *resultative* usage as in (598). I would argue that the base value corresponds to its conceptual content, which is shared with the PS and the PC. The two pragmatic values reflect the contextual value received by the [±narrativity] procedural feature encoded by this verbal tense, value inferred based on contextual information.

(597) Hier, j'*ai perdu* ma clef et j'*ai dormi* à l'hotel.
Yesterday, loose.PC my key and I sleep.PC at the hotel
'Yesterday, I lost my key and I slept at the hotel.'

(598) *As*-tu *trouvé* ta clef?
Find.PC your key?
'Have you found your key?'

Thirdly, it is assumed in the literature that the PS encodes by default the instruction for temporal progression and that this instruction is blocked if contextual information allows it. My suggestion is that the PS encodes the instruction to determine a contextual value of the [±narrativity] procedural feature but it does not impose the narrative value. The results of Experiment 1 indicated that the PS was judged as having a narrative usage in 92% of the cases as in (599) and non-narrative in 8% of the cases as in (600).

(599) Marie *étudia* jour et nuit. Elle *réussi* tous ses examens.
Mary study.PS day and night. She pass.PS all her exams.
'Mary studied day and night. She passed all her exams.'

(600) Bianca *chanta* le recitative et Ygor l'*accompagna* au piano.
Bianca sing.PS the recitative and Ygor accompany.PC her at the piano
'Bianca sung the recitative and Ygor accompanied her at the piano.'

As pointed out in section 7.2.1.2, the consideration whether the PS encodes or not by default the narrative value of the [±narrativity] procedural feature must be addressed in experimental work targeting online processing. The predictions for the two possibilities, i.e. narrative by default vs. encoding the instruction to assign a contextual value to the [± narrativity] feature, tested in an online experiment with self-paced reading task are the following:

- If the PS encodes by default the narrative interpretation, then non-narrative interpretations should produce longer reading times.
- If the PS encodes the instruction to assign a contextual value to the [± narrativity] feature, then narrative and non-narrative interpretations should produce similar reading times.

Fourthly, it is assumed in the literature that the IMP encodes a null directional instruction as in (601), where the IMP expresses a situation holding before the situation introduced with the PS. Under the pressure of contextual information the null directional instruction can be changed in an instruction for temporal ordering, especially for the narrative IMP as in (602) (see for example Saussure (2003); as discussed in section 2.5.1.2.2). In (602), the adverbial *une seconde plus tard* provides to the IMP the reference point required and the IMP allows temporal sequencing.

(601) Paul *entra* dans le bar. Marie *buvait* un café.
Paul enter.PS in the bar. Mary drink.IMP a coffee
'Paul entered in the bar. Mary was drinking a coffee.'

(602) Paul *entra* dans le bar. Une seconde plus tard, Marie *partait*.
Paul enter.PS in the bar. One second later, Mary leave.IMP
'Paul entered in the bar. One second later, Mary left.'

Similarly to the PS, the model defended in this thesis suggests that the IMP does not encode a null directional instruction by default. On the contrary, it encodes the instruction to determine a contextual value of the [±narrativity] procedural feature. This theoretical position will be verified experimentally in future work. My prediction is that narrative and non-narrative usages of the IMP will result in similar reading times.

Finally, my proposition is that the PS, PC and IMP are interchangeable only when they share, besides conceptual information, procedural information. For example, the PS, PC and IMP are interchangeable in their narrative usages, not only in FR but also cross-linguistically, as shown in the following examples. Precisely, example (603) is the original text written in EN where a SP form is used, example in (604) is the translation into FR of the EN text where a narrative IMP is used, example in (605) is the into IT of the EN text where a narrative PS is used and, finally, example in (606) is the into IT of the EN text where a narrative PS is used. The narrative IMP used in (604) could be replaced with a narrative PS as in (607) or a narrative PC as in (608). However, according to Grisot and Moeschler's model (2014), one would argue that only the narrative IMP provides a subjective perspective on the eventuality expressed. This brings into discussion the notion of *subjectivity*, which was not validated in the empirical work carried out in Grisot (in preparation). I think however that it deserves further experimental investigation.

(603)   But when the Rabbit actually took a watch out of its waistcoat-pocket and looked at it and then hurried on, Alice started to her feet, […] and, burning with curiosity, she ran across the field after it and was just in time to see it pop down a large rabbit-hole, under the hedge. In another moment, down *went* Alice after it!

(604)   Cependant, lorsque le Lapin tira bel et bien une montre de la poche de son gilet, regarda l'heure, et se mit à courir de plus belle, Alice se dressa d'un bond, […]. Dévorée de curiosité, elle traversa le champ en courant à sa poursuite, et eut la chance d'arriver juste à temps pour le voir s'enfoncer comme une flèche dans un large terrier placé sous la haie. Un instant plus tard, elle y *pénétrait* à son tour.

(605)   Ma quando il Coniglio trasse un oriuolo dal taschino del panciotto, e vi affisò gli occhi, e scappò via, Alice saltò in piedi, […] e divorata dalla curiosità, traversò il campo correndogli appresso, e giunse proprio a tempo di vederlo slanciarsi in una spaziosa conigliera, di sotto alla siepe. In un altro istante, giù Alice *scivolò*.

(606)   Dar când iepurele, imediat după asta, scoase din buzunarul veste un ceas, îl privi şi începu să se grăbească, Alice sări în picioare […] şi, arzând de curiozitate, o luă la fugă peste câmp după el chiar la timp pentru a-l putea vedea sărind într-o gaură de iepure mare de sub gardul viu. Într-o clipă Alice *sări* după el.

(607)   Un instant plus tard, elle y *pénétra* à son tour.

(608)   Un instant plus tard, elle y *a pénétré* à son tour.

However, the interchangeability rate depends also on the combination of Tense with Aspect and Aktionsart. I will address this issue in the following lines.

Aspect and Aktionsart

Currently, in the literature it is assumed that the FR PS and PC are perfective (Martin 1971; Tahara 2000) whereas the IMP is imperfective (Martin 1971; Guillemin-Flescher 1981; Vetters 1996, among others) even if in some cases it can remain underdetermined with respect to Aspect. According to the model defended in this thesis, all verbal tenses in Romance and in EN provide information about Tense and Aspect as they are applied to Aktionsart. In other words, each verbal tense expresses temporal location (i.e. Tense) and the speaker's viewpoint (i.e. Aspect) on eventualities (i.e. Aktionsart).

In this thesis, the relation between Tense and Aspect for FR verbal tenses was not investigated directly. It is possible, however, to make some observations based on the results of Experiment 5 carried out on data randomly selected from a translation corpus. This experiment was carried out on SP items, which were translated into Serbian where Aspect is morphologically expressed. The results of this experiment showed that the perfective viewpoints expressed with a SP correspond to a translation through a PC or PS and imperfective viewpoint expressed with a SP correspond to a translation with an IMP in 78% of the cases. In 22% of the cases, the contrary combination of features takes place: perfective viewpoints expressed with a SP correspond to a translation through an IMP and imperfective viewpoint expressed with a SP correspond to a translation with a PC or PS.

Based on these results, I assume that each of these verbal tenses is not perfective or, respectively imperfective by default as it is suggested in the literature. According to the model defended in this thesis, Tense combines with Aspect and all four combinations are possible: *narrative perfective* as in (609), *narrative imperfective* as in (610) where the lexical

paraphrase *être en train de* 'be+*ing*' expresses explicitly the imperfective viewpoint, *non-narrative perfective* as in (611) and *non-narrative imperfective* as in (612)[206]. There are however combinations which are more frequent than others and they are associated with one or another verbal tense. For instance, the narrative perfective combination is more frequently associated with the PC and the PS whereas the non-narrative imperfective is more frequently associated with the IMP.

(609) Il *toqua* à la porte et *entra* dès qu'il y eut une réponse.
He knock.PS a the door and he enter.PS as soon as he have.PS an answer.
'He knocked at the door and entered as soon as he has an answer.'

(610) Dans son rêve, il *était en train de* chercher sa sœur. Ensuite il s'*arretait* et l'*appelait* de toutes ses forces.
In his dream, he be.*ing*.look for his sister. Then, he stop.IMP and call.IMP with all his strength.
'In his dream, he looked for his sister. Then he stoped and called her with all his strength.'

(611) Marie *ferma* les yeux et s'*imagina* être une princesse.
Mary close.PS her eyes and imagine.PS to be a princess
'Mary closed her eyes and imagined she was a princess.'

(612) Marie entra dans la chambre. Jean la *était en train de* la chercher et il l'*appelait* par son prénom.
Mary enter.PS the room. John be.*ing* look for her and he call.IMP by her name.
'Mary entered the room. John was looking for her and was calling her name.'

Similar observations can be made with respect to the relation between Tense and Aktionsart. In this thesis, this relation was not investigated directly for FR verbal tenses. Experiment 4 targeted the usage of the SP with telic and atelic situations, which were operationalized in terms of [±boundedness]. The cross-linguistic analysis of the results of this experiment indicated that bounded eventualities expressed with a SP correspond to a translation through a PC or a PS and unbounded eventualities correspond to a translation with an IMP in 82% of the cases. In 18% of the cases SP unbounded eventualities correspond are translated a PC or a PS and SP bounded eventualities correspond are translated through an IMP. Consequently, the FR PC and PS can express unbounded eventulities as in (613) and the IMP can express bounded eventualities as in (614). In other words, each verbal tense can be associated with one or the other type of eventuality. There are however correlations that are more frequent than others, such as bounded eventualities expressed with a PC or a PS and unbounded eventualities expressed with an IMP.

(613) Il *a* toujours *été* très poli.
He be.PC always very polite
'He has always been very polite.'

(614) Il *atteignait* le sommet quand l'orage commença.
He reach.IMP the top of the mountain when the storm begin.PS
'He was reaching the top of the mountain when the storm began.'

---

[206] The four combinations are easier to grasp in aspect-prominent languages, where Aspect is morphologically expressed.

These conclusions are inferred based on the analysis of translation corpora (from a tense-prominent language into an aspect-prominent language). The *cross-linguistic transfer of properties* method was used for transferring aspectual information from Serbian to the EN SP (as discussed in section 4.2.3 and in Experiment 4). Further experimental investigations carried on FR data should be carried out in order to validate these conclusions and to determine if the PC, the PS and the IMP are or not associated by default with one or the other value of the [±boundedness] feature.

Conclusion

In this section I provided a reanalysis of the FR PS, PC and IMP verbal tenses in the light of the empirical work carried out in this research. I argued that these verbal tenses encode both conceptual and procedural information and that they combine felicitously with Aspect and Aktionsart. They share conceptual information not only among themselves but also with the EN SP and the IT and RO PS, PC and IMP. As for their instructional content, it was operationalized as the [±narrativity] feature. Each of these verbal tenses encodes the instruction to locate eventualities with respect to one another without encoding by default one of the values of this feature. Each of these verbal tenses combines felicitously with Aspect and Aktionsart.

7.4.3 Italian

Three IT verbal tenses were investigated empirically in this thesis: the PC, the PS and the IMP. The PRES is accounted for from a theoretical point of view, and principally, with respect to its opposition to the PC, the PS and the IMP. It is generally assumed that verbal tenses in Romance languages have similar functions at the discoursive level. The model defended in this research is a cross-linguistic model representing a *tertium comparationis* (see section 4.1 for the methodology used in Contrastive Analysis field) that allows comparison among verbal tenses in several languages. The [±narrativity] feature was tested in annotation experiments for IT and RO data. At this stage of the research, no experiments were carried out with aspectual information. In the following lines I will discuss the conceptual and procedural information encoded by the IT PC, PS and IMP.

Conceptual information

As pointed out in section 7.2.1.1 verbal tenses encode conceptual information in the form of a pro-concept TIME, which can be operationalized as the localization of E with respect to S via R. The IT PS, PC and IMP share the same conceptual meaning most frequently expressed as the ad hoc concept E<S (i.e. *pastness*). At this level of encoded information, from a theoretical point of view, these three verbal tenses are interchangeable as shown in the following examples. There is however an important regional variation between the Northern and the Southern parts of Italy. Precisely, in Northern IT the PC reached the end of the aoristicization process and it took over the discourse functions played by the PS. As a consequence, the PS, which is less frequently used than the PC, would occur only in literature at the beginning of a novel for example. In southern IT, the situation is the

contrary. The PC did not develop the aoristic function and, consequently, the PS is used very frequently. According to a northern IT speaker only the PC in (616) can replace the IMP in (615). A hypothesis can be made that southern speakers of IT would accept the PS in (617) to replace the IMP in (615).

(615) Alla sommità della collina la terra nericcia era indurita dal gelo, e il freddo *mi faceva* rabbrividire. (Literature Corpus)
'On that bleak hill-top the earth was hard with a black frost, and the air made me shiver through every limb.'

(616) Alla sommità della collina la terra nericcia era indurita dal gelo, e il freddo *mi ha fatto* rabbrividire.

(617) ?Alla sommità della collina la terra nericcia era indurita dal gelo, e il freddo *mi fece* rabbrividire.

At this point of the discussion, the question of the possibility of usage of the PRES arises. Example in (618) shows that the PRES cannot be used because the ad hoc concept encoded (i.e. E=S) is incompatible with the ad hoc concept encoded by the IMP *era* in the first clause.

(618) *Alla sommità della collina la terra nericcia era indurita dal gelo, e il freddo *mi fa* rabbrividire.

The concept of *reference to past* for the IMP in Romance languages has been rejected in the literature (see for example Saussure 2003, 2011 for the FR IMP) mainly because of the requirement that its conceptual content should account for all usages of the IMP, such as temporal, counterfactual, hypocoristic, conditional and mitigation among others. In a recent study, Baranzini and Ricci (to appear) propose a common interpretative procedure of the semantic and pragmatic values of IT IMP. They suggest three parameters that are relevant for characterizing the uses of the IMP:

- Past temporal deixis: it can be direct (eventuality is located in the past) or indirect (reference to the decision-making or the conceptualization of its realization);
- Context dependence: the IMP is dependent on other verbal tenses (the IMP is used in alternation with a perfective verbal tense) or not (the IMP can autonomously convey the whole course of events. Its usages can be either perfective or imperfective, producing a 'flattening' of its aspectual (i.e. grammatical aspect) content.
- Realization of the process: the eventuality can be realized, not realized in the present but it can be realized in the future or not realized and the possibility to be realized in the future is excluded.

In the following lines I will focus on Baranzini and Ricci's first parameter, which they explain as it follows:

We consider that only one, underdetermined semantic feature can be associated with the imperfect form, this feature being past temporal reference. At this level, the scope of the imperfect is not specified: when no further contextual indications are available, what is referred to the past is the event itself; when a specific context clearly does not allow for this interpretation, the hearer must search for a preliminary, past situation related to the event- a planning or conception phase of it, for instance. […] This scheme may be read in the following terms: the imperfect characterizes an event e or a moment/phase in time associated with E

preceding S. (pp. 15, 17)

In their view, all possible usages of the IT IMP can be accounted for starting with this basic semantics through pragmatic enrichment due to contextual indications[207]. As far as Aspect is concerned, they suggest that the IMP is imperfective by default, i.e. 'the moment of observation of the event is simultaneous to a moment internal to the course of the event' (p. 16).

Baranzini and Ricci point to a key notion regarding the semantics and pragmatics of the IT IMP, specifically reference to past time for all usages of the IMP. In my view this basic semantics corresponds to its conceptual content, the ad-hoc concept of pastness, built by the hearer based on contextual information. As for Aspect, I suggest in this thesis that the IMP correlates positively with the imperfective value of Aspect, without acquiring this value by default. I think that the default imperfective value defended by Baranzini and Ricci, as well as the hypothesis made in this thesis, should be tested in online experiments. The predictions are that default values should trigger longer reading times in incongruent contexts.

Procedural information

The model defended in this thesis assumes that IT verbal tenses encode also the instruction to temporally related eventualities with respect to one another. Procedural information encoded by the IT PS, PC and IMP was operationalized as the [±narrativity] feature. This feature was tested and validated in Experiment 9. The results of this experiment indicated that the IMP was judged as non-narrative in 84% of the cases as in example (620), the PC was judged as narrative in 88% of the cases and the PS was judged as narrative in 96% of the cases as in example (621).

(619)   A tale proposito ricordo la risoluzione del 15 settembre scorso, in cui si *raccomandava* di presentare la proposta il più rapidamente possibile. (EuroParl Corpus)

(620)   Il 5 ottobre il governo irlandese ha perso di misura (51,7 per cento contro 48,3) un referendum per l'abolizione del Seanad Éireann, la camera alta del parlamento. La sorprendente sconfitta *ha colpito* duramente la coalizione di governo tra Fine Gael e Labour. (Journalistic Corpus)
'On October 5, the Irish government lost (51.7 percent vs. 48.3) a referendum to abolish the Seanad Éireann, the upper house of parliament. The surprising defeat hit hard the government coalition between Fine Gael and Labour.

(621)   Diego Garcia fu consegnata al Pentagono nel 1973. [...] gli abitanti di Diego Garcia *furono* sbrigativamente sfrattati e *spediti* a Mauritius e nelle Seychelles. (Journalistic Corpus)

---

[207] Such as for example, the neutralization of the alternation between imperfect and perfective tenses in a limited subpart of a broader narrative context including imperfect-perfective tenses alternation, which triggers the narrative IMP interpretation. Another example is the neutralization of the alternation between imperfect and perfective tenses where the IMP is used perfectively and covers all phases of the narrative sequence. In this case, pragmatic enrichment leads to an evidential interpretation closely linked to the speaker's commitment (p. 17). Baranzini and Ricci note that 'in all these cases, past reference concerns not only the event itself, but also the moment in which the speaker is exposed to the narrative experience of the event (dream, ludic scenario planning, reading a book, watching a movie, listening to a testimony).' (p. 17).

'Diego Garcia was delivered to the Pentagon in 1973. [...] the inhabitants of Diego Garcia were summarily evicted and sent to Mauritius and the Seychelles.'

There is an important regional difference regarding the usages of the PC and PS in IT. Squartini and Bertinetto (2000) provide the results of an offline experiment, where participants had to fill in sentences with the PC and the PS (for a detailed description of the experiment, see Bertinetto 1996). Participants were originating from towns situated in Northern, Centre and Southern parts of Italy, as well as Sicily and Sardinia. The results of their questionnaire indicated that there is a small variation for the usage of the PC with respect to its perfectal functions, consisting of the following functions: *inclusivity, co-occurrence with temporal adverbs orientated towards S, persistent result* and *experientialism* (i.e. operationalized as non-narrative usages in this thesis). There is an important regional variation for the narrative usages of the PC investigated in personal, impersonal and historical narrations: the PC is more frequently used as narrative in the North (62.5%) than in the Centre (45.5%) and in the South (23.5). The frequent usage of the narrative PS in the Centre is statistically significant.

Experiment 9 was carried out in Napoli, a town representing Southern IT. The results of this experiment indicated that the PS is clearly recognized as having most frequently narrative usages (96%) as in (622). However, the PC also was judged as narrative more often than non-narrative (88%), as in the pair of examples (623) and (624).

(622) Diego Garcia fu consegnata al Pentagono nel 1973. [...] gli abitanti di Diego Garcia *furono* sbrigativamente sfrattati e *spediti* a Mauritius e nelle Seychelles. (Journalistic Register)
'Diego Garcia was demanded by and handed to the Pentagon in 1973. [...] and the Diego Garcians were summarily evicted to Mauritius and the Seychelles.

(623) [...] il rinvio deve essere effettuato non oltre i 60 giorni successivi al giorno in cui l'autorita competente *ha deciso* la destinazione della partita, a meno che non sia stata intrapresa un'azione legale. (EuroParl Register)
'[...] a re-dispatch shall take place no more than 60 days after the day on which the competent authority *decided* on the destination of the consignment, unless legal action has been undertaken.'

(624) L'esperimento sul campo [...] è stato accompagnato da misurazioni in quattro stazioni fisse, con l'ausilio di quindici unite mobili, un aereo e palloni aerostatici, e *ha incluso* calcoli di modello sulla base di un inventario dettagliato delle emissioni. (Journalistic Register)
'The field experiment […] was accompanied by measurements at four fixed stations, with 15 mobile units, with an aircraft and balloons and *included* model calculations on the basis of a detailed emission inventory.'

I think that this result can indicate two things. The first is that compared to the IMP, the PC has more frequently narrative usages than non-narrative ones. The second is that the PC is evolving in the aoristic drift also in the South of Italy (recall that Squartini and Bertinetto's study was carried out in the late nineties).

The results of Experiment 9 indicated that the IT IMP has more frequently non-narrative than narrative usages as shown in the pair of examples (625) and (626).

(625) Non *era* questo l'intento di Cameron quando ha pronunciato il suo discorso (a lungo rinviato) sul futuro del paese in Europa. (Journalistic Register)

'That was not the prime minister's purpose when he delivered his much-delayed speech on Britain's future in Europe.'

(626) Non c'è niente di meglio di una cannoniera. La Illustrious di Sua Altezza Reale il 12 agosto è scivolata silenziosamente fuori da Portsmouth, ha superato la HMS Victory e la folla plaudente di patrioti. Nel giro di una settimana *era* al largo di Gibilterra, a un tiro di cannone da Capo Trafalgar. I petti della nazione si sono gonfiati, sono spuntate alcune lacrime. Lo spirito olimpico si è dileguato per strinare la barba del re di Spagna. (Literature Corpus)

'Nothing beats a gunboat. HMS Illustrious glided out of Portsmouth on August 12, past HMS Victory and cheering crowds of patriots. Within a week it was off Gibraltar, a mere cannon shot from Cape Trafalgar. The nation's breast heaves, the tears prick. The Olympic spirit is off to singe the king of Spain's beard.'

As far as aspectual information (Aspect and Aktionsart) is concerned, the PC and the PS are traditionally described as being perfective whereas the IMP is imperfective. More specifically, the PS has the aoristic aspect, the PC the perfect aspect and the IMP the imperfective aspect (Bertinetto 1986). As pointed out above, the PC is fairly advanced in the aoristic drift, and consequently, it diminishes the extent of its perfectal aspect and all the semantic-pragmatic interpretations associated with it. As for the IMP, there are two opposing standpoints: one the one hand, it is assumed to be imperfective (Bertinetto 1986; Baranzini and Ricci 2015) and, on the other hand, neutral (i.e. underdetermined) with respect to Aspect (Blücher 1974). According to the model defended in this thesis, a verbal tense underdetermines the content communicated by a speaker (i.e. both Tense and Aspect) and it must be contextually worked out at the level of explicatures. Accordingly, I assume that the IT IMP, similarly to the FR IMP, does not encode any of the possible values of Tense and Aspect by default. Frequencies of usage have however been observed, such as the non-narrative and imperfective usages for the IMP. This assumption should be tested for IT in further experimental investigations.

With respect to Aktionsart, Bertinetto (1986) argued that inclusive interpretations of the PC (i.e. non-narrative usages) are restricted to atelic situations. The prediction therefore being that the non-narrative PC expresses unbounded eventualities and that the narrative PC expresses bounded eventualities. Experiment 4 indicated that for the EN SP the situation is not as clear-cut as they theory suggests it. I have not, at least as far as this thesis is concerned, explored empirically the relation between verbal tenses and Aktionsart in IT. My prediction would however be that there are no straightforward correlations. Similarly, the assumption that the PS expresses bounded eventualities and the IMP unbounded eventualities should be tested experimentally in further work.

Conclusion

In this section, it was shown that the IT verbal tenses PC, PS and PC encode conceptual and procedural information. They share conceptual information, which is most frequently instantiated as the ad hoc concept E<S. They differ with respect to the contextual value of the [±narrativity] procedural feature computed most frequently for each verbal tense.

Additionally, the IT PRES instantiates most frequently the ad hoc concept E=S. It does, however, similarly to the FR PRES, instantiate an ad-hoc concept E<S constrained by the narrative value of the [±narrativity] procedural feature.

7.4.4 Romanian

Three RO verbal tenses were investigated empirically in this thesis: the PC, the PS and the IMP. The PRES is accounted for from a theoretical point of view, and principally, with respect to its opposition to the PC, the PS and the IMP. The model defended in this thesis is a cross-linguistic valid model, predicting both similar and dissimilar behaviours of these verbal tenses in FR, IT and RO. Precisely, my suggestion is that verbal tenses in FR, IT and RO behave similarly with respect to the conceptual information encoded. Moreover, they all encode the [±narrativity] procedural feature. The contextual values of this feature depend cross-linguistically on the stage where each language is situated in the aoristic drift.

Conceptual information

The RO PS, PC and IMP encode conceptual information in the form of a pro-concept TIME, which can be operationalized as the localization of E with respect to S. These three verbal tenses share the same conceptual meaning in terms of E<S. Verbal tenses used in examples (627)-(629) (IMP, PC and, respectively PS) share conceptual meaning, however this is not the case with the PRES in (630).

(627) Pe acel vârf de deal pleşuv, pământul era îngheţat bocnă, iar aerul, tăios, *mă făcea* să dârdâi din tot trupul.
'On that bleak hill-top the earth was hard with a black frost, and the air made me shiver through every limb.'

(628) Pe acel vârf de deal pleşuv, pământul era îngheţat bocnă, iar aerul, tăios, *m-a făcut* să dârdâi din tot trupul.

(629) Pe acel vârf de deal pleşuv, pământul era îngheţat bocnă, iar aerul, tăios, *mă făcu* să dârdâi din tot trupul.

(630) *Pe acel vârf de deal pleşuv, pământul era îngheţat bocnă, iar aerul, tăios, *mă face* să dârdâi din tot trupul.

Procedural information

The PC, the PS and the IMP encode the instruction to relate eventualities with respect to one another. This instruction was operationalized in this thesis as the [±narrativity] feature, which was tested in Experiment 9. The results of this experiment indicated that the IMP was judged as non-narrative 71% of the cases as in (631), the PC was judged as narrative in 83% of the cases as in (632) and the PS was judged as narrative in 93% of the cases as in (633).

(631) Poliţia austriacă a oprit un camion care transporta 137 de căţeluşi. Vehiculul prezenta defecţiuni grave, iar paşapoartele animalelor *erau* false, deoarece câinii nu ajunseseră la vârsta legală pentru transport. (EuroParl Corpus)

'The Austrian police stopped a truck containing 137 puppies. The vehicle manifested serious defects and the animals' passports were forgeries, as the dogs had not reached the legally required age for transport.'

(632) În faţa acestui spectacol m-am retras imediat, mi-am luat pălăria şi, după un drum de patru mile, *am ajuns* la poarta grădinii lui Heathcliff, tocmai la timp pentru a mă adăposti de primii fulgi de zăpadă ai viscolului ce se dezlănţuise. (Literature Corpus)
'This spectacle drove me back immediately; I took my hat, and, after a four-miles' walk, arrived at Heathcliff's garden-gate just in time to escape the first feathery flakes of a snow-shower.'

(633) Totuşi, la a doua încercare, dădu peste o perdea pe care n-o observase mai înainte, şi, în spatele ei era o uşiţă cam de cincisprezece ţoli înălţime; *încercă* cheiţa la încuietoare şi, spre marea-i încântare, se potrivi! (Literature Corpus)
'However, on the second time 'round, she came upon a low curtain she had not noticed before, and behind it was a little door about fifteen inches high. She tried the little golden key in the lock, and to her great delight, it fitted!'

These results indicate also the IMP was judged as narrative in 29% of the cases as in (634), the PC was judged as non-narrative in 17% of the cases as in (635) (note the usage of the IMP in (636), which is the translation into FR of the original EN text, pointing to the overall non-narrative interpretation of the discourse) and the PS was judged as non-narrative in (637).

(634) Nimic nu bate o canonieră. HMS Illustrious a lunecat falnic din portul Portsmouth pe 12 august, trecând pe lângă HMS Victory, sub aplauzele mulţimii de patrioţi. Într-o săptămână *ajungea* în largul Gibraltarului, la o lovitură de tun de Capul Trafalgar. (Journalistic Corpus)
'Nothing beats a gunboat. HMS Illustrious glided out of Portsmouth on August 12, past HMS Victory and cheering crowds of patriots. Within a week it was off Gibraltar, a mere cannon shot from Cape Trafalgar. The nation's breast heaves, the tears prick. The Olympic spirit is off to singe the king of Spain's beard.'

(635) Experimentul de teren [...] a fost însoţit de măsurători la patru staţii fixe [...]. Studiul *a avut scopul* de a răspunde următoarelor întrebări, luând ca exemplu un episod tipic de smog pe timpul verii. (Legislation Corpus)
'The field experiment [...] was accompanied by measurements at four fixed stations [...]. The study *was designed* to answer the following question, taking a typical summer smog episode as an example.'

(636) L'expérience sur le terrain [...] a été accompagnée par des mesures dans quatre stations fixes [...]. Cette étude *tendait* à répondre aux questions suivantes en prenant comme exemple un épisode de smog estival typique.

(637) Aşa că continuă să reflecteze, pentru sine, [...] dacă plăcerea de a împleti o coroniţă din margarete merită efortul de a se apleca şi de a culege margaretele, când, repede, un Iepure Alb cu ochii roz *trecu* foarte aproape de ea. (Literature Corpus)
'So she was considering in her own mind [...] whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.'

Turning now to the combination of Tense, Aspect and Aktionsart, the empirical work carried out in this thesis did consider this matter for RO. Some hypotheses can however be

made based on the existing descriptions of verbal tenses in RO (see section 2.5.1.4). As far as Aspect is concerned, it is a grammatical category only recently recognized by Romanian grammars (such as GLR edited by Guțu-Romalo, 2005). Margan (2009) suggested that RO verbal tenses marks the [±perfectivity]: the PS and the PC are [+perfective], and the IMP is [-perfective], in other words, imperfective. In GLR (2005, 428-431) it is suggested that aspectual modifications can occur, such the slip from imperfective to perfective readings when the IMP applies to punctual situations and also due to contextual interpretation, as in in example (638), where the opening of the door is expressed perfectively.

(638) Lucian tocmai *deschidea* uşa, când deodată *a auzit* pe cineva strigând.
Lucian just open.IMP, when suddely hear.PC someone screaming.
'Lucian was just opening the door, when suddenly he heard someone screaming.'

Since the RO PC is fairly advanced in the aorist drift, it expresses both the perfect and the aorist aspects. These two functions are observed both in child (up to 3 years old) and in adult language (Stoicescu 2012).

Finally the PRES correlates most frequently with the imperfective aspect as in (639) both in children and adult language (Stoicescu 2012). It may however express also habitual as in (640) and generic aspects as in (641) from Stoicescu (2012, 134).

(639) *Vine* maşina.
Car arrive.PRES
'The car is arriving.'
(640) Câinele meu *muşcă*.
My dog bite.PRES
'My dog bites.'
(641) Păsările *zboară*.
Birds fly.PRES
'Birds fly.'

With respect to Tense and Aktionsart, Stoicescu (2010) found that in 70% of the cases, atelic situations (states and activities) were used with the PRES verbal tense whereas more telic situations are used wit the PC (76%). These patterns decrease with age; therefore they are not as strong in adult language as in child language.

The model defended in this thesis makes the following hypothesis, which must be investigated experimentally in further research. Tense, Aspect and Aktionsart co-occur and their respective values are not correlated by default. Correlations of values (past vs. non-past, perfective vs. imperfective, bounded vs. unbounded) may be observed, however, they are not exhaustive. My hypothesis against the correlation by default assumption is that in online experiments, similar values of reactions times will be found for cases when a verbal tense has narrative compared to cases when it has non-narrative usages.

Conclusion

In this section, it was shown that the RO verbal tenses PC, PS and PC encode conceptual and procedural information. They share conceptual information, which is most frequently

instantiated as the ad hoc concept E<S. They differ with respect to the contextual value of the [±narrativity] procedural feature computed most frequently for each verbal tense. The RO PC being fairly advanced in the aoristic drift, it expresses frequently the aoristic aspect and it is interchangeable with the PS. The IMP has both non-narrative (most frequent) and narrative values. Finally, Tense, Aspect and Aktionsart co-occur and their respective values are not correlated by default.

7.4.5 A cross-linguistically valid framework

In section 2.5.2, I pointed out the lack of a common frame that would allow a consistent contrastive comparison of verbal tenses. My aim was hence to propose a cross-linguistic valid framework that would be both theoretically and empirically grounded. The features included in the model developed in this thesis originate in the specialized literature targeting EN, FR, IT and RO verbal systems, as well as the inflectional categories that verbs take in tensed languages. These features were validated experimentally and the model was developed based on translation corpora using methods as cross-linguistic transfer of properties.

This model of temporal reference holds for English and three Romance languages. The question that arises at this point of the discussion concerns the possible application of this model for aspect-prominent languages and tenseless languages. I suggest a positive answer to this question due to the fact that temporal reference was investigated by separating the categories of Tense, Aspect and Aktionsart. This allows a possible application to aspect-prominent languages, where Aspect is grammaticalized to a greater degree than Tense. This makes it more necessary, more systematic and more pervasive than Tense. The model defended here is a medium-grained model enabling adjustments in accordance with the properties of the target language(s).

Regarding tenseless languages, the model is applicable with the specification that a significant place should be given to temporal adverbials (such as *jiangyao* for future and *cengjing* for past time reference, as well as *yijing* 'already' and *cai* 'just' for recency). This research did not incorporate temporal adverbials and aspectual particles and their interrelations with verbal tenses but this should definitely be studied in further research. Studies on Mandarin Chinese indicated that it presents all other temporal devices except what has been defined as Tense (as discussed in section 2.1). Finally, features linked to Aspect and Aktionsart can give account of the rich aspectual system (such as perfective morphemes *-le* and *–guo*, resultative and perfective verb complements and imperfective morphemes *zai* and *–zhe*).

Finally, the model defended in this thesis has proven its efficacy for automatic treatment of language and application to machine translation for the pair of languages EN and FR (as discussed in section 6.3). My suggestion is that it could be used for other pairs of languages, including aspect-prominent and tenseless languages.

7.5   Conclusive remarks

This chapter has given an account of the model defended in this thesis. It was suggested

that the global interpretation of temporal reference at the discursive level is determined, on the one hand, by the linguistic means existent in a language, and by their ad-hoc inferential contextual saturation. In tensed languages, as EN, FR, IT and RO temporal reference is expressed linguistically through TAM markers, temporal connectives and temporal adverbials. Linguistic expressions in general, including TAM markers, underdetermine the content communicated by a speaker both at the level of explicature and implicatures. In the interpretation process, their meaning is worked out contextually. My suggestion was that the human brain tends to treat temporal information coming from different sources (Tense, Aspect, Aktionsart, temporal connectives and temporal adverbials) in a coherent manner. I hypothesized that temporal coherence takes places at the cognitive and discursive levels.

Moreover, a reanalysis of Tense, Aspect and Aktionsart was proposed in the light of the empirical work carried out in this research. Firstly, I proposed a conceptualist view of Tense. Specifically, I argued that Tense encodes both conceptual and procedural information. Tense encodes a pro-concept TIME, which is semantically incomplete, is inferentially worked out and contributes to the truth-conditions of an utterance. I suggested that hearers build an ad hoc concept of pastness (E<S) or non-pastness (E≥S) based on contextual information, which are neuro-linguistically valid categories. Tense encodes the instruction to relate temporally eventualities with respect to one another. This information was operationalized as the [±narrativity] feature, which has two possible values specified in the interpretation process. It was argued that a verbal tense does not encode one of the values of the [±narrativity] feature by default. On the contrary, it represents a contextual value determined equally depending on other parameters, such as Aspect and Aktionsart.

Furthermore, it was argued the grammatical category of Aspect represents procedural information constraining the formulation of hypotheses about the explicit content of an utterance. The [±perfectivity] feature operationalizes the speaker's viewpoint on the eventuality expressed. Verbal tenses do not correlate by default with one of the two possible values of the [±perfectivity] feature. Additionally, the category of Aktionsart represents conceptual information contributing to the truth-conditions of an utterance. This information was operationalized as the [±boundedness] feature, which represents the actual realization of an eventuality.

Finally, a reanalysis of verbal tenses in EN, FR, IT and RO based on the theoretical model defended in this thesis is proposed both in a monolingual and in a cross-linguistic perspective. It was assumed that the SP in EN, the PC, the PS and the IMP in IT, FR and RO share conceptual meaning and the instruction to relate temporally eventualities with respect to one another. This procedural information is a cross-linguistic valid feature, which is materialized in a dissimilar way in the languages considered. Using Squartini and Bertinetto's hypothesis about the aoristicization process, it was argued that there is a positive correlation between the degree of advancement of the PC in the aoristic drift and the frequency of its narrative usages.

A series of suggestions discussed in this chapter were tested and validated in the empirical work carried out in this thesis. Additionally, a number of hypotheses have been formulated which will be addressed in further experimental work. I will discuss more in detail future experimental work in section 8.3.

8 Conclusion

8.1 Summary

The research presented in this thesis subscribes to two multidisciplinary projects aiming at improving the results of machine translation (SMT) systems in terms of textual coherence. This investigation had a two-fold purpose. The first aim was directly linked to the SMT projects and was to investigate the translation of verbal tenses in parallel corpora and to create human annotated data to be used for training NLP softwares. The second aim was to investigate temporal reference and its processing at the discursive level, investigation carried out from theoretical and empirical perspectives on data from four languages. Temporal reference is defined broadly and takes place at two levels: location of an eventuality with respects to S and location of eventualities with respect to one another. In tensed languages, temporal reference is expressed linguistically through verbal tenses (containing in EN and Romance languages both Tense and Aspect), Aktionsart, temporal connectives and adverbials.

From a theoretical point of view, the research was carried out in the RT framework according to which the linguistic expressions that a speaker utters underdetermine the content that she communicates not only at the level of implicatures but also the propositional contents she communicates explicitly (that is the explicature of the utterance). The hearer must therefore recover inferentially the speaker's intended meaning, at the level of explicatures and implicatures. This interpretative process is guided by the expectation of relevance and the quest for cognitive effects.

From an empirical point of view, three main types of methodology were used. Firstly, translation corpora containing natural instances of human communication were explored in order to investigate the usages and the translation of verbal tenses, as well as to identify translation divergences. Secondly, corpus data was investigated using the methodology proposed in the Contrastive Analysis field, according to which systematic language comparison must be carried out in three steps: monolingual description followed by the actual comparison based on a cross-linguistic valid *tertium comparationis*. Finally, experiments with linguistic judgment task were carried out in order to test theoretical hypotheses suggested in the literature. The Kappa coefficient was used to measure inter-annotator agreement rate. Furthermore, two novel techniques making use of the properties of translation corpora were used for fine-grained investigations, precisely *translation spotting* and *cross-linguistic transfer of properties.*

The findings of this research can be summarized as follows.

Firstly, it was shown with the example of the English SP and its translation into FR, IT and RO that Tense encodes both conceptual and procedural information, precisely, a pro-concept TIME and procedural information operationalized as the [±narrativity] feature. Aspect expressing the speaker's viewpoint represents procedural information whereas Aktionsart expressing temporal information inherent to situation types represents conceptual information. Procedural information constrains the formulation of hypotheses about the

explicit and the implicit content of an utterance whereas conceptual information contributes to the truth-conditional content of an utterance.

Secondly, a verbal tense (Tense and Aspect applied to Aktionsart) has contextual usages, associated with the various combinations of conceptual and procedural contents. Cross-linguistically, there is little variation with respect to conceptual information. Procedural information encoded by Tense and Aspect is cross-linguistically valid, however, contextual values that they receive are not (narrative vs. non-narrative, perfective vs. imperfective). Consequently, procedural information triggers important cross-linguistic variation.

Thirdly, the human brain treats temporal information coming from Tense, Aspect and Aktionsart in a coherent manner. Temporal coherence takes places at the cognitive level with respect to the hearer's processing of the utterance and, the discursive level, with respect to temporal discourse relations, their linguistic marking and inferential nature. Multifactorial analyses of the data have shown that the features [±narrativity], [±boundedness] and [±perfectivity] are correlated and that the values of one feature can be predicted based on the values of the other features. Moreover, they are statistically significant factors for predicting the verbal tense used in a target language.

Finally, NLP experiments with the features [±narrativity] and [±boundedness] showed that this information can be dealt with automatically for producing reliable data. Moreover, SMT systems aware of the temporal information provided by [±narrativity] and [±boundedness] features produce better translations than systems unaware of this pragmatic information.

In the following two sections I will speak about the main contributions of this thesis and propose suggestions for further research.

8.2 Main contributions

The research carried out in this thesis brought the following contributions to the current state of the art:

(1) It proposes two quantitative measures for the conceptual/procedural distinction suggested in relevance theoretic pragmatics.

(2) Regarding the semantics and pragmatics of temporal reference in natural language and its ingredients in tensed languages, it suggests the need to separate the categories of Tense, Aspect and Aktionsart in order to have an accurate picture, as well as to deemphazise the role played by Tense.

(3) It provides a quantitative analysis of features encoded by Tense, Aspect and Aktionsart and their contextual values. Their interaction and their correlations were quantified based on human annotated data.

(4) It provides a cross-linguistic valid *tertium comparationis* for verbal tenses as well as a reanalysis of individual verbal tenses in EN and three Romance languages: FR, IT and RO.

(5) This thesis is a plea in favour of the necessity of having empirical basis for linguistic models and in favour of the complementarity of corpus and experimental work.

(6) It shows that SMT systems have better results when they are aware of pragmatic information, such as the features encoded by Tense, Aspect and Aktionsart and their

contextual values.

(7) It points to the granularity of linguistic features to be included in a cross-linguistically valid model of temporal reference. The medium-coarse grained features [±narrativity] and [±boundedness] were not only successfully implemented for automatic processing but also their implementation in NLP and their application to MT produced significant improvements of the results of these systems. In other words, these ameliorations represent an empirical indirect but solid validation of the theoretical model proposed.

In the following lines, I will discuss more in detail some of these aspects. The thesis contributes to advancing research on the conceptual/procedural distinction by providing arguments in favour of a dualistic view of linguistic expressions, precisely, of verbal tenses. Moreover, two quantitative measures are suggested for investigating the conceptual and procedural information encoded by linguistic expressions. In the last few years, researchers renewed their interest in this distinction by searching an empirical basis for the existent theoretical assumptions. Empirical evidence comes from corpus work in a cross-linguistic perspective, experimentation with online and offline tasks, diachrony and language acquisition (Zufferey 2012, 2014; Cartoni et al. 2013; Grisot and Moeschler 2014). Due to this empirically oriented investigation, researchers argue nowadays that linguistic expressions can encode both conceptual and procedural information, that procedural information constrains (i.e. disregards other possible interpretative paths) the interpretative process and that diachronically speaking, a linguistic expression can develop from fully conceptual to fully procedural (Wilson 2011, 2015; Blochowiak 2015a). For example, Wilson (2015) argued that since temporal connectives clearly contribute to the truth-conditional content of utterances in which they occur, the standard arguments for procedural treatment do not apply.

In this thesis, I suggest that quantitative measures are needed, next to the existent qualitative features, to have a better and more objective understanding of the conceptual/procedural distinction and its application to linguistic expressions, such as connectives and verbal tenses among others. I propose two quantitative measures: the *inter-judge agreement rate*, measured in this research with the Kappa coefficient (Carletta 1996, Artstein and Poesio 2008), and the quantity of *cross-linguistic variation* measured in translation corpora.

The results of experiments with linguistic judgment task performed by two or more native speakers were evaluated with the Kappa coefficient. It has values from 0 meaning lack of agreement other than expected to occur by chance to 1 signalling perfect agreement. Conceptual information results in consistent inter-annotator agreement and high Kappa values (linked by the notion of accessibility to consciousness for native speakers of an easily graspable intuitive concept) while procedural information results in inconsistent inter-annotator agreement and low Kappa values (procedural calculus depends on non-guaranteed pragmatic inferences).

As for the second measure, I suggest that translating conceptual information triggers little cross-linguistic variation whereas translating procedural information is source of important variation. In this research it was shown that for translating the SP into three Romance

languages, verbal tenses expressing past time are used in more than 70% of the cases (precisely, 73% in FR, 72% in IT and 83% in RO) whereas the PRES is used in less than 8% of the cases (precisely, 8% in FR, 5% in IT and in RO). Moreover, an important variability was identified with respect to the usage of specific verbal tenses expressing past time, which can be discriminated based on procedural information encoded by Tense and Aspect. More precisely, for translating the SP in FR the PC, IMP and PS had the following frequencies of usage: 34%, 23% and respectively 16%. Similar values were found for IT (33%, 17% and respectively 22%) and RO (49%, 15% and 18%).

I showed based on corpus and experimental data, that verbal tenses encode both conceptual and procedural information. Specifically, I suggest that the neurolinguistically valid distinction between pastness (E<S) vs. non-pastness (E≥S) is of a conceptual nature. Native speakers have a rich ability to consciously evaluate the conceptual information encoded by a verbal tense in offline tasks. Furthermore, the analysis of a parallel translations corpus showed that when translated into three Romance languages, the EN SP is translated in 87% of the cases through verbal tenses expressing past time and only in 5% of the cases though verbal tenses expressing non-past. Moreover, when dealing with the instruction encoded by verbal tenses, native speakers have a poor ability to consciously evaluate the status of temporal relations holding among eventualities in offline tasks. However, when they are asked to provide a connective that would render explicit this implicit instruction, their agreement rate increases (i.e. the Kappa value approaches the maximum value of 1).

Based on current theoretical knowledge and on the experimental work carried on EN, FR, IT and RO data, a cross-linguistic *tertium comparationis* was proposed. It represents a hierarchical structure of procedural and conceptual information encoded by TAM markers (Tense, Aspect, Aktionsart and Mood) that can be applied to tensed languages. Based on this model, predictions can be made with respect to the processing of temporal reference in tenses languages. Tenseless languages function in a similar manner, with the specificity that the function played by Tense is taken over by the other markers. As far as this thesis is concerned, this model is only partly validated empirically. Further research should consider it entirely.

I propose an analysis of four RO verbal tenses made on a relatively reduced translation corpus and on offline experimentation with two Romanian judges. The research is however carried out in a multilingual perspective aiming at identifying a language independent and cross-linguistically valid *tertium comparationis* for verbal tenses. With this thesis, I hope to have contributed to advancing our knowledge of the pragmatics of verbal tenses in RO. It should however be followed and developed by more consistent empirical and experimental work targeting Romanian language in a monolingual context.

It was questioned how current semantic and pragmatic theories account for the neurolinguistic findings discussed in 3.3.1. Temporal reference through E related to S via R is either procedural or conceptual information encoded by Tense. I argue that the current assumptions of the procedural account are insufficient for explaining these findings. Firstly, if temporal reference is procedural information encoded by Tense, then the status of temporal reference through Aspect, Aktionsart and a rich 'Other means' category in tenseless languages should be considered. Precisely, are temporal adverbs and aspectual particles in Chinese also procedural markers? How is it possible that several types of linguistic means

encode the same procedure as that encoded by Tense in tensed languages. If, on the contrary, temporal location with E and S is conceptual, then it is the TIME concept that is expressed by a great variety of linguistic and non-linguistic means in tenseless languages. Secondly, if Tense encodes only instructions about the temporal location of eventualities (therefore no specification of the type past or non-past), then there should be no significant difference in the processing of reference to past, present or future times. Neurolinguistic studies brought evidence pointing to a difference in processing past vs. non-past temporal location in healthy and in aphasic speakers.

In this research, I explored and defended the second account: reference through E related to S is a conceptual type of information. Concepts are language independent and languages use different linguistic means to express them. Precisely, Tense is the major source in tensed languages, whereas grammatical, lexical aspect, temporal adverbs and other linguistic and non-linguistic means are primary in tenseless languages. My suggestion was that it is reference to the concept of pastness that it is more impaired then reference to the concept of non-pastness (present and future). Past time verbal tenses in EN, FR, IT, RO among many other languages, encode an ad hoc concept of pastness, which is determined contextually through the contextual saturation of the temporal variables S and E.

8.3 Perspectives

The series of offline experiments described in Chapter 6 privileged two aspects that are essential for linguistic empirical research targeting pragmatic interpretation. The first is *preserving the naturalness of the data* and the second is *the usage of a linguistic expression in its original context*. The preservation of naturalness and of the original context is opposed to the artificial data used in the experimental design strictly speaking. However, privileging these two aspects lead to one limitation: the impossibility of controlling for confounding variables.

My suggestion is that in future experimental research the hypotheses tested in this thesis, where naturalness and original context have been preserved, should be further tested through on-line experimentation and using classical experimental paradigms. If on-line experimentation validated only partly the model developed in this thesis, it would also provide additional evidence necessary for having a more accurate understanding of temporal reference in tensed languages. I suggest that expanding the present research (corpus-based and experimental) to Slavic languages and to tenseless languages is needed in order to develop a comprehensive theory of temporal reference in natural language. Undoubtedly, adopting also a neurolinguistic perspective would add substantially to the picture.

For example, online experiments using a self-paced reading task, where reaction times and accuracy (i.e. number of correct answers and number of errors) are evaluated, could investigate the humans' processing of verbal tenses. I suggest a design consisting of the following experimental conditions: narrative vs. non-narrative usages, presence or absence of a connective dedicated to temporal ordering such as *then* and its FR counterparts *puis* and *ensuite*, presence vs. absence of a context. My first prediction is that similar reaction times will be found for both narrative and non-narrative usages indicating that a verbal tense does not encode by default one of these values. My second prediction is that shorter reaction times will be found in the condition presence of the temporal connective than in the condition

absence of temporal connective. This is due to the facilitating effect of the connective, which encodes procedural information, as shown for causal connectives by Zufferey (2014).

Moreover, it was suggested that the [±narrativity] procedural feature operates both at the syntactic and the interpretation levels. EEG experimental technique would be very useful to test its function at the syntactic. For example, Dragoy et al.'s (2012) designed a study that focused on processing of time reference violations in which verbal tenses do not match a time frame previously set by adverbial by healthy speakers. They found that processing of a past time context disrupted by a present tense verb produced a P600 response triggered by the targeted verb. In contrast, processing present time context disrupted by a past time verb did not produce an immediate brain response. However, both past and present time reference produced sentence final negativity, which is a typical response to referential violations (as discussed in 3.3.1). With respect to the [±narrativity] procedural feature, on would expect that processing incongruent contexts as in example (642), where the eventualities of *sleeping* and *painting* can not occur simultaneously for the same person, would produce a P600 response due to the impossibility of assigning a contextual value of the [± narrativity] feature. On the contrary, processing the sentence in (643) would not produce the same brain response since it does not represent an incongruent context.

(642) Yesterday from 2PM to 5PM John slept. *He finished painting the wall.
(643) Yesterday from 2PM to 5PM John slept. He then finished painting the wall.

It was proposed that Tense encodes conceptual information behaving as a pro-concept TIME, which is must be contextually determined. Currently, there is no experimental design able to test conceptual vs. procedural information encoded by the same expressions. In this thesis, I suggested two quantitative measures (the inter-judge agreement rate measured with the Kappa coefficient and the quantity of cross-linguistic variation in translation corpora) based on the qualitative measures proposed by Wilson and Sperber (1993). These two measures are a first step towards a more objective approach of the conceptual/procedural distinction but additional evidence is required in order to confirm these two measures and also to help our understanding of this distinction.

Another issue that will be considered in further empirical research is the relation between temporality and causality. In one of the experiments where speakers were asked to propose a connective that would render explicit the temporal relations holding among eventualities, the causal connectives *because* and *thus* were suggested. The correlation between temporal and causal relations is however not exhaustive: there are cases where a temporal relation occurs without a causal relation, and the other way around (see for example Moeschler 2011 for a comparison between the French *parce que* 'because' and *donc* 'thus'). In other words, both narrative and non-narrative values of the [±narrativity] feature may be associated with causality.

Finally, Grisot (in preparation) brought evidence through an offline experiment that subjectivity is not linked to Tense. Two suggestions were made. The first is that subjectivity is procedural information encoded by Aspect rather than Tense, and it constrains the formulation of implicatures. The second is that subjectivity is processed at the level of general implicatures. These two hypotheses must be explored in further experimental work.

To conclude, the research carried out in this thesis has highlighted many questions in

need of further investigation.

Appendix

A1: Description of the corpora and their sources

Bilingual data: Literature Register

1. *The portrait of Mr. W. H.,* O. Wilde, French translation by J. Gattgno, Editions Gallimard, 2000. Electronic version and bilingual alignment by C. Grisot.

2. *Sense and sensibility*, J. Austen, French translation available on http://www.gutenberg.org/. Bilingual alignment by C. Grisot. Retrived on 30-04-2011.

3. *Le petit prince*, A. St. Exupery. Available at http://srogers.com/books/little_prince/contents.asp. Bilingual alignment by C. Grisot. Retrived on 30-04-2011.

4. *Cinq semaines en ballon*, J. Verne, Ch. 1. Corpus built by the FORELL laborytory in collaboration with Philippe Rivière. Available at www.cabal.rezo.net. Retrived on 30-04-2011.

5. *Vingt mille lieues sous les mers*, J. Verne, Ch. 1. Corpus built by the FORELL laborytory in collaboration with Philippe Rivière. Available at www.cabal.rezo.net. Retrived on 30-04-2011.

Bilingual data: Journalistic Register

6. *News Commentaries.* Translation corpus built for the International Workshop on Spoken Language and Translation. Available at http://iswlt2010.fbk.eu/node/34. Retrived on 30-04-2011.

7. *Time Magazine.* Corpus built by the FORELL laborytory in collaboration with Philippe Rivière. Available at www.cabal.rezo.net. Retrived on 30-04-2011.

8. *Presseurop* Website. http://www.presseurop.eu/fr. Bilingual alignment by C. Grisot. Retrieved on 30-04-2011.

9. *Le monde diplomatique*. Corpus built by the FORELL laboratory in collaboration with Philippe Rivière. Available at www.cabal.rezo.net. Retrived on 30-04-2011.

Bilingual data: Legislation and EuroParl Registers

10. *The JRC-Acquis Multilingual Parallel Corpus.* Built by J. Tiedemann (2009, 2012). Available at http://opus.lingfil.uu.se/JRC-Acquis.php Retrieved on 30-04-2011.

11. *EuConst Corpus.* Built by J. Tiedemann (2009). Available at http://opus.lingfil.uu.se/EUconst.php Retrieved on 30-04-2011.

12. *EuroParl* Corpus. Built by Philipp Koen (2005, 2012). Available at www.opus.lingfil.uu.se/ Retrieved on 30-04-2011.

Multilingual data:

13. *Alice in Wonderland*, L. Carol (e-book). French translation by Henry Bué (e-book), Italian translation by Pietrocola-Rossetti (e-book), Romanian translation by Popescu Bogdan (e-book). Multilingual alignment by M. Costagliola and C.

Grisot. Retrived on 30-03-2013.

14. *Presseurop* Website. http://www.presseurop.eu/fr. Multilingual alignment by M. Costagliola and C. Grisot. Retrieved on 30-09-2013.

15. *EuConst Corpus.* Built by J. Tiedemann (2009). Available at http://opus.lingfil.uu.se/EUconst.php Retrieved on 30-09-2013.

16. *EuroParl* Corpus. Built by Philipp Koen (2005). Available at http://www.statmt.org/europarl/. Retrieved on 30-09-2013.

A2: Annotation guidelines used in Experiment 1

Directives d'annotation

Introduction

Dans chaque extrait qui se trouve dans le fichier joint, vous pouvez identifier au moins deux événements ou états (nous allons les appeler des éventualités) qui sont présentés comme l'exemple suivant :

    a)   Jean est tombé [e2]. Sa jambe est cassée [e3]. Max l'a poussé. [e1]

Dans cet exemple, e1 et e2 sont des événements et e3 est un état.

Explication

Ces éventualités peuvent ou pas être liés. Il peut y avoir deux cas.

Cas 1:

Les éventualités sont liées temporellement. Cela veut dire qu' e1 a eu lieu avant e2. La relation peut être explicitée dans la phrase ou peut être implicite (comprise dans le contexte).

Exemple:

    b)   Son mariage a eu lieu le samedi. Ils sont partis en voyage de noces heureux et amoureux comme jamais.

Dans cet exemple, les deux événements sont 'le mariage qui a eu lieu' et 'ils sont partis. The deuxième événement est présenté en relation avec le premier (d'abord il s'est marié et ensuite sont partis en voyage de noces). Dans ce cas, la relation temporelle est implicite et elle peut être rendue explicite par l'insertion du connecteur *puis*. Après avoir inséré le connecteur, le sens de la phrase ne change pas.

Cas 2 :

Les éventualités ne sont pas liées temporellement. Cela veut dire qu'e1 et e2 ont eu lieu soit en même temps soit ne sont pas liées (cas opposé au cas 1).

Exemple :

    c)   Le propriétaire de cette maison était un homme seul, qui a vécu jusqu'un âge avancé et qui, pendant des années, a eu sa sœur comme ami fidèle et ménagère.

Dans cet exemple, il y a trois éventualités qui décrivent le propriétaire: 'était un homme seul', 'a vécu' et 'a eu'. Ces éventualités ne sont pas liées temporellement et donc on ne peut pas insérer le connecteur *puis* sans changer le sens de la phrase.

Tache

Pour chaque extrait, nous sommes intéressés dans le verbe écrit en *italique*. Merci de lire l'extrait et de juger si le verbe en italique a un usage narratif ou non-narratif, selon les définitions données.

A3: Annotation guidelines used in Experiment 2

Directives d'annotation

Introduction

Dans chaque extrait qui se trouve dans le fichier joint, vous pouvez lire des phrases qui contiennent un verbe conjugué à Imparfait écrits en *italique*. A chaque fois le verbe peut avoir deux usages : *narratif* et *non-narratif.*

Explication

Dans le premier cas (usage narratif) l'éventualité exprimée par le verbe à l'imparfait succède temporellement l'éventualité exprimée par le verbe précédent. Cela veut dire qu' e1 a eu lieu avant e2.

Exemple :

    a)   Elle a fini par fuguer à Kaboul, où elle a été recueillie par une femme généreuse. Quelques mois plus tard, elle *épousait* un jeune cousin de sa bienfaitrice dont elle était tombée amoureuse.

Dans cet exemple, l'éventualité *épousait* suit temporellement les éventualités précédentes, notamment 'a fini par fuguer' et 'a été recueillie'. Cela est aussi montré dans le texte par l'adverbe temporel 'quelques mois plus tard'. La relation temporelle pourrait aussi être rendue explicite par l'insertion du connecteur *puis*. Après avoir inséré le connecteur, le sens de la phrase ne change pas.

Dans le deuxième cas (usage non-narratif) l'éventualité exprimée par le verbe à l'imparfait a lieu simultanément que l'éventualité exprimée par le verbe précédent.

Exemple:

    b)   Il y a une heure Max *boudait* dans son coin, et ça n'est pas près de changer.

Les éventualités ne se succèdent pas temporellement. Elles sont simultanées. Cela veut dire qu'e1 et e2 ont eu lieu en même temps (cas opposé au premier cas).

Tache

Pour chaque extrait, nous sommes intéressés dans le verbe écrit en *italique.* Merci de lire l'extrait et de juger si le verbe conjuguer à l'imparfait écrit en italique a un usage narratif ou non-narratif, selon les définitions données.

A4: Annotation guidelines used in Experiment 3

Annotation guidelines

Introduction

In each of the excerpts in the attached file, you can identify at least two events or states (let's call them *eventualities*) that are presented as in the next example:

    a)   John fell [e2]. His leg is broken [e3]. Max pushed him [e1].

In this example *john fell* is an event, *his leg is broken* is a state and *max pushed him* is another event.

Explanation

These eventualities can or not be related each other. There are two cases:

Case 1:

The eventualities are temporally linked. This means that e1 happened before e2. The relation may be explicitly expressed in the sentence or it may be implicit (you understand it based on the context).

Example:

    b)   By his own marriage, likewise, which happened soon afterwards, he *added* to his wealth.

In this example, the two events are "the marriage that *happened*" and "the wealth which was *added*". The second event is presented in relation to the first one (first he got married and then he added to his wealth). In this case, the temporal relation is implicit and you can render it explicit by inserting the connective *and then*. After inserting the connectives, the meaning of the sentence remains the same.

Case 2:

The eventualities are not temporally linked. This means that e1 and e2 happened either in the same time (simultaneously) or they are not temporally linked (opposite to case 1 above).

Example:

c) The late owner of this estate was a single man, who lived to a very advanced age, and who for many years of his life, *had* a constant companion and housekeeper in his sister.

In this example, there are three eventualities that describe the owner of the estate. These states are "*was* a single man", "*lived*" and "*had* a companion". These eventualities are not temporally ordered but simultaneous. In this case, you cannot insert the connective *and then* without changing the meaning.

Task

For each excerpt, we are interested in the verb written in *italics*. Please read each excerpt and decide if the verb in italics has a narrative or a non-narrative usage, as they are described above.

A5: Annotation guidelines used in Experiment 4

Annotation guidelines

Introduction

In each of the excerpts in the attached file, you can find a verb written in italics. The verb expresses a situation that is viewed as being bounded or unbounded.

Explanation

Firstly, bounded situations are situations that have attained their natural endpoint as in example a), where the running of the one-mile race is finished. The same is the case for situations that do not have a natural endpoint but which have viewed as the finished as in example b).

a) Max ran the one-mile race.
b) I have lived in Paris from June to December 1998.

Secondly, bounded situations are situations that have not attained their natural endpoint as in example c) where the running of the one-mile race is not finished. The same is the case of situations that don't have a natural endpoint as in example d) where the living in Paris does not have a natural endpoint.

c) Max is running the one-mile race.
d) I have lived in Paris.

There are three linguistic tests that can help you judge if an eventuality is bounded or unbounded:

Test	Bounded	Unbounded
in/for adverbials	*in* adverbials	*for* adverbials
homogeneity	-	+
entailment with progressive	-	+

Task

For each excerpt, we are interested in the verb written in *italics*. Please read each excerpt and decide if the verb in italics expresses a bounded or an unbounded situation, as they are described above.

A6: Annotation guidelines used in Experiment 5

Annotation guidelines

Introduction

In each of the excerpts in the attached file, you can find a verb written in italics. The verb expresses a situation that is viewed as being perfective or imperfective.

Explanation

Firstly, perfective situations are viewed as being finished and the situation is a completed whole as in example a), where the letter was finished when John entered the president's office.

   a)   John entered the president's office. The president *wrote* a letter.

Secondly, imperfective situations are viewed as being in progress and the situation is not completed as in example b), where the letter was not finished when John entered the president's office.

   b)   John entered the president's office. The president *was writing* a letter.

<u>Task</u>

For each excerpt, we are interested in the verb written in *italics*. Please read each excerpt and decide if the verb in italics expresses a perfective or an imperfective situation, as they are described above.

A7: Annotation guidelines used in Experiment 6

Annotation guidelines

<u>Introduction</u>

In each of the excerpts in the attached file, you can find a verb written in italics. The verb expresses a situation that is located in the past. The most salient information expressed by the excerpt is the present result of the past situation or the past situation itself.

<u>Explanation</u>

The first is the case when the most salient information expressed by the excerpt is the present result of the past situation as in example a) where the most salient information is the false declaration.

   a)   And instead of full cooperation and transparency, Iraq *has filed* a false declaration to the United Nations that amounts to a 12,200-page lie.

The second is the case when most salient information expressed by the excerpt is the past situation itself as in example b) where the most salient information is the lack of choice of Musharraf.

   b)   In a historic ruling that Musharraf *had* little choice but to accept, the Supreme Court itself reinstated the Chief Justice in July.

<u>Task</u>

For each excerpt, we are interested in the verb written in *italics*. Please read each excerpt and decide if the verb in italics expresses the present result of a past situation or the past situation itself, as they are described above.

A9: Annotation guidelines used in Experiment 7

Directives d'annotation

<u>Introduction</u>

Dans cette expérience nous sommes intéressés à la conjugaison des verbes.

<u>Explication</u>

La conjugaison d'un verbe dépend du contexte. Par exemple, dans la phrase en a) le verbe au présent *pensons* est un indice pour conjuguer le verbe entre parenthèses au présent aussi.

   a)   Bien que nous soutenions pleinement les initiatives visant à lutter contre la fraude, nous avons décidé de voter contre le rapport, car nous ne pensons pas que la création d'un ministère public pour les questions financières (constituer) une solution appropriée au problème.

Un autre exemple est donné avec la phrase en b), où l'adverbe temporel *dans les années 1870* est un indice pour conjuguer le verbe entre parenthèses au passé composé ou au passé simple.

   b)   Quand Otto von Bismarck, le Chancelier de Fer, (s'attaquer) au pouvoir de l'Eglise catholique dans l'Allemagne unifiée depuis peu, dans les années 1870, l'affrontement fut

baptisé le Kukturkampf — la lutte pour la culture.

<u>Tache</u>

Dans chaque extrait qui se trouve dans le fichier joint, vous pouvez lire des phrases qui contiennent un verbe entre parenthèses qui est à l'infinitif. Vous devez le conjuguer afin que cela ne change pas le sens de l'extrait.

A11: Annotation guidelines used in Experiment 8

Guida per le annotazioni

<u>Introduzione</u>

In ciascuno degli estratti proposti nel file allegato, possono essere individuati almeno due eventi o stati (i.e. *eventualità*) così come mostrato nel seguente esempio:

    a)   Maria è caduta [e2]. La sua gamba è rotta [e3]. Antonio l'ha spinta [e1].

Nell' esempio (a) *Maria è caduta* è un evento, *la sua gamba è rotta* è uno stato e *Antonio l'ha spinta* è un ulteriore evento.

<u>Spiegazione</u>

Tali eventualità possono o meno essere correlate tra loro. Possono verificarsi i seguenti due casi.
Caso 1:

Le eventualità sono temporalmente/causalmente collegate. Ciò significa che e1 ha avuto luogo prima di e2 e/o e1 rappresenta la causa di e2. Nella frase, la relazione può essere espressa in modo esplicito o implicito (lo si evince dal contesto)
Esempio:

    b)   Scartò con ansia il pacchetto trovato sul tavolo; vi trovò una sorpresa inaspettata.

Nell'esempio riportato, i due eventi sono "scartò il pacchetto" e "vi trovò una sorpresa". Il secondo evento è in relazione col primo (dapprima scartò il pacchetto e poi vi trovò una sorpresa inaspettata). In questo caso, la relazione temporale/causale è implicita e si può esplicitare inserendo il connettore *e poi/prima* o *poiché/di conseguenza*. Dopo l'inserimento dei connettori il significato della frase resta invariato.
Caso 2

Le eventualità non sono temporalmente/causalmente collegate. Ciò significa che e1 e e2 hanno avuto luogo allo stesso tempo (simultaneamente) oppure che i due eventi non sono temporalmente/causalmente collegati (contrariamente al caso riportato *supra*)
Esempio:

    c)   Guardò tutta la sera un film in cui l'attore principale era Massimo Troisi che interpretava il ruolo di un postino.

In quest'esempio ci sono tre eventualità che non sono temporalmente/causalmente collegate; non si possono dunque aggiungere i connettori *poi/prima* o *poiché/di conseguenza* senza cambiare il senso della frase.

<u>Attività</u>

Per ogni esempio, considerare il verbo scritto in grassetto. Leggi con attenzione ogni esempio e decidi se il verbo in grassetto ha un uso narrativo (caso 1) o non narrativo (caso 2).

A12: Annotation guidelines used in Experiment 9

Indicații de annotare
<u>Introducere</u>

In fiecare din pasajele următoare, poți indentifica cel puțin două evenimente sau stări care s-au

317

întamplat, ca în exemplul următor.

  a) Ionut a căzut [e1]. Și-a rupt piciorul [e2]. Marc l-a împins [e3].

Explicație

Aceste evenimente pot sau nu avea legătura între ele. Există două cazuri:

Cazul 1:

Evenimentele sunt legate printr-o relație *temporală* si/sau *cauzală*. Aceasta înseamnă că e1 s-a întamplat înainte de e2 și/sau este cauza lui e2. Relația poate fi exprimată explicit în pasaj sau e implicită (este înțeleasă in context).

Exemplu:

  b) Prin căsătoria sa, de asemenea, care a avut loc [e1] după aceea, *s-a imbogățit* [e2] și mai mult.

In acest exemplu, sunt două evenimente: *căsătoria care a avut loc* și *imbogațirea*. Aceste două evenimente sunte legate printr-o relație temporală (mai întai s-a căsătorit si apoi s-a îmbogațit) și o relație causală (datorită căsătoriei, s-a imbogațit). În acest caz, aceste două relații sunt exprimate în mod implicit. Ele pot fi exprimate explicit dacă introducem un conector între cele două evenimente, ca *și apoi* sau *datorită*. Sensul propoziției nu se schimbă daca introducem acești conectori. Acest caz este numit *uzaj narativ*.

Cazul 2:

Evenimentele nu sunt legate printr-o relație *temporală* si/sau *cauzală*. Aceasta inseamnă ca e1 si e2 au avut loc in același timp (simultan) sau evenimentele nu sunt legate (caz opus cazului nr. 1).

Exemplu:

  c) Ultimul proprietar al acestei bogații *era* un om singur [e1], care trăise [e2] până la o vârstă
      înaintată, și care de-a lungul anilor, o avuse [e2] pe sora sa ca insoțitoare, prietenă și
      gospodină.

In acest exemplu, proprietarul e descris prin trei evenimente/stari: *era un om singur, înintat in vârstă* și *avuse o pe sora sa ca gospodină de-a lungul anilor*. Aceste evenimente/stări nu sunt legate printr-o relație temporală de succesiune si de simaltaneitate. În acest caz, nu putem introduce conectorii *și apoi* sau *datorită* fara să schimbăm sensul propoziției. Acest caz este numit *uzaj non-narativ*.

Sarcină

In fiecare propoziție sau pasaj, ne interesează verbul scris în *italic*. Citește te rog fraza și decide dacă verbul scris în italic est folosit in uzajul narativ sau non-narativ, cum le-am descris mai sus.

A12: Annotation guidelines used in Experiment 10

Directives d'annotation

Introduction

Dans chaque extrait qui se trouve dans le fichier joint, vous pouvez identifier au moins deux événements ou états (nous allons les appeler des éventualités) qui sont présentés comme l'exemple suivant :

  a) Jean est tombé [e2]. Sa jambe est cassée [e3]. Max l'a poussé. [e1]

Dans cet exemple, e1 et e2 sont des événements et e3 est un état.

Explication

Ces éventualités peuvent ou pas être liées. Il peut y avoir deux cas.

Cas 1 :

Les éventualités sont liées temporellement et causalement. Cela veut dire qu' e1 a eu lieu avant e2 et/ou e1 est la cause de e2 (e2 est la conséquence de e1). La relation peut être explicitée dans la phrase ou peut être implicite (comprise dans le contexte).

Exemple :

b)   Son mariage a eu lieu le samedi. Ils sont partis en voyage de noces heureux et amoureux comme jamais.

Dans cet exemple, les deux événements sont 'le mariage qui a eu lieu' et 'ils sont partis. The deuxième événement est présenté en relation avec le premier (d'abord il s'est marié et ensuite sont partis en voyage de noces). Dans ce cas, la relation temporelle est implicite et elle peut être rendue explicite par l'insertion du connecteur *puis* et la relation causale par le connecteur *alors.* Après avoir inséré le connecteur, le sens de la phrase ne change pas.

Cas 2 :

Les éventualités ne sont pas liées temporellement et causalement. Cela veut dire qu'e1 et e2 ont eu lieu soit en même temps soit ne sont pas liées (cas opposé au cas 1).

Exemple :

c)   Le propriétaire de cette maison était un homme seul, qui a vécu jusqu'un âge avancé et qui, pendant des années, a eu sa sœur comme ami fidèle et ménagère.

Dans cet exemple, il y a trois éventualités qui décrivent le propriétaire : 'était un homme seul', 'a vécu' et 'a eu'. Ces éventualités ne sont pas liées temporellement et donc on ne peut pas insérer le connecteur *puis* sans changer le sens de la phrase.

<u>Tache</u>

Pour chaque extrait, nous sommes intéressés dans le verbe écrit en *italique*. Merci de lire l'extrait et de juger si le verbe en italique a un usage narratif ou non-narratif, selon les définitions données.

References

Aarts, B. (2011). *Oxford Modern English Grammar*. Oxford: Oxford University Press.

Ahern, A. (2010). Propositional attitudes in relevance theory: An overview. In E. Wałaszewska, M. Kisielewska-Krysiuk, & A. Piskorska (Eds.), *In the Mind and across Minds: A Relevance-Theoretic Perspective on Communication and Translation* (pp. 147–167). Newcastle upon Tyne: Cambridge Scholars Publishing.

Ahern, A., & Leonetti, M. (2004). The Spanish subjunctive. In R. Márquez-Reiter & M. E. Placencia (Eds.), *Current trends in the pragmatics of Spanish* (pp. 35–56). Amsterdam: John Benjamins.

Aijmer, K., & Altenberg, B. (1996). Introduction. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast: papers from a symposium on text-based cross-linguistic studies, Lund, 4-5 March 1994*. Lund: Lund University Press.

Allen, R. (1966). *The Verb System of Present-Day American English*. The Hague: Mouton.

Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. In B. Altenberg & S. Granger (Eds.), *Lexis in Contrast: corpus-based approaches* (pp. 3–50). Amsterdam: John Benjamins.

Amenós-Pons, J. (2011). Cross-linguistic variation in procedural expressions: Semantics and pragmatics. In V. Escandell-Vidal, M. Leonetti, & A. Ahern (Eds.), *Procedural meaning: Problems and perspectives* (pp. 235–266). Bingley: Emerald Group Publishing.

Anscombre, J.-C., & Ducrot, O. (1983). *L'argumentation dans la langue*. Bruxelles: Editions Mardaga.

Ariel, M. (1994). Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics*, *30*(01), 3–42.

Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, *34*(4), 555–596.

Asher, N. (1993). *Reference to abstract objects in discourse*. Dordrecht: Kluwer Academic Publishers.

Asher, N. (2013). Temporal modification. In K. M. Jaszczolt & L. Saussure de (Eds.), *Time: Language, Cognition & Reality* (pp. 15–36). Oxford: Oxford University Press.

Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.

Asher, N., & Morreau, M. (1991). Commonsense entailment: A modal theory of nonmonotonic reasoning. *Logics in AI*, 1–30.

Atlas, J. D., & Levinson, S. C. (1981). It-clefts, informativeness and logical form: radical pragmatics. In P. Cole (Ed.), *Radical pragmatics* (pp. 1–62). New York: Academic Press.

Austin, J. L. (1957). A plea for excuses. *Philosophical Papers*, 175–204.

Avrutin, S. (2000). Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. *Language and the Brain: Representation and Processing*, 295–313.

Avrutin, S. (2006). Weak syntax. In K. Amunts & Y. Grodzinsky (Eds.), *Broca's region* (pp. 49–62). New York: Oxford University Press.

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge: Cambridge University Press.

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, *9*(1), 5–16.

Bach, K. (1994). Conversational impliciture. *Mind & Language*, *9*(2), 124–162.

Bach, K. (2001). Semantically speaking. In I. Kecskes & R. M. Harnish (Eds.), *Perspectives on Semantics, Pragmatics, and Discourse: A Festschrift for Franc Kiefer* (pp. 147–172). Amsterdam: John Benjamins.

Baggio, G. (2008). Processing temporal constraints: An ERP study. *Language Learning*, *58*(s1), 35–55.

Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233–253). Amsterdam: John Benjamins.

Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, *7*(2), 223–243.

Banfield, A. (1982). *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan Paul.

Baranzini, L., & Ricci, C. (2015). Semantic and pragmatic values of the Italian imperfetto: towards a common interpretative procedure. *Catalan Journal of Linguistics*, to appear.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization* (pp. 101–140). Cambridge: Cambridge University Press.

Bastiaanse, R. (2008). Production of verbs in base position by Dutch agrammatic speakers: Inflection versus finiteness. *Journal of Neurolinguistics*, *21*(2), 104–119.

Bastiaanse, R., Bamyaci, E., Hsu, C.-J., Lee, J., Duman, T. Y., & Thompson, C. K. (2011). Time reference in agrammatic aphasia: A cross-linguistic study. *Journal of Neurolinguistics*, *24*(6), 652–673.

Bastiaanse, R., Hugen, J., Kos, M., & van Zonneveld, R. (2002). Lexical, morphological, and syntactic aspects of verb production in agrammatic aphasics. *Brain and Language*, *80*(2), 142–159.

Bastiaanse, R., & Thompson, C. K. (2003). Verb and auxiliary movement in agrammatic Broca's aphasia. *Brain and Language*, *84*(2), 286–305.

Bayerl, P. S., & Paul, K. I. (2011). What determines inter-coder agreement in manual annotations? a meta-analytic investigation. *Computational Linguistics*, *37*(4), 699–725.

Behrens, B., & Fabricius-Hansen, C. (2003). Translation Equivalents as Empirical Data for Semantic/Pragmatic Theory. The Case of Elaboration. *Meaning through Language Contrast*, *2*, 463–476.

Belletti, A. (1990). *Generalised verb movement: Aspects of verb syntax*. Torino: Rosenberg & Sellier.

Benveniste, E. (1959). *Les relations de temps dans le verbe français*. Berlin: Klincksieck.

Benveniste, E. (1966). *Problèmes de linguistique générale*. Paris: Gallimard.

Berthonneau, A.-M., & Kleiber, G. (1993). Pour une nouvelle approche de l'imparfait: l'imparfait, un temps anaphorique méronomique. *Langages*, 55–73.

Bertineto, P. M. (1986). *Tempo, aspetto e azione nel verbo italiano*. Firenze: Presso L'Accademia della Crusca.

Bertinetto, P. M. (1996). La distribuzione del Perfetto Semplice e del Perfetto Composto nelle diverse varieta di italiano. *Romance Philology*, *49*, 383–419.

Bertinetto, P. M., Bianchi, V., Dahl, Ö., & Squartini, M. (1995). *Temporal reference: Aspect and actionality*. Torino: Rosenberg & Sellier.

Bhat, D. S. (1999). *The prominence of tense, aspect and mood*. Amsterdam: John Benjamins Publishing.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, *8*(4), 243–257.

Binnick, R. I. (1991). *Time and the verb: A guide to tense and aspect*. New York: Oxford University Press.

Binnick, R. I. (2009). Tense and aspect. In F. Brisard, J.-O. Östman, & J. Verschueren (Eds.), *Grammar, meaning and pragmatics* (pp. 268–288). Amsterdam: John Benjamins.

Binnick, R. I. (2012). *The Oxford handbook of tense and aspect*. Oxford: Oxford University Press.

Bittar, A. (2010). *Building a TimeBank for French: a reference corpus annotated according to the ISO-TimeML standard* (PhD Dissertation). Paris 7, Paris.

Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Blackwell.

Blakemore, D. (1988). So as a constraint on relevance. In R. Kempson (Ed.), *Mental representation: The interface between language and reality* (pp. 183–195). Cambridge: Cambridge University Press.

Blakemore, D. (1992). *Understanding utterances: An introduction to pragmatics*. Oxford: Blackwell.

Blakemore, D. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.

Blass, R. (1989). Grammaticalisation of interpretive use: the case of re in Sissala. *Lingua*, *79*(4), 299–326.

Blochowiak, J. (2009). La relation causale, ses relata et la négation. *Nouveaux Cahiers de Linguistique Française*, *29*, 149–172.

Blochowiak, J. (2014a). A Presuppositional Account of Causal and Temporal Interpretations of and. *Topoi*, 1–15.

Blochowiak, J. (2014b). *A theoretical approach to the quest for understanding. Semantics and pragmatics of whys and becauses* (PhD Dissertation). Université de Genève, Genève.

Blochowiak, J. (2015a). *Connectives: concepts, procedures or both*. Presented at the 48th Annual Meeting of the Societas Linguistica Europaea, University of Leiden, Netherlands.

Blochowiak, J. (2015b). *Logical meaning of non-logical connectives and non-logical meaning of logical connectives*. Presented at the 1st International LogPrag Workshop, Les Diablerets.

Blücher, K. (1974). L'imperfetto italiano è veramente un imperfetto? In *Actes du XIIIe Congrès International de Linguistique et Philologie Romanes* (pp. 397–402). Québec: Les Presses de l'Université Laval.

Bohnemeyer, J. (2009). Temporal anaphora in a tenseless language. In W. Klein & P. Li (Eds.), *The expression of time in language* (pp. 83–128). Berlin: Mouton de Gruyter.

Bohnemeyer, J., & Swift, M. (2004). Event realization and default aspect. *Linguistics and Philosophy*, *27*(3), 263–296.

Boogaart, R. (2007). The past and perfect of epistemic modals. In L. Saussure de, J. Moeschler, & G. Puskas (Eds.), *Recent advances in the syntax and semantics of tense, mood and aspect.* (Vol. 185, pp. 47–70). Berlin: Walter de Gruyter.

Boogaart, R. J. U. (1999). *Aspect and temporal ordering: A constrastive analysis of Dutch and English* (PhD Dissertation). Holland Academic Graphics, The Hague.

Bos, L. S., Dragoy, O., Stowe, L. A., & Bastiaanse, R. (2013). Time reference teased apart from tense: Thinking beyond the present. *Journal of Neurolinguistics*, *26*(2), 283–297.

Bras, M. (2008). *Entre relations temporelles et relations de discours* (Rapport). Toulouse: Université de Toulouse II-Le Mirail.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, *100*(3), 434–463.

Brown, P. F., Cocke, J., Pietra, S. A. D., Pietra, V. J. D., Jelinek, F., Lafferty, J. D., … Roossin, P. S. (1990). A statistical approach to machine translation. *Computational Linguistics*, *16*(2), 79–85.

Brown, P. F., Pietra, S. A. D., Pietra, V. J. D., & Mercer, R. L. (1991). Word-sense disambiguation using statistical methods. In *Proceedings of the 29th annual meeting on Association for Computational Linguistics* (pp. 264–270).

Brown, P. F., Pietra, V. J. D., Pietra, S. A. D., & Mercer, R. L. (1993). The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, *19*(2), 263–311.

Brunot, F. (1922). *La pensée et la langue: méthode, principes et plan d'une théorie nouvelle du langage appliquée au français*. Paris: Masson.

Burchert, F., Swoboda-Moll, M., & Bleser, R. D. (2005). Tense and agreement dissociations in German agrammatic speakers: Underspecification vs. hierarchy. *Brain and Language*, *94*(2), 188–199.

Caenepeel, M. (1989). *Aspect, temporal ordering and perspective in narrative fiction* (PhD Dissertation). University of Edinburgh, Edinburgh.

Caenepeel, M. (1995). Aspect and text structure. *Linguistics*, *33*(2), 213–254.

Calarașu, C. (1992). Quelques significations des temps verbaux roumains en perspective romane (avec applications aux langues roumaine et française). *Revue Roumaine de Linguistique*, (37), 137–143.

Carletta, J. (1996). Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, *22*(2), 249–254.

Carl, M., & Way, A. (2003). *Recent advances in example-based machine translation*. Dordrecht: Kluwer Academic Publishers.

Carruthers, P. (2006). *The architecture of the mind*. Oxford: Oxford University Press.

Carston, R. (1988). Implicature, explicature, and truth-theoretic semantics. In R. Kempson (Ed.), *Mental representations: The interface between language and reality* (pp. 155–181). Cambridge: Cambridge University Press.

Carston, R. (1996). Metalinguistic negation and echoic use. *Journal of Pragmatics*, *25*(3), 309–330.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and Implications* (pp. 179–238). Amsterdam: John Benjamins.

Carston, R. (2000). Explicature and semantics. *UCL Working Papers in Linguistics*, *12*(1), 44–90.

Carston, R. (2002a). Metaphor, ad hoc concepts and word meaning–more questions than answers. *UCL Working Papers in Linguistics*, *14*, 83–105.

Carston, R. (2002b). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford:

Blackwell.

Carston, R. (2004). Relevance theory and the saying/implicating distinction. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 633–656). Oxford: Blackwell Publishing.

Carston, R. (2009). The explicit/implicit distinction in pragmatics and the limits of explicit communication. *International Review of Pragmatics*, *1*(1), 35–62.

Carston, R. (2010a). Explicit communication and "free"pragmatic enrichment. In S. Belén (Ed.), *Explicit communication: Robyn Carston's pragmatics* (pp. 217–285). London: Palgrave Macmillan.

Carston, R. (2010b). Lexical pragmatics, ad hoc concepts and metaphor: from a relevance theory perspective. *Italian Journal of Linguistics*, *22*(1), 153–180.

Carston, R. (2012). Word meaning and concept expressed. *The Linguistic Review*, *29*(4), 607–623.

Carston, R., & Hall, A. (2012). Implicature and explicature. In H.-J. Schmid (Ed.), *Cognitive pragmatics* (p. 47). Berlin, Boston: Mouton de Gruyter.

Carston, R., & Powell, G. (2006). Relevance theory–new directions and developments. In E. Lepore & B. C. Smith (Eds.), *The Oxford Handbook of Philosophy of Language* (pp. 279–299). Oxford: Oxford University Press.

Cartoni, B., Zufferey, S., & Meyer, T. (2013). Annotating the meaning of discourse connectives by looking at their translation: The translation-spotting technique. *Dialogue & Discourse*, *4*(2), 65–86.

Chen, W. (2008). Dimensions of subjectivity in natural language. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers* (pp. 13–16). Association for Computational Linguistics.

Chesterman, A. (1998). *Contrastive functional analysis*. Amsterdam: John Benjamins Publishing.

Chevallier, C., Bonnefond, M., Van der Henst, J.-B., & Noveck, I. A. (2010). Using ERPs to capture inferential processes guided by prosodic cues. *Italian Journal of Linguistics*, *22*(1), 125–152.

Chilton, P. (2005). Vectors, viewpoint and viewpoint shift. *Annual Review of Cognitive Linguistics*, *3*(1), 78–116.

Chilton, P. (2007). Geometrical concepts at the interface of formal and cognitive models: Aktionsart, aspect, and the English progressive. *Pragmatics & Cognition*, *15*(1), 91–114.

Chilton, P. (2014). *Language, Space and Mind: The Conceptual Geometry of Linguistic Meaning*. Cambridge: Cambridge University Press.

Chomsky, N. (1957). *Syntactic Structures*. Berlin: Walter de Gruyter.

Chomsky, N. (1965). *Aspectsof the theory of syntax*. Cambridge MA: MIT Press.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham, MA: Blaisdell.

Chomsky, N. (1981). *Lectures on government and binding*. Cambridge MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89–155). Cambridge MA: MIT Press.

Clahsen, H., & Ali, M. (2009). Formal features in aphasia: tense, agreement, and mood in English agrammatism. *Journal of Neurolinguistics*, *22*(5), 436–450.

Cohen, J., & others. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cohen, L. J., & Bar-Hillel, Y. (1971). Some remarks on Grice's views about the logical particles of natural language. In *Pragmatics of natural languages* (pp. 50–68). Dordrecht: Springer.

Coles, E. (1677). *Syncrisis*. Menston: Scolar Press.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Comrie, B. (1981). *The languages of the Soviet Union*. Cambridge: Cambridge University Press.

Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.

Cowper, E. (2002). *Finiteness* (PhD Dissertation). University of Toronto, Toronto.

Cowper, E. (2005). The geometry of interpretable features: Infl in English and Spanish. *Language*, 10–46.

Dahl, O. (1981). On the Definition of the Telic-Atelic (Bounded–Nonbounded). In P. J. Tedeschi & A. Zaenen (Eds.), *Tense and Aspect* (pp. 79–90). New York: Academic Press.

Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Blackwell.

Dahl, Ö., & Velupillai, V. (2013a). Perfective/Imperfective Aspect. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/65

Dahl, Ö., & Velupillai, V. (2013b). Tense and Aspect. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzing: Max Planck Institute for Evolutionary Anthropology.

Damourette, J., & Pichon, E. (1911). *Des mots à la pensée. Essai de grammaire de la langue française* (Vols. 1–7). Paris: D'Artrey.

Davidson, D. (1967). Causal relations. *The Journal of Philosophy*, *64*(21), 691–703.

Davidson, D. (1980). *Essays on actions and events*. Oxford: Clarendon Press.

Declerck, R. (1979). Aspect and the bounded/unbounded (telic/atelic) distinction. *Linguistics*, *17*(9-10), 761–794.

Declerck, R. (1986). From Reichenbach (1947) to Comrie (1985) and beyond: towards a theory of tense. *Lingua*, *70*(4), 305–364.

Declerck, R. (1989). Boundedness and the structure of situations. *Leuvense Bijdragen*, *78*, 275–308.

Declerck, R. (1991a). *A comprehensive Descriptive Grammar of English*. Tokyo: Kaitakusha.

Declerck, R. (1991b). *Tense in English: its structure and use in discourse*. London: Routledge.

Declerck, R. (2006). *The grammar of the English verb phrase*. Berlin: Mouton de Gruyter.

Demirdache, H., & Uribe-Etxebarria, M. (2007). Economy constraints on temporal subordination. In L. Saussure de, J. Moeschler, & G. Puskas (Eds.), *Recent advances in the syntax and semantics of tense, mood and aspect.* (pp. 169–192). Berlin: Walter de Gruyter.

Depraetere, I. (1995a). On the necessity of distinguishing between (un) boundedness and (a) telicity. *Linguistics and Philosophy*, *18*(1), 1–19.

Depraetere, I. (1995b). The effect of temporal adverbials on (a) telicity and (un) boundedness. In P. M. Bertineto, V. Bianchi, Ö. Dahl, & M. Squartini (Eds.), *Temporal reference, aspect and actionality* (pp. 43–53). Torino: Rosenberg & Sellier.

De Swart, H. (1998). Aspect shift and coercion. *Natural Language & Linguistic Theory*, *16*(2),

Cohen, J., & others. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cohen, L. J., & Bar-Hillel, Y. (1971). Some remarks on Grice's views about the logical particles of natural language. In *Pragmatics of natural languages* (pp. 50–68). Dordrecht: Springer.

Coles, E. (1677). *Syncrisis*. Menston: Scolar Press.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Comrie, B. (1981). *The languages of the Soviet Union*. Cambridge: Cambridge University Press.

Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.

Cowper, E. (2002). *Finiteness* (PhD Dissertation). University of Toronto, Toronto.

Cowper, E. (2005). The geometry of interpretable features: Infl in English and Spanish. *Language*, 10–46.

Dahl, O. (1981). On the Definition of the Telic-Atelic (Bounded–Nonbounded). In P. J. Tedeschi & A. Zaenen (Eds.), *Tense and Aspect* (pp. 79–90). New York: Academic Press.

Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Blackwell.

Dahl, Ö., & Velupillai, V. (2013a). Perfective/Imperfective Aspect. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/65

Dahl, Ö., & Velupillai, V. (2013b). Tense and Aspect. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzing: Max Planck Institute for Evolutionary Anthropology.

Damourette, J., & Pichon, E. (1911). *Des mots à la pensée. Essai de grammaire de la langue française* (Vols. 1–7). Paris: D'Artrey.

Davidson, D. (1967). Causal relations. *The Journal of Philosophy*, *64*(21), 691–703.

Davidson, D. (1980). *Essays on actions and events*. Oxford: Clarendon Press.

Declerck, R. (1979). Aspect and the bounded/unbounded (telic/atelic) distinction. *Linguistics*, *17*(9-10), 761–794.

Declerck, R. (1986). From Reichenbach (1947) to Comrie (1985) and beyond: towards a theory of tense. *Lingua*, *70*(4), 305–364.

Declerck, R. (1989). Boundedness and the structure of situations. *Leuvense Bijdragen*, *78*, 275–308.

Declerck, R. (1991a). *A comprehensive Descriptive Grammar of English*. Tokyo: Kaitakusha.

Declerck, R. (1991b). *Tense in English: its structure and use in discourse*. London: Routledge.

Declerck, R. (2006). *The grammar of the English verb phrase*. Berlin: Mouton de Gruyter.

Demirdache, H., & Uribe-Etxebarria, M. (2007). Economy constraints on temporal subordination. In L. Saussure de, J. Moeschler, & G. Puskas (Eds.), *Recent advances in the syntax and semantics of tense, mood and aspect.* (pp. 169–192). Berlin: Walter de Gruyter.

Depraetere, I. (1995a). On the necessity of distinguishing between (un) boundedness and (a) telicity. *Linguistics and Philosophy*, *18*(1), 1–19.

Depraetere, I. (1995b). The effect of temporal adverbials on (a) telicity and (un) boundedness. In P. M. Bertineto, V. Bianchi, Ö. Dahl, & M. Squartini (Eds.), *Temporal reference, aspect and actionality* (pp. 43–53). Torino: Rosenberg & Sellier.

De Swart, H. (1998). Aspect shift and coercion. *Natural Language & Linguistic Theory*, *16*(2),

347–385.

De Swart, H. (2003). Coercion in a cross-linguistic theory of aspect. In E. Francis & L. A. Michaelis (Eds.), *Mismatch: Form-function incongruity and the architecture of grammar* (pp. 231–258). Stanford CA: CSLI Publications.

De Swart, H. (2011). Mismatches and coercion. In C. Maieborn, K. von Heusinger, & P. Portner (Eds.), *International Handbook of Natural Language Meaning.* (pp. 574–598). Berlin: Mouton de Gruyter.

Dîndelegan, G. P. (2013). *The grammar of Romanian.* Oxford: Oxford University Press.

Doron, E. (1991). Point of view as a factor of content. In S. Moore & A. Z. Wyner (Eds.), *Proceedings of SALT* (pp. 51–64). New York: Cornell University Press.

Dorr, B. J., & Gaasterland, T. (1995). Selecting tense, aspect, and connecting words in language generation. In *IJCAI* (pp. 1299–1307). Montreal.

Dorr, B. J., & Gaasterland, T. (2002). *Constraints on the generation of tense, aspect, and connecting words from temporal expressions* (Technical Report). Computer Science Department, University of Maryland. Retrieved from http://drum.lib.umd.edu/handle/1903/1218

Dowty, D. R. (1972). *Studies in the logic of verb aspect and time reference in English* (PhD Dissertation). University of Texas, Austin.

Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ.* Dordrecht: Kluwer Academic Publishers.

Dowty, D. R. (1982). Tenses, time adverbs, and compositional semantic theory. *Linguistics and Philosophy*, *5*(1), 23–55.

Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and Philosophy*, *9*(1), 37–61.

Dragoy, O., & Bastiaanse, R. (2013). Aspects of time: Time reference and aspect production in Russian aphasic speakers. *Journal of Neurolinguistics*, *26*(1), 113–128.

Dragoy, O., Stowe, L. A., Bos, L. S., & Bastiaanse, R. (2012). From time to time: Processing time reference violations in Dutch. *Journal of Memory and Language*, *66*(1), 307–325.

Dry, H. A. (1981). Sentence aspect and the movement of narrative time. *Text*, *1*(3), 233–240.

Dry, H. A. (1983). The movement of narrative time. *Journal of Literary Semantics*, *12*(2), 19–53.

Ducrot, O. (1979). L'imparfait en français. *Linguistische Berichte Braunschweig*, (60), 1–23.

Dyvik, H. (1998). A translational basis for semantics. In S. Johansson & S. Oksefjell (Eds.), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies* (pp. 51–86). Amsterdam: Rodopi.

Enç, M. (1996). Tense and modality. In S. Lappin (Ed.), *The Handbook of Contemporary Semantic Theory* (pp. 345–358). Oxford: Blackwell.

Escandell-Vidal, V., & Leonetti, M. (2011). On the rigidity of procedural meaning. In V. Escandell-Vidal, M. Leonetti, & A. Ahern (Eds.), *Procedural Meaning: Problems and Perspectives* (pp. 81–102). Bingley: Emerald Group Publishing.

Faroqi-Shah, Y., & Dickey, M. W. (2009). On-line processing of tense and temporality in agrammatic aphasia. *Brain and Language*, *108*(2), 97–111.

Faroqi-Shah, Y., & Thompson, C. K. (2007). Verb inflections in agrammatic aphasia: Encoding of tense features. *Journal of Memory and Language*, *56*(1), 129–151.

Ferro, L., Mani, I., Sundheim, B., & Wilson, G. (2001). *TIDES Temporal Annotation Guidelines-Version 1.0. 2* (MITRE Technical Report). McLean-VG-USA: The MITRE Corporation.

Fillmore, C. J. (1992). Corpus linguistics or computer-aided armchair linguistics. In J. Svartvik (Ed.), *Directions in corpus linguistics* (pp. 35–60). Berlin: Mouton de Gruyter.

Fiorin, G. (2010). *Meaning and dyslexia: a study on pronouns, aspect, and quantification* (PhD Dissertation). Netherlands Graduate School of Linguistics, Utrecht.

Fisiak, J. (1981). *Some introductory notes concerning contrastive linguistics*. Oxford: Pergamon.

Fleischman, S. (1985). Discourse functions of tense-aspect oppositions in narrative: toward a theory of grounding. *Linguistics*, *23*(6), 851–882.

Fleischman, S. (1990). *Tense and narrativity*. London: Routledge.

Fodor, J. A. (1975). *The language of thought*. Cambridge MA: Harvard University Press.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge MA: MIT press.

Friedmann, N. 'ama, & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: Pruning the syntactic tree. *Brain and Language*, *56*(3), 397–425.

Friedmann, N. (2008). Traceless relatives: Agrammatic comprehension of relative clauses with resumptive pronouns. *Journal of Neurolinguistics*, *21*(2), 138–149.

Gagarina, N. (2004). Does the acquisition of aspect have anything to do with aspectual pairs. *ZAS Papers in Linguistics*, *33*, 39–61.

Garey, H. B. (1957). Verbal aspect in French. *Language*, *33*(2), 91–110.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.

Gellerstam, M. (1996). Translations as a source for cross-linguistic studies. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast: papers from a symposium on text-based cross-linguistic studies* (pp. 53–62). Lund: Lund Universtity Pres.

Genette, G. (1972). *Discours du recit*. Paris: Seuil.

Gibbs Jr, R. W., & Moise, J. F. (1997). Pragmatics in understanding what is said. *Cognition*, *62*(1), 51–74.

Gibbs, R. W., & Matlock, T. (1999). Psycholinguistics and mental representations. *Cognitive Linguistics*, *10*(3), 263–270.

Gilquin, G., & Gries, S. T. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory*, *5*(1), 1–26.

Giorgi, A., & Pianesi, F. (1997). *Tense and Aspect: Form Semantics to Morphosyntax*. Oxford: Oxford University Press.

Gong, Z., Zhang, M., Tan, C. L., & Zhou, G. (2012a). Classifier-Based Tense Model for SMT. In *COLING (Posters)* (pp. 411–420).

Gong, Z., Zhang, M., Tan, C., & Zhou, G. (2012b). N-gram-based tense models for statistical machine translation. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning* (pp. 276–285).

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in Contrast. Text-based cross-linguistic studies* (pp. 37–51). Lund: Lund University Press.

Granger, S. (2003). The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies. In S. Granger, J. Lerot, & S. Petch-Tyson (Eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* (pp. 17–29).

Amsterdam: Rodopi.

Granger, S., Hung, J., & Petch-Tyson, S. (2002). *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching*. Amsterdam: John Benjamins.

Grevisse, M. (1980). *Le bon usage* (11th edition). Bruxelles: De Boeck Duculot.

Grice, H. P. (1967). *William James Lectures*. Cambridge MA: Harvard University Press.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3: Speech Acts). New York: Academic Press.

Grice, H. P. (1989). *Study in the way of words*. Cambridge: Cambridge University Press.

Gries, S. T. (2002). Evidence in linguistics: Three approaches to genitives in English. *The LACUS Forum*, *28*, 17–31.

Gries, S. T. (2009). *Quantitative corpus linguistics with R: a practical introduction*. Oxon: Routledge.

Gries, S. T. (2010). Useful statistics for corpus linguistics. In A. Sanchez & M. Almela (Eds.), *A mosaic of corpus linguistics: selected approaches* (pp. 269–291). Frankfurt am Main: Peter Lang.

Gries, S. T. (2014). Frequency tables: Tests, effect sizes, and explorations. In D. Glynn & J. A. Robinson (Eds.), *Human Cognitive Processing*. Amsterdam: John Benjamins.

Grisot, C. (2015). Tense, Aspect and Subjectivity: experimental approach. *In Preparation*.

Grisot, C., & Cartoni, B. (2012). Une description bilingue des temps verbaux: étude contrastive en corpus. *Nouveaux Cahiers de Linguistique Française*, *30*, 101–117.

Grisot, C., & Costagliola, M. (2014). *What do translation corpora tell us about the semantics and pragmatics of tense? The case of the English preterit*. Presented at the Chronos 11, Scuola Normale Superiore, Italy.

Grisot, C., & Meyer, T. (2014). Cross-linguistic annotation of narrativity for English/French verb tense disambiguation. In *9th Edition of the Language Resources and Evaluation Conference* (pp. 963–966). Reykjavik, Island.

Grisot, C., & Moeschler, J. (2014). How do empirical methods interact with theoretical pragmatics? The conceptual and procedural contents of the English Simple Past and its translation into French. In J. Romero-Trillo (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 2014: New Empirical and Theoretical Paradigms* (pp. 7–33). Switzerland: Springer International Publishing.

Grivaz, C. (2012). *Automatic extraction of causal knowledge from natural language texts* (PhD Dissertation). Université de Genève, Genève.

Grote, B. (1998). Representing temporal discourse markers for generation purposes. In *Coling/ACL Workshop on Discourse Relations and Discourse Markers* (pp. 22–28).

Grote, B. (2003). *Signaling coherence relations in text generation: A case study of German temporal discourse markers* (PhD Dissertation). University of Bremen, Bremen.

Guenthner, F., Hoepelman, J., & Rohrer, C. (1978). A note on the passé simple. In C. Rohrer (Ed.), *Papers on Tense. Aspect and Verb Classification* (pp. 18–36). Tüurbingen: Narr.

Guéron, J. (1993). Sur la syntaxe du temps. *Langue Française*, 102–122.

Guéron, J. (2007). On tense and aspect. *Lingua*, *117*(2), 367–391.

Guéron, J. (2008). On the difference between telicity and perfectivity. *Lingua*, *118*(11), 1816–1840.

Guillaume, G. (1929). *Temps et verbe: théorie des aspects, des modes, et des temps*. Paris: Champion.

Guillaume, G. (1971). *Leçons de linguistique de Gustave Guillaume* (Vols. 1–12). Québec: Presses

Université Laval.

Guillemin-Flescher, J. (1981). *Syntaxe comparée du français et de l'anglais: problèmes de traduction*. Paris: Editions Ophrys.

Guțu-Romalo, V. (2005). *Gramatica Limbii Române (GLR)* (Vols. 1–2). București: Editura Academiei Române.

Hackmack, S. (n.d.). *Reichenbach's Theory of Tense and it's Application to English*. Retrieved from http://www.fb10.uni-bremen.de/khwagner/tempus/pdf/reich.pdf

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hall, R. A. (1964). *Introductory linguistics*. New York: Chilton Books.

Harris, M. (1982). The "past simple" and the "present perfect" in Romance. *Studies in the Romance Verb*, 42–70.

Hewes, J. (1624). *A Perfect Survey of the English Tongue*. London: William Garret.

Hinrichs, E. (1986). Temporal anaphora in discourses of English. *Linguistics and Philosophy*, *9*(1), 63–82.

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, *3*(1), 67–90.

Hobbs, J. R. (1985). *On the coherence and structure of discourse* (No. CSLI-85-37). Center for the Study of Language and Information: Stanford University.

Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context* (pp. 11–42). Washington DC: Georgetown University Press.

Horn, L. (1992). The said and the unsaid. *Ohio State University Working Papers in Linguistics*, *40*, 163–192.

Horn, L. (2004). Implicature. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 3–28). Oxford: Blackwell.

Horn, L. R. (2006). The border wars: A neo-Gricean perspective. In K. von Heusinger & K. Turner (Eds.), *Where Semantics Meets Pragmatics: The Michigan Papers* (pp. 21–48). Amsterdam: Elsevier.

Horn, L. R. (2007). Toward a Fregean pragmatics: Voraussetzung, Nebengedanke, Andeutung. In I. Kecskes & L. Horn (Eds.), *Explorations in Pragmatics: Linguistic, Cognitive and Intercultural Aspects*. Berlin: Mouton de Gruyter.

Hornstein, N. (1990). *As time goes by*. Cambridge MA: MIT Press.

Howell, J. (1662). *A new English grammar prescribing as certain rules as the languages will bear, for forreners to learn English*. London: H. Brome and H. Marsh.

Huddleston, R. D. (2006). *A student's introduction to English grammar* (2nd edition). Cambridge: Cambridge University Press.

Hume, D. (1738). *A treatise of human nature*. London: Millar.

Ifantidou, E. (2000). Procedural encoding of explicatures by the Modern Greek particle taha. In G. Andersen & T. Fretheim (Eds.), *Pragmatic Markers and Propositional Attitude* (pp. 119–144). Amsterdam: John Benjamins.

Ifantidou, E. (2001). *Evidentials and relevance*. Amsterdam: John Benjamins.

Imbs, P. (1960). *L'emploi des temps verbaux en français moderne*. Paris: Klincksieck.

Ingria, R., & Pustejovsky, J. (2002). *TimeML Specification 1.0*.

James, C. (1980). *Contrastive Analysis*. London: Longman.

Jaszczolt, K. (2005). *Default semantics: Foundations of a compositional theory of acts of communication*.

Oxford: Oxford University Press.

Jaszczolt, K. M. (2009). *Representing time: an essay on temporality as modality*. Oxford: Oxford University Press.

Jaszczolt, K. M. (2010). Defaults in semantics and pragmatics. Retrieved February 10, 2015, from http://stanford.library.usyd.edu.au/entries/defaults-semantics-pragmatics/

Jaszczolt, K. M. (2012). Cross-linguistic differences in expressing time and universal principles of utterance interpretation. In L. Filipovic & K. M. Jaszczolt (Eds.), *Space and time in languages and cultures: Linguistic diversity* (pp. 95–122). Amsterdam: John Benjamins.

Jespersen, O. (1961). *A Modern English Grammar on historical principles* (reprinted ed.). London: George Allen & Unwin Limited.

Jespersen, O. (1965). *The philosophy of grammar*. Chicago: University of Chicago Press.

Johansson, S. (1998). On the role of corpora in cross-linguistic research. In S. Johansson & S. Oksefjell (Eds.), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies* (pp. 3–24). Amsterdam: Rodopi.

Johansson, S. (2003). Contrastive linguistics and corpora. In S. Granger, J. Lerot, & S. Petch-Tyson (Eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* (pp. 31–44). Amsterdam: Rodopi.

Johansson, S. (2007). *Seeing through multilingual corpora: on the use of corpora in contrastive studies* (Vol. 26). Amsterdam: John Benjamins.

Johansson, S., & Oksefjell, S. (1998). *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Amsterdam: Rodopi.

Johnson, K. (2008). *Quantitative Methods In Linguistics*. Malden MA: Wiley-Blackwell.

Kamp, H. (1979). Events, instants and temporal reference. In R. Bauerle, U. Egli, & A. von Stechow (Eds.), *Semantics from different points of view* (pp. 376–418). Amsterdam: Springer.

Kamp, H. (1981). A theory of truth and semantic representation. In P. Portner & B. Partee (Eds.), *Formal semantics-the essential readings* (pp. 189–222). Oxford: Blackwell.

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Amsterdam: Springer.

Kamp, H., & Rohrer, C. (1983). Tense in Texts. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language* (pp. 250–269). Berlin: Walter de Gruyter.

Katsos, N., & Breheny, R. (2008). Two experiments and some suggestions on the meaning of scalars and numerals. In E. Németh & K. Bibok (Eds.), *The Role of Data at the Semantics–Pragmatics Interface* (pp. 125–160). Berlin/New Work: Mouton de Gruyter.

Katsos, N., & Cummins, C. (2010). Pragmatics: from theory to experiment and back again. *Language and Linguistics Compass*, *4*(5), 282–295.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 170–210.

Kehler, A. (2004). Discourse Coherence. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 241–265). Oxford: Blackwell.

Kennedy, G. D. (1998). *An introduction to corpus linguistics*. London: Longman.

Kepser, S., & Reis, M. (2005). *Linguistic evidence: empirical, theoretical and computational perspectives*. Berlin: Walter de Gruyter.

Kiparsky, P. (1968). Linguistic universals and linguistic change. In E. Bach & R. Harms (Eds.), *Universals in Linguistic Theory* (pp. 170–202). New York: Holt, Rinehart, and

Winston.

Kleiber, G. (1994). *Anaphores et pronoms*. Bruxelles: Duculot.

Kleiber, G. (2003). Entre les deux mon coeur balance ou l'imparfait entre aspect et anaphore. *Langue Française*, 8–19.

Kleiber, G., & Riegel, M. (1989). Une sémantique qui n'a pas de sens n'a vraiment pas de sens. *Linguisticae Investigationes*, *13*(2), 405–417.

Kleiber, G., & Riegel, M. (1991). Sens lexical et interprétations référentielles: un écho à la réponse de D. Kayser. *Linguisticae Investigationes*, *15*(1), 181–201.

Klein, W. (1992). The present perfect puzzle. *Language*, 525–552.

Klein, W. (1994). *Time in language*. London and New York: Routledge.

Klum, A. (1961). *Verbe et adverbe*. Upsala: Almqvist and Wiksell.

Knott, A., & Sanders, T. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, *30*(2), 135–175.

Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. In *MT summit* (Vol. 5, pp. 79–86).

Koehn, P. (2010). *Statistical Machine Translation*. Cambridge: Cambridge University Press.

Koehn, P., & Hoang, H. (2007). Factored Translation Models. In *EMNLP-CoNLL* (pp. 868–876).

Kolaiti, P., & Wilson, D. (2014). Corpus analysis and lexical pragmatics: an overview. *International Review of Pragmatics*, *6*(2), 211–239.

Kozlowska, M. (1996). "Ensuite" et l'ordre temporel. *Cahiers de Linguistique Française*, *18*, 243–274.

Kozlowska, M. (1998). Bornage, télicité et ordre temporel. In J. Moeschler, J. Jayez, J.-M. Luscher, L. Saussure de, & B. Sthioul (Eds.), *Le temps des événements* (pp. 221–244). Paris: Kimé.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, *30*(3), 411–433.

Krzeszowski, T. P. (1990). *Contrasting languages: The scope of contrastive linguistics* (Vol. 51). Berlin: Walter de Gruyter.

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts*. Seattle WA: University of Washington Press.

Lakoff, G. (1965). *On the nature of syntactic irregularity* (PhD Dissertation). Indiana University, Bloomington.

Langacker, R. W. (1991). *Concept, Image, and Symbol. The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Langacker, R. W. (1999). *Grammar and conceptualization* (Vol. 14). Berlin: Mouton de Gruyter.

Langacker, R. W. (2002). Remarks on the English grounding systems. In F. Brisard (Ed.), *Grounding: The epistemic footing of deixis and reference* (pp. 29–38). Berlin: Mouton de Gruyter.

Langacker, R. W. (2006). Subjectification, grammaticization, and conceptual archetypes. In A. Athanasiadou, C. Canakis, & C. Bert (Eds.), *Subjectification: various paths to subjectivity* (Vol. 31, pp. 17–40). Berlin: Mouton de Gruyter.

Lapata, M., & Lascarides, A. (2004). Inferring Sentence-internal Temporal Relations. In *HLT-NAACL* (pp. 153–160).

Lascarides, A., & Asher, N. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, *16*(5), 437–493.

Laviosa, S. (2002). *Corpus-based translation studies: theory, findings, applications*. Amsterdam: Rodopi.

Leech, G. (2005). Adding linguistic annotation. In M. Wynne (Ed.), *Developing linguistic corpora : a guide to good practice* (pp. 17–29). Oxford: Oxbow Books.

Leech, G. N. (2004). *Meaning and the English verb* (3rd Edition). London: Routledge.

Leech, G., & Svartvik, J. (1975). *A Communicative Grammar of English* (1st edition). London: Longman.

Leech, G., & Svartvik, J. (2002). *A Communicative Grammar of English* (3rd edition). New York: Routledge.

Lee, J., & Tonhauser, J. (2010). Temporal interpretation without tense: Korean and Japanese coordination constructions. *Journal of Semantics*, ffq005.

Lefer, M.-A. (2009). *Exploring lexical morphology across languages: a corpus-based study of prefixation in English and French writing* (PhD Dissertation). Universtité Catolique de Louvain, Louvain.

Leonetti, M., & Escandell-Vidal, V. (2003). On the quotative readings of Spanish imperfecto. *Cuadernos de Lingüística Del IU Ortega Y Gasset*, *10*, 135–154.

Leonetti, M., & Escandell-Vidal, V. (2004). Semántica conceptual/semántica procedimental. In *Actas del V Congreso de Lingüística General: León 5-8 de marzo de 2002* (pp. 1727–1738). Arco Libros.

Lepschy, A. L., & Lepschy, G. (1998). *The Italian language today* (2nd Edition). London: Routledge.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge Textbooks in Linguistics.

Levinson, S. C. (1987). Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics*, *23*(02), 379–434.

Levinson, S. C. (1989). A review of Relevance. *Journal of Linguistics*, *25*(02), 455–472.

Levinson, S. C. (1995). Interactional biases in human thinking. In E. N. Goody (Ed.), *Social intelligence and interaction* (pp. 221–260). Cambridge: Cambridge University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge MA: MIT Press.

Lin, J.-W. (2012). Tenselessness. In R. I. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 669–695). Oxford: Oxford University Press.

Li, W., Wong, K.-F., Cao, G., & Yuan, C. (2004). Applying machine learning to Chinese temporal relation resolution. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics* (pp. 582–588).

Li, W., Wong, K.-F., & Yuan, C. (2001). A model for processing temporal references in Chinese. In *Proceedings of the workshop on Temporal and spatial information processing* (Vol. 13, p. article no. 5).

Loáiciga, S., & Grisot, C. (2015). Predicting the Lexical Aspect of Simple Past Verbs for English-to-French Machine Translation. *Linguistic Issues in Language Technology (LiLT)*, *submitted*.

Loáiciga, S., Meyer, T., & Popescu-Belis, A. (2014). English-French Verb Phrase Alignment in Europarl for Tense Translation Modeling. In *The Ninth Language Resources and Evaluation Conference* (pp. 674–681). Reykjavik, Island.

Ludlow, P. (2013). Tensism. In K. M. Jaszczolt & L. Saussure de (Eds.), *Time, Language, Cognition and Reality* (pp. 175–192). Oxford: Oxford University Press.

Luscher, J.-M. (1998). Procédure d'interprétation du Passé Composé. In J. Moeschler, J. Jayez, M. Kozlowska, J.-M. Luscher, L. Saussure de, & B. Sthioul (Eds.), *Le temps des événements* (pp. 181–196). Paris: Kimé.

Luscher, J.-M., & Sthioul, B. (1996). Emplois et interprétations du Passé Composé. *Cahiers de Linguistique Française*, *18*, 187–217.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Maat, H. P., & Degand, L. (2001). Scaling causal relations and connectives in terms of Speaker Involvement. *Cognitive Linguistics*, *12*(3), 211–246.

Maat, H. P., & Sanders, T. (2000). Domains of use or subjectivity? The distribution of three Dutch causal connectives explained. *Topics in English Linguistics*, *33*, 57–82.

Maat, H. P., & Sanders, T. (2001). Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics*, *12*(3), 247–274.

Manea, D. (2008). Modul, Timpul. In *Gramatica limbii romane* (pp. 358–448). Bucuresti: Editura Academiei Române.

Manning, C., & Klein, D. (2003). Optimization, maxent models, and conditional estimation without magic. In *Proceedings of the 2003 Conference of the North American Chapter Association for Computational Linguistics on Human Language Technology: Tutorials* (Vol. 5, pp. 8–8). Stroudsburg, USA: Association for Computational Linguistics.

Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. *Discourse Processes*, *9*(1), 57–90.

Marcu, D. (2000). *The theory and practice of discourse parsing and summarization*. Cambridge MA: MIT press.

Margan, M. (2009). Câteva considerații asupra relației timp-aspect în limbile română și engleza. *Filologia Banatică*, *1*, 49–58.

Martin, R. (1971). *Temps et aspect. Essai sur l'emploi des temps narratifs en moyen français*. Paris: Klincksieck.

McCawley, J. D. (1971). Tense and time reference in English. In C. J. Fillmore & T. D. Langendoen (Eds.), *Studies in Linguistic Semantics* (pp. 96–113). Dunfermline: Irvington.

McEnery, T., & Wilson, A. (1996). *Corpus linguistics*. Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies*. London: Routledge.

McTaggart, J. E. (1908). The unreality of time. *Mind*, 457–474.

Meibauer, J., & Steinbach, M. (2011). *Experimental Pragmatics/Semantics*. Amsterdam: John Benjamins.

Meyer, T. (2014). *Discourse-level features for statistical machine translation* (PhD Dissertation). EPFL and Idiap Research Institute, Lausanne and Martigy.

Meyer, T., Grisot, C., & Popescu-Belis, A. (2013). Detecting narrativity to improve English to French translation of simple past verbs. In B. Webber, A. Popescu-Belis, K. Markert, & J. Tiedemann (Eds.), *Proceedings of the 1st DiscoMT Workshop at ACL 2013 (51st Annual Meeting of the Association for Computational Linguistics)* (pp. 33–42). Sofia, Bulgaria.

Milner, J. C. (1982). *Ordres et raisons de langue*. Paris: Seuil.

Moens, M. (1987). *Tense, aspect and temporal reference* (PhD Dissertation). University of Edinburgh, Edinburgh.

Moens, M., & Steedman, M. (1987). Temporal ontology in natural language. In *Proceedings of the 25th annual meeting on Association for Computational Linguistics* (pp. 1–7). California, USA: Association for Computational Linguistics.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, *14*(2), 15–28.

Moeschler, J. (1985). *Argumentation et conversation*. Paris: Hatier.

Moeschler, J. (1989a). *Modélisation du dialogue: représentation de l'inférence argumentative*. Paris: Hermès.

Moeschler, J. (1989b). Pragmatic connectives, argumentative coherence and relevance. *Argumentation*, *3*(3), 321–339.

Moeschler, J. (1993). Relevance and conversation. *Lingua*, *90*(1), 149–171.

Moeschler, J. (1994). *Anaphore et déixis temporelles: sémantique et pragmatique de la référence temporelle*. Nancy: Presses Universitaires de Nancy.

Moeschler, J. (1996). *Théorie pragmatique et pragmatique conversationnelle*. Paris: Armand Colin.

Moeschler, J. (1998). Pragmatique de la référence temporelle. In J. Moeschler, J. Jayez, M. Kozlowska, J.-M. Luscher, L. Saussure de, & B. Sthioul (Eds.), *Le temps des événements* (pp. 157–180). Paris: Kimé.

Moeschler, J. (2000a). Le modèle des inférences directionnelles. *Cahiers de Linguistique Française*, *22*, 57–100.

Moeschler, J. (2000b). L'ordre temporel est-il naturel? In J. Moeschler & M.-J. Béguelin (Eds.), *Référence temporelle et nominale* (pp. 71–105). Berne: Peter Lang.

Moeschler, J. (2001). Speech act theory and the analysis of conversation. In D. Vanderveken & S. Kubo (Eds.), *Essays in speech act theory* (pp. 239–262). Amsterdam: John Benjamins.

Moeschler, J. (2002a). Connecteurs, encodage conceptuel et encodage procédural. *Cahiers de Linguistique Française*, *24*, 265–292.

Moeschler, J. (2002b). Pragmatics and linguistic encoding. Evidence from the conceptual/procedural distinction. *Draft*. Retrieved from http://www.researchgate.net/publication/228740741

Moeschler, J. (2003). Economy and pragmatic optimality: the case of directional inferences. *Generative Grammar Geneva*, *3*, 1–20.

Moeschler, J. (2007a). Discours causal, chaîne causale et argumentation. *Cahiers Chronos*, *18*, 69–86.

Moeschler, J. (2007b). The role of explicature in communication and in intercultural communication. *Explorations in Pragmatics*, 73–94.

Moeschler, J. (2007c). Why are there no negative particulars? Horn's conjecture revisited. *Generative Grammar Geneva*, *5*, 1–13.

Moeschler, J. (2011). Causal, inferential and temporal connectives: Why parce que is the only causal connective in French. In S. Hancil (Ed.), *Marqueurs discursifs et subjectivité* (pp. 97–114). Mont-Saint-Aignan: Publications des universités de Rouen et du Havre.

Moeschler, J. (2012a). Conversational and conventional implicatures. In H.-J. Schmid (Ed.), *Cognitive Pragmaitics* (pp. 405–436). Berlin, Boston: Mouton de Gruyter.

Moeschler, J. (2012b). Pourquoi le sens est-il structuré? Une approche vériconditionnelle de la signification linguistique et du sens pragmatique. *Nouveaux Cahiers de Linguistique Française*, *30*, 53–71.

Moeschler, J. (2013). Is a speaker-based pragmatics possible? Or how can a hearer infer a speaker's commitment? *Journal of Pragmatics*, *48*(1), 84–97.

Moeschler, J. (2014). Subjectivité et langage: l'exemple du présent historique. In B. Pavelin Lesic (Ed.), *Francontraste: L'affectivité et la subjectivité dans le langage* (pp. 29–40). Mons: Ed. du CIPA.

Moeschler, J. (2015a). Argumentation and connectives. In A. Capone & J. L. Mey (Eds.), *Interdisciplinary Studies in Pragmatics, Culture and Society, Perspectives in Pragmatics* (pp. 405–436). Switzerland: Springer International Publishing.

Moeschler, J. (2015b). La frontière sémantique-pragmatique existe-t-elle? La question des préssupositions et des implicatures révisée. In A. Rabatel, A. Ferrara-Léturgie, & A. Léturgie (Eds.), *La sémantique et ses interfaces* (pp. 263–288). Limoges: Lambert-Lucas.

Moeschler, J. (2015c). Présupposition et implicature: où passe la frontière. In A. Biglari & M. Bonhomme (Eds.), *La présupposition entre théorisation et mise en discours*. Paris: Editions L'Harmattan.

Moeschler, J. (2015d). *With and without procedural information*. Presented at the 48th Annual Meeting of the Societas Linguistica Europaea, University of Leiden, Netherlands.

Moeschler, J., Grisot, C., & Cartoni, B. (2012). Jusqu'où les temps verbaux sont-ils procéduraux? *Nouveaux Cahiers de Linguistique Française*, *30*, 119–139.

Moeschler, J., Jayez, J., Kozlowska, M., Luscher, J.-M., Saussure de, L., & Sthioul, B. (1998). *Le temps des événements: pragmatique de la référence temporelle*. Paris: Kimé.

Moeschler, J., & Reboul, A. (1994). *Dictionnaire encyclopédique de pragmatique*. Paris: Seuil.

Molendijk, A. (2002). La structuration logico-temporelle du texte: le passé simple et l'imparfait du français. In E. Labeau & P. Larivée (Eds.), *Les temps du passé français et leur enseignement* (pp. 91–104). Amsterdam: Rodopi.

Molendijk, A. L. (1990). *Le passé simple et l'imparfait: une approche reichenbachienne*. Amsterdam: Rodopi.

Montague, R. (1973). The proper treatment of quantification in ordinary English. In K. J. J. Hintikka, J. M. E. Moravcsik, & P. Suppes (Eds.), *Approaches to Natural Language* (pp. 221–242). Dordrecht: Reidel.

Montague, R. (1974). *Formal philosophy; selected papers of Richard Montague*. (R. H. Thomason, Ed.). New Haven: Yale University Press.

Mourelatos, A. P. (1978). Events, processes, and states. *Linguistics and Philosophy*, *2*(3), 415–434.

Mourelatos, A. P. (1981). Events, processes and states. In P. J. Tedeschi & A. Zaenen (Eds.), *Tense and Aspect* (Vol. 14, pp. 191–212). New York: Academic Press.

Mucha, A. (2013). Temporal interpretation in Hausa. *Linguistics and Philosophy*, *36*(5), 371–415.

Nanousi, V., Masterson, J., Druks, J., & Atkinson, M. (2006). Interpretable vs. uninterpretable features: Evidence from six Greek-speaking agrammatic patients. *Journal of Neurolinguistics*, *19*(3), 209–238.

Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., & Neville, H. J. (2007). An ERP study of regular and irregular English past tense inflection. *Neuroimage*, *34*(1), 435–445.

Nicolle, S. (1997). Conceptual and procedural encoding: Criteria for the identification of

linguistically encoded procedural information. In M. Groefsema (Ed.), *Proceedings of the Hertford-shire relevance theory workshop* (pp. 45–56). Chelmsford: Peter Thomas and Associates.

Nicolle, S. (1998). A relevance theory perspective on grammaticalization. *Cognitive Linguistics*, *9*(1), 1–35.

Noël, D. (2003). Translations as evidence for semantics: an illustration. *Linguistics*, *41*(4), 757–785.

Novakov, P., & Lazović, M. (2009). Aspect and boundedness in English and Romanian. *Bucharest Working Papers in Linguistics*, (2), 41–50.

Noveck, I. A., & Sperber, D. (2004). *Experimental Pragmatics*. London: Palgrave Macmillan.

Noveck, I. A., & Sperber, D. (2007). Th why and how of experimental pragmatics. The case of scalar inferences. In N. Roberts (Ed.), *Advances in Pragmatics*. London: Palgrave Macmillan.

Nuyts, J. (2001). *Epistemic modality, language, and conceptualization: A cognitive-pragmatic perspective*. Amsterdam: John Benjamins Publishing.

Nuyts, J. (2008). Pattern versus process concepts of grammar and mind: A cognitive-functional perspective. *Jezikoslovlje*, (9.1-2), 87–107.

Olsen, M., Traum, D., Van Ess-Dykema, C., & Weinberg, A. (2001). *Implicit cues for explicit generation: using telicity as a cue for tense structure in a Chinese to English MT system* (Technical Report). College Park, USA: University of Maryland.

Olsen, M., Traum, D., Van Ess-Dykema, C., Weinberg, A., & Dolan, R. (2000). Telicity as a cue to temporal and discourse structure in Chinese-English machine translation. In *Proceedings of the 2000 NAACL-ANLP Workshop on Applied interlinguas: practical applications of interlingual approaches to NLP* (Vol. 2, pp. 34–41). Stroudsburg, USA: Association for Computational Linguistics.

Padilla-Cruz, M. (2009). Might interjections encode concepts? More questions than answers. *Lodz Papers in Pragmatics*, *5*(2), 241–270.

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics* (pp. 311–318). Philadelphia,USA: Association for Computational Linguistics.

Parsons, T. (1989). The progressive in English: Events, states and processes. *Linguistics and Philosophy*, *12*(2), 213–241.

Parsons, T. (1990). *Events in the Semantics of English*. Cambridge MA: MIT Press.

Partee, B. (1984). Compositionality. *Varieties of Formal Semantics*, *3*, 281–311.

Partee, B. H. (1973). Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy*, 601–609.

Partee, B. H. (1989). Binding implicit variables in quantified contexts. In *Proceedings of the Chicago Linguistics Society 25* (pp. 342–365). Chicago: University of Chicago Press.

Partee, B., Ter Meulen, A., & Wall, R. (1990). *Mathematical methods in linguistics*. Dordrecht: Kluwer Academic Publishers.

Passonneau, R. J. (1988). A computational model of the semantics of tense and aspect. *Computational Linguistics*, *14*(2), 44–60.

Pit, M. (2003). *How to express yourself with a causal connective: subjectivity and causal connectives in*

Dutch, German and French. Amsterdam: Rodopi.

Pit, M. (2006). Determining subjectivity in text: The case of backward causal connectives in Dutch. *Discourse Processes*, *41*(2), 151–174.

Pit, M. (2007). Cross-linguistic analyses of backward causal connectives in Dutch, German and French. *Languages in Contrast*, *7*(1), 53–82.

Pollock, J.-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry*, 365–424.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. K., & Webber, B. L. (2008). The Penn Discourse TreeBank 2.0. In *Proceedings of the 6th International conference on Language Resources and Evaluation*. Marrakech. Retrieved from http://www.lrec-conf.org/proceedings/lrec2008/

Prasad, R., Miltsakaki, E., Joshi, A., & Webber, B. (2004). Annotation and data mining of the Penn Discourse TreeBank. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation* (pp. 88–97). Barcelona, Spain.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar*. New York: John Wiley & Sons.

Prior, A. N. (1957). *Time and modality*. Oxford: Oxford University Press.

Prior, A. N. (1967). *Past, present and future*. Oxford: Clarendon Press.

Prior, A. N. (1968). *Papers on time and tense*. (P. Hasle, P. Ohrstrom, T. Braüner, & J. Copeland, Eds.) (2nd ed.). Oxford: Oxford University Press.

Prosser, S. (2013). Experience, thought, and the metaphysics of time. In K. M. Jaszczolt & L. Saussure de (Eds.), *Time, Language, Cognition and Reality* (pp. 157–174). Oxford: Oxford University Press.

Puskas, G. (2013). *Initiation au Programme Minimaliste*. Berne: Peter Lang.

Pustejovsky, J., Ingria, B., Sauri, R., Castano, J., Littman, J., Gaizauskas, R., … Mani, I. (2005). The specification language TimeML. In I. Mani, J. Pustejovsky, & R. Gaizauskas (Eds.), *The language of time: A reader* (pp. 545–557). Oxford: Oxford University Press.

Pustejovsky, J., Knippen, R., Littman, J., & Saurí, R. (2005). Temporal and event information in natural language text. *Language Resources and Evaluation*, *39*(2-3), 123–164.

Qiu, Y., & Zhou, X. (2012). Processing temporal agreement in a tenseless language: An ERP study of Mandarin Chinese. *Brain Research*, *1446*, 91–108.

Quirk, R., Greenbaum, S. L., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Radden, G., & Dirven, R. (2007). *Cognitive English Grammar*. Amsterdam: John Benjamins Publishing.

Rebotier, A. (2009). Le Futur de l'allemand en comparaison avec les langues romanes: esquisse d'une définition d'une catégorie translinguistique de Futur. *Faits de Langues*, (33), 69–78.

Reboul, A. (1992). *Rhétorique et stylistique de la fiction*. Nancy: Presses Universitaires de Nancy.

Reboul, A. (1994). L'anaphore pronominale: le problème de l'attribution des référents. In J. Moeschler, J.-M. Luscher, & J. Jayez (Eds.), *Langage et pertinence. Référence temporelle, anaphore, connecteurs et métaphore* (pp. 105–173). Nancy: PUN.

Reboul, A. (1996). Le paradoxe de l'imperfectif: evenements, causalite et etats de faits. In R. Landheer & P. J. Smith (Eds.), *Le paradoxe en linguistique et en littérature* (pp. 39–58). Genève:

Droz.

Reboul, A. (2007). *Langage et cognition humaine*. Grenoble: Presses Universitaires de Grenoble.

Reboul, A. (2012). Language: Between cognition, communication and culture. *Pragmatics & Cognition*, *20*(2), 295–316.

Reboul, A., Delfitto, D., & Fiorin, G. (2015). The Syntactic and Semantic Properties of Free Indirect Discourse. *Annual Review of Linguistics, submitted*.

Reboul, A., & Moeschler, J. (1998). *Pragmatique du discours. De l'interprétation de l'énoncé à l'interprétation du discours*. Paris: Armand Colin.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Recanati, F. (2010). *Truth-conditional pragmatics*. Oxford: Oxford University Press Oxford.

Reichenbach, H. (1947). *Elements of symbolic logic*. New York: Mcmillan.

Riddle, E. (1986). The meaning and discourse function of the past tense in English. *TESOL Quarterly*, *20*(2), 267–286.

Riegel, M., Pellat, J.-C., & Rioul, R. (1994). *Grammaire méthodique du français*. Paris: PUF.

Rothstein, S. (2004). *Structuring Events: a study in the semantics of lexical aspect*. Oxford: Blackwell Publishing.

Roulet, E., Auchlin, A., & Moeschler, J. (1985). *L'articulation du discours en français contemporain*. Berne: Peter Lang.

Roze, C., Danlos, L., & Muller, P. (2010). LEXCONN: a French lexicon of discourse connectives. *Proceedings of Multidisciplinary Approaches to Discourse (MAD 2010), Moissac, France*.

Sadock, J. M. (1978). On testing for conversational implicature. In P. Cole (Ed.), *Syntax and Semantics: Pragmatics.* (pp. 281–297). New York: Academic Press.

Salkie, R. (2002). How can linguists profit from parallel corpora? *Language and Computers*, *43*(1), 93–109.

Salvi, G., & Vanelli, L. (2004). *Nuova grammatica italiana*. Bologna: Il Mulino.

Samardzic, T. (2013). *Dynamics, causation, duration in the predicate-argument structure of verbs: a computational approach based on parallel corpora* (PhD Dissertation). Université de Genève, Genèva.

Sanders, J., & Redeker, G. (1996). Perspective and the representation of speech and thought in narrative discourse. In G. Fauconnier & E. Sweetser (Eds.), *Spaces, worlds and grammar* (pp. 290–317). Chicago: University of Chicago Press.

Sanders, J., Sanders, T., & Sweetser, E. (2012). Responsible subjects and discourse causality. How mental spaces and perspective help identifying subjectivity in Dutch backward causal connectives. *Journal of Pragmatics*, *44*(2), 191–213.

Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, *24*(1), 119–147.

Sanders, T. (2005). Coherence, causality and cognitive complexity in discourse. In *Proceedings/Actes SEM-05, First International Symposium on the exploration and modelling of meaning* (pp. 105–114).

Sanders, T. J., & Noordman, L. G. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*(1), 37–60.

Sanders, T. J., Spooren, W. P., & Noordman, L. G. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*(1), 1–35.

Sanders, T. J., Spooren, W. P., & Noordman, L. G. (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, *4*(2), 93–133.

Sanders, T., & Spooren, W. (2009). Causal categories in discourse–Converging evidence from language use. In T. Sanders & E. Sweetser (Eds.), *Causal categories in discourse and cognition* (pp. 205–246). Berlin: Mouton de Gruyter.

Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language: Explorations of comprehension beyond the sentence*. New York: Wiley New York.

Saurí, R., Littman, J., Gaizauskas, R., Setzer, A., & Pustejovsky, J. (2006). *TimeML annotation guidelines, version 1.2. 1*. Retrieved from http://www.timeml.org/site/publications/timeMLdocs/timeml_1.2.1.html

Saussure de, F. (1967). *Cours de Linguistique Générale*. Paris: Payot.

Saussure de, L. (1997). Passé simple et encapsulation d'événements. *Cahiers de Linguistique Française*, *19*, 323–344.

Saussure de, L. (1998). L'approche référentielle: de Beauzée à Reichenbach. In J. Moeschler, J. Jayez, M. Kozlowska, J.-M. Luscher, L. Saussure de, & B. Sthioul (Eds.), *Le temps des événements* (pp. 19–44). Paris: Kimé.

Saussure de, L. (2000a). *Pragmatique temporelle des énoncés negatives* (PhD Dissertation).Université de Genève, Genève.

Saussure de, L. (2000b). Quand le temps ne progresse pas avec le passé simple. In A. Carlier, V. Lagae, & C. Benninger (Eds.), *Passé et parfait* (pp. 37–48). Amsterdam: Rodopi.

Saussure de, L. (2003). *Temps et pertinence: éléments de pragmatique cognitive du temps*. Bruxelles: De Boeck Duculot.

Saussure de, L. (2007). Procedural pragmatics and the study of discourse. *Pragmatics & Cognition*, *15*(1), 139–159.

Saussure de, L. (2011). On some methodological issues in the conceptual/procedural distinction. In V. Escandell-Vidal, M. Leonetti, & A. Ahern (Eds.), *Procedural meaning: Problems and perspectives* (pp. 55–79). Bingley: Emerald Group Publishing.

Saussure de, L. (2013). Perspectival interpretations of tenses. In K. M. Jaszczolt & L. Saussure de (Eds.), *Time, Language, Cognition and Reality* (pp. 46–69). Oxford: Oxford University Press.

Saussure de, L., & Morency, P. (2012). A cognitive-pragmatic view of the French epistemic future. *Journal of French Language Studies*, *22*, 207–223.

Saussure de, L., & Sthioul, B. (1998). L'approche psychologique: Damourette et Pichon. In J. Moeschler, J. Jayez, M. Kozlowska, L. Saussure de, B. Sthioul, & J.-M. Luscher (Eds.), *Le temps des événements* (pp. 67–86). Paris: Kimé.

Saussure de, L., & Sthioul, B. (1999). L'imparfait narratif: point de vue (et images du monde). *Cahiers de Praxématique*, *32*, 167–188.

Saussure de, L., & Sthioul, B. (2005). Imparfait et enrichissement pragmatique. In P. Larivée & E. Labeau (Eds.), *Nouveaux développements de l'imparfait* (pp. 103–120). Amsterdam: Rodopi.

Savic, M. D. (1979). L'uso dei tempi passati nei quotidiani pubblicati nelle lingue romanze con particolare riguardo all'italiano. *Linguistica Ljubljana*, *19*, 171–197.

Schlenker, P. (2004). Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind & Language*, *19*(3), 279–304.

Scott, W. A. (1955). Reliability of Content Analysis: The Case of Nominal Scale Coding. *Public Opinion Quarterly*, *19*(3), 321.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Searle, J. R., & Vanderveken, D. (1985). *Foundations of illocutionary logic*. Cambridge: Cambridge University Press.

Setzer, A. (2001). *Temporal information in newswire articles: an annotation scheme and corpus study* (PhD Dissertation). University of Sheffield, Sheffield.

Shi, M. (2011). *A corpus based contrastive approach for the analysis of tense and aspect in translation from English into Mandarin Chinese* (PhD Dissertation). Münster, Westfälische Wilhelms-Universität Münster.

Simonsen, H. G., & Lind, M. (2002). Past tense expression in a Norwegian man with Broca's aphasia. In F. Windsor, M. L. Kelly, & N. Hewitt (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 45–56). Mahwah: Erlbaum.

Smith, C. (1986). A speaker-based approach to aspect. *Linguistics and Philosophy*, *9*(1), 97–115.

Smith, C. S. (1994). Aspectual viewpoint and situation type in Mandarian Chinese. *Journal of East Asian Linguistics*, *3*(2), 107–146.

Smith, C. S. (1997). *The Parameter of Aspect* (2nd edition). Dordrecht: Kluwer Academic Publishers.

Smith, C. S. (2003). *Modes of discourse: The local structure of texts*. Cambridge: Cambridge University Press.

Smith, C. S. (2007). Reference time without tense. In L. Saussure de, J. Moeschler, & G. Puskas (Eds.), *Recent advances in the syntax and semantics of tense, mood and aspect.* (pp. 229–250). Berlin: Walter de Gruyter.

Smith, C. S. (2008). Time with and without tense. In J. Guéron & J. Lecarme (Eds.), *Time and modality* (pp. 227–249). Berlin: Springer.

Smith, C. S., & Erbaugh, M. S. (2005). Temporal interpretation in Mandarin Chinese. *Linguistics*, *43*(4), 713–756.

Smith, N. (1990). Observations on the pragmatics of tense. *UCL Working Paoers in Linguistics*, *2*, 113–46.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. In *Proceedings of association for machine translation in the Americas* (pp. 223–231).

Sperber, D. (2005). Modularity and relevance. In P. Carruthers (Ed.), *The Innate Mind : Structure and Contents: Structure and Contents* (pp. 53–68). New York: Oxford University Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (1998). The mapping between the mental and the public lexicon. In P. Carruthers & J. Boucher (Eds.), *Language and thought: Interdisciplinary themes* (pp. 184–200). Cambridge: Cambridge University Press.

Spooren, W., & Degand, L. (2010). Coding coherence relations: reliability and validity. *Corpus Linguistics and Linguistic Theory*, *6*(2), 241–266.

Sporiș, V. (2012). Valențele semantico-stilistice ale timpurilor verbale în limba română. *Studia Universitatis Petru Maior-Philologia*, (13), 64–73.

Squartini, M., & Bertinetto, P. M. (2000). The simple and compound past in Romance

languages. In Ö. Dahl (Ed.), *Tense and Aspect in the languages of Europe* (pp. 403–440). Berlin: Mouton de Gruyter.

Stavrakaki, S., & Kouvava, S. (2003). Functional categories in agrammatism: Evidence from Greek. *Brain and Language*, *86*(1), 129–141.

Steedman, M. (1997). Temporality. In J. Benthem van & A. Meulen ter (Eds.), *Handbook of logic and language* (pp. 895–938). Amsterdam: Elsevier Science.

Steinhauer, K., & Ullman, M. T. (2002). Consecutive ERP effects of morpho-phonology and morpho-syntax. *Brain and Language*, 62–65.

Sthioul, B. (1998). La conceptualisation du temps: Guillaume. In J. Moeschler, J. Jayez, J.-M. Luscher, L. Saussure de, & B. Sthioul (Eds.), *Le temps des événements* (pp. 45–66). Paris: Kimé.

Sthioul, B. (2000). Passé simple, imparfait et sujet de conscience. In A. Carlier, V. Lagae, & C. Benninger (Eds.), *Passé et parfait* (Vol. 6, pp. 79–93). Amsterdam: Rodopi.

Stoicescu, I. (2010). Aspectual classes in adult and child Romanian. *Bucharest Working Papers in Linguistics*, (1), 171–193.

Stoicescu, I. (2012). Temporal deixis in early child Romanian. *Bucharest Working Papers in Linguistics*, (1), 133–152.

Stoicescu, I. (2013). *Coercion Costs: Situation Type Aspect and Verbal Inflection in Child Romanian*. Presented at the GALA Conference, Oldenburg, Germany.

Stoll, S. (1998). The role of Aktionsart in the acquisition of Russian aspect. *First Language*, *18*(54), 351–376.

Stowell, T. (1995). The phrase structure of tense. In L. Zaring & J. Rooryck (Eds.), *Phrase structure and lexicon* (pp. 277–291). Dordrecht: Kluwer Academic Publishers.

Stowell, T. (2007). The syntactic expression of tense. *Lingua*, *117*(2), 437–463.

Stowell, T. (2012). Syntax. In R. I. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 184–211). Oxford: Oxford University Press.

Stowell, T. A. (1981). *Origins of phrase structure* (PhD Dissertation). Massachusetts Institute of Technology, Cambridge MA.

Stukker, N., & Sanders, T. (2012). Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics*, *44*(2), 169–190.

Taboada, M. (2006). Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics*, *38*(4), 567–592.

Taboada, M., & Mann, W. C. (2006). Applications of rhetorical structure theory. *Discourse Studies*, *8*(4), 567–588.

Tahara, I. (2000). Le passé simple et la subjectivité. *Cahiers de Linguistuque Francaise 22*, 189–218.

Tahara, I. (2004). *Usage descriptif et usage interprétatif des temps du passé et des adverbes temporels dans le discours de fiction* (PhD Dissertation). Université de Genève, Genève.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*(1), 49–100.

Tasmowski-De Ryck, L. (1985). L'imparfait avec et sans rupture. *Langue Française*, 59–77.

Tatevosov, S. (2002). The parameter of actionality. *Linguistic Typology*, *6*(3), 317–401.

Taylor, B. (1977). Tense and continuity. *Linguistics and Philosophy*, *1*(2), 199–220.

Tedeschi, P. J., & Zaenen, A. (Eds.). (1981). *Tense and Aspect* (Vol. 14). New York: Academic Press.

Ter Meulen, A. G. (1997). *Representing time in natural language: The dynamic interpretation of tense and aspect* (2nd ed.). Cambridge MA: MIT Press.

Tiedemann, J. (2009). News from OPUS-A collection of multilingual parallel corpora with tools and interfaces. In N. Nicolov, G. Angelova, & R. Mitkov (Eds.), *Recent Advances in Natural Language Processing* (pp. 237–248). Amsterdam: John Benjamins.

Tonhauser, J. (2015). Cross-Linguistic Temporal Reference. *Annual Review of Linguistics*, *1*(1).

Traugott, E. C. (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language*, 31–55.

Traugott, E. C. (1995). Subjectification in grammaticalization. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation* (pp. 31–54). Cambridge: Cambridge University Press.

Traugott, E. C. (1999). The rhetoric of counter-expectation in semantic change: a study in subjectification Elizabeth Closs Traugott. In A. Blank & P. Kock (Eds.), *Historical semantics and cognition* (pp. 177–196). Berlin/New Work: Mouton de Gruyter.

Trnavac, R. (2006). *Aspect and subjectivity in modal constructions* (PhD Dissertation). LOT, Utrecht.

Van Eijck, J., & Kamp, H. (1997). Representing discourse in context. In J. Benthem van & A. Meulen ter (Eds.), *Handbook of logic and language* (pp. 179–237). Cambridge: MIT Press.

Varlokosta, S., Valeonti, N., Kakavoulia, M., Lazaridou, M., Economou, A., & Protopapas, A. (2006). The breakdown of functional categories in Greek aphasia: Evidence from agreement, tense, and aspect. *Aphasiology*, *20*(8), 723–743.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 143–160.

Vendler, Z. (1967). *Linguistics in philosophy*. Cornell: Cornell University Press.

Verhagen, A. (1995). Subjectification, syntax, and communication. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: linguistic perspectives* (pp. 103–128). Cambridge: Cambridge University Press.

Verkuyl, H. J. (1972). *On the compositional nature of the aspects*. Dordrecht: Reidel.

Verkuyl, H. J. (1996). *A theory of aspectuality: The interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press.

Véronis, J., & Langlais, P. (2000). Evaluation of parallel text alignment systems. In J. Véronis (Ed.), *Parallel text processing* (Vol. 27, pp. 369–388). Dordrecht: Kluwer Academic Publishers.

Vet, C. (1980). *Temps, aspects et adverbes de temps en français contemporain: essai de sémantique formelle*. Genève: Droz.

Vetters, C. (1992). *L'opposition passé simple-imparfait: une queston d'aspect ou de structuration textuelle?* (PhD Dissertation). Universitaire Instelling Antwerpen, Antwerpen.

Vetters, C. (1996). *Temps, aspect et narration*. Amsterdam: Rodopi.

Vişan, N. (2006). Between the Perfect and the Preterite–A DRT Analysis of the Romanian and the English Complex Past. *Revue Roumaine de Linguistique*, (1), 55–71.

Vuillaume, M. (1990). *Grammaire temporelle des récits*. Paris: Éditions de Minuit.

Wagner, R. L., & Pinchon, J. (1962). *Grammaire du français classique et moderne*. Paris: Hachette.

Webber, B. L. (1988). Tense as discourse anaphor. *Computational Linguistics*, *14*(2), 61–73.

Weinrich, H. (1973). *Le temps*. Paris: Seuil.

Wenzlaff, M., & Clahsen, H. (2004). Tense and agreement in German agrammatism. *Brain and Language*, *89*(1), 57–68.

Wenzlaff, M., & Clahsen, H. (2005). Finiteness and verb-second in German agrammatism. *Brain and Language*, *92*(1), 33–44.

Wiebe, J. M. (1990). Identifying subjective characters in narrative. In *Proceedings of the 13th conference on Computational linguistics* (Vol. 2, pp. 401–406).

Wiebe, J. M. (1994). Tracking point of view in narrative. *Computational Linguistics*, *20*(2), 233–287.

Wiebe, J., & Mihalcea, R. (2006). Word sense and subjectivity. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics* (pp. 1065–1072). Association for Computational Linguistics.

Wilson, A. (2005). Corpus linguistics. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics* (pp. 448–463). Oxford: Oxford University Press.

Wilson, D. (2015). *Relations between temporal and inferential uses of connectives*. Personal communication.

Wilson, D. (2011). The conceptual procedural-distinction: Past, Present and Future. In V. Escandell-Vidal, M. Leonetti, & A. Ahern (Eds.), *Procedural Meaning: Problems and Perspectives* (pp. 3–32). Bingley: Emerald Group Publishing.

Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 230–259). Basingstoke: Palgrave Macmillan.

Wilson, D., & Sperber, D. (1988). Mood and the analysis of non-declarative sentences. In J. Dancy, J. Moravcsik, & C. Taylor (Eds.), *Human agency: Language, duty and value* (pp. 77–101). Stanford CA: Stanford University Press.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua*, *90*(1), 1–25.

Wilson, D., & Sperber, D. (1998). Pragmatics and time. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and implications* (pp. 1–22). Amsterdam: John Benjamins.

Wilson, D., & Sperber, D. (2004). Relevance theory. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 607–632). Oxford: Blackwell Publishing.

Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge: Cambridge University Press.

Yap, F. H., Chu, P. C. K., Yiu, E. S. M., Wong, S. F., Kwan, S. W. M., Matthews, S., … Shirai, Y. (2009). Aspectual asymmetries in the mental representation of events: Role of lexical and grammatical aspect. *Memory & Cognition*, *37*(5), 587–595.

Ye, Y., Fossum, V. L., & Abney, S. (2006). Latent features in automatic tense translation between Chinese and English. In *Proceedings of the Fifth SIGHAN Workshop on Chinese Language Processing* (pp. 48–55).

Ye, Y., Schneider, K.-M., & Abney, S. (2007). Aspect marker generation in English-to-Chinese machine translation. *Proceedings of MT Summit XI*, 521–527.

Zafiu, R. (2013). Mood, tense and aspect. In G. P. Dindelegan (Ed.), *The grammar of Romanian* (pp. 24–64). Oxford: Oxford University Press.

Zagona, K. (1990). *Times as temporal argument structure*. University of Washington, Seattle.

Zagona, K. (1995). Temporal argument structure: Configurational elements of construal. In P. M. Bertineto, V. Bianchi, & M. Squartini (Eds.), *Temporal reference, aspect and actionality* (Vol. 1, pp. 397–410). Torino: Rosenberg & Sellier.

Zagona, K. (2003). Tense and anaphora: Is there a tense-specific theory of coreference. In A.

Barrs (Ed.), *Anaphora: A reference guide* (pp. 140–171). Oxford: Blackwell Publishing.

Zagona, K. (2013). Tense, aspect and modality. In M. Dikken den (Ed.), *The Cambridge handbook of generative syntax* (pp. 746–792). Cambridge: Cambridge University Press.

Zegarac, V. (1991). *Tense, aspect and relevance*. (PhD Dissertation). University of London.

Zufferey, S. (2010). *Lexical pragmatics and theory of mind: the acquisition of connectives*. Amsterdam: John Benjamins.

Zufferey, S. (2012). Car, parce que, puisque revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics*, *44*(2), 138–153.

Zufferey, S. (2014). Givenness, procedural meaning and connectives. The case of French puisque. *Journal of Pragmatics*, *62*, 121–135.

Zufferey, S., & Cartoni, B. (2012). English and French causal connectives in contrast. *Languages in Contrast*, *12*(2), 232–250.