



Master

2021

Open Access

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

La post-édition monolingue des textes de spécialité dans le domaine de la
musique

Salamanca Zarate, Diego Steven

How to cite

SALAMANCA ZARATE, Diego Steven. La post-édition monolingue des textes de spécialité dans le domaine de la musique. Master, 2021.

This publication URL: <https://archive-ouverte.unige.ch/unige:156672>



**UNIVERSITÉ
DE GENÈVE**

**FACULTÉ DE TRADUCTION
ET D'INTERPRÉTATION**

DIEGO SALAMANCA ZARATE

**LA POST-ÉDITION MONOLINGUE
DES TEXTES DE SPÉCIALITÉ
DANS LE DOMAINE DE LA MUSIQUE**

Directrice : Pierrette Bouillon

Jurée : Johanna Gerlach

Mémoire présenté à la Faculté de Traduction et d'Interprétation
pour l'obtention de la Maîtrise universitaire en Traduction et Technologies

Université de Genève
Août 2021

Déclaration attestant le caractère original du travail effectué

J'affirme avoir pris connaissance des documents d'information et de prévention du plagiat émis par l'Université de Genève et la Faculté de traduction et d'interprétation (notamment la Directive en matière de plagiat des étudiants, le Règlement d'études des Maîtrises universitaires en traduction et du Certificat complémentaire en traduction de la Faculté de traduction et d'interprétation ainsi que l'Aide-mémoire à l'intention des étudiants préparant un mémoire de Ma en traduction).

J'atteste que ce travail est le fruit d'un travail personnel et a été rédigé de manière autonome.

Je déclare que toutes les sources d'information utilisées sont citées de manière complète et précise, y compris les sources sur Internet.

Je suis conscient que le fait de ne pas citer une source ou de ne pas la citer correctement est constitutif de plagiat et que le plagiat est considéré comme une faute grave au sein de l'Université, passible de sanctions.

Au vu de ce qui précède, je déclare sur l'honneur que le présent travail est original.

Salamanca Zárate Diego
Genève, le 24 août 2021

A handwritten signature in black ink that reads "Diego Salamanca Zárate". The signature is written in a cursive style and is enclosed within a light gray rectangular border.

Remerciements

La rédaction de ce mémoire, et la formation dont il représente la dernière étape, ont été encadrées dans un contexte marqué par l'incertitude, l'isolement et les basculements dans les modalités d'enseignement. Je voudrais donc remercier les personnes qui m'ont permis d'avancer dans cette tâche et qui, d'une façon ou d'une autre, ont rendu possible ce projet :

Mme. Pierrette Bouillon, directrice de ce mémoire, pour son accompagnement solide et sa patience, ainsi que pour tout le soutien qu'elle m'a apporté tout au long du processus de conception et d'exécution du projet. Son professionnalisme et son soutien permanent ont été des facteurs essentiels pour mener à bien ce mémoire.

Lucile Ortholan, pour tout, et très particulièrement pour le temps et les efforts investis dans les révisions linguistiques et stylistiques en langue française. Ce travail et cette page sont lisibles grâce à elle, et ont été écrits à côté d'elle.

L'Unité d'espagnol de la FTL, pour m'avoir accueilli et m'avoir formé en tant que traducteur et usager privilégié de notre langue ; au département de Traitement Informatique Multilingue, pour m'avoir offert de nouvelles voies de réflexion.

Sylvie Baulo de l'Université de Toulouse et Ángela Rodríguez du Conservatoire de Musique de l'Université Nationale de la Colombie, car sans elles je n'aurais jamais eu le privilège d'arriver à Genève.

Tous mes collègues, linguistes et musiciens, qui ont accepté de participer dans ce projet, pour leur générosité et leur bonne volonté. Je remercie particulièrement les post-éditeurs et l'annotatrice qui ont effectué un travail de grande qualité.

Julie Svortkova et Igor Sergeev pour leur aide précieuse dans les transcriptions des textes russes.

A mis padres, por su incansable apoyo y generosidad.

Table des matières

<i>Déclaration attestant le caractère original du travail effectué</i>	<i>I</i>
<i>Remerciements</i>	<i>II</i>
<i>Index des Figures, des Tableaux et des Illustrations</i>	<i>VII</i>
1. Introduction	1
1.1 Contexte	1
1.2 La post-édition monolingue	2
1.3 Les traités de musique	3
1.4 Questions de recherche, méthodologie et hypothèses	5
1.5 Plan	7
2. La traduction automatique	8
2.1 Introduction	8
2.2 La traduction automatique : histoire et architectures	8
2.3 Parcours historique de la TA	10
2.3.1 Le débuts	10
2.3.2 Le mémorandum de Weaver	11
2.3.3 La TA dans le contexte de la guerre froide	12
2.3.4 La TA au cours des années 1960 – 1990	13
2.3.5 Émergence de l'approche de la TA à partir de corpus	15
2.4 Architectures des systèmes de TA	16
2.4.1 Les systèmes linguistiques	16
2.4.2 Les systèmes statistiques	19
2.4.3 Les systèmes neuronaux	22
3. Post-édition et utilisabilité de la TA	26
3.1 Introduction	26
3.2 La post-édition et ses formes	26
3.2.1 Types de post-édition	26
3.2.2 La post-édition monolingue	28
3.2.3 Profil des post-éditeurs	31
3.3 Utilisabilité de la TA et des post-éditions	34

4. La post-édition monolingue appliquée aux traités de musique	42
4.1 Introduction	42
4.2 La musique et ses textes : les traités de musique	42
4.2.1 Caractéristiques propres aux traités de musique	42
4.2.2 Traduction spécialisée des traités de musique	45
4.3 Décoder les textes de spécialité à l'aide des systèmes neuronaux	48
4.3.1 La TA comme outil d'aide à la traduction spécialisée des traités de musique	48
4.3.2 L'utilisabilité comme critère d'évaluation de la post-édition monolingue	50
4.3.3 Rôle du post-éditeur monolingue : spécialiste ou traducteur ?	52
4.4 Conclusion	53
5. Méthodologie	54
5.1 Objectif, questions de recherche et hypothèses	54
5.2 Choix méthodologiques	55
5.2.1 Plan expérimental	55
5.2.2 Choix des textes	56
5.2.2.1 <i>Le texte non spécialisé (Texte 1)</i>	56
5.2.2.2 <i>Le texte hautement spécialisé (Texte 2)</i>	56
5.2.3 Choix des participants	58
5.2.3.1 <i>Post-éditeurs</i>	58
5.2.3.2 <i>Évaluateurs de l'utilisabilité</i>	59
5.2.4 Choix des outils de post-édition, d'évaluation et de questionnaire	60
5.2.5 Choix des métriques qualitatives	62
5.2.5.1 <i>Erreurs de post-édition (Phase 4)</i>	62
5.2.5.2 <i>Évaluation de l'utilisabilité par les juges (Phase 5)</i>	62
5.3 Préparation et plan de l'expérience	63
5.3.1 Première phase : préparation et traitement des textes	63
5.3.2 Deuxième phase : annotation des erreurs de la TA brute	65
5.3.3 Troisième phase : post-édition monolingue et questionnaire post-tâche	66
5.3.4 Quatrième phase : évaluation des fautes corrigés par les post-éditeurs	69
5.3.5 Cinquième phase : évaluation de l'utilisabilité des post-éditions par les juges	71
5.4 Conclusion	73
6. Résultats	74
6.1 Introduction	74

6.2 Évaluation des fautes corrigées par les post-éditeurs monolingues	75
6.2.1 Taux de correction des post-éditions monolingues	76
6.2.1.1 <i>Texte non spécialisé (Texte 1)</i>	76
6.2.1.2 <i>Texte hautement spécialisé (Texte 2)</i>	77
6.2.1.3 <i>Conclusion</i>	79
6.2.2 Évaluation qualitative des post-éditions monolingues et réponses aux questionnaires post-tâche	80
6.2.3 Conclusion : incidence de l'expertise des post-éditeurs sur la correction d'erreurs de la TA	84
6.3 Résultats de l'évaluation qualitative de l'utilisabilité par les juges experts	86
6.3.1 Utilisabilité des post-éditions monolingues	86
6.3.1.1 <i>Texte non spécialisé (Texte 1)</i>	86
6.3.1.2 <i>Texte hautement spécialisé (Texte 2)</i>	87
6.3.1.3 <i>Conclusion</i>	89
6.3.2 Évaluations quantitative et qualitative des post-éditions monolingues à la lumière de l'utilisabilité estimée par les juges	90
6.3.3 Conclusion : incidence de l'expertise des post-éditeurs sur l'utilisabilité des post-éditions	92
6.4 Comparaison des systèmes de TA	92
6.4.1 Quantité et types d'erreurs annotées dans les TA brutes	93
6.4.2 Choix du système de TA effectué par les post-éditeurs	95
6.4.3 Conclusion de la comparaison entre les systèmes	97
6.5 Conclusion	98
7. Conclusion	103
7.1 Synthèse et résultats	103
7.2 Limites et perspectives	104
<i>Bibliographie</i>	107
<i>Annexes</i>	114
Annexe 1 : Tableau Récapitulatif – État de l'art	115
Annexe 2 : Textes source en langue russe	117
Annexe 3 : TA effectuées par Google Translate et DeepL	121
Annexe 4 : Format d'annotation	129

Annexe 5 : Format de post-édition	140
Annexe 6 : Exemple de post-édition effectué lors de la troisième phase	149
Annexe 7 : Schéma de composition des questionnaires pour les experts	157
Annexe 8 : Récapitulatif des réponses aux questionnaires post-tâche	168

Index des Figures, des Tableaux et des Illustrations

Figure 1 : Taux de correction en fonction des post-éditeurs (musiciens ou linguistes) – Texte 1	76
Figure 2 : Taux de correction en fonction des post-éditeurs (musiciens ou linguistes) – Texte 2	78
Figure 3 : Taux de correction par texte	79
Figure 4 : Total d’erreurs corrigées et non corrigées	80
Figure 5 : Taux global de correction	81
Figure 6 : Utilisabilité des post-éditions – Texte 1	87
Figure 7 : Utilisabilité des post-éditions – Texte 2	88
Figure 8 : Utilisabilité des post-éditions pour les deux textes	89
Tableau 1 : Standard TAUS de post-édition	27
Tableau 2 : Exemples du style hybride technique-littéraire	44
Tableau 3 : Nombre de mots par texte et par segment	58
Tableau 4 : Groupes de post-éditeurs et identifiant correspondant à chaque post-éditeur	59
Tableau 5 : Échelle d’évaluation des post-éditions monolingues	62
Tableau 6 : Exemple de tâche de post-édition complétée	67
Tableau 7 : Taux de correction par texte et par groupe de post-éditeurs	70
Tableau 8 : Exemple d’évaluation avec les numéros de question	71
Tableau 9 : Récapitulatif des questions de recherche et hypothèses	75
Tableau 10 : Nombre d’erreurs identifiées par l’annotatrice et corrigées par les post-éditeurs pour le Texte 1	77
Tableau 11 : Nombre d’erreurs identifiées et corrigées par les post-éditeurs pour le Texte 2	78
Tableau 12 : Valeur des taux de correction pour les musiciens et les linguistes	82
Tableau 13 : Bilan des réponses des experts – Texte 1	87
Tableau 14 : Bilan des réponses des experts – Texte 2	89
Tableau 15 : Nombre d’erreurs par système – Texte non spécialisé	93

Tableau 16 : Nombre d'erreurs par système – Texte hautement spécialisé	94
Tableau 17 : Nombre total d'erreurs par système des deux textes	95
Tableau 18 : Total et pourcentage des segments choisis par les musiciens	96
Tableau 19 : Total et pourcentage des segments choisis par les linguistes	96
Tableau 20 : Total et pourcentage des segments choisis par les deux groupes de post-éditeurs	98
Tableau 21 : Bilan total de favorabilité non nuancée	100
Tableau 22 : Bilan des « Non utilisables » par texte	101
Illustration 1 : Carte disciplinaire de Holmes-Toury	1
Illustration 2 : Branche appliquée de la traductologie	2
Illustration 3 : Triangle de Vauquois	17
Illustration 4 : Exemple musical	45
Illustration 5 : Exemple de question sur Limesurvey	61

1. Introduction

1.1 Contexte

Le croisement des chemins entre les technologies langagières, les humanités et les arts ouvre des voies d'exploration et de recherche favorables à l'essor de dialogues parfois innovants entre les disciplines. Dans ce mémoire, le croisement entre la traduction automatique, la pratique de la post-édition et les écrits théoriques sur la musique est étudié avec l'intention d'explorer son potentiel et ses limites dans un contexte pédagogique.

Le sujet d'étude principal de ce travail est la post-édition monolingue d'une typologie de textes particulière, à savoir les écrits sur la musique. En tant que domaine de recherche à part entière, la post-édition est un phénomène relativement récent (Poibeau 2017) ; c'est particulièrement vrai pour la post-édition monolingue.

Du point de vue traductologique, ce mémoire s'inscrit donc dans la branche théorique de la carte disciplinaire de Holmes-Toury citée par Munday (2016) :

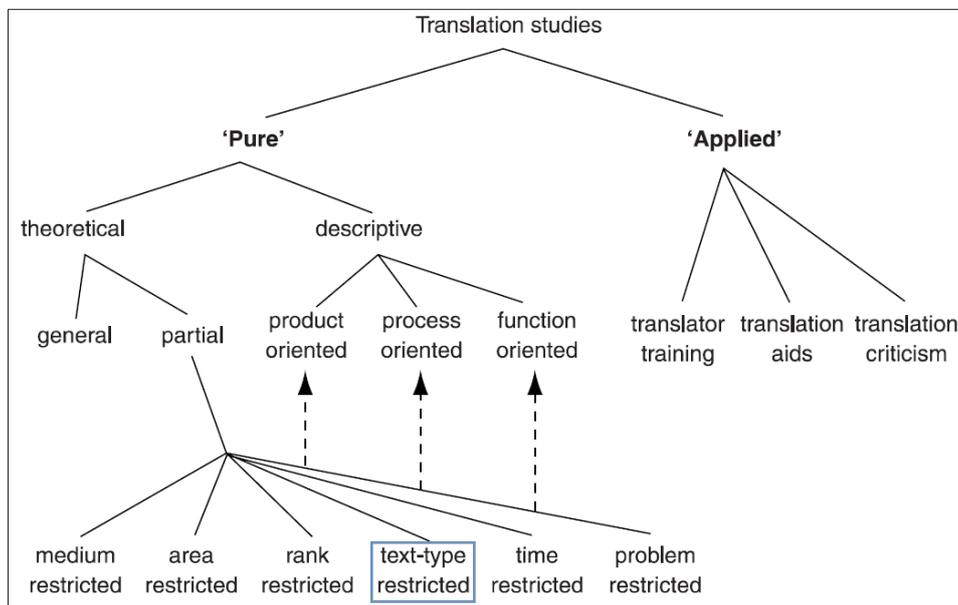


ILLUSTRATION 1. CARTE DISCIPLINAIRE DE HOLMES-TOURY (MUNDAY 2016, 17)

Comme montré dans l'illustration 1, il s'agit d'une recherche qui explore un type de texte en particulier, le traité de musique.

Du point de vue disciplinaire, cette recherche s'inscrit également dans la branche appliquée de la carte disciplinaire de Holmes-Toury, telle qu'elle a été élargie par Munday (2016) (voir Illustration 2) :

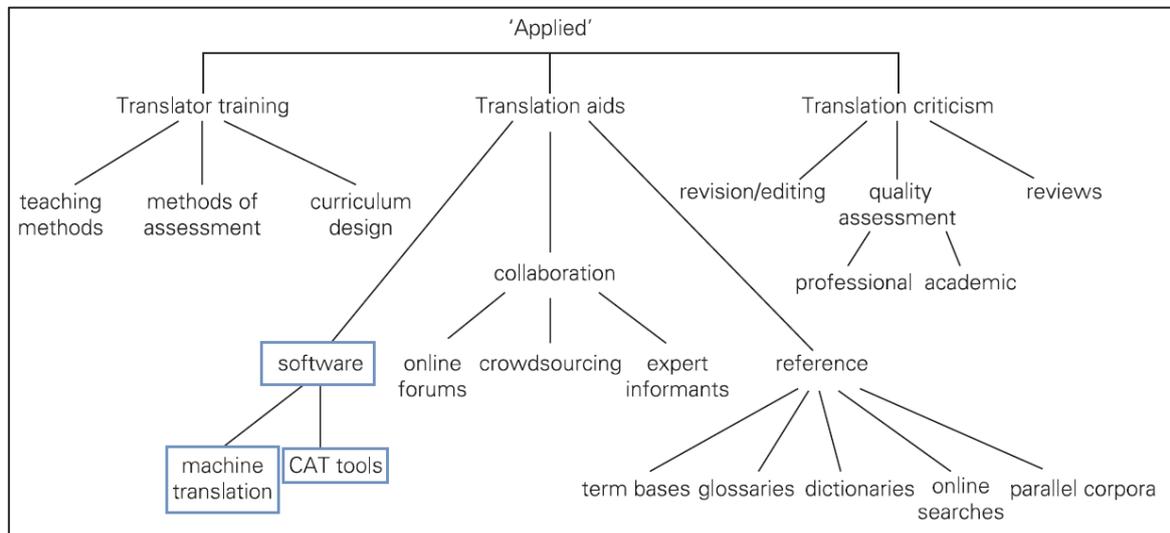


ILLUSTRATION 2. BRANCHE APPLIQUEE DE LA TRADUCTOLOGIE (MUNDAY 2016, 20)

Il s'agit en effet d'une recherche en technologies de la traduction qui vise à explorer la pratique de la post-édition monolingue à travers une étude expérimentale menée avec des participants humains dont les résultats seront le fondement pour confirmer ou réfuter les hypothèses et la problématique de ce travail.

1.2 La post-édition monolingue

La post-édition peut être définie en décrivant la tâche effectuée par les post-éditeurs : « The task of the post-editor is to edit, modify and/or correct pre-translated text that has been processed by an MT system from a source language into (a) target language(s) » (Allen 2003, 298). Cette tâche de post-édition est, dans la plupart des cas, pratiquée dans un scénario textuel bilingue. Pour Nunes Vieira (2019), les risques associés au non-recours au texte source impliquent en effet que la post-édition bilingue soit aujourd'hui la norme dans les milieux professionnels.

Naturellement, la normalisation de la post-édition bilingue dans les milieux professionnels se reflète aussi dans les recherches en post-édition, majoritaires dans le domaine. Toutefois, il est possible de trouver des recherches en post-édition monolingue appliquée à des textes issus des divers domaines comme le journalisme

(Koponen et Salmi 2015), la sociologie (Nitzke 2016), l'informatique (Van Egdom et Pluymaekers 2019) ou la biologie (Schwartz 2014).

Certaines de ces recherches ont démontré que les post-éditeurs sans connaissance de la langue source peuvent arriver à améliorer la qualité de la traduction automatique. Schwartz (2014), par exemple, a démontré qu'une spécialiste peut post-éditer correctement plus de 96% des phrases d'un texte de spécialité sans recourir au texte source. Koehn (2010) estime dans une autre expérience que les post-éditeurs monolingues ont post-édité correctement un tiers des phrases, cette fois pour des textes journalistiques. Dans les deux cas, les connaissances spécialisées ou thématiques des post-éditeurs ont été des conditions déterminantes pour la qualité des post-éditions monolingues.

Par ailleurs, les recherches en post-édition monolingue posent aussi des questions tout-à-fait pertinentes autour des notions telles que : l'utilisabilité (Doherty et O'Brien 2014; Van Egdom et Pluymaekers 2019), la qualité de la traduction automatique donnée pour effectuer la tâche de post-édition (Nitzke 2016), et le profil des post-éditeurs (Nitzke 2016; Barbosa de Lima Fonseca 2019; Schwartz 2014). C'est pourquoi ce travail s'intéresse à ces questions en mettant l'accent sur l'incidence du profil des post-éditeurs monolingues et sur la qualité du résultat final.

1.3 Les traités de musique

Nous nous intéresserons ici à l'impact de la post-édition monolingue sur une typologie textuelle particulière : le traité de musique. Si on adopte la définition du « traité » proposée par Le Robert, le traité de musique peut être défini comme un ouvrage didactique où est exposé de manière systématique un sujet ou un ensemble de sujets concernant la musique¹.

Du point de vue de la traduction automatique, ce type de textes possède des caractéristiques linguistiques qui peuvent représenter un défi intéressant pour les systèmes actuels, telles que la terminologie spécialisée du domaine ou encore le

¹ © Dictionnaires Le Robert - Le Petit Robert de la langue française (2020). Entrée « Traité ». <https://petitrobert.lerobert.com/robert.asp>, 04 avril 2021.

caractère à la fois technique et littéraire du langage employé par les spécialistes dans ce type de textes. De plus, la combinaison linguistique adoptée dans ce mémoire (russe > espagnol) n'a pas non plus été étudiée dans les recherches en post-édition monolingue.

Comme il a été montré par Palisca et Bent (2001), les traités de musique sont, en règle générale, regroupés autour de la discipline appelée « théorie de la musique ». Néanmoins, ils déploient une diversité de contenus, objectifs et publics cibles qui démontre l'extension et la richesse du concept de théorie dans la musique occidentale. À travers les siècles, les traditions musicales byzantine, arabe, hébraïque, chinoise et indienne ont produit une grande quantité de ce type de textes (Palisca et Bent 2001). Néanmoins, la notion de « traité de musique » adoptée dans cette étude se limitera exclusivement aux traités issus de la tradition européenne occidentale du XX^{ème} siècle à nos jours.

L'absence de travaux de recherche sur la traduction de ce type de textes dans la littérature est presque totale. En effet, il a été signalé par Benoit-Otis (2013) il y a quelques années que « même si de nombreux travaux (...) ont déjà été consacrés aux problèmes liés à la traduction du texte d'œuvres vocales, en particulier les livrets d'opéra, la traduction des écrits théoriques sur la musique n'a été que très peu étudiée jusqu'à maintenant ». Ce constat est toujours d'actualité aujourd'hui, d'autant plus dans le domaine des technologies de la traduction.

Le seul travail concernant en effet les conséquences traductologiques de la fonction textuelle des textes didactiques de musique est celui de Vella Ramírez et Muñoz Muñoz (2010), dans lequel les auteurs montrent la présence simultanée de traits du langage technique et littéraire. En général, les travaux qui explorent des liaisons entre la traduction et la musique appartiennent le plus souvent à l'une des trois catégories suivantes : la traduction de chansons, la traduction de livrets d'opéra ou la traduction multimédia et audiovisuelle. De ce fait, l'originalité de ce projet est justifiable du point de vue traductologique.

En ce qui concerne la recherche en TA, malgré l'absence de ce type de textes dans la littérature, il est possible de trouver des travaux dans ce domaine consacrés à la

traduction automatique du langage technique ou littéraire. Cependant, notre recherche propose d'appliquer la TA à une typologie textuelle qui transgresse fréquemment la frontière entre ces deux langages et qui pose pour la première fois la question sur la possibilité d'employer la TA et la post-édition monolingue au lieu de la traduction humaine pour ce type de textes. Cela justifie aussi l'originalité du projet du point de vue de la TA.

En dehors de la littérature scientifique, deux travaux académiques précédents de l'auteur de ce mémoire ont pris le traité de musique comme objet d'étude. Le premier a adopté une perspective bibliométrique pour déterminer dans quelle mesure ce type de textes a été l'objet de traductions vers l'espagnol (Salamanca 2018) ; le deuxième a proposé une traduction inédite d'un texte du XVIIIème siècle, issu du corpus de la première étude, ainsi qu'une réflexion sur la pratique de la traduction spécialisée des traités de musique (Salamanca 2019). Aucun d'entre eux n'a développé de lien avec le domaine des technologies de la traduction comme possible outil d'aide pour traduire cette typologie textuelle.

Par conséquent, l'originalité de ce projet est justifiable par le fait de voir l'impact de la post-édition monolingue sur un type de texte très peu étudié. Cette approche permet donc d'explorer les possibilités offertes par la traduction automatique et plus précisément par la post-édition monolingue. Elle s'interroge sur l'incidence de deux variables dans la qualité et l'utilisabilité des post-éditions : le profil des post-éditeurs et le degré de spécialité des textes.

1.4 Questions de recherche, méthodologie et hypothèses

L'expertise des spécialistes est-elle plus utile pour effectuer une post-édition monolingue des textes spécialisés que la connaissance linguistique et l'expérience des traducteurs ? Cette question a déjà été explorée dans des recherches en post-édition bilingue appliquée à des domaines comme la médecine (Rigouts Terryn et al. 2019) et l'ingénierie (Temizöz 2013), ou encore en post-édition monolingue dans le domaine de la biologie (Schwartz 2014). Similaire à ces travaux, ce mémoire vise à répondre à cette question en comparant la performance des musiciens et des

traducteurs dans la post-édition monolingue de deux textes avec des niveaux de technicité différents.

Dans ce sens, le degré de spécialité de ce type de textes peut-il aussi influencer la performance des post-éditeurs monolingues ? Autrement dit, est-il possible de rendre utilisables en langue cible des textes hautement techniques, tout comme des textes dédiés à des notions de base de la spécialité, à travers un processus de post-édition monolingue ? Ces questions visent directement les contextes d'utilisation des traités de musique et leur caractère pédagogique.

En résumé, et tout en considérant les divers éléments évoqués par ces questionnements, la question fondamentale de ce mémoire de maîtrise est :

- *dans quelle mesure l'expertise du post-éditeur monolingue dans le domaine de spécialité, ainsi que le degré de spécialisation des textes, sont des variables déterminantes pour l'utilisabilité des post-éditions produites ?*

Par ailleurs, nous cherchons à répondre à deux questions subsidiaires concernant les systèmes de traduction automatique employés lors de notre expérience, Google Translate et DeepL :

- *lequel des deux systèmes a une meilleure performance d'après les annotations de l'annotatrice bilingue ?*
- *lequel des deux systèmes est d'avantage choisi par les post-éditeurs pour effectuer la tâche qui leur est présentée ?*

Pour répondre à ces questions, ce mémoire propose une méthodologie de recherche en cinq phases. Dans la première phase, nous allons effectuer toute la préparation et le traitement des textes nécessaires pour mener l'expérience. Dans la deuxième phase, une annotatrice bilingue, experte dans le domaine, marquera des erreurs dans les traductions automatiques brutes avec le texte de départ comme référence. Dans la troisième phase, deux groupes de post-éditeurs, le premier composé de spécialistes dans le domaine de la musique, et le deuxième de traducteurs, effectueront la tâche de la post-édition monolingue. Dans la quatrième phase, nous allons effectuer une évaluation quantitative et qualitative des fautes corrigées par

les post-éditeurs. Dans la cinquième phase, un groupe de jurés hispanophones, experts dans le domaine de la musique, évaluera l'utilisabilité des post-éditions en prenant en compte le caractère didactique de ce type de textes.

Il est possible de formuler comme hypothèse que plus le post-éditeur est expert dans le domaine, plus il est possible d'améliorer l'utilisabilité de la traduction automatique brute d'un texte de spécialité. De même, dans la pratique de la post-édition monolingue, il est possible que les experts voient plus de fautes de fidélité que les traducteurs.

1.5 Plan

Le deuxième chapitre de ce mémoire sera dédié à la traduction automatique. Il présentera un bref parcours historique de la traduction automatique, ainsi qu'une description des architectures propres aux systèmes de traduction automatique.

Le troisième chapitre développera les notions de post-édition et d'utilisabilité de la traduction automatique. Nous y présenterons les études les plus récentes dans ces domaines qui sont en lien avec notre recherche.

Le quatrième chapitre, quant à lui, exposera les principales caractéristiques de notre typologie textuelle, ainsi que la possibilité d'appliquer la post-édition monolingue comme outil d'aide à la traduction de ce type de textes.

Le cinquième chapitre de ce mémoire présentera une description détaillée de la procédure employée lors de l'expérience menée pour ce mémoire, tout en précisant les choix méthodologiques et le déroulé.

Finalement, le sixième chapitre présentera les résultats de l'expérience, à la lumière de nos hypothèses et questions de recherche, avant de passer à la conclusion dans le septième et dernier chapitre de ce travail.

2. La traduction automatique

2.1 Introduction

Avant de développer en détail la question de l'utilisabilité des post-éditions monolingues dans le domaine de la musique, il est nécessaire de s'interroger sur la pratique de la post-édition en soi, ainsi que sur les formes qu'elle peut prendre. Cela permettra de mieux comprendre les enjeux propres à l'expérience qui sera menée pour répondre à nos questions de recherche.

La question sur la post-édition mène à la technologie utilisée pour la traduction : la traduction automatique. Celle-ci est définie comme le processus entièrement automatisé qui part d'un texte dans une langue source et en produit un autre de même sens en langue cible (Melby 2019, 419). Cette section propose, d'abord, un bref résumé de son histoire tout en mettant l'accent sur la notion de décodage, ainsi qu'une description des différentes architectures des systèmes de traduction automatique. Ensuite, elle développe les caractéristiques des différents types de post-édition et les avancées méthodologiques et empiriques proposées par des travaux récents dans ce domaine.

Pour l'état de l'art développé dans ce chapitre, deux approches seront suivies : (a) une approche chronologique pour rendre compte de l'évolution historique de la traduction automatique jusqu'à nos jours (section 2.3) ; (b) une approche thématique pour analyser et définir les différentes architectures des systèmes, ainsi que les recherches récemment développées dans le domaine (section 2.4).

Dans ce chapitre, nous allons présenter un aperçu de l'histoire et du développement des systèmes de traduction automatique (section 2.2). Ensuite, nous allons brièvement présenter les différentes étapes de l'histoire de la traduction automatique (section 2.3). Enfin, nous allons décrire chacune des trois architectures des systèmes de traduction automatique (section 2.4).

2.2 La traduction automatique : histoire et architectures

Les débuts de la traduction automatique (TA) ont été marqués, entre autres, par les enjeux de communication pendant et après la Seconde Guerre mondiale. La

cryptologie, l'un des domaines en plein développement à cette période, a permis à Warren Weaver, père de la TA, de concevoir un parallélisme entre l'acte de traduire et le décodage : un document en langue source peut être considéré comme « crypté » et, de ce fait, la traduction aurait la fonction de le rendre intelligible, c'est-à-dire de le « décoder » pour retrouver la langue source qui a la plus grande probabilité de générer la langue cible (Poibeau 2019, 39). Cette idée a été très influente dans l'histoire de la TA.

Il est également intéressant de remarquer la vitesse des progrès dans l'histoire de la TA : « A total of some seventeen major events took place during the history of human translation, which may be 3135 years old. (...) In comparison with human translation, machine translation has advanced enormously since its inception in the 1940s » (Chan 2015, 23). Cette vitesse est sans doute liée au rythme de l'évolution technologique depuis la deuxième moitié du XX^{ème} siècle, mais peut aussi être expliquée par cette vieille « pulsion humaine » de traduire (Berman 1995) qui, à partir des années 40, trouve dans l'informatique une source de réponses possibles aux problèmes liés à la communication interlinguistique.

Comme indiqué par Hutchins (2015), l'histoire de la TA peut être divisée en trois périodes. Chacune de ces périodes a impliqué des évolutions considérables dans la conception et les architectures des systèmes de TA :

- Les débuts (1949–1966), marqués par la naissance des premiers systèmes opérationnels d'IBM aux États-Unis, ainsi que par le développement d'une grande diversité d'approches.
- La période des systèmes de TA linguistiques (1967–1989), caractérisée par une approche linguistique, ainsi que par l'application de la TA à des fins commerciales, le développement du langage contrôlé et des premiers outils d'aide à la traduction.
- La période des systèmes basés sur les corpus (1990–aujourd'hui), marquée par le développement des systèmes statistiques, l'usage des ressources numériques liées à l'internet, la croissance exponentielle de la puissance des ordinateurs et, finalement, le développement des systèmes neuronaux.

Dans la suite de ce chapitre, nous allons présenter une esquisse historique qui suit de près ces étapes, des débuts jusqu'à l'apparition des systèmes de traduction neuronaux. Ensuite, nous allons décrire les architectures des systèmes linguistiques, statistiques et neuronaux, en mettant particulièrement l'accent sur cette dernière.

2.3 Parcours historique de la TA

2.3.1 Le débuts

Il est possible de tracer l'apparition des idées autour de l'automatisation du processus de traduction avant la Seconde Guerre mondiale. Chan (2015) mentionne, par exemple, la machine créée par l'ingénieur franco-arménien Georges Artsrouni entre 1932 et 1937 ; la machine du chercheur russe Petr Petrovič Smirnov-Troyanskij, qui date de la même époque, s'inscrit aussi dans les premières explorations dans le domaine pendant la période d'avant-guerre (Ibid., 23). Ces outils ont un fonctionnement purement mécanique, mais ils peuvent être considérés comme les premiers essais d'automatiser la traduction à l'aide d'une machine.

La naissance de l'idée de la TA est en général attribuée à la collaboration entre Andrew D. Booth, ingénieur électrique anglais, et Warren Weaver, cybernéticien et mathématicien américain. En 1946, Booth, motivé par l'apparition des premières calculatrices électroniques, sollicite la collaboration de Warren Weaver pour développer le premier ordinateur britannique (León 2008, 2767). Cette collaboration donne lieu à la première application de l'informatique à un processus de traduction automatisé.

Du point de vue historique, Booth et Weaver développent pour la première fois, non seulement l'idée d'utiliser les ordinateurs pour effectuer une tâche essentiellement humaine (traduire), mais aussi l'idée de traiter automatiquement les langues naturelles (Chan 2015, 24). Il s'agit donc des débuts de la TA, mais aussi de la recherche dans des domaines liés comme la linguistique computationnelle ou l'intelligence artificielle.

2.3.2 Le mémorandum de Weaver

C'est pour ces raisons que Warren Weaver est considéré aujourd'hui comme le père de la TA et du traitement automatique des langues (Poibeau 2019, 41). Son influence dans le domaine est allée au-delà de sa collaboration avec Booth et d'autres chercheurs, particulièrement avec la rédaction en 1949 de son *Mémorandum*. Ce texte est souvent considéré comme le « point de départ des recherches » dans le domaine de la TA (Poibeau 2019, 43). Dans son Mémorandum, Weaver propose quatre principes pour dépasser le simple transfert mot-à-mot dans la conception des systèmes de TA. Bien que tous ces principes ont eu une grande influence dans le développement des futurs systèmes de TA (Melby 2019), un de ces principes s'avère particulièrement pertinent pour notre recherche : celui de la cryptographie.

Dans ce contexte, la cryptographie est le troisième principe proposé par Weaver dans son Mémorandum. Selon lui, « la traduction automatique peut s'inspirer des techniques développées en cryptographie » (Poibeau 2019, 43). Celle-ci avait en effet accompli des progrès significatifs pendant la Seconde Guerre mondiale, et Weaver voyait des parallèles entre la décodification des messages et l'acte de traduire. Dans son texte, il cite notamment le travail qu'il avait développé avec Claude Shannon.

En 1949, Weaver et Shannon ont travaillé sur une modélisation mathématique de la communication : un message est encodé par une source, puis le message est émis pour, finalement, être décodé par un récepteur (Poibeau 2019, 41). Il s'agit du processus de base de la cryptographie appliqué à la communication. De ce fait, l'acte de traduire un texte source peut être conçu comme une codification en langue étrangère d'un texte « original » en langue cible ; ce texte devrait être « décodé » pour le rendre à nouveau intelligible grâce au processus de traduction (Poibeau 2019, 41; Melby 2019, 423).

Cette idée de la décodification des langues est devenue « très populaire » selon Poibeau (2019, 44), a été qualifiée comme « completely backwards to many people in the 1950s » selon Melby (2019, 43). Ceci peut s'expliquer par le fait que, à cette époque, les textes source écrits dans une langue étrangère avec un système

alphabétique différent de celui de la langue cible, sont perçus comme « codifiés ». Cette idée aura un impact et sera prise très au sérieux par les chercheurs à partir des années 1990 dans le domaine de la TA.

2.3.3 La TA dans le contexte de la guerre froide

Mais l'histoire de la naissance de la TA est aussi en lien étroit avec le contexte politique de l'époque. Bien que l'invention des ordinateurs a rendu possible la TA après la Seconde Guerre mondiale, en lien avec des facteurs tels que l'apparition de la théorie de l'information et les avancées en cryptolog, il existait, en réalité, un besoin applicatif direct de la TA dans le contexte de la guerre froide.

Les relations conflictuelles entre la Russie et les États-Unis ont en effet stimulé l'intérêt des deux parties dans la possibilité de traduire automatiquement la langue de l'autre, mais aussi d'autres langues. Les projets initiaux de traduction automatique, de l'allemand vers l'anglais et de l'allemand vers le russe, coïncident notamment avec le partage des connaissances sur la technologie des missiles (Ayvazyan et Pym 2017, 238), ce qui suggère que les deux puissances voyaient la TA comme un moyen de répondre à leurs besoins d'information dans la course aux armements.

Comme le notent (Ayvazyan et Pym 2017, 233), la recherche dans ce domaine était tout aussi intense dans l'Union Soviétique qu'aux États-Unis, et les deux pays ont été pionniers dans les recherches en TA (Hutchins 2015, 121). La première conférence sur la TA a été organisée en 1952 au MIT par Joshua Bar-Hillel, chercheur d'origine israélienne qui a joué un rôle clé dans le développement de la TA aux États-Unis (Poibeau 2019, 45-46). À Moscou, le premier congrès de TA a eu lieu en 1958 et a rassemblé plus de 340 participants (León 2008).

À cette conférence, le premier système de TA (qui traduisait du russe vers l'anglais) a été dévoilé : le projet Georgetown-IBM, capable de traduire quelques phrases et qui comptait seulement six règles de transfert (Chan 2015, 24). C'est à partir de ce moment que la TA s'est développée.

Cette démonstration a aussi attiré l'attention des chercheurs soviétiques qui, par manque de moyens et d'ordinateurs disponibles pour développer leur recherche,

produisent à partir de 1955 des travaux théoriques et des stratégies autour de l'analyse syntaxique et sémantique (Poibeau 2019, 51). Mais il s'agit seulement des débuts de la recherche dans les systèmes linguistiques à base de règles, car les études dans le domaine s'étendront jusqu'à la fin des années 1980.

Cette première décennie du développement a vu naître les trois approches de base de la TA linguistique (Hutchins 2015). Ces modèles montreront leur limite. Très vite, les chercheurs ont reconnu que leur travail avait finalement été confronté à la « barrière sémantique ». Produire des traductions automatiques de qualité nécessiteraient que les machines soient capables de « comprendre » ce qu'elles traduisent (Hutchins 2015, 121). Cette conclusion est encore d'actualité aujourd'hui et alimente les réflexions d'une partie des chercheurs qui s'intéressent aux systèmes de TA neuronaux.

2.3.4 La TA au cours des années 1960 – 1990

C'est en partie à cause de ces limites, et malgré les efforts des chercheurs et les avancées dans le développement des systèmes de TA, que Joshua Bar-Hillel rédige en 1959 le *Rapport sur l'état de la traduction automatique aux États-Unis et en Grande-Bretagne* (Poibeau 2019, 52-53). Tout en laissant de côté le fait qu'il s'agissait d'une discipline très récente, il émet un jugement défavorable sur les travaux consacrés à la TA. Il préconise donc de remplacer cette recherche par le développement des outils d'aide à la traduction qui pourraient augmenter la productivité des traducteurs humains (Ibid., 54). Cette publication a eu un impact négatif sur la perception de la recherche en TA et a ouvert les portes pour la rédaction de l'ALPAC, un rapport commandé en 1966 par le Département de Défense des États-Unis.

Ce rapport a déclaré qu'il n'y avait pas de perspectives immédiates ou prévisibles d'un système de TA utile (Chan 2015, 4) et, étant donné que la TA était beaucoup plus chère que la traduction humaine, le rapport recommandait d'arrêter les ressources destinées à soutenir la recherche en TA. Ce rapport a marqué la fin de cette étape de l'histoire de la TA, ainsi que la reconversion de nombreux chercheurs vers d'autres sujets de recherche. Par la suite, nous allons voir plus en détail que les travaux développés lors des décennies suivantes ont prouvé que la TA était

encore, et malgré tout, un champ d'étude et de recherche qui continuait à susciter de l'attention.

Entre les années 1960 et 1971, Bernard Vauquois, chercheur à l'Université de Grenoble, développe avec son équipe un système interlingue pour traduire des textes spécialisés en mathématiques et physique du russe vers le français (Hutchins 2015, 124). Il est tout-à-fait intéressant de remarquer l'importance de la traduction du russe vers les langues européennes pendant cette période, qui s'explique pour le contexte de la guerre froide. Mais cela s'explique aussi par un intérêt authentique pour la connaissance développée par des auteurs russophones dans des spécialités diverses. C'est en effet le cas de notre recherche.

La TA pouvait donc être aussi conçue par les chercheurs comme un vecteur de connaissances. Quant à Vauquois et son équipe de Grenoble, malgré les résultats peu encourageants de leur système interlingue, le processus d'expérimentation et de recherche a continué. Ils ont développé le système Ariane qui a eu une énorme influence sur la recherche de l'époque malgré le fait que, comme beaucoup de leurs contemporains, ce système n'a jamais été opérationnel (Hutchins 2015, 124). La recherche sur le modèle d'Ariane s'est arrêtée à la fin des années 1980.

Les systèmes linguistiques sont quasiment inutilisés aujourd'hui. D'abord à cause de leur complexité et le coût impliqué dans leur fonctionnement, mais aussi pour l'évolution des systèmes à base de corpus qui sont entrés en scène dans les années 1990. Mais il existe encore un autre facteur qui mérite d'être pris en compte, car un des piliers épistémologiques des systèmes linguistiques et de leurs règles est la notion des universaux du langage.

Weaver a proposé cette idée dans son Mémoire comme l'une des voies possibles pour l'évolution des systèmes de TA, et elle a été au centre de la grammaire générative de Chomsky, très influente à l'époque. Néanmoins, la linguistique s'est éloignée de la quête des universaux du langage, tout comme de la question philosophique sur l'existence même de ces universaux (Melby 2019, 428). Cela a eu pour conséquence le questionnement des fondements autour de l'architecture et du fonctionnement de ces systèmes.

En résumé, cette période de l'histoire de la TA a permis de comprendre la complexité de l'analyse et du traitement des langues naturelles, l'automatisation du processus de traduction et l'importance des niveaux sémantiques et pragmatiques pour la qualité des traductions automatiques, et cela en raison des limites des systèmes linguistiques (Poibeau 2019, 55). C'est pourquoi la recherche et les avancées de cette période ne sont pas inutiles et se sont montrées importantes pour les développements postérieurs des systèmes de TA.

2.3.5 Émergence de l'approche de la TA à partir de corpus

L'émergence des méthodes basées sur des corpus de données à partir de 1989 a marqué la fin de la dominance de l'approche linguistique et des systèmes de TA basés sur les règles (Hutchins 2015, 128; Melby 2019, 419). Comme le mentionne Hutchins (2015), les méthodes statistiques avaient été déjà explorées aux débuts de la recherche sur la TA, mais les résultats s'étaient avérés décourageants. À l'époque, l'accent avait été mis sur les liens entre les mots des langues source et cible, plutôt que les segments (Ibid., 128-29).

Du point de vue contextuel, le passage à ce nouveau paradigme a été rendu possible par quatre facteurs : l'évolution de l'apprentissage automatique, la possibilité d'accéder à des ordinateurs puissants plus abordables et de traiter de grandes quantités de textes, et une nette augmentation de la demande des traductions (Melby 2019, 425). Les premiers modèles des systèmes statistiques développés par l'équipe de recherche de l'IBM aux États-Unis ont eu beaucoup de succès et la plupart de systèmes de TA statistiques développés postérieurement se sont construits sur ce paradigme (Poibeau 2017), du moins jusqu'à l'apparition des systèmes neuronaux.

Tout comme les systèmes statistiques, le développement des systèmes neuronaux repose aussi sur des corpus de textes informatisés (Melby 2019, 429). Leur développement a également été stimulé par de fortes demandes de traductions à une époque où la quantité des documents à traduire fait l'objet d'une croissance exponentielle. Depuis son apparition en 2016, l'approche neuronale a été très rapidement adoptée et est aujourd'hui utilisée notamment par Google, Facebook, Bing, Systran et DeepL (Poibeau 2017).

Dans tous les cas, il est clair que les systèmes de traduction automatique neuronaux représentent de tout point de vue un énorme progrès par rapport aux systèmes apparus depuis 1949. La notion proposée par Weaver, père de la TA, de « décoder » le texte source en texte cible, a eu un impact bien au-delà des années 50. Encore aujourd’hui, elle reste, d’une manière ou d’une autre, l’un des principes de l’automatisation du processus de traduction. Par la suite, nous allons présenter plus en détail l’architecture des systèmes linguistiques (section 2.4.1), statistiques (section 2.4.2) et neuronaux (section 2.4.3).

2.4 Architectures des systèmes de TA

2.4.1 Les systèmes linguistiques

Les systèmes linguistiques sont souvent appelés en anglais « ruled based translation systems » (RBTS) étant donné qu’ils modélisent les connaissances avec différents types de règles développées par des humains. Ces règles opèrent sur différents niveaux linguistiques et peuvent être de type lexical ou morphologique, syntaxique, sémantique ou pragmatique (Hutchins 2015, 128; Melby 2019, 419).

Les systèmes de TA linguistiques suivent la conception du processus de traduction proposée par la théorie interprétative de Saleskovitch et Lederer. D’après cette théorie, les traducteurs humains passent par trois phases dans le processus de traduction : compréhension du sens, déverbalisation (transfert du sens) et réexpression en langue cible (Munday 2016, 101).

L’apparition en 1954 du système Georgetown-IBM, mentionné dans la section précédente, est le premier exemple de ce type de systèmes. En 2000, les systèmes linguistiques étaient les seuls systèmes commerciaux disponibles. Il existe trois modèles de ce type de système : les modèles de traduction *directe*, les modèles du *transfert*, et les modèles *interlangue* (Hutchins 2015, 120). Ces trois modèles ont été résumés dans un schéma proposé par Vauquois en 1968 qui montre les différentes possibilités pour passer de la langue source à la langue cible (Seligman 2019, 56).

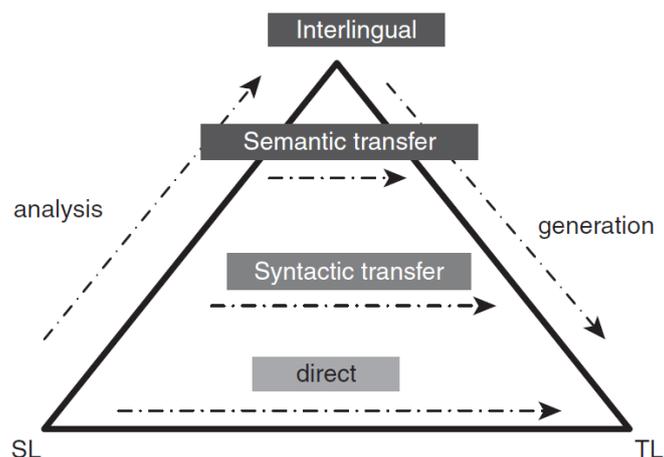


ILLUSTRATION 3
TRIANGLE DE VAUQUOIS (SELIGMAN 2019, 56)

Les systèmes *directs* traduisent le texte source mot à mot. De ce fait, la traduction directe s'appuie sur un dictionnaire bilingue dans lequel chaque mot ou expression de la langue source a un mot équivalent en langue cible (Jurafsky et James 2009, 886). Ce système utilise aussi des informations monolingues d'ordre grammatical comme le nombre, le genre ou la catégorie des mots. Par ailleurs, les systèmes directs permettent le développement de règles spécifiques de traduction propres à chaque combinaison linguistique, plus ou moins complexe, et une réorganisation syntaxique minimale (Hutchins 2015, 120) qui a pour but de rendre idiomatique le résultat de la traduction en langue cible. Ces systèmes se situent à la base du triangle de Vauquois.

Bien que ce type de systèmes puisse réorganiser correctement des mots dans le cas de certaines combinaisons linguistiques, et même traduire correctement des phrases simples, il ne possède en réalité aucune connaissance de la structure grammaticale des langues (Jurafsky et James 2009, 887). De ce fait, les réorganisations syntaxiques complexes exigées par certaines combinaisons linguistiques ne peuvent pas être traitées correctement par ces algorithmes. Finalement, l'élaboration des dictionnaires et des règles implique un effort très important qui n'est pas toujours justifié par la qualité des traductions produites par les systèmes directs.

Les systèmes *indirects* présentent une architecture avec deux structures intermédiaires pour effectuer le transfert (Seligman 2019, 57). La première est la structure syntaxique du texte source (elle est plus profonde car les systèmes directs

restent au niveau lexical). La seconde est une structure chargée de représenter le sens du texte source en langue cible. Cette représentation sert comme point de départ pour la génération du texte cible (Ibid.). Il s'agit d'une architecture plus complexe que les systèmes directs étant donné qu'elle repose sur une analyse syntaxique de la phrase à traduire (Poibeau 2019, 22), ce qui dépasse la réorganisation simple de mots proposée par les systèmes directs. De ce fait, ce qui est mis en relation dans les systèmes indirects ne sont pas les mots, mais la correspondance des représentations. Ces systèmes se situent au milieu du triangle de Vauquois.

Les systèmes linguistiques commerciaux présentaient souvent une architecture mixte entre les systèmes directs et indirects avec une phase d'analyse superficielle, une phase de transfert et une phase finale de synthèse. De ce fait, ces systèmes emploient des dictionnaires bilingues et effectuent des réorganisations de mots, comme les systèmes directs, mais aussi mettent en relation des représentations syntaxiques, comme les systèmes indirects (Jurafsky et James 2009, 890-91).

Finalement, les systèmes *interlingues* sont les plus ambitieux car leur approche vise à représenter le sens des phrases source de façon abstraite afin de générer une phrase cible qui véhicule le sens d'une telle représentation (Poibeau 2019, 22). Cette représentation peut être un langage abstrait ou un langage « pivot » existant comme l'espéranto ou l'anglais. Ce modèle présuppose donc l'existence d'une représentation du sens indépendant des langues cible et source produit une analyse sémantique profonde des phrases à traduire (Jurafsky et James 2009, 891). Ces systèmes se situent au sommet du triangle de Vauquois.

En ce qui concerne l'analyse sémantique, Poibeau avance qu'elle reste « une tâche très difficile (...) dans la mesure où il est quasi impossible de prévoir tous les contextes dans lesquels une expression peut apparaître » (2019, 24). Cela oblige, par exemple, à faire des distinctions qui ne sont pas nécessaires dans toutes les langues et à définir et associer correctement les concepts, ce qui implique une analyse exhaustive d'un domaine déterminé et la création d'une ontologie pour formaliser l'analyse (Jurafsky et James 2009, 892), et qui réduit son application effective à des domaines très spécifiques.

En résumé, le modèle de *traduction directe* propose le développement de règles spécifiques de traduction propres à chaque combinaison linguistique, avec une analyse et une réorganisation minimale lexicale et syntaxique. Dans le modèle du *transfert*, la conversion passe par une analyse et par une étape de transfert des représentations abstraites des textes en langue source vers les représentations équivalentes en langue cible, avant de générer les textes cible. Finalement, le modèle *interlingue* est celui dans lequel la traduction du texte source passe par des représentations abstraites, comme des codes ou des symboles indépendants des langues impliquées, ou des langues « pivot », avant de produire le texte cible (Hutchins 2015, 120).

Si l'on considère les niveaux linguistiques classiques (morphologique, syntaxique, sémantique et pragmatique), ces systèmes ont mis l'accent sur le transfert au niveau des mots et de la syntaxe, raison pour laquelle leur performance à ces deux niveaux était généralement acceptable (Melby 2019, 428). Cette caractéristique impliquait aussi que les niveaux sémantique et pragmatique n'étaient certainement pas le point le plus fort de ces systèmes. Or, toute communication humaine se joue aussi (et surtout) à ces niveaux-là. Néanmoins, ces systèmes fonctionnent pour des domaines restreints dans lesquels définir toutes les règles est une tâche réalisable, mais le développement de ces règles sera toujours coûteux.

Un des avantages des systèmes fondés sur des corpus, que nous explorerons dans les sections suivantes, est le fait que toute la connaissance linguistique se trouve dans les données, ce qui élimine la difficulté liée au développement manuel des règles, mais qui implique aussi d'autres enjeux.

2.4.2 Les systèmes statistiques

Les méthodes statistiques reposent sur des corpus bilingues alignés et des corpus monolingues de la langue cible. De ce fait, l'information contenue dans ces corpus, qui sont des traductions et des textes produits par des humains, est utilisée pour traduire automatiquement des textes inédits pour le système (Hearne et Way 2011, 205). Cette approche établit des modèles de probabilité concernant la fidélité et la fluidité, puis combine les deux modèles pour choisir la traduction la plus probable (Jurafsky et James 2009, 893).

Il existe deux générations de systèmes statistiques dérivés de la question de l'alignement. D'une part, les systèmes de TA statistiques par mots ; ce type de système cherche à trouver les correspondances de chaque mot de la langue source dans la langue cible, mais pose le problème des correspondances « asymétriques » entre les mots de certaines combinaisons linguistiques dans lesquelles un mot de la langue source ou cible « correspond à n mots dans l'autre langue » (Poibeau 2019, 92). D'autre part, les systèmes par segments ou « phrases », dans lesquelles les unités de traduction peuvent être aussi des séquences de mots ; dans ce type de systèmes, la probabilité est donc calculée pour trouver le segment en langue cible qui a la plus grande possibilité d'être la traduction d'un groupe de mots en langue source (Jurafsky et James 2009, 895).

L'architecture de ces systèmes est constituée de deux processus différents : l'entraînement et le décodage. Dans la phase d'entraînement, les modèles statistiques de langue et de traduction sont extraits des corpus monolingue et parallèle, respectivement ; dans la phase de décodage s'effectue la recherche de la phrase en langue cible avec la plus grande probabilité d'être la traduction de la phrase en langue source, c'est-à-dire, s'effectue la traduction proprement dite (Hearne et Way 2011, 205). Dans la suite, nous allons décrire plus en détail ces deux modèles.

D'une part, le modèle de langue de la phase d'entraînement « gives the likelihood that a target string T is actually a valid sentence in the target language » (Hearne et Way 2011, 208). De ce fait, le modèle de langue détermine les fréquences relatives d'un segment dans le corpus monolingue et, sur la base de ces fréquences, assigne un score de fluidité à une nouvelle phrase. Celui-ci peut être calculé sur la base des séquences d'un mot (unigramme), de deux mots (bigramme), ou de trois mots (trigramme). Le plus souvent, les scores de ces trois types de séquences sont combinés, avec un poids statistique différent, pour déterminer une probabilité finale qui combine la flexibilité des unigrammes avec la sensibilité au contexte des séquences de deux ou trois mots (Hearne et Way 2011, 211).

Le but du modèle de traduction est d'assigner une probabilité pour déterminer la traduction la plus fidèle en langue cible d'une phrase donnée en langue source

(Jurafsky et James 2009, 895). Le point de départ est un corpus aligné dans lequel chaque phrase en langue cible est couplée avec une phrase en langue source.

Plusieurs méthodes d'alignement (par des mots et intersections) impliquent l'identification d'équivalences entre mots individuels parmi les phrases alignées dans le corpus parallèle (Hearne et Way 2011, 213). Ce processus peut produire des alignements de mots « 1-1 » ou « 1 – plusieurs », mais il n'est pas possible de générer des alignements du type « plusieurs – 1 » ou « plusieurs – plusieurs », et, en plus, les alignements se réalisent uniquement dans le sens langue cible – langue source (Ibid. 2011, 217), ce qui représente les deux contraintes principales de l'algorithme. En revanche, les alignements par segments permettent d'extraire les alignements du type « plusieurs – plusieurs » et réaliser des alignements dans les deux sens (source – cible et cible – source). Une fois l'alignement terminé, chaque phrase et la probabilité de chaque phrase peuvent être « stockées » dans un tableau de traductions (Jurafsky et James 2009, 909). Finalement, dans la dernière étape du processus, l'algorithme de décodage emploie ce tableau pour déterminer la meilleure solution de traduction du point de vue de la fluidité et de la fidélité.

En résumé, le système calcule les fréquences dans le corpus des mots et des segments (« phrases ») à traduire et « décode » les équivalents qui ont le meilleur score en fluidité et fidélité (Hutchins 2015, 128). Dans ce contexte, la notion de « décodage » est différente à celle proposée par Weaver lorsqu'il évoquait les principes de la cryptographie comme voie possible pour « décoder » le texte en langue étrangère. Pour les systèmes statistiques, il s'agit de dériver la traduction la plus fluide et la plus fidèle d'après les calculs produits à partir de ce qui se trouve dans le corpus qui alimente le système.

Ce type de système a conduit à des améliorations notables dans la qualité des traductions automatiques et une réduction des efforts investis par des linguistes computationnels lors de la formulation de règles dans les systèmes linguistiques. Mais la raison principale pour laquelle les systèmes statistiques sont devenus le paradigme dominant, comme le remarque Hutchins (2015), est économique, étant donné qu'il peut produire un système de TA fondé sur une base de données composée de mémoires de traduction. Cela implique, donc, la réutilisation du

travail des traducteurs humains pour améliorer la qualité des traductions automatiques.

Paradoxalement, le système est aussi limité au contenu de son corpus. C'est la raison pour laquelle les mots inconnus par les systèmes ne sont pas toujours traduits. Les mots inconnus sont aussi des limites des systèmes neuronaux, qui seront décrites dans la prochaine section.

Finalement, les besoins informatiques des systèmes statistiques au niveau de la vitesse et du stockage sont plus importants que ceux des systèmes linguistiques (Melby 2019, 429). Néanmoins, les demandes à ce niveau sont encore plus exigeantes dans le cas des systèmes neuronaux, qui demandent beaucoup plus de puissance informatique et des corpus d'entraînement encore plus importants.

2.4.3 Les systèmes neuronaux

L'approche computationnelle des systèmes neuronaux est tout-à-fait différente des systèmes statistiques. Le processus de traduction dans ces systèmes est en effet effectué par un réseau composé de centaines d'unités artificielles ou *neurones* (Koehn et Knowles 2017, 301). L'activation ou l'inhibition de ces unités dépendent du stimulus reçu par l'activation d'autres unités et de son intensité (Forcada 2017, 292). Il s'agit donc d'un système d'unités qui travaillent dans un réseau complexe, stimulé par un input en langue source, pour produire une traduction, ou output, en langue cible.

Dans un premier temps, l'information de la phrase source est traitée et codifiée par les réseaux neuronaux du système. Dans cette étape, le système se base sur les « plongements de mots » (Poibeau 2019, 124) pour encoder le sens de la phrase (plongement contextuel). Pour arriver à ce regroupement des mots, le système procède à une analyse des apparitions en contexte de chaque mot qui, en conséquence, est liée « à un ensemble plus large » de mots sémantiquement homogènes (Ibid.). L'espace contextuel de recherche se limite à une fenêtre d'un nombre de mots déterminé qui se trouvent à gauche et à droite du mot considéré.

Les co-occurrences résultant de cette analyse forment une matrice qui permet de déduire les occurrences les plus significatives et les plus rares. Le système ignore

celles qui sont inférieures à un seul déterminé et regroupe les contextes plus semblables et ainsi produire progressivement les plongements de mots. Le système est donc capable de « vectoriser » les résultats et d'assigner une valeur numérique aux mots qui correspond à son empreinte sémantique (Poibeau 2019, 125).

Dans ce sens, les systèmes neuronaux effectuent un encodage de la phrase source pour, progressivement, générer une représentation sémantique sur la base des plongements déjà appris par le système ; la phase de décodage correspond donc à la même opération progressive de représentation sémantique, mais basée sur les plongements contextuels du système (Forcada 2017, 297-98). C'est pourquoi, pour les systèmes de TA neuronaux, la traduction implique un processus d'encodage et décodage.

Il est intéressant de remarquer comment la notion de décodage a été présente depuis les débuts de la TA. Néanmoins, avec l'utilisation de l'intelligence artificielle dans les systèmes basés sur des corpus, la notion a pris une toute nouvelle dimension.

Une des questions qui se posent avec les systèmes neuronaux, et qui est pertinente pour cette étude, est celle de l'adaptation des systèmes aux domaines de spécialité (Koehn et Knowles 2017, 29). En effet, à cause de la dimension des corpus nécessaires pour entraîner ces systèmes, les systèmes commerciaux les plus populaires comme Google Translate ou DeepL sont des systèmes entraînés avec des données générales.

Cette question de la spécialisation des systèmes neuronaux révèle aussi une caractéristique importante à signaler : lorsqu'ils sont confrontés à des conditions qui diffèrent sensiblement des conditions d'entraînement, la réponse de ces systèmes n'est pas optimale (Koehn et Knowles 2017, 37). Cela doit être pris en compte pour la suite de cette recherche, notamment lors de l'analyse des résultats dans la première et la deuxième étape de notre expérience. En effet, il est tout-à-fait possible que les défis de traduction imposés par les traités de musique (qui seront explorés plus tard), puissent différer des conditions d'entraînement qui viennent d'être évoquées.

En résumé, le réseau artificiel de *neurones* de ces systèmes et ses connexions sont entraînés avec des corpus multilingues de grandes dimensions ; ces réseaux, une fois confrontés à un texte source, codifient leurs informations, effectuent des calculs et décodent ce qu'on présume être la traduction la plus probable du texte source. Le problème, remarqué par Melby (2019, 430), est que, comme dans le cas de la traduction effectuée par des humains, les systèmes neuronaux font aussi des erreurs de traduction.

La différence, d'après cet auteur, est que ces architectures n'offrent aucune méthode pour savoir pourquoi le système a fait telle ou telle erreur, ou pour corriger les erreurs sur la base d'une explication de pourquoi elles se sont produites. C'est le problème du « black-box » posé par l'architecture des systèmes neuronaux, qui implique l'impossibilité de « demander » au système pourquoi il a traduit un mot ou une phrase de la façon qu'il l'a fait (Melby 2019, 431). Ce qui signifie que le fonctionnement interne des systèmes reste obscur pour les chercheurs et les développeurs, tout comme pour les usagers de ces systèmes.

Dans ce sens, Forcada (2017) avance que même si les erreurs de traduction faites par les systèmes neuronaux sont d'ordre sémantique, la traduction produite par ces systèmes est souvent très fluide. Contrairement aux systèmes statistiques, les systèmes neuronaux traduisent les mots qu'ils ne connaissent pas ; le problème est qu'ils proposent une traduction des mots inconnus qui peut être complètement sans rapport avec le texte source (Forcada 2017, 305; Koehn et Knowles 2017, 30), et que, là encore, le *black-box* ne permet pas de trouver une explication ou de corriger ces erreurs.

Ce souci se révèle particulièrement important lors d'un processus de post-édition monolingue comme celui proposé dans cette étude, étant donné que les participants n'auront accès qu'aux textes cible. Leur expérience repose entièrement sur ce qui va être proposé par la machine. Et ce n'est pas un fait anodin de se questionner s'il est important que les systèmes comprennent ou pas ce qu'ils traduisent. Ceci n'est pas toujours la priorité pour les développeurs car ils s'intéressent uniquement à la qualité des traductions produites par les systèmes, selon Melby (2019, 432), et laissent de côté les questions philosophiques.

Reste à savoir dans quelle mesure les systèmes neuronaux offrent des résultats fiables pour les post-éditeurs. Il faut rappeler que, malgré les progrès évidents des dernières années avec ces systèmes, « claims that near-human performance is just around the corner are misleading and create unrealistic expectations » (Koehn 2021). Reste à prouver aussi s'ils peuvent être considérés comme des points de départ solides pour une pratique comme la post-édition monolingue. Dans le chapitre suivant, nous explorerons les différents types de post-édition et les recherches récentes effectuées dans le domaine de la post-édition, en lien avec les notions centrales de la problématique de cette recherche.

3. Post-édition et utilisabilité de la TA

3.1 Introduction

La post-édition est définie selon la norme ISO 18589 (2017) comme l'acte de « modifier et corriger un texte résultant d'une traduction automatique ». Dans l'industrie de la communication interlinguistique, cette intervention humaine devrait normalement être effectuée par un traducteur, ou, dans certains contextes, par des experts dans le domaine, en suivant des directives établies par le standard de la Translation Automation User Society (TAUS 2016) qui normalise les critères de qualité de cette pratique.

Néanmoins, et comme signalé par Hu et Cadwell (2016), le standard proposé par TAUS n'est pas le seul. Parfois, les entreprises établissent leurs propres directives et critères de qualité. Hu et Cadwell analysent également plusieurs propositions issues de la recherche (Ibid. 2016, 347-48). Dans cette section, nous suivrons le standard de TAUS pour expliquer les caractéristiques des différents types de post-édition.

Par la suite, nous définirons d'abord les différents types de post-édition (section 3.2.1), pour ensuite explorer les notions de post-édition monolingue (section 3.2.2) ainsi que du profil des post-éditeurs (section 3.2.3), tout en mettant l'accent sur les recherches récentes qui concernent ces deux notions. Nous nous consacrerons également à l'utilisabilité de la TA et des post-éditions dans la section 3.3.

3.2 La post-édition et ses formes

3.2.1 Types de post-édition

Le standard TAUS et la norme ISO 18587 distinguent deux types de post-édition : la post-édition « légère » (ou « superficielle ») et la post-édition « complète ».

Connue également comme post-édition « superficielle », la post-édition légère est définie par la norme ISO 18587 (2017) comme le processus de post-édition « permettant d'obtenir un produit qui soit simplement compréhensible, sans tenter de parvenir à un produit comparable à celui obtenu par une traduction humaine ». La priorité dans ce type de post-édition est de véhiculer l'information et le sens du

texte source, tandis que les aspects linguistiques et stylistiques sont considérés comme secondaires.

Pour sa part, la post-édition complète est définie par la norme ISO 18587 (2017) comme le processus de post-édition « permettant d'obtenir un produit comparable à un produit obtenu par une traduction humaine ». Dans ce type de post-édition, l'information, le sens du texte source et les aspects linguistiques et stylistiques sont importants (même si, pour ce dernier, la qualité humaine ne pourrait pas toujours être atteinte). Le Tableau 1 présente ci-dessous une synthèse des indications de TAUS pour chaque type de post-édition. Ce tableau utilise les catégories proposées par Hu et Cadwell (2016, 349-50).

Critère	Post-édition Légère	Post-édition Complète
Précision	La PE communique le sens du texte source	La PE communique le sens du texte source
Information	Entièrement transmise	Entièrement transmise
Terminologie	N/A	La terminologie clé est correcte
Grammaire	Peut ne pas être correcte	Correcte
Sémantique	Correcte	Correcte
Ponctuation	N/A	Correcte
Orthographe	Respecte les normes essentielles	Respecte les normes essentielles
Syntaxe	Peut être inhabituelle	Normale
Style	Non	Épuré
Restructuration des phrases	Non nécessaire si le sens est correct	N/A
Facteurs culturels	Éditer si nécessaire	Éditer si nécessaire
Format	N/A	Correct
Autres	Éviter les corrections autant que possible	Qualité de la traduction humaine

TABLEAU 1. STANDARD TAUS DE POST-ÉDITION

Pour la traduction professionnelle, ces directives assument un scénario de post-édition bilingue (TAUS 2016, 14). En effet, selon Nunes Vieira (2019), le non-recours au texte source est considéré comme un risque dans des contextes commerciaux, raison pour laquelle la post-édition bilingue est la norme. Il existe

cependant de plus en plus d'intérêt pour la post-édition monolingue dans d'autres contextes (*gisting*, etc.), comme nous allons le voir par la suite.

3.2.2 La post-édition monolingue

En partant des définitions que la norme ISO 18587 (2017) propose pour les post-éditions légère et complète, la post-édition monolingue pourrait être définie comme le processus de post-édition qui, sans recourir au texte source, permet d'obtenir soit un produit compréhensible, soit un produit de qualité comparable à un produit obtenu par une traduction humaine.

Plusieurs auteurs se sont intéressés à la notion de la post-édition monolingue. Dans cette section, nous allons explorer cinq études (Mitchell, Roturier, et O'Brien 2014 ; Koehn 2010 ; Koponen et Salmi 2015 ; C. Hu et al. 2012 ; Schwartz et al. 2014) qui présentent un intérêt méthodologique et conceptuel, et proposent diverses approches du sujet.

Mitchell, Roturier, et O'Brien (2014) s'intéressent à l'évaluation de la post-édition bilingue en contraste avec la post-édition monolingue et, plus spécifiquement, à la production participative (« crowdsourcing ») de post-éditions dans le contexte d'un forum d'assistance en ligne dédié aux usagers d'un produit informatique. Les combinaisons linguistiques explorées sont l'anglais, langue source, et le français et l'allemand, langues cible. Les huit participants de cette étude sont des usagers de ce forum, spécialistes du domaine, quatre d'entre eux de langue maternelle française et quatre de langue maternelle allemande.

Les textes à post-éditer sont composés des trois éléments du forum : (a) des questions posées par les usagers ; (b) le sujet dans lequel s'inscrivent ces questions ; (c) et la réponse donnée par la communauté du forum. La post-édition a été effectuée dans un environnement spécifiquement développée pour effectuer ce type de tâches. Les participants dans cette étude ont complété les deux tâches de post-édition, bilingue et monolingue. L'évaluation s'est focalisée sur les critères de fluidité, compréhensibilité et fidélité à travers une échelle Likert.

Les résultats de cette recherche montrent que la post-édition monolingue a été légèrement mieux évaluée, en termes de fluidité et de compréhensibilité, que la

post-édition bilingue. Néanmoins, les scores de fidélité sont similaires entre les deux types de post-édition, et sont les mêmes en ce qui concerne la perception de la qualité de la compréhensibilité qui, dans les deux cas, a été évaluée négativement. D'autre part, la post-édition bilingue a enregistré une meilleure évaluation en termes de fidélité que la post-édition monolingue. D'après les résultats de cette étude, la post-édition monolingue n'est pas un exercice irréaliste dans ce type de contexte.

La recherche de Koehn (2010) travaille avec des langues source (le chinois et l'arabe) qui sont indéchiffrables pour un natif de la langue cible (l'anglais). Il s'agit d'un des critères qui ont été pris en compte pour le choix de la combinaison linguistique explorée dans ce mémoire, russe > espagnol.

La tâche des participants dans l'expérience de Koehn (2010), traducteurs en formation, consistait à post-éditer des textes journalistiques. L'objectif était de déterminer dans quelle mesure la qualité des systèmes de l'époque permettait aux post-éditeurs de s'appuyer sur la TA brute pour recréer le texte à travers une post-édition monolingue. Bien-sûr, le système statistique de l'époque était moins performant que les systèmes neuronaux d'aujourd'hui. Koehn a structuré une évaluation humaine de la qualité des post-éditions avec des traductions de référence. Les résultats de ces évaluations indiquent que les post-éditions ont été jugées correctes dans un tiers des cas, notamment pour la combinaison arabe > anglais. Il note aussi que la connaissance des post-éditeurs dans le domaine du texte est un facteur clé pour ce type de post-édition, même si parfois cela n'est pas suffisant pour produire un texte de qualité.

Une autre recherche qui se révèle intéressante est celle de Koponen et Salmi (2015). Ces deux chercheuses ont proposé de faire post-éditer des TA brutes de textes journalistiques en anglais traduits vers le finlandais et d'évaluer dans quelle mesure les participants parvenaient à « reconstruire » le sens de l'original sans avoir recours à d'autres ressources.

Les résultats de la recherche de Koponen et Salmi indiquent que des erreurs comme les omissions, les mots mal traduits et les expressions idiomatiques peuvent rendre opaque le sens du texte source ou même le déformer complètement

en rendant impossible de reconstruire le sens des textes source uniquement à partir de la TA brute. D'autres erreurs moins importantes, par exemple d'ordre morphologique, n'ont pas empêché les post-éditeurs de déduire le sens du passage concerné grâce au contexte.

La recherche menée par Koponen et Salmi pose des questions intéressantes autour de la possibilité de reconstruire un texte dans une langue cible uniquement à partir de la TA brute. Elle permet également d'envisager d'utiliser la post-édition monolingue pour mesurer la qualité des traductions produites par les systèmes de TA (voir aussi Doherty et O'Brien 2014). Les chercheuses ont utilisé un système statistique et un système linguistique, ce qui contraste avec la recherche de ce mémoire qui se basera sur des TA brutes produites par un système neuronal.

D'autres recherches se sont concentrées sur l'évaluation de systèmes statistiques conçus pour des scénarios où seule la post-édition monolingue est possible. C'est le cas de Hu et al. (2012) et Schwartz et al. (2014). La première, évalue le système *MonoTrans2* dans un scénario de traduction et retraduction entre participants monolingues dans la combinaison haïtien-créole > anglais. La deuxième évalue les résultats de la post-édition monolingue sur la TA brute d'un système entraîné dans les combinaisons linguistiques hindi > anglais et russe > anglais.

Ce qui est intéressant dans les deux recherches est le contexte d'utilisation dans lequel est intégré le processus de post-édition monolingue. Hu et al. (2012) ont eu l'intention de développer un flux de travail autour du système *MonoTrans2* pour aider à franchir la barrière linguistique dans des situations d'urgence, comme dans le cas du tremblement de terre qui a affecté Haïti en 2010. Schwartz et al. (2014) appartiennent au « Air Force Research Laboratory » des États-Unis, donc leur recherche est sensée servir à des fins de communication interlinguistique dans des contextes en relation avec les forces armées.

En résumé, les recherches explorées dans cette section ont abordé la post-édition monolingue dans une approche comparative avec la post-édition bilingue (Mitchell, Roturier, et O'Brien 2014), ou pour évaluer la possibilité de la TA. Cette dernière approche peut avoir comme objectif soit de mesurer la qualité de la TA brute (tout en l'identifiant comme un facteur déterminant pour la qualité de la post-édition

monolingue) (Koehn 2010), soit de comparer la performance des systèmes de TA avec des architectures différentes (Koponen et Salmi 2015). Par ailleurs, la post-édition monolingue a aussi été explorée comme un moyen pour franchir les barrières linguistiques imposées par des combinaisons linguistiques rares (et par les différences entre les systèmes alphabétiques de certaines langues) dans des situations d'urgence (C. Hu et al. 2012) ainsi que dans un contexte militaire (Schwartz et al. 2014).

3.2.3 Profil des post-éditeurs

Plusieurs recherches en post-édition s'interrogent sur l'incidence du profil des post-éditeurs sur la qualité finale des produits. Bien que moins nombreuses, ces recherches ont un point commun : elles visent à comparer la performance des traducteurs professionnels avec celle des spécialistes ou traducteurs non-professionnels, moins expérimentés ou en formation.

C'est le cas de l'étude de Nitzke (2016), qui consistait à proposer des tâches de post-édition bilingue et post-édition monolingue à deux groupes de participants : un groupe de traducteurs professionnels et un autre de traducteurs décrits comme « semi-professionnels ». Certes, il ne s'agit pas d'un contraste entre spécialistes et non spécialistes comme celui proposé par Temizöz (2013), mais la recherche de Nitzke reste très pertinente pour ce mémoire. Son objectif est de déterminer dans quelle mesure le comportement et le processus de post-édition monolingue change dans les deux cas.

Le chercheur a employé plusieurs méthodes pour recueillir les données, telles que l'évaluation de la qualité du produit final, le « eye-tracking », l'enregistrement d'écran et un questionnaire à la fin de la tâche pour mesurer les opinions et attitudes des participants. Les textes, en anglais, étaient des articles de presse et de sociologie, et ont été traduits vers l'allemand par un système de TA. Sa recherche a démontré que la qualité des produits de la post-édition monolingue n'est pas suffisamment acceptable pour que les textes soient publiés. Ceci s'explique par les fautes sémantiques qui, d'après le chercheur, étaient induites par la TA brute et par l'intervention des traducteurs. En comparaison à la post-édition bilingue ou à la traduction humaine, la post-édition monolingue présentait

beaucoup plus ce type de fautes, malgré le fait que la qualité linguistique était la même ou supérieure à celle des produits des deux autres modalités.

Les participants ont aussi exprimé davantage de frustration et d'insatisfaction face à la tâche de post-édition monolingue. Le chercheur indique dans sa conclusion que cela est sûrement dû au fait que les traducteurs ne sont pas entraînés pour donner du sens à un texte avec les caractéristiques de la TA brute sans avoir accès au texte en langue source. Quant aux profils des post-éditeurs, les professionnels se sont toujours montrés plus frustrés et critiques envers leur propre travail que les semi-professionnels.

Il s'agit donc d'un travail qui explore l'attitude des post-éditeurs face aux défis et risques de la post-édition monolingue, ce qui est aussi l'objet de la recherche développée dans ce mémoire et qui sera également observée à travers les questionnaires fournis aux post-éditeurs lors de l'expérience.

Une autre recherche s'est intéressée au profil des post-éditeurs dans le domaine de la post-édition monolingue, cette fois par rapport à l'incidence des protocoles de verbalisation, effectués par les participants, sur l'effort de post-édition. Dans son expérience, Barbosa de Lima Fonseca (2019) a donné la tâche de post-édition à trois groupes de participants : des étudiants en traduction, des traducteurs semi-professionnels et des traducteurs professionnels. L'objet central de sa recherche était de déterminer l'impact des protocoles de verbalisation sur l'effort temporel, cognitif et technique investi par les trois groupes de participants dans la tâche de post-édition des articles de vulgarisation scientifique de l'anglais vers le portugais brésilien.

D'après ses résultats, les protocoles de verbalisation n'ont pas été affectés par le profil. Une des hypothèses de sa recherche était que la verbalisation pourrait demander plus d'efforts de la part des étudiants qui ont moins d'expérience. Toutefois, ce n'était pas le cas, car l'effort des trois groupes a été similaire avec l'inclusion des protocoles de verbalisation dans le processus de post-édition monolingue. Il s'agit d'une recherche très riche et intéressante, notamment du point de vue de sa diversité méthodologique.

En ce qui concerne la post-édition bilingue, l'étude de Rigouts Terryn et al. (2019) met l'accent sur le profil des post-éditeurs dans le domaine médical. L'objectif de cette étude est de déterminer la stratégie optimale pour organiser le flux de travail dans la traduction du néerlandais vers le français de textes destinés au grand public. Ce type de textes cherche à faciliter l'accès des patients au contenu technique et scientifique concernant des informations médicales.

La méthodologie décrite dans cette recherche consiste dans un premier temps dans une post-édition suivie d'une étape de révision. Tous les participants étaient francophones natifs, professionnels de la médecine ou de la traduction. L'objectif était de comparer deux flux de traduction différents. Dans le premier, les traductions ont été effectuées par des médecins et les révisions par des traducteurs (« DT ») ; dans le second, les traducteurs ont effectué la traduction et les médecins la révision (« TD »).

Cette approche post-édition/révision a été adoptée pour assurer la qualité des textes étant donné que l'exercice proposé aux participants consistait à produire des post-éditions publiables. Dans une étape préalable, les erreurs de la TA brute ont été annotées et une traduction de « référence » a été produite par un spécialiste bilingue, ce qui a aidé à évaluer les post-éditions produites lors de l'expérience et à interpréter les résultats.

Les résultats de cette recherche démontrent que la différence entre les versions issues des deux flux de travail ne sont pas significativement différentes. Néanmoins, le premier flux de travail (« DT ») a conduit à moins d'erreurs que l'autre. Le flux de travail avec des spécialistes traducteurs et des traducteurs réviseurs semble donc légèrement plus prometteur. Il est toutefois intéressant de relever que les traducteurs ont détecté plus d'erreurs dans l'étape de révision que les experts, mais ont également effectué plus de corrections jugées non nécessaires.

La recherche de Rigouts Terryn et al. (2019) est aussi très pertinente du point de vue méthodologique pour notre recherche et s'intéresse au contraste entre le profil des spécialistes et traducteurs pour effectuer la tâche de post-édition. La procédure adoptée par les chercheurs consiste à annoter des erreurs dans la TA brute et

effectuer une comparaison entre les annotations et les post-éditions. Cette procédure a été intégrée dans notre recherche.

L'impact du profil des post-éditeurs sur l'utilisabilité des post-éditions a également intéressé d'autres recherches dans le domaine. Temizöz (2013) à l'Université Rovira I Virgili en Espagne étudie les différences entre les processus de post-édition effectués par des traducteurs professionnels et ceux effectués par des experts dans le domaine de spécialité concerné (l'ingénierie, dans ce cas). Le chercheur a intégré dans son étude des ingénieurs et des traducteurs pour déterminer les flux de travail post-édition/révision les plus efficaces en termes de temps et de qualité, ce qui rappelle la recherche de Rigouts Terryn et al. (2019). Il a employé des méthodes comme l'enregistrement d'écran et des questionnaires.

Les résultats de son étude montrent que l'expérience et les connaissances dans le domaine de spécialité ont une incidence positive très importante sur la qualité des post-éditions. Mais, contrairement à ce qui a été démontré par la recherche de Rigouts Terryn et al., le flux de travail qui a obtenu de meilleurs résultats est celui des traducteurs qui traduisent et des spécialistes qui révisent.

En résumé, les recherches explorées dans cette section s'intéressent au profil des post-éditeurs et à son incidence dans la qualité des post-éditions, tant dans des scénarios monolingues (Nitzke 2016 ; Barbosa de Lima Fonseca 2019) que bilingues (Rigouts Terryn et al. 2019 ; Temizöz 2013). Il est intéressant de remarquer que les recherches dédiées à la post-édition monolingue s'intéressent au contraste entre post-éditeurs professionnels et semi-professionnels, tandis que les recherches en post-édition bilingue s'intéressent à différencier post-éditeurs et spécialistes. Dans notre recherche en post-édition monolingue, tous cherchent à contraster les post-éditions des experts avec celles des post-éditeurs linguistes, ce qui implique de combiner les deux approches observées ici.

3.3 Utilisabilité de la TA et des post-éditions

Une autre notion importante est celle de l'utilisabilité des post-éditions, un concept central pour la problématique de cette recherche. La norme ISO 9241-11 (2018) définit l'utilisabilité comme « le degré selon lequel un produit peut être utilisé par

des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié ». Dans cette étude, le « produit » correspond aux post-éditions monolingues qui seront effectuées lors de l'expérience, et leur « contexte d'utilisation » est défini par le caractère pédagogique de la typologie textuelle annoncée dans l'introduction : le traité de musique (section 4.2).

Plusieurs chercheurs se sont récemment penchés sur des questions liées à l'utilisabilité de la TA ou des post-éditions (Forcada et al. 2018 ; Doherty et O'Brien 2014 ; Van Egdom et Pluymaekers 2019 ; Castilho et al. 2014 ; Castilho et O'Brien 2016 ; Screen 2019). Dans cette section, nous allons présenter d'abord une recherche autour de l'utilisabilité de la TA brute, puis deux études sur l'utilisabilité de la post-édition monolingue et, finalement, trois autres recherches sur l'utilisabilité des post-éditions bilingues (légères et complètes).

Tout d'abord, Forcada et al. (2018) se sont intéressés aux méthodes pour évaluer l'utilisabilité de la TA brute. Dans ce cas, l'utilisabilité a été mesurée par rapport à la notion de *gisting*, c'est-à-dire, à la capacité de la TA brute de transmettre en langue cible l'essentiel de l'information du texte source. Cette étude cherchait à comparer deux méthodes d'évaluation de l'utilisabilité : les questionnaires de compréhension et le *gap filling*. Les chercheurs proposaient cette deuxième méthode comme une alternative moins coûteuse, qui consiste à effacer des segments d'une traduction de référence pour évaluer dans quelle mesure les participants sont capables de compléter les segments manquants à l'aide des TA brutes ou d'autres traductions.

Dans leur expérience, les chercheurs ont préparé des documents de référence en anglais qui avaient été traduits de l'allemand par des systèmes de TA ou par un traducteur humain. Ils ont conçu différents types de questions et ont appliqué d'autres variables pour déterminer, par exemple, l'impact de la longueur du segment effacé ou du choix de segment à effacer sur la performance des participants. Les résultats de cette recherche ont démontré que les deux méthodes sont susceptibles de déterminer l'utilisabilité de la TA brute, notamment car les deux scores coïncident substantiellement. En conséquence, les deux méthodes

d'évaluation sont considérées comme valides pour évaluer l'utilisabilité d'une traduction (ou d'une post-édition) dans ce type de contexte.

Du point de vue méthodologique, cette étude présente un intérêt pour notre recherche étant donné que les participants n'ont pas eu accès au texte source en allemand. Lors de notre expérience, les participants ont effectué une évaluation monolingue des traductions, ce qui sera aussi le cas pour les juges des post-éditions dans notre étude.

Dans le contexte de la post-édition monolingue, la notion d'utilisabilité a aussi été explorée par Doherty et O'Brien (2014), mais cette étude s'intéresse uniquement à l'évaluation de l'utilisabilité des TA brutes. En effet, le but de la recherche était de déterminer l'impact de la TA sur l'utilisabilité d'un document qui présentait des instructions pour manipuler un produit en ligne. Le texte source, en anglais, a été traduit vers le japonais, l'espagnol, l'allemand et le français avec un système commercial statistique.

Pour évaluer l'utilisabilité des TA brutes, les chercheurs ont employé des techniques de « eye-tracking » pendant une expérience contrôlée dans laquelle les participants, natifs des langues concernées, devaient effectuer les tâches décrites dans le document. La mesure du temps pendant l'exécution de la tâche, ainsi qu'un questionnaire de satisfaction visant à évaluer la qualité des TA brutes, ont aussi fait partie de la méthode employée dans cette recherche.

Du point de vue des résultats, les TA brutes ont eu des scores d'utilisabilité plus bas que le document original. Néanmoins, le degré de satisfaction des participants semble être associé à la compréhensibilité de la TA brute, ce qui varie dans chaque langue. L'espagnol, par exemple, a eu des scores plutôt favorables dans ce sens, tandis que le japonais a eu les scores les plus bas des quatre langues. Dans tous les cas, les TA brutes ont révélé un effort cognitif supérieur que le texte source.

La recherche de Doherty et O'Brien (2014) explore la notion d'utilisabilité de la TA dans un contexte d'évaluation monolingue. Du point de vue méthodologique, la recherche a recours à des questionnaires pour mesurer la satisfaction des usagers par rapport à la TA brute, ce qui sera également le cas dans ce mémoire.

Dans la même veine, la recherche de Van Egdome et Pluymaekers (2019) évalue aussi la notion d'utilisabilité, mais cette fois celle-ci est axée sur les divers degrés de post-édition. Dans ce cas, tout comme dans la recherche précédente (Doherty et O'Brien 2014), l'intérêt pour la perception de l'utilisateur final implique une évaluation monolingue de l'utilisabilité des textes présentés aux participants.

La tâche des participants était de post-éditer deux textes, un texte informatif sur un sujet informatique adressé au public en général, l'autre, plus spécialisé, qui portait des instructions pour manipuler un logiciel et était destiné à des informaticiens. Les deux textes ont été traduits de l'anglais vers le néerlandais. Les consignes données aux post-éditeurs décrivaient les différents niveaux de post-édition qui devaient être appliqués pour traduire les textes.

Dans une deuxième étape, les textes ont été évalués par des jurés natifs de la langue cible avec la méthode du jugement intuitif en utilisant une échelle d'utilisabilité prédéfinie. Les participants étaient face à des textes avec des degrés de post-édition variable, choisis au hasard, dans la langue cible et sans le texte source en anglais. Les résultats de cette recherche démontrent que, en effet, le degré de post-édition influence la perception de la qualité des textes, ainsi que celle des produits concernés. Néanmoins, et contrairement à ce qui était attendu, la post-édition complète n'est pas toujours la mieux évaluée par les usagers, ce qui suggère que, dans ce type de contexte, la capacité du texte à transmettre l'information est plus appréciée que sa qualité linguistique.

Tout comme pour la recherche de Doherty et O'Brien (2014), la recherche de Van Egdome et Pluymaekers (2019) s'intéresse à l'utilisabilité perçue par les utilisateurs des textes en question. Les deux travaux s'interrogent sur l'évaluation de l'utilisabilité, tant de la TA brute que des post-éditions. Dans ce mémoire, les deux axes seront adoptés : d'une part, l'évaluation de la perception de la TA brute à travers un questionnaire dédié aux post-éditeurs monolingues ; de l'autre, des évaluations destinées aux utilisateurs potentiels des post-éditions monolingues produites.

En ce qui concerne l'évaluation de l'utilisabilité dans un scénario de post-édition bilingue, plusieurs recherches ont évalué l'utilisabilité des post-éditions bilingues légères (Castilho et al. 2014 ; Castilho et O'Brien 2016) et complètes (Screen 2019).

Dans leur recherche, Castilho et al. (2014) s'intéressent à l'utilisabilité des post-éditions légères en comparaison avec celle de la TA brute. Pour déterminer dans quelle mesure la post-édition légère a un impact sur l'utilisabilité des instructions, les chercheurs utilisent un texte en anglais qui présentait une série d'instructions concernant un produit informatique. Ce texte a été traduit en portugais brésilien par Microsoft Bing. Les chercheurs ont divisé les participants en deux groupes égaux. Leur tâche consistait à lire les instructions et à effectuer les tâches décrites dans les textes. Le premier groupe travaillait avec la TA brute et le deuxième avec la post-édition.

Les critères d'utilisabilité pris en compte dans cette étude sont « effectiveness, efficiency and satisfaction », repris de la norme ISO/TR 16982 de 2002 (Castilho et al. 2014, 184). Pour mesurer l'utilisabilité, les chercheurs ont utilisé trois méthodes principales : enregistrement d'écran, « eye-tracking » et un questionnaire donné aux participants après avoir effectué la tâche. Les résultats de cette recherche montrent que, même si les deux groupes ont effectué correctement les tâches, un meilleur niveau de satisfaction et d'efficacité est noté chez les personnes qui ont lu les instructions post-éditées.

Ces résultats confirment que la post-édition, même légère, apporte une valorisation du contenu traduit par un système de TA. Les mesures d'utilisabilité proposées dans cette recherche ont été appliquées par les usagers dans un contexte d'utilisation spécifique. Cette procédure peut être utilisée pour explorer d'autres combinaisons linguistiques, et d'autres types de contenus, dans des situations différentes. C'est le cas de la recherche développée dans ce mémoire qui s'intéresse à la combinaison linguistique russe > espagnol dans le contexte d'utilisation des traités de musique. L'étude de Castilho et al. (2014) a cependant été effectuée avec un système statistique, tandis que l'expérience qui sera menée ici utilisera des systèmes neuronaux.

Castilho et O'Brien (2016) comparent l'utilisabilité de la TA brute et de la post-édition légère avec celle du texte source. Pour cette étude, elles ont choisi des articles d'aide destinés aux usagers d'une application informatique. La TA vers l'allemand a été produite avec Microsoft Bing depuis l'anglais.

Pour leur expérience, les participants ont été divisés en trois groupes. Un groupe d'anglophones natifs a eu accès au texte source, tandis que les deux autres groupes, germanophones, ont respectivement eu accès à la TA brute et à la post-édition. Les mesures d'utilisabilité de l'étude précédente ont été adoptées ici, ainsi que la méthode de « eye-tracking » et du questionnaire à la fin de la tâche.

Comme dans l'expérience menée précédemment par les chercheuses, les participants ont reçu l'instruction de lire les textes et d'effectuer les tâches décrites. Les résultats ont montré que, du point de vue de l'effort cognitif des participants, il n'y a pas eu une différence significative entre ceux qui ont travaillé avec la post-édition et la TA brute. En revanche, une différence a été constatée entre les participants qui ont effectué les tâches avec seulement les post-éditions et ceux possédant le texte source. De plus, les participants avec la post-édition ont manifesté plus de satisfaction et ont effectué les tâches plus rapidement que le groupe avec la TA brute.

Dans cette étude, les chercheuses ont changé de combinaison linguistique, mais maintenu l'essentiel de leur démarche précédente (voir Castilho et al. 2014). Ici, une nouvelle perspective est offerte en comparant les résultats du groupe du texte source, avec les groupes des textes en langue cible. Ainsi, il y a trois degrés d'utilisabilité : (1) le texte source, (2) la post-édition légère et (3) la TA brute.

Finalement, l'article de Screen (2019) présente une recherche en post-édition complète qui s'intéresse à la notion d'utilisabilité et à la manière dont les usagers perçoivent les post-éditions. Son travail consiste à comparer la lisibilité et la compréhensibilité de deux textes traduits vers le gallois, l'un produit de la traduction humaine, et l'autre de la post-édition. Le texte source, en anglais, provient d'une institution publique. Cette étude avait pour but de déterminer dans quelle mesure le processus de post-édition a un impact sur la qualité perçue par les lecteurs du produit final.

Pour ce faire, Screen a employé la technique du « eye-tracking », ainsi qu'un questionnaire de satisfaction avec une échelle de type Likert. Les deux méthodes ont été appliquées à deux groupes de participants. Tous étaient locuteurs natifs du gallois. Le texte source a été post-édité et traduit par des professionnels. Le premier groupe a lu le texte traduit, et le second celui post-édité. Les consignes données aux post-éditeurs correspondaient aux caractéristiques de la post-édition complète (section 1.1.2.1) étant donné que la comparaison portait notamment sur la lisibilité des textes, ce qui inclut les aspects stylistiques et linguistiques.

Les résultats de cette étude ont montré que (a) le niveau de fixation capté par le système de « eye-tracking » était identique pour les deux groupes, et (b) qu'il n'y a pas de différences importantes au niveau de la lisibilité et compréhensibilité de la traduction et la post-édition. Ces résultats ont amené le chercheur à suggérer une utilisation plus massive de la post-édition complète pour la traduction des documents du secteur public vers le gallois.

La pertinence de la recherche de Screen (2019) pour notre étude s'explique par trois facteurs : (a) le focus sur les usagers des post-éditions ; (b) le fait de ne pas réduire la post-édition à un simple vecteur d'information ; et (c) l'exploration d'une combinaison linguistique relativement peu explorée, comme celle de l'anglais vers le gallois. Ces éléments sont également en relation avec la recherche développée dans ce mémoire.

En résumé, la recherche autour de la notion d'utilisabilité peut être consacrée à la TA brute (Forcada et al. 2018), et à la post-édition dans un scénario monolingue (Doherty et O'Brien 2014 ; Van Egdom et Pluymaekers 2019) ou bilingue (Castilho et al. 2014 ; Castilho et O'Brien 2016 ; Screen 2019). Ces travaux constituent l'état de l'art dans le domaine de l'utilisabilité.

Les travaux cités tout au long des sections 1.1.2 et 1.1.3 constituent un échantillon représentatif des recherches qui explorent des notions comme le profil des post-éditeurs (Rigouts Terryn et al. 2019; Temizöz 2013; Nitzke 2016; Barbosa de Lima Fonseca 2019), la post-édition monolingue (Mitchell, Roturier, et O'Brien 2014 ; Koehn 2010 ; Koponen et Salmi 2015 ; C. Hu et al. 2012 ; Schwartz et al. 2014) et l'utilisabilité (Forcada et al. 2018 ; Doherty et O'Brien 2014 ; Van Egdom et

Pluymaekers 2019 ; Castilho et al. 2014 ; Castilho et O'Brien 2016 ; Screen 2019). Ces trois notions sont fondamentales pour la problématique de cette recherche.

Du point de vue méthodologique, une grande majorité de ces études se fonde sur des questionnaires pour mesurer la qualité ou l'utilisabilité des produits issus, principalement, de la post-édition, à l'exception d'un travail qui mesure l'utilisabilité de la TA brute dans un contexte monolingue (Doherty et O'Brien 2014) et d'un travail consacré à la validité du *gap filling* pour évaluer aussi l'utilisabilité de la TA brute (Forcada et al. 2018). D'ailleurs, des questionnaires ont été utilisés dans la recherche de Doherty et O'Brien (2014) pour évaluer l'utilisabilité des post-éditions et pour avoir un aperçu de l'expérience de post-édition des participants, ce qui sera employé dans notre recherche. D'autre part, l'annotation d'erreurs dans la TA effectuée par un annotateur bilingue a été repéré uniquement dans un cas (Rigouts Terryn et al. 2019) ; cette méthode sera également adoptée dans cette recherche.

Une autre caractéristique est que pratiquement toutes les recherches utilisent des systèmes de TA statistiques, à l'exception de deux travaux en post-édition bilingue qui, comme cette recherche, utilisent des systèmes neuronaux (Rigouts Terryn et al. 2019; Screen 2019). Du point de vue des langues explorées dans cette recherche, seule une étude n'utilise pas l'anglais comme langue cible ou source (Rigouts Terryn et al. 2019), tandis que le russe et l'espagnol se trouvent uniquement dans Schwartz et al. 2014 et Doherty et O'Brien 2014.

De ce fait, les méthodes employées dans cette recherche (questionnaires, tâches de post-édition monolingue et annotation d'erreurs dans la TA brute), font partie de l'ensemble des recherches citées. L'Annexe 1 résume les recherches mentionnées tout au long des sections 3.2 et 3.3.

Aucune de ces recherches ne s'est occupée des textes de spécialité en relation directe ou indirecte avec la musique, ce qui constitue l'originalité de notre travail. Le caractère de cette typologie textuelle, ainsi que la possibilité d'appliquer la post-édition monolingue à ce type de textes, sera le point principal développé dans le chapitre suivant de ce mémoire.

4. La post-édition monolingue appliquée aux traités de musique

4.1 Introduction

La TA peut être conçue comme un outil de « décodage » dans le sens qui vient d'être exposé dans les chapitres précédents. La post-édition monolingue représente quant à elle la possibilité de rendre utilisable une TA dans le contexte spécifié par les traités de musique. Dans ce chapitre, nous explorerons les caractéristiques principales des traités de musique (section 4.2), ainsi que les enjeux liés à la pratique de la post-édition monolingue appliquée à cette typologie textuelle (section 4.3).

4.2 La musique et ses textes : les traités de musique

Dans ce mémoire, nous définissons les traités de musique comme un ouvrage à caractère fondamentalement didactique dans lequel est exposé un sujet ou un ensemble de sujets concernant la musique. Pour mener l'expérience décrite dans le chapitre 5, nous avons choisi deux textes en langue russe qui répondent à cette définition, l'un décrivant une méthode de piano pour les débutants de l'instrument et l'autre un traité d'orchestration adressé aux étudiants avancés ou à des compositeurs.

Dans cette section, nous analyserons d'abord les caractéristiques générales de ce type de textes (section 4.2.1) pour ensuite étudier plus en détail les principaux enjeux traductologiques dérivés de ces caractéristiques (section 4.2.2).

4.2.1 Caractéristiques propres aux traités de musique

Les traités de musique possèdent trois caractéristiques principales : (a) une fonction pédagogique, qui détermine la situation communicative ; (b) un discours hybride composé de traits du langage littéraire et du langage technique, qui détermine ses caractéristiques linguistiques ; et, (c) le recours à l'écriture musicale comme moyen d'illustration ou d'exemplification, ce qui détermine leur caractère transtextuel.

a. Situation communicative :

Pour analyser comment la fonction pédagogique de ce type de textes détermine la situation communicative, nous allons recourir aux catégories proposées par Ciapuscio et Kuguel (2002). Ces catégories ont pour but de classer les textes selon leur degré de spécialité à travers quatre niveaux ou dimensions des textes. Néanmoins, nous nous intéressons ici uniquement à deux niveaux : fonctionnel et situationnel.

Au niveau des fonctions textuelles, les traités de musique ont deux fonctions principales, informative (qui sert à décrire ou définir les phénomènes musicaux) et instructive ou didactique (qui sert à guider des procédures d'écriture ou exécution de la musique), ainsi qu'une fonction expressive, qui est secondaire. La fonction informative est la plus importante étant donné que le traité de musique est une exposition systématique d'un sujet ou d'un ensemble de sujets. Les auteurs ont en effet pour but de transmettre aux lecteurs des connaissances concernant le domaine de la musique. La fonction instructive ou didactique est aussi importante étant donné que, dans certains cas, il s'agit de textes qui indiquent comment jouer d'un instrument, ou comment lire ou écrire de la musique. La fonction expressive est secondaire puisque, dans ce type de textes, il est possible de trouver des appréciations esthétiques ou des opinions des auteurs sur un sujet ou une façon de procéder. Néanmoins, la fonction expressive repose sur le cadre établi par ces deux fonctions principales.

Au niveau de la situation communicative, et tout en considérant les fonctions textuelles, les traités de musique établissent une relation pédagogique entre l'auteur et ses potentiels lecteurs. De ce fait, la relation entre les interlocuteurs est celle établie entre un spécialiste et un apprenti, une relation fondamentalement asymétrique du point de vue des connaissances. Cette relation asymétrique caractérise ce type de textes.

b. Langue littéraire et technique

Les traités de musique présentent à la fois des traits du langage littéraire et du langage technique. Cette particularité s'explique en partie pour les fonctions qui

viennent d'être exposées : la fonction informative et la directive sont mieux exprimées en langage technique, tandis que la fonction expressive s'associe plutôt au langage littéraire. Mais, si on va plus loin dans l'analyse, ce discours hybride peut être expliqué par la nature de ce type de textes, étant donné que les traités exposent des sujets concernant la musique. Ceci peut expliquer l'hybridation du discours car, d'un côté, la musique est une expression artistique et, de l'autre, un domaine avec divers degrés de spécialisation qui vont de la vulgarisation à la haute technicité.

Pour exemplifier cette double présence du style littéraire et technique dans cette typologie textuelle, le Tableau 2 présente deux extraits de traités en français et en anglais. Le texte marqué en bleu signale des termes qui évoquent des réalités très concrètes et spécifiques, ce qui est propre au langage technique. Le texte marqué en vert illustre des aspects du langage littéraire tels que des comparaisons, des références culturelles et des personnifications.

Exemple 1

[...] Alors, on se servira des **parties intermédiaires** pour écrire une **harmonie aussi plaine, aussi logique, aussi musicale que possible**. C'est parfois assez malaisé. Il serait encore beaucoup plus difficile d'écrire, combiné avec **le canon** de la **Basse et du Soprano**, *un canon* entre le **Ténor** et le **Contralto** ; **ce sont là des tours de force propres à J.S Bach**. Alors, la difficulté semble exciter sa puissance créatrice, exalter son imagination harmonique : d'où résultent des morceaux admirables... Mais nous ne proposons pas un tel but à l'élève.

Source : Charles Koechlin, *Traité de l'harmonie Vol. II*, p. 36.

Exemple 2

[...] When **measure 414** arrives, the **tempo suddenly gets even faster**, the **meter** changes to **duple time**, and the **octave leap, which was part of the previous motive, suddenly cascades through several repetitions** into a **cadence**, in **measure 416**. But this **cadence** itself is quite brusque, and the **bassoons immediately enter with new, contrasting material**, which will form part of the **Trio section**. The **whole-note D also turns out to be the start** of a **new theme**.

Source : Alan Belkin, *Music Composition: Craft and Art*, p. 169

TABLEAU 1. EXEMPLES DU STYLE HYBRIDE TECHNIQUE-LITTÉRAIRE

c. Notation musicale et transtextualité

Les traités de musique utilisent l'écriture musicale comme moyen d'illustration ou d'exemplification de leurs propos. En réalité, cette caractéristique permet effectivement au contenu des traités et aux sujets exposés d'être plus concrets pour

les musiciens, mais limite la compréhension des personnes qui ne sont pas familiarisées avec la notation musicale car beaucoup d'informations restent codifiées et le texte, la plupart de temps, fait référence à l'exemple musical choisi par l'auteur (voir Illustration 4).

Thickening the pizzicato texture in the accompaniment going into measures 4–5, by having the bass leap back down into the lower register, creates a gentle accent.⁵

Figure 20.9. Exercise 3: a model solution

ILLUSTRATION 4
SOURCE : ALAN BELKIN, MUSIC COMPOSITION: CRAFT AND ART, P. 209

Le fait que le texte du traité de musique fasse référence à la notation musicale, sous forme de description ou de commentaire, établit une relation transtextuelle entre le traité de musique et la musique proprement dite ; Gérard Genette l'appelle métatextualité (Genette 1982, 4). C'est-à-dire que le traité de musique est le métatexte de la musique ou de l'entité musicale qu'il commente, explique, analyse, cite ou évoque.

En conclusion, il est raisonnable d'affirmer que, pour arriver à une compréhension et une appréciation ce type de textes, il est nécessaire de tenir compte de ses fonctions, du caractère hybride de son langage (qui oscille entre le style technique et le littéraire), et de sa relation métatextuelle avec la musique. Ces considérations nous amènent donc à la question de la traduction spécialisée de ce type de textes.

4.2.2 Traduction spécialisée des traités de musique

La traduction spécialisée des traités de musique est un sujet très peu exploré dans la littérature traductologique et musicologique (Benoit-Otis 2013). Au 25.05.2021, la recherche du mot « music » dans la base de données spécialisée « Translation

Studies Bibliography» donne un total de 310 entrées bibliographiques. Une analyse attentive des entrées montre que l'article de Vella Ramírez et Muñoz Muñoz (2010) est le seul consacré au sujet. La plupart des autres documents traitent de sujets comme la traduction des paroles de chansons ou des livrets d'opéra et offrent des réflexions autour de notions théoriques comme la traduction intersémiotique. Dans le domaine de la musicologie, la question sur la traduction d'œuvres vocales ou de traités est mentionnée comme un possible sujet de recherche (Leßmann 2020), mais celles-ci restent rares.

En conséquence, pour caractériser la traduction spécialisée de notre typologie textuelle, nous reviendrons sur les caractéristiques des traités exposées dans la section précédente pour les mettre en relation avec la notion de traduction spécialisée proposée par Scarpa (2010).

En premier lieu, Scarpa indique que la théorie fonctionnaliste du *skopos* reflète un des aspects essentiels de la traduction spécialisée :

« À la recherche d'équivalences entre le texte de départ et le texte d'arrivée succède l'accent mis sur le texte d'arrivée et, en traduction spécialisée, sur la fonction de dissémination du savoir spécialisée, à l'intérieur d'une nouvelle situation socioculturelle » (Scarpa 2010, 104).

De ce fait, si on considère le caractère pédagogique de notre typologie textuelle, on peut en déduire que la fonction principale de la traduction spécialisée de traités de musique est la dissémination du savoir musical.

En deuxième lieu, Scarpa assume une distinction entre la traduction littéraire et la traduction spécialisée :

« S'il [est] vrai qu'il est impossible de distinguer nettement la traduction littéraire de la traduction spécialisée, parce qu'il existe un continuum entre ces deux sous-catégories de la traduction et parce que la gamme de procédés de traduction à la disposition du traducteur technoscientifique et du traducteur littéraire est la même, il reste qu'on n'aborde pas la traduction d'un article d'astrophysique, ou encore d'ethnolinguistique, de la façon qu'on

aborderait la traduction d'un roman de Dickens ou d'un poème de John Donne » (Scarpa 2010, 95).

En effet, pour Scarpa, la traduction littéraire s'occupe des textes ouverts à l'interprétation des lecteurs et des traducteurs, tandis que les textes spécialisés ont tendance à être fermés de ce point de vue (Scarpa 2010, 96). De ce fait, si on considère que le langage employé dans les traités de musique présente des traits du langage littéraire et du langage technique, on peut déduire que la traduction de cette typologie textuelle est caractérisée par une approche mixte entre l'ouverture interprétative de la traduction littéraire et la précision technique de la traduction spécialisée.

Finalement, Scarpa avance qu'il existe une dimension culturelle inévitablement liée à la traduction spécialisée :

« Font partie du contexte socioculturel de l'activité traductive les variables sociales liées à la situation et aux participants à l'acte de communication... qui influencent inévitablement la traduction d'un texte et qui entrent à leur tour dans l'ensemble des liens imposées par le contexte. [...] Cette vision du texte comme structure multidimensionnelle à l'intérieur d'une situation essentiellement extralinguistique... s'inscrit à son tour dans un contexte socioculturel » (Scarpa 2010, 109).

Rappelons ici que les traités de musique auxquels s'intéresse notre travail sont des produits culturels issus de la traduction musicale occidentale, et qu'une de leurs caractéristiques est le recours à l'écriture musicale comme moyen d'illustration ou d'exemplification. Cette notation musicale est en effet l'expression écrite de la musique développée et employée en occident. Le traité de musique est donc un texte multidimensionnel lié à une situation extralinguistique, la musique en soi, qui existe en dehors du traité. Ainsi, la traduction spécialisée de traités de musique est aussi inévitablement liée à cette tradition musicale.

En résumé, la traduction spécialisée de traités de musique se caractérise par sa fonction de dissémination du savoir musical, par l'approche mixte entre la traduction littéraire et la traduction spécialisée (exigée par les caractéristiques du

langage employé dans les traités) et, finalement, par une dimension culturelle extralinguistique qui est à l'origine même de la typologie textuelle : la musique.

Dans ce mémoire, nous nous intéressons à la possibilité de rendre utilisables des post-éditions monolingues de traités de musique à l'aide de la TA. Les trois caractéristiques de la traduction spécialisée de ce type de textes qui viennent d'être explorées dans cette section seront sans doute un facteur déterminant pour les résultats de notre étude. Dans la section suivante (section 4.3), nous allons explorer la possibilité d'intégrer la technologie comme outil d'aide à la traduction de notre typologie textuelle, ainsi que les notions d'utilisabilité et du profil de post-éditeur.

4.3 Décoder les textes de spécialité à l'aide des systèmes neuronaux

Compte tenu des caractéristiques des traités de musique et de la traduction spécialisée de cette typologie textuelle, il est maintenant pertinent de s'interroger sur le possible rôle de la TA comme outil d'aide à la traduction dans ce contexte. Pour ce faire, la section 4.3.1 propose une analyse du rôle de la TA comme outil d'aide à la traduction des traités de musique, plus particulièrement dans le contexte établi par la combinaison linguistique choisie dans cette recherche, ainsi que par les caractéristiques des traités de musique. Nous allons explorer aussi les difficultés dérivées de ces caractéristiques concernant l'application de la TA à notre type de texte. Les sections 4.3.2 et 4.3.3 traiteront respectivement des notions d'utilisabilité et du profil des post-éditeurs dans ce contexte.

4.3.1 La TA comme outil d'aide à la traduction spécialisée des traités de musique

La notion du *décodage* présentée tout au long de l'histoire de la TA (section 2.3) peut être conçue comme un outil d'aide à la traduction de ce type de textes si on considère la combinaison linguistique choisie dans cette étude. Dans ce sens, il convient de rappeler la justification donnée par Koehn (2010) sur les langues impliquées dans sa recherche sur la post-édition monolingue :

The... motivation for choosing Arabic–English and Chinese–English is that they are undecipherably foreign for a typical European or American speaker of English. The fact that both languages are written in a different script already makes it impossible to spot cognates, except for the occasional number. In our study, the test subjects had to practically exclusively rely on the given sentence translations or phrase translations options. (Koehn 2010, 540).

Dans le cas de notre recherche, la combinaison linguistique russe > espagnol permet d'adopter la même perspective. Par conséquent, il nous est possible de concevoir et d'intégrer la TA comme outil de décodage d'un message en langue source écrit avec un système alphabétique différent de celui de la langue cible. En effet, il est intéressant de rappeler ici que le principe du « décodage » à l'origine des systèmes de TA répondait aussi à une situation similaire du point de vue linguistique, notamment avec la combinaison russe > anglais.

Par ailleurs, la combinaison linguistique russe > espagnol présente un intérêt du point de vue de la spécialité. La Russie a joué un rôle central dans le développement de la musique occidentale depuis le XVIIIème siècle (Frolova-Walker et al. 2001). De ce fait, dans ce domaine de spécialité, la production artistique, pédagogique et intellectuelle en langue russe est vaste et a une valeur culturelle incalculable. Néanmoins, sauf de notables exceptions, les traductions de ce type de textes sont très souvent effectuées vers l'anglais, ce qui, dans ce contexte, fait de notre combinaison linguistique une rareté (Salamanca 2018).

Pour analyser la possibilité d'employer la TA comme outil d'aide à la traduction de notre typologie textuelle, il est nécessaire de mettre en avant une des caractéristiques des traités de musique que nous avons développée dans la section 4.2.1 : la double présence du langage spécialisé et du langage littéraire.

En ce qui concerne le langage spécialisé, la TA et la pratique de la post-édition ont été employées comme outils d'aide à la traduction de domaines de spécialité divers comme l'économie (Läubli et al. 2019) et le droit (Wiesmann 2019). Dans le cas de l'étude de Läubi et al. (2019), l'utilisation de la TA combinée avec la post-édition a impliqué une réduction substantielle du temps et une amélioration légère de la

qualité des traductions ; dans le cas de Wiesmann (2019), la qualité de la TA évaluée ne permet pas l'utilisation des traductions automatiques à des fins de post-édition d'après les critères établis par la chercheuse. Dans les deux études, les chercheurs ont employé un système de traduction neuronal.

En ce qui concerne le langage littéraire, la TA et la pratique de la post-édition ont également été employées pour traduire de la prose (Toral, Wieling, et Way 2018) et même de la poésie (Lee 2015). Dans la première étude, qui combine la TA avec la post-édition, les chercheurs ont constaté une amélioration de la productivité mais suggèrent aussi un effort de post-édition considérable. Pour sa part, Lee (2015) avance qu'il serait possible d'intégrer les technologies de la traduction au processus de traduction de la poésie si on adopte une approche différente de la manière dont le transfert de sens est accompli en traduction littéraire. On observe donc que la littérature et le langage littéraire ont aussi été objet de réflexions dans le domaine des technologies de la traduction.

De ce fait, les traités de musique nous mettent dans une position privilégiée pour explorer ces deux voies. De ce point de vue, ce qui est inédit dans notre étude est le fait d'appliquer la TA et, plus spécifiquement, la post-édition monolingue, à un type de texte qui intègre les deux types de langage, spécialisé et littéraire. Par la suite, nous explorerons la notion d'utilisabilité comme un concept central de notre étude qui dialogue à la fois avec les caractéristiques des traités (section 4.2.1) et avec leur traduction spécialisée (section 4.2.2), ce qui complète le panorama ouvert par l'utilisation des technologies de la traduction pour traduire notre typologie textuelle.

4.3.2 L'utilisabilité comme critère d'évaluation de la post-édition monolingue

L'expérience qui sera menée dans ce mémoire vise à déterminer dans quelle mesure il est possible de produire des post-éditions monolingues utilisables de traités de musique. C'est-à-dire, la notion d'utilisabilité, telle qu'elle est définie par la norme ISO 9241-11 de 2019 (section 3.1), est la notion centrale de notre étude. La définition donnée par la norme ISO de l'utilisabilité met en avant deux facteurs d'importance pour nous : le contexte d'utilisation et les utilisateurs. Ces deux

facteurs de la définition seront appliqués à notre typologie textuelle et, de fait, aux post-éditions monolingues qui seront produites.

Tout d'abord, le contexte d'utilisation est dérivé d'une des caractéristiques principales des traités de musique, à savoir leur fonction pédagogique (section 4.2.1). Cela veut dire qu'un traité de musique est susceptible d'être utilisable dans tout contexte défini par l'enseignement ou l'apprentissage de la musique. En conséquence, les post-éditions monolingues produites par les participants de notre étude seront évaluées par rapport à leur utilisabilité potentielle dans un contexte pédagogique.

L'évaluation de l'utilisabilité des post-éditions nous mène au profil des utilisateurs. En réalité, l'utilisateur potentiel d'un traité de musique est déterminé par le degré de spécialisation du texte. C'est-à-dire, un texte très technique et complexe sera utilisable pour des spécialistes (musiciens professionnels), tandis qu'un texte pédagogique d'initiation à la pratique instrumentale ou à la lecture de partitions sera utilisable pour une ample gamme d'utilisateurs allant des professeurs-spécialistes aux débutants dans le domaine.

C'est la raison pour laquelle, dans ce mémoire, nous nous intéressons également à la question du degré de spécialité des textes, ainsi qu'à l'incidence de ce degré de spécialité sur l'utilisabilité des post-éditions et la performance des post-éditeurs. Nous avons donc choisi deux textes : un texte avec un haut degré de spécialité et un texte d'initiation aux notions fondamentales de la lecture musicale (section 5.2.3). De ce fait, les utilisateurs qui pourront trouver utilisables les post-éditions monolingues de notre étude, indépendamment de leur degré de spécialité, sont des musiciens professionnels. Ils seront donc les juges pour déterminer dans quelle mesure les post-éditions produites par nos participants pourront être considérées comme utilisables dans un contexte d'enseignement et d'apprentissage de la musique.

Ainsi, l'utilisabilité des post-éditions monolingues sera jugée par des spécialistes dans le domaine. Le degré de spécialité des textes est la première variable que nous allons considérer pour interpréter les résultats de ces évaluations. La deuxième

variable qui nous intéresse de prendre en compte dans cette étude est l'objet de la section suivante : le profil du post-éditeur.

4.3.3 Rôle du post-éditeur monolingue : spécialiste ou traducteur ?

La question du profil du traducteur spécialisé a été traitée, entre autres, par Scarpa (2010). Pour l'auteure, l'adoption d'une stratégie de traduction passe d'abord par un acte de lecture qui a pour but de diagnostiquer les problèmes posés par le texte spécialisé, ces problèmes pouvant être d'ordre pragmatique ou culturel (Scarpa 2010, 147-50). En post-édition monolingue, ce diagnostic des problèmes devra se faire sans avoir recours au texte source, ce qui rend la tâche plus délicate. C'est sur ce principe que, dans l'introduction, nous avons formulé l'hypothèse selon laquelle, dans la pratique de la post-édition monolingue, les experts verront plus de fautes de fidélité que les traducteurs puisqu'ils sont plus à même d'identifier les problèmes.

Ceci explique également pourquoi, selon Koehn (2010), « a good monolingual translator has good language skills in the target language and understands the domain ». En effet, le diagnostic de problèmes mentionné par Scarpa pourrait être problématique si le post-éditeur n'a pas les connaissances pour aller au-delà des erreurs de langue, ou si, tout en comprenant le sujet du texte, il n'arrive pas à corriger les fautes linguistiques. Comme nous l'avons vu, ceci risque d'être encore plus problématique avec les systèmes neuronaux étant donné que leurs traductions ont tendance à être plus « idiomatiques ».

Dans ce mémoire, nous adoptons le principe proposé par Bonet Heras (2004). Selon ce principe, plus le texte à traduire appartient à un des domaines qui bénéficient d'un langage très contrôlé, plus il est nécessaire que le « traducteur technique » évolue vers la composante technique (Bonet Heras 2004, 41). Autrement dit, dans le cas de notre étude, plus le degré de spécialité du texte est haut, plus il devient nécessaire que le post-éditeur monolingue soit un spécialiste. C'est sur ce principe que nous avons également formulé l'hypothèse selon laquelle plus le post-éditeur est expert dans le domaine, plus il est possible d'améliorer l'utilisabilité de la traduction automatique brute d'un texte de spécialité.

En résumé, le profil du post-éditeur est une variable centrale pour notre étude dans la mesure où la possibilité d'effectuer une post-édition monolingue d'un traité de musique peut être déterminée par le fait d'être ou non un spécialiste dans le domaine. Celle du degré de spécialité des textes à post-éditer (section 5.2.3) nous intéresse également car la performance des post-éditeurs est également susceptible de varier en fonction des difficultés posées par les textes. Finalement, la littérature démontre des études sur la possibilité d'employer des technologies de la traduction et la post-édition comme aides à la traduction des textes à caractère spécialisé, tout comme des textes à caractère littéraire. De ce point de vue, les traités de musique nous permettent d'étudier ces deux variables dans un seul type de texte qui réunit des caractéristiques des deux langages, littéraire et spécialisé.

4.4 Conclusion

Tout au long des chapitres 2, 3 et 4, nous avons d'abord abordé l'état de la question avec un bref aperçu historique de la traduction automatique et la littérature la littérature au sujet de la post-édition et ses formes, ainsi qu'au sujet de l'utilisabilité dans ce contexte. Nous avons ensuite approfondi sur l'application de la post-édition monolingue à notre typologie textuelle en explorant les caractéristiques des traités de musique et de la traduction spécialisée de cette typologie textuelle. Nous avons finalement exposé quelques applications antérieures de la TA et de la post-édition à des textes avec les mêmes caractéristiques que les traités de musique ; en ce sens, nous avons signalé l'importance de la notion d'utilisabilité et du profil du post-éditeur pour notre recherche et pour interpréter les résultats dérivés de l'expérience qui sera menée.

Dans la suite, nous présenterons l'objectif, les questions de recherche et les hypothèses de notre étude (section 5.1). Nous allons justifier aussi les choix méthodologiques (section 5.2) concernant les formats, métriques, textes et profil des participants de notre expérience. Pour finir, nous allons décrire en détail le déroulé de notre expérience (section 5.3).

5. Méthodologie

5.1 Objectif, questions de recherche et hypothèses

Dans ce chapitre, nous présentons la méthode de recherche appliquée lors de notre expérience. Nous nous pencherons sur la justification et l'explication des choix qui ont déterminé notre méthode de recherche.

L'expérience vise à répondre aux deux questions principales suivantes :

- *dans quelle mesure l'expertise du post-éditeur monolingue dans le domaine de spécialité est un facteur déterminant pour l'utilisabilité des post-éditions de traités de musique ?*
- *quel est l'impact du degré de spécialisation du texte sur l'utilisabilité des post-éditions produites par les participants de notre étude ?*

Par ailleurs, nous voulons voir l'impact des systèmes de TA utilisés et répondre à deux questions subsidiaires : (a) lequel des deux systèmes de TA, Google Translate ou DeepL, fait le moins d'erreurs d'après les annotations de l'annotatrice bilingue ? ; (b) lequel des deux systèmes est d'avantage choisi par les post-éditeurs pour effectuer la tâche qui leur est présentée ?

Nous prenons comme point de départ les deux hypothèses suivantes :

- Plus le post-éditeur est expert dans le domaine, plus il peut rendre utilisable la traduction automatique brute d'un texte de spécialité.
- Dans la pratique de la post-édition monolingue, les experts voient plus de fautes de fidélité que les traducteurs.

Pour infirmer ou confirmer ces hypothèses, ainsi que pour répondre à notre question de recherche, nous avons organisé une expérience en cinq phases :

1. Préparation et traitement des textes
2. Annotation des erreurs de la TA brute par un expert bilingue
3. Post-édition monolingue et questionnaire post-tâche

4. Évaluation des fautes corrigées par la post-édition sur la base de l'annotation
5. Évaluation de l'utilisabilité des post-éditions par des juges monolingues

Dans ce sens, la phase 4 permet de prouver la deuxième hypothèse, tandis que la phase 5 permet de prouver la première.

Dans la suite de ce chapitre, nous allons décrire nos choix méthodologiques (section 5.2), ainsi que la préparation et le déroulement des cinq phases de notre expérience (section 5.3).

5.2 Choix méthodologiques

Dans cette section, nous allons exposer les choix méthodologiques qui nous ont guidé tout au long de l'expérience. Pour ce faire, nous allons présenter le plan expérimental (section 5.2.1), ainsi que les choix que nous avons privilégiés dans notre expérience par rapport aux textes (section 5.2.2), aux outils employés pour la post-édition, l'évaluation et les questionnaires (section 5.2.3), aux métriques qualitatives (section 5.2.4), et aux participants (section 5.2.5).

5.2.1 Plan expérimental

Dans notre expérience, nous avons deux groupes de participants : les post-éditeurs (section 5.2.5.1) et les évaluateurs (section 5.2.5.2). Le groupe des post-éditeurs est divisé en experts et linguistes, tandis que le groupe d'évaluateurs est homogène, composé d'experts dans le domaine de la musique. Tous les participants des deux groupes sont des locuteurs natifs de l'espagnol.

Les participants ont post-édité deux textes : un texte non spécialisé (section 5.2.2.1) et un texte hautement spécialisé (section 5.2.2.2). Nous avons demandé aux évaluateurs de comparer les post-éditions monolingues faites par les experts avec celles faites par les linguistes. Le jugement des évaluateurs nous permettra de déterminer leur perception sur l'utilisabilité des post-éditions faites par les deux groupes de post-éditeurs.

De ce fait, les variables de l'expérience sont : (a) le degré de spécialité des textes (non spécialisé et hautement spécialisé), et (b) le profil du post-éditeur (musicien

ou linguiste). Dans la suite de ce chapitre, nous allons décrire ces deux variables, ainsi que les aspects méthodologiques et les phases de l'expérience.

5.2.2 Choix des textes

5.2.2.1 *Le texte non spécialisé (Texte 1)*

Le premier texte est la méthode *The Russian School of Piano Playing* (« Школа игры на фортепиано » en russe), un texte d'initiation à la pratique du piano écrit par Alexander Nikolayev, pianiste et pédagogue russe (Nikolayev 2004). Ce texte se caractérise par l'emploi d'un langage clair dans ses explications et par une progression très méthodique qui part des rudiments du langage musical, et va jusqu'à des notions avancées d'interprétation et de lecture des partitions. La méthode s'adresse principalement aux enfants mais peut être utilisée par toute personne souhaitant s'initier à la pratique instrumentale et à la lecture de la musique.

Nous avons choisi un fragment de cinq segments (226 mots au total) du premier volume de la méthode qui est un bon exemple de texte peu spécialisé et équivaut au niveau de vulgarisation des textes scientifiques. Celui-ci explique les notions les plus fondamentales de l'écriture musicale et il s'adresse à tout public. De ce fait, l'utilisation des symboles ou de la terminologie spécialisée est toujours progressive et bien expliquée par l'auteur, et aucune connaissance de base n'est exigée aux lecteurs de ce texte.

Cette méthode, très populaire dans les écoles de musique russophones à l'époque de l'Union Soviétique, a été depuis traduite vers l'allemand, l'anglais et l'espagnol. Encore aujourd'hui, elle est largement utilisée pour initier les enfants à la pratique de l'instrument. Elle a été réimprimée à plusieurs reprises, mais nous avons choisi l'édition la plus récente en langue russe (Nikolayev 2004) pour notre étude.

5.2.2.2 *Le texte hautement spécialisé (Texte 2)*

Le deuxième texte, les Principes d'orchestration (« Основы оркестровки » en russe) a été écrit par le compositeur russe Nikolai Rimski-Korsakov (Rimsky-Korsakov 2019). Il s'agit d'un texte dans une discipline de la musique qui étudie la façon d'écrire pour divers instruments et familles d'instruments de musique en prenant en compte leurs particularités et limites. Cette discipline étudie également

comment combiner les différents instruments et familles d'instruments lorsqu'on écrit de la musique pour des ensembles instrumentaux ou l'orchestre. De par la nature de la discipline, ce texte s'adresse aux spécialistes dans le domaine ou aux étudiants avancés de composition musicale. Son langage se caractérise par l'utilisation de termes et technicisms, ainsi que des symboles et des exemples tirés de la littérature musicale, étant donné que l'auteur assume une connaissance très avancée de la part de ses potentiels lecteurs.

Nous avons choisi un fragment de cinq segments (226 mots au total) du premier volume de ce traité d'orchestration. Ce fragment explique comment combiner les sons des instruments, provenant de deux familles d'instruments de l'orchestre, pour réussir une écriture techniquement correcte et musicalement expressive. L'auteur illustre ses propos avec des exemples musicaux qui ont un haut niveau de technicité. Le texte dialogue avec ces illustrations en créant un réseau textuel complexe et hautement spécialisé.

Ce texte a été l'un des traités les plus importants dans ce sujet. D'abord, parce que son auteur est une figure d'autorité dans l'histoire de la musique. Mais également par la qualité des explications, par sa capacité d'illustration des propos avec ses propres œuvres et par la vaste connaissance de l'auteur dans le domaine de spécialité. Ce traité a été traduit en plusieurs langues, dont l'anglais, le français, l'espagnol et l'allemand. Il est encore aujourd'hui réimprimé et consulté par les étudiants de composition et de musique du monde entier. Comme dans le cas du premier texte, nous avons choisi l'édition la plus récente en langue russe pour notre expérience (Rimsky-Korsakov 2019).

Le Tableau 3 montre le nombre de mots par segment des deux textes source choisis pour notre expérience.

Segment	Texte 1	Texte 2
	Nombre de mots	Nombre de mots
1	70	63
2	23	43
3	59	43
4	47	41
5	27	16
Total	226 mots	206 mots

TABLEAU 3. NOMBRE DE MOTS PAR TEXTE ET PAR SEGMENT

5.2.3 Choix des participants

5.2.3.1 Post-éditeurs

Le profil des post-éditeurs est une des variables de notre question de recherche. Pour mener notre expérience, nous avons recruté un total de 6 post-éditeurs qui ont participé à la troisième phase de notre étude. Tous les post-éditeurs sont locuteurs natifs de l'espagnol provenant de 6 pays hispanophones différents : Colombie, Espagne, Mexique, Bolivie, Equateur et Pérou. Ce groupe de 6 post-éditeurs a été divisé en deux groupes de 3 post-éditeurs selon leur domaine de spécialité.

Le premier groupe de post-éditeurs est composé de 3 musiciens professionnels sans aucune formation particulière dans le domaine de la traduction ou de la post-édition et avec les compétences linguistiques d'un locuteur natif avec une éducation universitaire. Ce groupe est composé d'un directeur d'orchestre, d'un violoniste et d'un docteur en musicologie. De ce fait, ce premier groupe de post-éditeurs possède une connaissance approfondie dans le domaine de spécialité.

Le deuxième groupe de post-éditeurs est composé de 3 linguistes, traducteurs professionnels ou en formation. Ce groupe a une connaissance approfondie de la langue espagnole, ainsi que de l'expérience en traduction, documentation et post-édition. Ils n'ont aucune formation musicale au niveau professionnel ou semi-professionnel, et seulement une des participantes de ce groupe a déclaré avoir une formation de base dans l'interprétation d'un instrument de musique et dans le déchiffrement des partitions.

Nous avons ainsi deux groupes de post-éditeurs : un premier groupe composé de musiciens, spécialistes dans le domaine, sans formation linguistique particulière (« Groupe A », en vert) , et un deuxième groupe composé de linguistes sans connaissances spécifiques dans le domaine de spécialité (« Groupe B », en bleu).

Groupes de post-éditeurs	
« Groupe A » : Musiciens	« Groupe B » : Linguistes
Post-éditeur A	Post-éditeur D
Post-éditeur B	Post-éditeur E
Post-éditeur C	Post-éditeur F

TABLEAU 4. GROUPES DE POST-EDITEURS ET IDENTIFIANT CORRESPONDANT A CHAQUE POST-EDITEUR

Les post-éditions produites par les participants ont été utilisées pour élaborer l'évaluation dans la quatrième phase de notre étude. Cette évaluation oppose, à chaque fois, les post-éditions produites par les deux groupes de participants avec l'objectif de déterminer leur degré d'utilisabilité.

5.2.3.2 *Évaluateurs de l'utilisabilité*

Le profil des évaluateurs qui ont participé à la quatrième phase de notre étude répond à notre objectif de déterminer l'utilisabilité des post-éditions. De ce fait, nous avons recruté 21 juges, spécialistes dans le domaine, c'est-à-dire, des musiciens professionnels. Tous les évaluateurs des post-éditions étaient des hispanophones natifs et avaient de l'expérience dans l'enseignement de la musique.

Cette dernière caractéristique était très importante car les traités de musique sont des textes fondamentalement didactiques, à caractère pédagogique. Par conséquent, les juges, en tant que musiciens pédagogues et hispanophones natifs, étaient en mesure de juger l'utilisabilité des textes destinés à l'enseignement et l'apprentissage de la musique car ils pouvaient juger tant la qualité linguistique, que la clarté et la précision des contenus.

5.2.4 Choix des outils de post-édition, d'évaluation et de questionnaire

La post-édition a été réalisée avec des formulaires d'annotation et de post-édition au format *.docx* de Microsoft Word. Ce choix est justifié pour trois raisons principales.

Tout d'abord, nos participants résidaient dans des pays différents (Suisse, France, Espagne, États-Unis, Mexique et Colombie). De ce fait, la seule option était de structurer un processus de collecte de données entièrement en ligne. Ceci était également imposé par les mesures de distanciation physique liées au contexte pandémique d'actualité au moment de l'expérience.

Ensuite, notre typologie textuelle se caractérise par la présence de notation musicale (sous forme d'exemples musicaux) et par l'inclusion dans le texte de signes dérivés de la notation musicale ; ces formes de notation « non-alphabétique » peuvent être traitées comme des images dans Microsoft Word à l'aide d'outils technologiques externes d'édition musicale (comme MuseScore) et d'édition d'image (comme GIMP), puis insérées dans les documents en format *.docx*.

Finalement, Microsoft Word est un outil informatique intuitif, amplement utilisé et avec des fonctionnalités multiples qui peuvent faciliter l'exécution des tâches. Ce logiciel permet aussi d'avoir un contrôle sur les espaces de texte éditables par les participants, ce qui nous a permis, dans la première phase de notre expérience (préparation et traitement des textes) d'établir un même formulaire de participation pour tous les post-éditeurs, mais qui pouvait être édité comme n'importe quel autre document de texte. À cet égard, le formulaire d'annotation employé lors de la deuxième phase était totalement protégé contre écriture, raison pour laquelle les annotations ont été effectuées avec la fonction « commentaire » par l'annotatrice.

En ce qui concerne le questionnaire post-tâche envoyé aux post-éditeurs, nous avons utilisé la plateforme Limesurvey. Ceci nous a permis d'automatiser et de centraliser la collecte des réponses des participants. Étant donné qu'il s'agissait de questionnaires anonymes, nous avons créé deux questionnaires identiques pour les deux groupes de participants (section 4.2.4.1).

Pour la quatrième phase de notre expérience (évaluation par les juges, section 5.3.4), le questionnaire a été également réalisé avec Limesurvey. Tout en considérant les caractéristiques de notre typologie textuelle (la présence de notation musicale et de symboles musicaux, principalement) nous avons décidé de composer les questionnaires d'évaluation avec des images des post-éditions, issues de captures d'écran, pour préserver la complexité de la typologie textuelle et être fidèles aux conditions de post-édition que nous avons proposées aux participants lors de la phase précédente, comme le montre l'illustration 5 :

Laquelle des deux versions convient le mieux à l'utilisation de ces textes dans des contextes d'enseignement et d'apprentissage de la musique ?

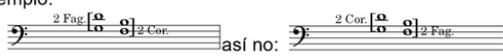
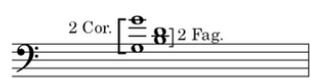
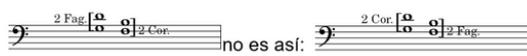
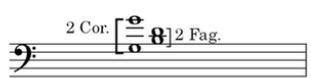
LAMBDA	SIGMA
<p>La combinación de dos trompas y dos fagotes en un acorde de cuatro partes generalmente se organiza en forma cruzada, para una mayor mezcla de timbres.</p> <p>* En este caso, es mejor asignar intervalos consonantes a las trompas francesas e intervalos disonantes a los fagot, por ejemplo:</p>  <p>La técnica del entorno también es aplicable, y los cuernos franceses rodean los fagotes, y no al revés:</p> 	<p>La combinación de dos trompas y dos fagotes en un acorde a cuatro voces suele disponerse de forma cruzada, para una mayor mezcla de tonos.</p> <p>* Al hacerlo, los intervalos consonantes se asignan mejor a las trompas y los disonantes a los fagotes, por ejemplo:</p>  <p>También se aplica la técnica del cerco, en la que las trompas rodean a los fagotes, y no al revés:</p> 
<p>ⓘ Veuillez sélectionner une réponse ci-dessous</p> <p><input type="radio"/> LAMBDA est beaucoup plus adapté à l'utilisation que SIGMA.</p> <p><input type="radio"/> LAMBDA est légèrement plus adapté à l'utilisation que SIGMA.</p> <p><input type="radio"/> Aucune des deux options n'est utilisable</p> <p><input type="radio"/> SIGMA est légèrement plus adapté à l'utilisation que LAMBDA.</p> <p><input type="radio"/> SIGMA est beaucoup plus adapté à l'utilisation que LAMBDA.</p>	

ILLUSTRATION 5
EXEMPLE DE QUESTION SUR LIMESURVEY

Il faut noter que ces caractéristiques vont au-delà des possibilités d'édition de textes dans la plateforme Limesurvey. Ces images ont été éditées et organisées pour assurer la lisibilité et faciliter la comparaison entre les versions. Nous avons répliqué l'ambiance de post-édition présentée aux participants de la troisième étape étant donné qu'elle a été conçue pour faciliter la comparaison de deux versions différentes d'un même texte.

5.2.5 Choix des métriques qualitatives

5.2.5.1 Erreurs de post-édition (Phase 4)

Pour évaluer les corrections effectuées dans la post-édition, nous avons adapté l'échelle d'évaluation des post-éditions employée par Schwartz et al. (2014) :

Taux de correction	Valeur qualitative
100 %	La totalité des erreurs a été corrigée. La traduction post-éditée est supérieure à la TA.
80 – 99 %	La plupart des erreurs ont été corrigées. Le sens est pleinement exprimé dans la traduction post-éditée.
50 – 79 %	L'essentiel du sens est transmis grâce aux corrections appliquées à la post-édition.
21 – 49 %	La traduction post-éditée déforme le sens de manière importante ou présente beaucoup de petites erreurs.
1 – 20 %	La plupart des erreurs n'ont pas été corrigées. Très peu de sens est véhiculé.
0 %	Aucune des erreurs de la TA n'ont été corrigées. La traduction post-éditée n'a aucun sens.

TABLEAU 5. ÉCHELLE D'ÉVALUATION DES POST-ÉDITIONS MONOLINGUES

Pour cette étape, nous nous servons des annotations produites dans la deuxième phase pour déterminer quelles erreurs de la TA, de celles qui ont été marqués par l'annotatrice, ont été effectivement corrigées par les post-éditeurs. Nous allons donc déterminer un taux de correction par rapport au pourcentage d'erreurs effectivement corrigées par les post-éditeurs. Cette évaluation nous permettra de confirmer ou d'infirmer notre deuxième hypothèse, c'est-à-dire, démontrer si les post-éditeurs experts voient davantage de fautes de fidélité que les post-éditeurs linguistes.

5.2.5.2 Évaluation de l'utilisabilité par les juges (Phase 5)

Pour effectuer l'évaluation de l'utilisabilité lors de la cinquième phase de notre étude, nous avons opté pour une évaluation comparative entre les post-éditions monolingues produites par les experts et les post-éditions monolingues produites par les linguistes. Nous avons construit une échelle Likert, de cinq degrés, adaptée à nos besoins. Pour chaque question, les experts seront mis en situation d'évaluer deux versions différentes d'une même fragment (qui seront appelés LAMBDA et SIGMA), comme dans le formulaire de post-édition. Pour ce faire, nous avons créé l'échelle suivante :

- LAMBDA est beaucoup plus adapté à l'utilisation que SIGMA
- LAMBDA est légèrement plus adapté à l'utilisation que SIGMA
- Aucune des deux options n'est utilisable
- SIGMA est légèrement plus adapté à l'utilisation que LAMBDA
- SIGMA est beaucoup plus adapté à l'utilisation que LAMBDA

Cette échelle permet de signaler une préférence et de nuancer son choix, ainsi que de signaler que, selon le critère des experts, aucune des deux versions proposées n'est utilisable dans un contexte d'enseignement et apprentissage de la musique. Ce type d'échelles a été également utilisé par Mitchell et al. (2014), Doherty et O'Brien (2014), et Van Egdom et Pluymaekers (2019).

5.3 Préparation et plan de l'expérience

Comme il a été mentionné auparavant, la préparation et le plan de notre expérience se divise en cinq phases, qui seront détaillées dans cette section.

5.3.1 Première phase : préparation et traitement des textes

La première phase du processus consistait à effectuer le traitement des textes nécessaire pour obtenir les traductions automatiques qui ont été employées dans cette recherche.

Tout d'abord, nous avons dû effectuer une sélection des fragments textuels représentatifs des textes originaux en russe (section 5.2.2). Avant de commencer, nous avons décidé que les fragments des deux textes employés pour notre étude devaient présenter une longueur entre 250 et 300 mots ; nous avons jugé que les échantillons textuels de cette longueur pourraient refléter suffisamment le niveau de spécialité de chacun, ce qui s'avérait important pour la suite de l'expérience.

Pour procéder au choix de ces fragments, et étant donné que le chercheur de ce mémoire n'a pas de compétences de lecture en langue russe, nous avons décidé de nous appuyer sur des versions anglaise et espagnole des textes. Une fois les fragments de texte choisis, nous les avons localisés dans les originaux en russe, ce qui a été facilité par les intitulés et la structure des textes (qui est clairement numérotée), ainsi que par les illustrations qui accompagnent ces textes (pratiquement invariables dans toutes les versions).

À ce stade, il est important de mentionner que nous avons travaillé avec des textes digitalisés au format PDF. Ces documents sont en libre accès sur internet où l'accès a été facilité par des services de numérisation des bibliothèques (notamment la Bibliothèque nationale d'Espagne BNE). La qualité des digitalisations permettait une lecture sans efforts. Naturellement, ces textes au format PDF ne pouvaient pas être présentés aux systèmes de TA. De plus, la qualité des digitalisations, ainsi que la complexité graphique des textes, ont empêché l'utilisation des technologies OCR pour extraire les textes à traduire. Nous avons donc eu recours à la transcription humaine tapuscrit des textes en russe.

Nous avons reçu l'aide de deux russophones natifs. La première personne, linguiste de profession, a effectué la transcription des fragments choisis des deux textes en russe dans un document au format *.docx*. Pour faciliter sa tâche, nous sommes intervenus sur les documents originaux en PDF pour qu'elle puisse travailler avec une version qui contenait uniquement les fragments textuels à transcrire. Une fois les deux fragments transcrits, la deuxième personne a effectué une révision des textes en confrontant les transcriptions et les originaux pour corriger de possibles fautes ou omissions.

Nous avons obtenu les deux fragments textuels au format *.docx* à la suite de ce processus de transcription. Néanmoins, les illustrations sont un facteur essentiel pour cette typologie textuelle car elles complètent le texte. Nous avons donc décidé d'éditer ces illustrations pour éviter d'adopter celles des originaux, affectées par la qualité des digitalisations. À cette fin, nous avons utilisé MuseScore©, un logiciel, libre et gratuit, d'écriture musicale et d'édition avancée des partitions. Nous avons finalement intégré les illustrations éditées dans les textes pour recréer dans notre étude des versions fidèles aux textes originaux. Le résultat de ce processus peut être consulté dans l'Annexe 2.

Une fois terminées la transcription et la mise en forme des documents en russe, nous étions en mesure de générer les TA en espagnol. Nous avons donc généré une version des textes en russe sans illustrations ni signes musicaux (remplacés par la suite de caractères « [...] »). De ce fait, nous étions en mesure de présenter aux systèmes de TA l'intégralité des textes choisis. Nous avons jugé cela important

étant donné qu'une TA générée phrase par phrase aurait pu altérer la cohésion du résultat des TA.

Pour notre expérience, nous avons choisi deux systèmes de TA commerciaux avec une architecture neuronale : DeepL et Google Translate. Nous avons donc généré deux versions en espagnol : une TA issue de DeepL et une TA issue de Google Translate. Par la suite, nous avons intégré les illustrations à chaque version pour obtenir des versions qui ressemblent aux textes originaux. Le résultat de ce processus peut être consulté dans l'Annexe 3. Les deux TA en espagnol générées lors de cette phase ont été utilisées tout au long des phases suivantes.

5.3.2 Deuxième phase : annotation des erreurs de la TA brute

La deuxième phase de notre méthode était l'annotation d'erreurs dans les TA. Cette phase a été fondamentale pour la suite de l'expérience étant donné que les versions annotées ont servi pour évaluer les post-éditions effectuées par nos participants lors de la phase suivante. C'est-à-dire, leur fonction principale est de permettre au chercheur d'identifier dans quelle mesure les post-éditeurs ont été capables d'identifier et corriger les erreurs dans les TA signalées par l'annotatrice.

Nous avons jugé indispensable que l'annotation soit effectuée par une personne qui possède deux caractéristiques : d'abord, une connaissance approfondie des langues russe et espagnole ; ensuite, cette personne devrait être une spécialiste dans le domaine de la musique. Le chercheur a donc contacté une personne qui corresponde à ces deux critères. À la suite des premiers contacts, la personne a accepté de participer dans notre étude. Elle est une professeure d'université avec une longue carrière dans l'enseignement de la théorie de la musique. Russophone native, elle a obtenu son titre de docteur en musique du Conservatoire de Musique Tchaïkovski de Moscou et elle habite depuis plus de dix ans en Colombie. Elle est professeur des universités depuis son arrivée au pays. De ce fait, ses compétences linguistiques en espagnol sont très avancées et sa condition de russophone native, sommée à son expertise dans le domaine, correspondaient parfaitement à nos besoins.

Pour faciliter la tâche, nous avons préparé un format d'annotation bilingue au format *.doc* avec les textes originaux en russe et les TA en espagnol. Dans ce

format, l'annotatrice devait employer la fonction « Nouveau commentaire » de Microsoft Word pour signaler les erreurs de traduction. Nous avons limité ce qui devait être identifié comme erreur par l'annotatrice à quatre types d'erreurs : sémantique, terminologique, addition et omission. Nous avons pris comme référence les guides de post-édition de TAUS (2016) pour choisir ces types d'erreurs. Ceci nous permettait de centrer les annotations sur le contenu étant donné que l'annotatrice n'est pas une linguiste et que, à ce stade, l'intérêt de l'annotation portait plus sur le contenu et moins sur les fautes de langue.

Dans les instructions données à l'annotatrice avec le format, nous avons inclus le lien vers une courte vidéo explicative de comment procéder pour annoter les erreurs. La disposition des textes a été organisée pour faciliter la comparaison des deux versions en espagnol entre elles et avec l'original en russe. Les textes ont été intégrés dans ce format avec leurs illustrations et ont été divisés en cinq fragments chacun. Tout au long de notre étude, nous avons décidé de fragmenter les textes par segments, et non par phrases, pour éviter une division artificielle et prioriser l'exposition d'idées et des notions proposées par les textes. Le format d'annotation peut être consulté dans l'Annexe 4.

Une fois que l'annotatrice a terminé d'effectuer la tâche, elle nous a fait parvenir le format avec ses annotations. La tâche avait été effectuée en suivant toutes les indications.

5.3.3 Troisième phase : post-édition monolingue et questionnaire post-tâche

La troisième phase de notre méthode consistait à faire post-éditer les TA produites par les deux systèmes à deux groupes de post-éditeurs différents : un Groupe A, composé de musiciens, spécialistes dans le domaine ; et un Groupe B, composé de linguistes (section 5.2.3.1). À la fin de la tâche, chaque participant a dû répondre à un questionnaire pour rendre compte de son expérience de post-édition monolingue.

Une fois le profil des post-éditeurs identifié, nous avons créé une liste de participants possibles qui ont ensuite été contactés par messagerie instantanée. Nous avons finalement reçu trois réponses positives pour chaque groupe. Une fois

l'intention de participation confirmée, nous avons envoyé le texte à post-éditer sans le texte source, avec un formulaire de consentement similaire à celui envoyé à l'annotatrice de la deuxième phase. Le format et les instructions étaient identiques pour les deux groupes.

Étant donné que nous avons divisé en 5 segments chacun des textes choisis pour notre étude (voir Tableau 3), chaque post-éditeur devait post-éditer un total de 10 segments : 5 segments du texte spécialisé, et 5 du texte non-spécialisé. Pour chaque segment, les post-éditeurs avaient le choix de post-éditer la TA de leur préférence, mais les instructions indiquaient de toujours choisir seulement une de ces deux versions pour post-éditer et la signaler avec un carré à cocher (☒) situé en bas, à gauche de chaque version. Le Tableau 6 illustre un exemple du schéma de la tâche de post-édition.

Texte 1 (non-spécialisé)		
Segment	LAMBDA (Google Translate)	SIGMA (DeepL)
1	☒	☐
2	☐	☒
3	☐	☒
4	☒	☐
5	☐	☒
Texte 2 (spécialisé)		
Segment	LAMBDA (Google Translate)	SIGMA (DeepL)
1	☐	☒
2	☐	☒
3	☐	☒
4	☒	☐
5	☐	☒
Total :	Dix segments post-édités	

TABLEAU 6. EXEMPLE DE TACHE DE POST-EDITION COMPLETEE

Comme ce tableau le montre, pour éviter que ce choix soit influencé par les préjugés des participants sur la qualité des systèmes de TA employés, chaque colonne a été

renommée avec des noms de lettres grecques : SIGMA (Google Translate) et LAMBDA (DeepL). L'ordre du texte a été respecté et la post-édition a donc été faite en contexte. Le format de post-édition peut être consulté dans l'Annexe 5.

Nous avons jugé essentiel de donner aux post-éditeurs le contexte de la tâche. La post-édition monolingue n'est pas une pratique courante dans l'industrie de la communication multilingue, comme nous l'avons exposé dans la première partie du mémoire, ce qui pouvait éventuellement mener les participants à se questionner sur le sens et la faisabilité de leurs efforts de post-édition dans ce contexte et à se décourager face à la tâche et biaiser les résultats. Le contexte que nous avons présenté permet d'établir un scénario d'utilisation concret des post-éditions effectuées par les participants (voir Annexe 5).

Comme dans la phase précédente, les instructions de post-édition incluaient également une courte vidéo explicative, destinée aux participants, pour mieux illustrer la procédure attendue lors du processus de post-édition. Nous avons inclus une définition opérationnelle de « post-édition monolingue » pour faciliter la compréhension de nos participants sur les aspects essentiels de cette pratique. Nous avons accordé un délai de 8 à 10 jours pour finaliser la tâche et retourner le format avec les post-éditions effectuées. Une fois la tâche effectuée et le formulaire de consentement renvoyé, nous avons procédé à l'envoi du lien vers le questionnaire de participation.

Le questionnaire de participation auquel les post-éditeurs devaient répondre à la fin de la tâche était composé des questions suivantes :

1. *Combien de temps avez-vous investi pour réaliser le travail de post-édition ?*
2. *Suite à cette expérience, quelle est votre opinion sur la pratique de la post-édition monolingue ?*
3. *Avez-vous trouvé la tâche de post-édition facile ou difficile ? Motivez votre réponse.*
4. *Avez-vous confiance dans les traductions automatiques qui vous ont été données pour accomplir la tâche de post-édition ?*
5. *Pensez-vous que vous avez réussi à produire des textes fiables ?*

6. *D'après le processus de post-édition et de votre perception de la qualité du résultat final, comment évaluez-vous votre expérience ?*
7. *Utiliserez-vous les résultats de cette expérience pour apprendre ou pour enseigner la musique ?*

Cette série de questions avait pour but de récolter l'opinion des post-éditeurs sur la pratique de la post-édition monolingue en particulier celle des extraits de traités de musique dans ce contexte. Les réponses des deux groupes de post-éditeurs seront analysées et contrastées dans la partie du mémoire dédiée aux résultats de l'expérience. Un exemple d'une des post-édition effectuées par un des participants lors de cette phase peut être consulté dans l'Annexe 6.

5.3.4 Quatrième phase : évaluation des fautes corrigés par les post-éditeurs

La quatrième phase de notre méthode consistait à évaluer les fautes corrigées dans les post-éditions effectuées par les linguistes et les musiciens lors de la troisième phase de notre étude. Pour effectuer cette évaluation nous avons pris les annotations produites par l'experte bilingue dans la deuxième phase de notre étude. Notre objectif était de déterminer les taux de correction de chaque texte enregistré, d'un côté, par l'ensemble des musiciens, et, de l'autre, par l'ensemble de linguistes. Cette procédure nous permettait d'associer les taux de correction à une valeur qualitative assignée aux post-éditions effectuées par les deux groupes (voir Tableau 5, section 5.2.5.1).

Dans un premier temps, nous avons déterminé la quantité d'erreurs dans les TA brutes qui ont été choisies par chaque post-éditeur lors de la troisième phase. Nous avons donc comparé les TA brutes avec les annotations d'erreurs de la deuxième phase. Nous avons marqué les types d'erreurs signalés par l'annotatrice (sémantique, terminologique, addition ou omission) avec des couleurs différentes, ce qui nous a permis d'identifier non seulement un total d'erreurs, mais aussi d'établir combien de fois chaque type d'erreur était présent dans les TA brutes.

Dans un deuxième temps, nous avons établi combien d'erreurs ont été effectivement corrigées par les post-éditeurs ; finalement, nous avons obtenu un

pourcentage de correction d'erreurs, ou « taux de correction », pour chaque type d'erreur et chaque texte, comme le montre le Tableau 7.

Musiciens			
Texte 1			
Type d'erreur	N° d'erreurs identifiées	N° d'erreurs corrigées	Taux de correction
Sémantique	$[(14) + (14) + (16)] = 44$	$[(10) + (11) + (14)] = 35$	79,5 %
Terminologique	$[(13) + (13) + (14)] = 40$	$[(11) + (13) + (7)] = 31$	77,5 %
Addition	$[(0) + (0) + (0)] = 0$	$[(0) + (0) + (0)] = 0$	NA
Omission	$[(1) + (0) + (1)] = 2$	$[(1) + (0) + (0)] = 1$	50 %
Total Texte 1 :	86	67	77,9 %
Texte 2			
Type d'erreur	N° d'erreurs identifiées	N° d'erreurs corrigées	Taux de correction
Sémantique	$[(5) + (8) + (8)] = 21$	$[(2) + (5) + (6)] = 13$	61,9 %
Terminologique	$[(11) + (13) + (12)] = 36$	$[(0) + (12) + (12)] = 24$	66,7 %
Addition	$[(1) + (1) + (1)] = 3$	$[(1) + (1) + (1)] = 3$	100 %
Omission	$[(1) + (0) + (1)] = 2$	$[(0) + (0) + (0)] = 0$	0 %
Total Texte 2 :	62	40	64,5 %
Total Global :	148	107	72,3 %
Linguistes			
Texte 1			
Type d'erreur	N° d'erreurs identifiées	N° d'erreurs corrigées	Taux de correction
Sémantique	$[(13) + (14) + (14)] = 41$	$[(4) + (6) + (12)] = 22$	53,6 %
Terminologique	$[(14) + (11) + (12)] = 37$	$[(3) + (4) + (7)] = 14$	37,8 %
Addition	$[(0) + (0) + (0)] = 0$	$[(0) + (0) + (0)] = 0$	NA
Omission	$[(1) + (0) + (1)] = 2$	$[(0) + (0) + (0)] = 0$	0 %
Total Texte 1 :	80	36	45 %
Texte 2			
Type d'erreur	N° d'erreurs identifiées	N° d'erreurs corrigées	Taux de correction
Sémantique	$[(5) + (6) + (6)] = 17$	$[(0) + (1) + (0)] = 1$	5,8 %
Terminologique	$[(11) + (11) + (11)] = 33$	$[(0) + (0) + (0)] = 0$	0 %
Addition	$[(1) + (1) + (1)] = 3$	$[(0) + (0) + (0)] = 0$	0 %
Omission	$[(1) + (1) + (1)] = 3$	$[(0) + (0) + (0)] = 0$	0 %
Total Texte 2 :	56	1	1,7 %
Total Global :	136	37	27,2 %

TABLEAU 7. TAUX DE CORRECTION PAR TEXTE ET PAR GROUPE DE POST-EDITEURS

Les éléments de réponse apportés par ces résultats seront analysés en détail dans le chapitre suivant (section 6.2).

5.3.5 Cinquième phase : évaluation de l'utilisabilité des post-éditions par les juges

La dernière phase de notre méthode de recherche consistait à faire évaluer les post-éditions effectuées lors de la phase précédente. L'évaluation a été menée par un total de 21 juges (section 5.2.3.2). Chacun des juges, spécialistes dans le domaine, a été contacté par messagerie instantanée pour être invité à participer à l'évaluation des post-éditions. Une fois l'invitation acceptée, nous avons procédé à l'envoi du lien vers l'évaluation Limesurvey en ligne. Le but de cette évaluation était de déterminer lequel des deux groupes de post-éditeurs avait réussi à produire les post-éditions monolingues les plus « utilisables ».

Pour cela, nous avons effectué une évaluation comparative dans laquelle chacun des segments post-édités par les post-éditeurs des groupes A (musiciens) et B (linguistes) a été évalué par les experts entre 5 et 8 fois. Dans l'exemple du Tableau 8, nous avons pris le premier segment du Texte 1 post-édité par le « post-éditeur B » musicien (segment « B1 »), et nous l'avons opposé au premier segment de ce même texte post-édité par le « post-éditeur F » linguiste (segment « F1 ») ; nous avons répliqué cette procédure pour chaque segment des deux textes.

À chaque fois, nous avons organisé notre formulaire d'évaluation, dans les colonnes LAMBDA et SIGMA, pour opposer systématiquement les post-éditions des groupes A et B, comme on peut le constater dans l'exemple du Tableau 8.

Groupes de post-éditeurs					
« Groupe A » : Musiciens			« Groupe B » : Linguistes		
Post-éditeur A			Post-éditeur D		
Post-éditeur B			Post-éditeur E		
Post-éditeur C			Post-éditeur F		
TEXTE 1			TEXTE 2		
S.	LAMBDA	SIGMA	S.	LAMBDA	SIGMA
1	B1	F1	1	E1	B1
2	C2	E2	2	A2	D2
3	B3	E3	3	D3	B3
4	F4	B4	4	A4	F4
5	F5	C5	5	F5	C5

TABEAU 8. EXEMPLE D'ÉVALUATION AVEC LES NUMEROS DE SEGMENT

De ce fait, pour chaque segment (qui équivaut à une question dans le format d'évaluation), nous avons choisi aléatoirement deux post-éditions : une issue du Groupe A, et une du Groupe B. En suivant ce modèle, nous avons pu construire un questionnaire différent pour chaque évaluateur, ce qui nous a permis d'utiliser toutes les post-éditions produites par les participants des deux groupes avec les deux systèmes de TA. Le schéma suivi pour l'élaboration des 21 questionnaires, utilisé pour interpréter et classer les réponses des juges, peut être consulté dans l'Annexe 7.

Cette organisation a été choisie pour éviter de faire évaluer la totalité des post-éditions produites dans notre expérience à tous les évaluateurs, étant donné que cela pourrait entraîner de la fatigue et, par conséquent, biaiser les réponses des experts. Néanmoins, notre schéma garantit que les groupes A et B seront toujours en contraste dans nos évaluations, et que les post-éditions de ces deux groupes ne suivent pas un patron fixe car ces dernières se trouvent réparties aléatoirement dans les colonnes LAMBDA et SIGMA. De plus, chaque évaluation a été composée et organisée différemment (voir Annexe 7), tout en suivant les mêmes principes, ce qui nous a permis de recueillir des données dans des conditions assez diverses de ce point de vue.

L'indication donnée aux évaluateurs dans le questionnaire était la suivante :

Cher participant :

Dans ce questionnaire, vous évaluerez différentes versions de fragments tirés de deux textes différents. Les questions de 1 à 5 correspondent à des extraits du « Texte 1 », et les questions de 6 à 10 à des extraits du « Texte 2 ». Dans chacune de ces questions, vous trouverez deux versions différentes de chaque extrait, situées respectivement dans les colonnes « LAMBDA » et « SIGMA ».

Dans chaque cas, vous répondrez à la question suivante : ***Laquelle des deux versions convient le mieux à l'utilisation de ces textes dans des contextes d'enseignement et d'apprentissage de la musique ?*** Vous pouvez choisir l'une des cinq options de réponse qui vous sont proposées.

Les textes sont tirés de traités musicaux. Ils ont une fonction essentiellement didactique et ont été traduits en espagnol au

moyen d'un processus mixte qui alterne traduction automatique et édition et correction humaines. Ce processus est connu sous le nom de « post-édition ». Notre objectif est d'évaluer si cette procédure permet de traduire ce type de textes de manière à ce qu'ils soient utilisables dans les contextes pour lesquels ils ont été conçus dans leur langue d'origine.

Merci d'avance pour vos réponses.

À la suite de ces indications, les participants pouvaient évaluer les différentes versions de chaque segment des textes pour exprimer leur jugement en choisissant un des cinq degrés de notre échelle Likert (section 4.2.2.2) pour répondre à la question.

Pour les évaluateurs, la participation dans notre étude terminait après avoir envoyé et enregistré la totalité des réponses du questionnaire. À la fin de cette phase, nous avons recueilli la totalité des données pour confirmer ou infirmer les hypothèses.

5.4 Conclusion

En résumé, ce chapitre a présenté l'objectif, les questions de recherche principales et subsidiaires et les hypothèses de départ de notre étude. Nous avons également exploré les choix méthodologiques concernant les textes choisis et le profil des participants, ainsi que les outils et les métriques employés. Nous avons présenté les cinq phases de notre expérience : (1) la préparation et le traitement des textes, (2) l'annotation des erreurs de la TA brute, (3) la post-édition monolingue et le questionnaire post-tâche, (4) l'évaluation des fautes corrigées par la post-édition sur la base de l'annotation, ainsi que (5) l'évaluation de l'utilisabilité des post-éditions monolingues par les juges experts.

Dans le chapitre suivant, nous allons présenter et interpréter les données récoltées lors de notre expérience pour répondre à nos questions de recherche et pour infirmer ou confirmer nos hypothèses. Nous en tirerons des conclusions qui nous permettront de conclure cette étude et d'accomplir l'objectif du mémoire.

6. Résultats

6.1 Introduction

Ce chapitre sera consacré à la présentation et analyse des résultats obtenus. Nous allons tout d'abord présenter les résultats pour la phase d'évaluation des fautes corrigées par les post-éditeurs (phase 4, section 5.3.4 du chapitre 5) (section 6.2). Ensuite, nous présenterons les résultats issus de l'évaluation comparative de l'utilisabilité des post-éditions effectuées par des experts du domaine (phase 5, section 5.3.5 du chapitre 5) (section 6.3). Pour rappel, la phase 4 cherchait à évaluer les fautes des post-éditions monolingues effectuées par les deux groupes de post-éditeurs qui ont participé à notre étude (linguistes et spécialistes) pour déterminer leur taux de correction, et la phase 5 consistait à faire évaluer ces post-éditions par des experts pour déterminer lequel des deux groupes avait produit les post-éditions les plus utilisables dans des contextes d'enseignement et d'apprentissage de la musique.

Pour finir, dans ce chapitre, nous allons effectuer une évaluation comparative entre les systèmes de TA employés dans notre étude (section 6.4) pour déterminer lequel des deux systèmes a produit le moins d'erreurs, et lequel a été davantage choisi par les post-éditeurs pour effectuer la tâche demandée lors de la troisième phase de notre étude.

À titre indicatif, le Tableau 9 récapitule les questions de recherche principales et subsidiaires, ainsi que les hypothèses, et signale dans quelle section de ce chapitre nous allons répondre à ces questions.

Questions principales de recherche	
Première question	Section
Dans quelle mesure l'expertise du post-éditeur monolingue dans le domaine de spécialité est un facteur déterminant pour l'utilisabilité des post-éditions de traités de musique ?	6.5
Deuxième question	Section
Quel est l'impact du degré de spécialisation du texte sur l'utilisabilité des post-éditions produites par les participants de notre étude ?	6.5
Questions subsidiaires	
Première question subsidiaire	Section
Lequel des deux systèmes de TA, Google Translate ou DeepL, fait le moins d'erreurs d'après les annotations de l'annotatrice bilingue ?	6.4.3
Deuxième question subsidiaire	Section
Lequel des deux systèmes est d'avantage choisi par les post-éditeurs pour effectuer la tâche qui leur est présentée ?	6.4.3
Hypothèses	
Première hypothèse	Section
Plus le post-éditeur est expert dans le domaine, plus il peut rendre utilisable la traduction automatique brute d'un texte de spécialité.	6.3.3
Deuxième hypothèse	Section
Dans la pratique de la post-édition monolingue, les experts voient plus de fautes de fidélité que les traducteurs.	6.2.3

TABLEAU 9. RECAPITULATIF DES QUESTIONS DE RECHERCHE ET HYPOTHESES

6.2 Évaluation des fautes corrigées par les post-éditeurs monolingues

Dans cette section, nous cherchons à valider notre seconde hypothèse : dans la pratique de la post-édition monolingue, les experts voient plus de fautes de fidélité que les traducteurs. Nous allons présenter d'abord les taux de correction des post-éditions monolingues en fonction des textes et post-éditeurs (section 6.2.1), puis les résultats de l'interprétation qualitative des post-éditions ainsi que les réponses aux questionnaires post-tâche données par les post-éditeurs (section 6.2.2). Cette interprétation qualitative est dérivée d'une évaluation effectuée par nous-mêmes selon l'échelle présentée dans la section 5.2.5.1 du chapitre 5 et sur la base des corrections faites par une experte bilingue. Finalement, nous concluons avec une synthèse des données présentées tout au long de la section (section 6.2.3).

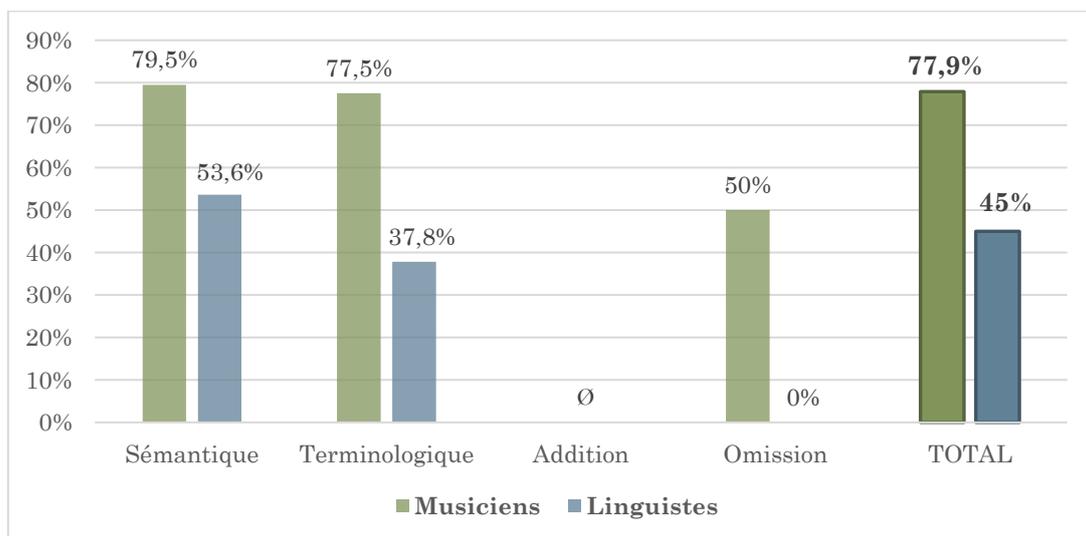
6.2.1 Taux de correction des post-éditions monolingues

6.2.1.1 Texte non spécialisé (Texte 1)

Concernant le texte non spécialisé, les résultats de notre évaluation montrent que, au total, les musiciens ont corrigé 77,9 % des erreurs signalées par l'annotatrice bilingue, tandis que les linguistes ont corrigé 45 %. Pour rappel, dans la tâche d'annotation d'erreurs (section 5.3.2 du chapitre 5) nous nous sommes limités à quatre types d'erreurs (sémantique, terminologique, additions et omissions) en suivant les guides de post-édition TAUS (2016). Ceci a permis à l'annotatrice, experte bilingue qui n'est pas linguiste, de centrer les annotations sur le contenu véhiculé par les TA et moins sur les fautes de langue liées à la syntaxe qu'il pourrait y avoir.

La Figure 1 montre le taux de correction des musiciens et des linguistes par type d'erreur, ainsi que le taux de correction total de chaque groupe pour le Texte 1. Nous pouvons constater que les musiciens présentent un taux de correction plus haut que les linguistes dans les trois types d'erreurs identifiées (étant donné que nous n'avons pas identifié d'additions). De ce fait, le taux de correction total des musiciens est nettement supérieur à celui des linguistes pour le Texte 1.

FIGURE 3. TAUX DE CORRECTION EN FONCTION DES POST-EDITEURS (MUSICIENS OU LINGUISTES) – TEXTE 1



Le Tableau 10 montre plus en détail le nombre d'erreurs identifiées dans les TA brutes par l'annotatrice, ainsi que le nombre d'erreurs corrigées par chaque groupe de post-éditeurs ; ceci pour chaque type d'erreur et en fonction des post-éditeurs (musiciens et linguistes). Les chiffres entre parenthèses, à l'intérieur des crochets,

indiquent le nombre d'erreurs identifiées et corrigées dans le format de post-édition de chaque participant. Le « Taux de correction » a été calculé, pour chaque groupe de post-éditeurs, selon la formule suivante :

$$\text{Taux de correction} = \frac{\text{Nombre d'erreurs corrigées par les post-éditeurs}}{\text{Nombre d'erreurs identifiées par l'annotatrice}} \times 100$$

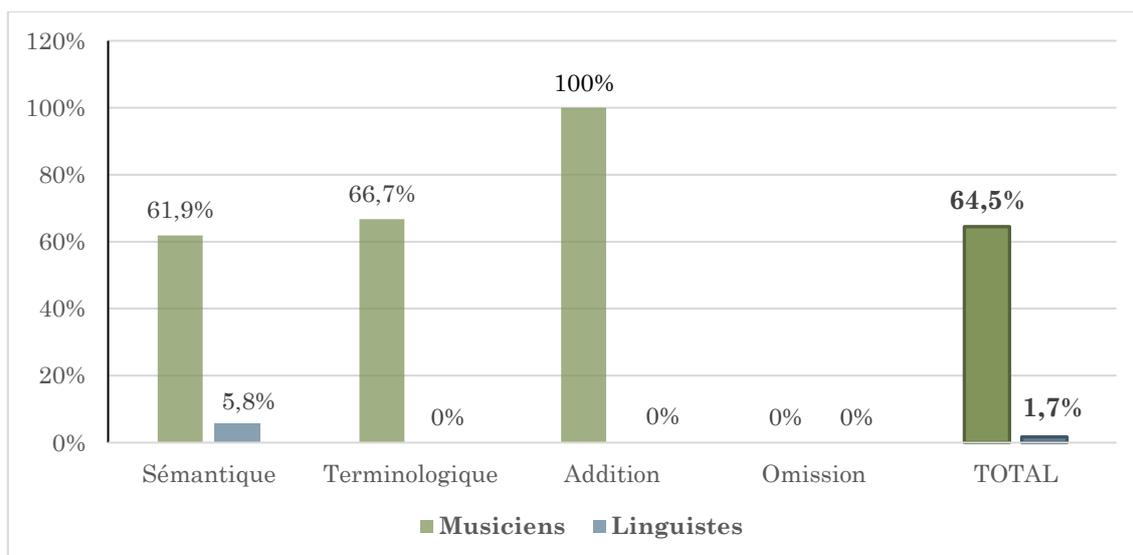
Musiciens			
Texte 1			
Type d'erreur	Nbr. d'erreurs identifiées	Nbr. d'erreurs corrigées	Taux de correction
Sémantique	[(14) + (14) + (16)] = 44	[(10) + (11) + (14)] = 35	79,5 %
Terminologique	[(13) + (13) + (14)] = 40	[(11) + (13) + (7)] = 31	77,5 %
Addition	Ø	Ø	Ø
Omission	[(1) + (0) + (1)] = 2	[(1) + (0) + (0)] = 1	50 %
Total Texte 1 :	86	67	77,9 %
Linguistes			
Texte 1			
Type d'erreur	Nbr. d'erreurs identifiées	Nbr. d'erreurs corrigées	Taux de correction
Sémantique	[(13) + (14) + (14)] = 41	[(4) + (6) + (12)] = 22	53,6 %
Terminologique	[(14) + (11) + (12)] = 37	[(3) + (4) + (7)] = 14	37,8 %
Addition	Ø	Ø	Ø
Omission	[(1) + (0) + (1)] = 2	[(0) + (0) + (0)] = 0	0 %
Total Texte 1 :	80	36	45 %

TABLEAU 10. NOMBRE D'ERREURS IDENTIFIÉES PAR L'ANNOTATRICE ET CORRIGÉES PAR LES POST-ÉDITEURS POUR LE TEXTE 1

6.2.1.2 Texte hautement spécialisé (Texte 2)

Le taux de correction du texte hautement spécialisé révèle une différence très marquée entre les deux groupes de participants. D'une part, le taux de correction total des musiciens, pour ce texte, est de 64,5 %. De l'autre, le taux de correction total des linguistes est de seulement 1,7 %. En effet, la Figure 2 montre que le taux de correction des musiciens a été supérieur à 60 % pour les erreurs sémantiques et terminologiques, et de 100 % pour les erreurs d'addition. Quant aux linguistes, le taux de correction est de 5,8 % pour les erreurs sémantiques. Toutefois, les deux groupes ont un taux de correction de 0 % pour les erreurs d'omission dans le texte hautement spécialisé. Il semblerait que le taux de correction de 0 % pour trois types d'erreurs (terminologique, addition et omission) chez les linguistes soit lié au degré de spécialité du texte.

FIGURE 4. TAUX DE CORRECTION – TEXTE 2



Tout comme le Tableau 10, le Tableau 11 montre en détail les chiffres concernant le nombre d'erreurs identifiées par l'annotatrice et corrigées dans les formats de post-édition des deux groupes de participants. En effet, nous avons identifié 62 erreurs dans les TA choisies par les musiciens pour effectuer la tâche de post-édition ; ils ont corrigé un total de 40 erreurs. En revanche, parmi les 56 erreurs identifiées dans les TA des linguistes, uniquement 1 erreur a été corrigée, ce qui explique la différence du taux de correction dans les deux groupes.

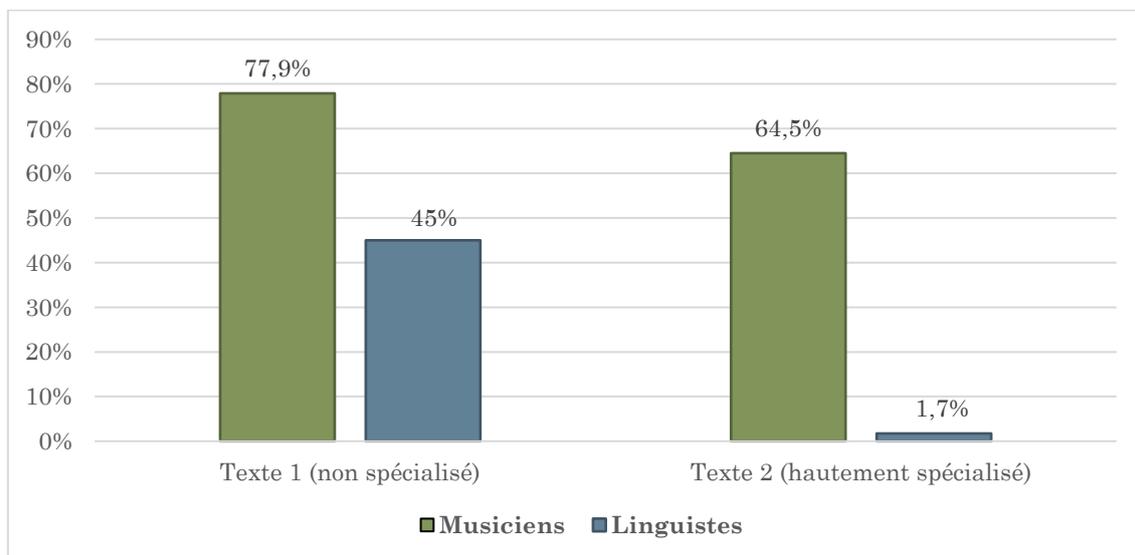
Musiciens			
Texte 2			
Type d'erreur	Nbr. d'erreurs identifiées	Nbr. d'erreurs corrigées	Taux de correction
Sémantique	$[(5) + (8) + (8)] = 21$	$[(2) + (5) + (6)] = 13$	61,9 %
Terminologique	$[(11) + (13) + (12)] = 36$	$[(0) + (12) + (12)] = 24$	66,7 %
Addition	$[(1) + (1) + (1)] = 3$	$[(1) + (1) + (1)] = 3$	100 %
Omission	$[(1) + (0) + (1)] = 2$	$[(0) + (0) + (0)] = 0$	0 %
Total Texte 2 :	62	40	64,5 %
Linguistes			
Texte 2			
Type d'erreur	Nbr. d'erreurs identifiées	Nbr. d'erreurs corrigées	Taux de correction
Sémantique	$[(5) + (6) + (6)] = 17$	$[(0) + (1) + (0)] = 1$	5,8 %
Terminologique	$[(11) + (11) + (11)] = 33$	$[(0) + (0) + (0)] = 0$	0 %
Addition	$[(1) + (1) + (1)] = 3$	$[(0) + (0) + (0)] = 0$	0 %
Omission	$[(1) + (1) + (1)] = 3$	$[(0) + (0) + (0)] = 0$	0 %
Total Texte 2 :	56	1	1,7 %

TABLEAU 11. NOMBRE D'ERREURS IDENTIFIEES ET CORRIGEEES PAR LES POST-EDITEURS POUR LE TEXTE 2

6.2.1.3 Conclusion

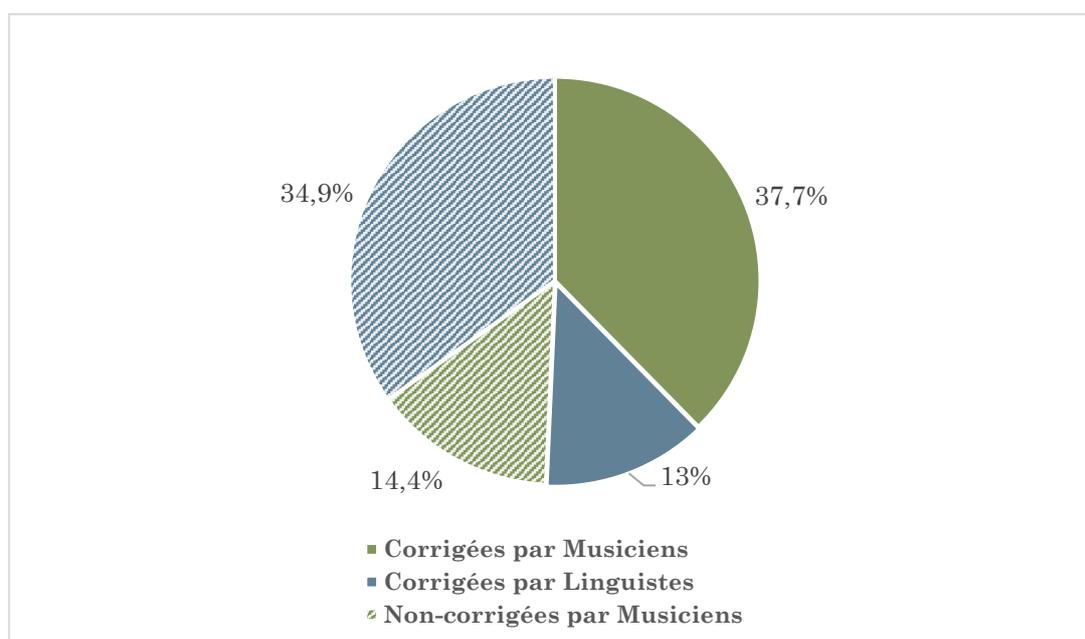
Dans cette section, nous avons montré les résultats de nos analyses du taux de correction des deux groupes de post-éditeurs (musiciens et linguistes) pour les deux textes, non spécialisé (Texte 1) et hautement spécialisé (Texte 2). Nous sommes en mesure d'affirmer que, pour ces deux textes, les post-éditions monolingues des musiciens ont un taux de correction nettement plus haut que celui des linguistes. Ainsi, pour le Texte 1, les musiciens ont corrigé 77,9 % des erreurs marquées dans leurs TA brutes par l'annotatrice, tandis que les linguistes en ont corrigé 45 %. Quant au Texte 2, la différence est davantage marquée étant donné que les musiciens ont corrigé 64,5 % des erreurs, contrairement aux linguistes qui en ont corrigé 1,7 %. Ces tendances sont résumées dans la Figure 3.

FIGURE 5. TAUX DE CORRECTION PAR TEXTE



De ce fait, si l'on adopte une perspective globale de la tâche de post-édition monolingue comprenant les erreurs identifiées dans les textes post-édités par nos participants, on constate que les musiciens ont corrigé 37,7 % des erreurs, tandis que les linguistes en ont corrigé 13 % (Figure 4). Ainsi, un peu plus de la moitié des erreurs (50,7 %) ont été corrigées par nos participants. Par conséquent, 49,3 % des erreurs n'ont pas été corrigées. Parmi celles-ci, 34,9 % n'ont pas été corrigées par les linguistes, et 14,4 % par les musiciens.

FIGURE 6. TOTAL D'ERREURS CORRIGÉES ET NON-CORRIGÉES



Dans la section suivante, nous allons analyser ces résultats quantitatifs à la lumière de l'échelle d'évaluation qualitative que nous avons adoptée, ainsi que des réponses données par les post-éditeurs dans leurs questionnaires post-tâche.

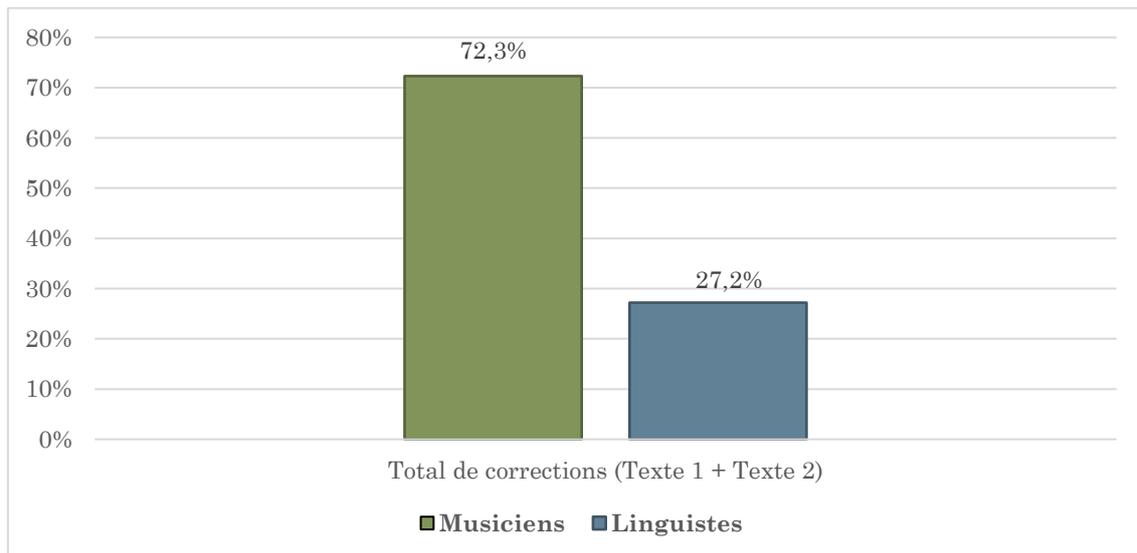
6.2.2 Évaluation qualitative des post-éditions monolingues et réponses aux questionnaires post-tâche

Les taux de correction issus de notre analyse des post-éditions monolingues produites par nos participants peuvent être interprétés à l'aide de l'échelle d'évaluation décrite dans le chapitre sur la Méthodologie (voir Tableau 5, section 5.2.5.1). À cet effet, la Figure 5 montre le taux global de correction. À la différence des taux de correction précédents, calculés séparément pour les Textes 1 et 2 (voir la formule dans la section 6.2.1.1), le « Taux global de correction » a été calculé, pour chaque groupe de post-éditeurs, selon la formule suivante :

$$Taux\ global\ de\ correction = \frac{\begin{array}{c} \text{Total du nombre d'erreurs corrigées} \\ \text{par les post - éditeurs} \\ (Texte\ 1 + Texte\ 2) \end{array}}{\begin{array}{c} \text{Total du nombre d'erreurs identifiées} \\ \text{par l'annotatrice} \\ (Texte\ 1 + Texte\ 2) \end{array}} \times 100$$

De ce fait, la Figure 5 montre que les musiciens ont corrigé 72,3 % de toutes les erreurs identifiées dans les textes 1 et 2. Quant aux linguistes, ils ont corrigé 27,2 % des erreurs identifiées dans les textes 1 et 2.

FIGURE 7. TAUX GLOBAL DE CORRECTION



Si l'on interprète le taux de correction global constaté dans la Figure 5 à la lumière de notre échelle, nous pourrions dire que l'essentiel du sens des post-éditions monolingues effectuées par les musiciens est transmis grâce aux corrections appliquées à la post-édition. Quant aux post-éditions monolingues des linguistes, d'après les valeurs de notre échelle, elles déforment le sens de manière importante ou présentent beaucoup de petites erreurs à cause des corrections non-appliquées lors de la phase de post-édition. Ces interprétations seront confrontées avec les résultats de la phase d'évaluation par des juges monolingues pour confirmer ou infirmer leur validité.

Finalement, le Tableau 12 montre le taux de correction de chaque texte, ainsi que le taux de correction global (qui vient d'être mentionné) pour chaque groupe de post-éditeurs, en relation à la valeur qualitative établie par notre échelle d'évaluation (voir Tableau 5, section 5.2.5.1).

Taux de correction	Musiciens	Linguistes	Valeur qualitative
50 – 79 %	Texte 1 (77,9 %)	∅	L'essentiel du sens est transmis grâce aux corrections appliquées à la post-édition.
	<u>Taux global</u> (72,3 %)		
	Texte 2 (64,5 %)		
21 – 49 %	∅	Texte 1 (45 %)	La traduction post-éditée déforme le sens de manière importante ou présente beaucoup de petites erreurs.
		<u>Taux global</u> (27,2 %)	
1 – 20 %	∅	Texte 2 (1,7 %)	La plupart des erreurs n'ont pas été corrigées. Très peu de sens est véhiculé.

TABLEAU 12. VALEUR DES TAUX DE CORRECTION POUR LES MUSICIENS ET LES LINGUISTES

Nous observons clairement que l'ensemble des taux de correction chez les musiciens se trouve dans la même catégorie, en haut du tableau, tandis que ceux des linguistes sont dans les deux catégories inférieures du tableau. Aucun des taux de correction de nos participants n'est tombé dans les deux catégories supérieures (80 % à 99 % et 100 %), et inférieure (0 %).

Les taux de correction, ainsi que leurs valeurs qualitatives assignées par notre échelle d'évaluation, peuvent être comparés avec les réponses données par les deux groupes de post-éditeurs dans les questionnaires post-tâche. Les sept questions posées aux participants (section 5.3.3) avaient pour objectif de recueillir leurs avis sur leur expérience de post-édition monolingue dans le contexte particulier imposé par notre étude. Nous avons pu recueillir les trois réponses des participants musiciens, mais uniquement deux réponses des linguistes. L'Annexe 8 récapitule les questions et les réponses données par nos participants à ces questionnaires.

En général, les musiciens ont déclaré avoir une opinion beaucoup plus favorable de la post-édition monolingue que les linguistes. Ils ont déclaré qu'il s'agissait d'une pratique intéressante, importante, voire nécessaire dans certains contextes académiques. Un des participants a affirmé, par exemple, que l'expérience démontre que, pour cette typologie textuelle, la TA ne peut pas remplacer la traduction humaine. Pour leur part, les linguistes ont déclaré que la post-édition monolingue peut en effet s'avérer utile, mais qu'elle est laborieuse et que pour bien effectuer ce type de tâche, le texte source sera toujours nécessaire. Le contraste

entre l'opinion des musiciens et des linguistes est cohérent avec les résultats exposés tout au long de ce chapitre.

En ce qui concerne la difficulté de la tâche, un des musiciens a déclaré que celle-ci lui semblait facile grâce à ses connaissances dans le domaine. Les deux autres musiciens ont déclaré qu'elle était difficile car elle demandait de la rigueur et que « l'original » devait être respecté. D'abord, il est intéressant qu'un des musiciens a identifié la TA brute comme un « original », ce qui démontre une certaine posture face à l'acte de traduire, mais aussi une perception tout-à-fait différente de la TA brute que celle des post-éditeurs linguistes. En effet, ces derniers ont déclaré que la tâche représentait une difficulté moyenne, voire qu'elle était difficile car il leur manquait des connaissances dans le domaine, ainsi que le texte source. On peut donc constater que les linguistes ne considéraient pas la TA comme un original et ont trouvé la tâche plus compliquée à cause de leur manque d'outils conceptuels dans le domaine, contrairement aux musiciens.

Quant à la fiabilité de la TA, les musiciens ont été partagés entre des réponses du type « pas fiable », « partiellement fiable » et « fiable » ; l'opinion des linguistes a été plus rassemblée autour des réponses comme « pas fiable » et « partiellement fiable ». De ce fait, le spectre de perception de la fiabilité de la TA est plus large chez les musiciens, tandis que les linguistes ne considèrent dans aucun cas que la TA est fiable.

Cette tendance s'est maintenue chez les linguistes en ce qui concerne la fiabilité des post-éditions produites lors de l'expérience. Les linguistes ont considéré que leurs post-éditions monolingues n'étaient pas fiables, principalement à cause du manque de connaissances dans le domaine et de ressources terminologiques. En revanche, deux musiciens ont considéré que leurs post-éditions étaient fiables, et l'autre les a estimées partiellement fiables. Une fois de plus, ces réponses sont cohérentes avec le taux global de correction (voir Figure 5) et le taux de correction par texte (voir Figure 3).

Nous avons aussi demandé aux post-éditeurs d'évaluer leur expérience. Les musiciens ont eu des opinions très favorables, déclarant que la sensation post-tâche était positive, malgré le fait d'être confronté à cette tâche pour la première

fois ; ils ont également déclaré que l'expérience en général avait été gratifiante, ludique et qui représentait un défi intellectuel qui invitait à mobiliser et amplifier leurs connaissances. Ceci contraste fortement avec la frustration exprimée par l'un des post-éditeurs linguistes, qui a déclaré que l'expérience manquait de ressources, d'outils d'aide à la traduction et que, dans ces conditions, la qualité ne pouvait pas être assurée. Néanmoins, un des linguistes a trouvé l'expérience également intéressante. Dans tous les cas, ces opinions divisées ne reflètent pas seulement une perception différente entre les deux groupes par rapport au processus de traduction, à la post-édition et à la qualité des textes, mais servent aussi pour expliquer les résultats que nous avons obtenu lors de l'analyse des post-éditions monolingues et de leurs taux de correction.

Finalement, les musiciens ont tous déclaré avoir une opinion très favorable de l'utilisabilité de leurs post-éditions monolingues dans des contextes d'enseignement et d'apprentissage de la musique ; une fois de plus, les linguistes se sont montrés beaucoup moins enthousiastes, l'un d'entre eux déclarant que le produit de l'expérience pourrait être utilisé pour saisir une idée générale du texte. Le deuxième linguiste qui a répondu au questionnaire a déclaré sans nuances que le produit de son travail n'était pas utilisable dans des contextes pédagogiques.

Comme nous pouvons le constater, tout au long du questionnaire les positions des deux groupes sont contrastées. Ceci peut être à l'origine d'une différence marquée dans la perception de l'expérience, des textes et, par conséquent, des post-éditions monolingues produites. Nous signalons une nouvelle fois que ces positions des deux groupes, si différentes entre elles, sont cohérentes avec les résultats obtenus lors de l'évaluation quantitative et qualitative des post-éditions monolingues.

6.2.3 Conclusion : incidence de l'expertise des post-éditeurs sur la correction d'erreurs de la TA

Suite aux évaluations effectuées sur les post-éditions monolingues des deux groupes de participants, nous sommes en mesure de reprendre la deuxième hypothèse que nous avons formulée :

- Dans la pratique de la post-édition monolingue, les experts voient plus de fautes de fidélité que les traducteurs.

Naturellement, nous avons associé la capacité des post-éditeurs de voir des fautes de fidélité avec leur taux de correction étant donné que, dans la tâche de post-édition monolingue que nous avons proposée aux participants, la correction des erreurs passe par leur détection sans avoir recours au texte source. De ce fait, si l'on considère le taux de correction pour le Texte 1 (Figure 1) et pour le Texte 2 (Figure 2), ainsi que le taux global de correction des post-éditeurs (Figure 5), nous pouvons confirmer que, en effet, les experts ont une capacité clairement supérieure à celle des traducteurs/linguistes pour voir et, par conséquent, corriger les fautes de fidélité dans les TA brutes.

Ce phénomène est sans doute associé aux connaissances dans le domaine. Dans les questionnaires post-tâche, nous avons observé chez les linguistes une reconnaissance de leur manque de connaissances spécialisées pour traiter ce type de texte dans le contexte imposé par la tâche de post-édition monolingue. Ils ont identifié et corrigé 27,2 % des erreurs marquées par l'experte bilingue. Chez les musiciens, en revanche, la connaissance dans le domaine leur a permis d'identifier et de corriger 72,3 % des erreurs signalées lors de la phase d'annotation (voir Figure 5). De ce fait, la deuxième hypothèse est validée.

En ce qui concerne nos questions de recherche, nous pourrions uniquement dégager des éléments de réponse concrets dans la phase d'évaluation par les juges qui sera présentée dans la section suivante. Néanmoins, les résultats de cette phase semblent confirmer que l'expertise du post-éditeur monolingue dans le domaine de spécialité peut être un facteur déterminant pour l'utilisabilité des post-éditions et que, en effet, le degré de spécialisation du texte peut avoir un impact sur l'utilisabilité de celles-ci. Cependant, cet impact semble être beaucoup plus marqué chez les linguistes que chez les musiciens (voir Tableau 12). Tout comme nos questions de recherche (et notre première hypothèse), ces éléments pourront uniquement être confirmés ou infirmés suite à l'analyse des réponses données par les experts dans la phase d'évaluation de l'utilisabilité des post-éditions par des juges monolingues.

6.3 Résultats de l'évaluation qualitative de l'utilisabilité par les juges experts

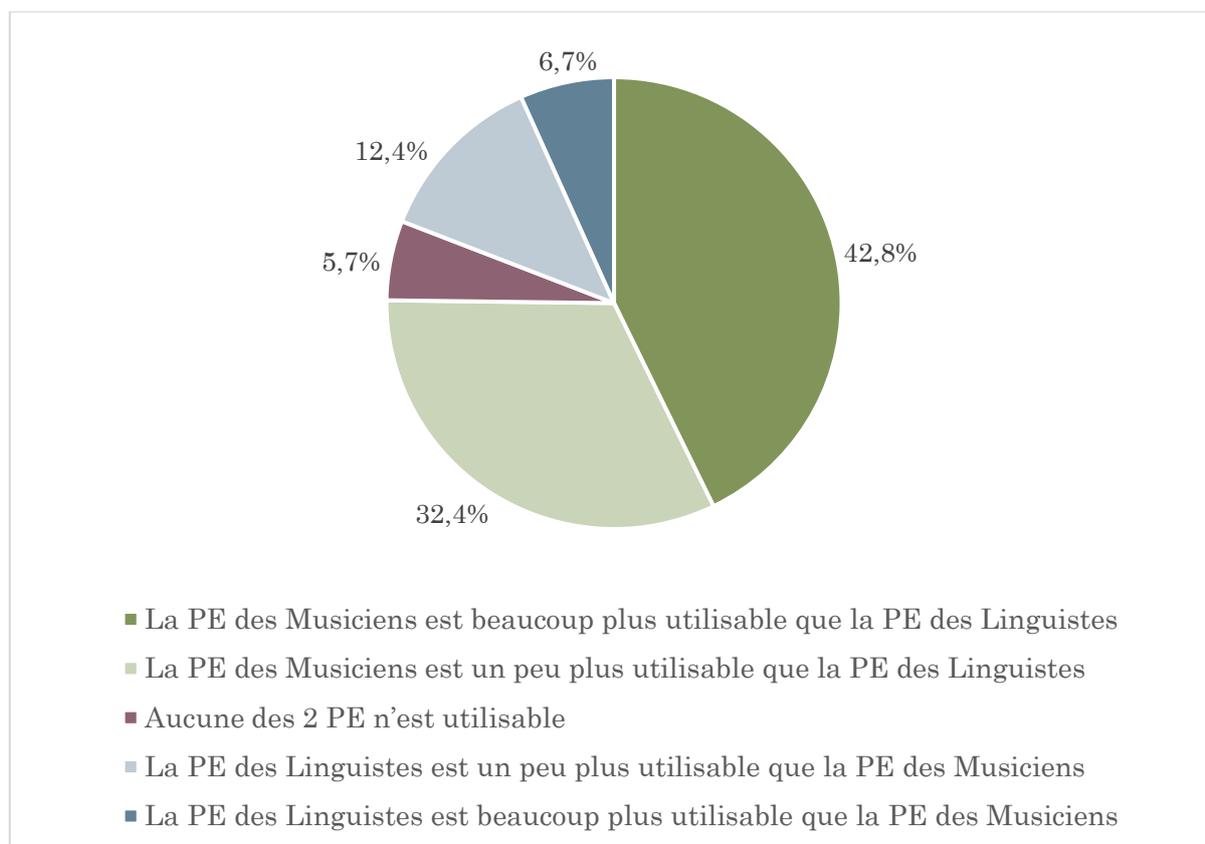
Dans cette section, nous allons d'abord présenter les résultats du jugement des experts sur l'utilisabilité des post-éditions monolingues réalisés par nos participants (section 6.3.1). Ensuite, nous allons analyser les évaluations des post-éditions effectuées lors de la quatrième phase (section 6.2) à la lumière de l'utilisabilité estimée par les juges (section 6.3.2). Finalement, nous reviendrons sur notre première hypothèse et nous offrirons des éléments de réponse à nos questions de recherche sur la base des données qui seront présentées et analysées dans cette section.

6.3.1 Utilisabilité des post-éditions monolingues

6.3.1.1 *Texte non spécialisé (Texte 1)*

Concernant le texte non spécialisé, 42,8 % des réponses des experts indiquait que les post-éditions (PE dans les Figures et les Tableaux) des musiciens étaient « beaucoup plus » utilisables que celles des linguistes, et 32,4 % « un peu plus » utilisable que celles des linguistes. Ces résultats indiquent une préférence marquée des experts pour les post-éditions effectuées par les musiciens. D'autre part, seulement 6,7 % des réponses ont indiqué une nette préférence des experts pour les post-éditions des linguistes par rapport à celles des musiciens ; 12,4 % des réponses ont avancé que les post-éditions des linguistes étaient « un peu plus » utilisables que celles des musiciens. Finalement, 5,7 % des réponses des experts n'ont validé aucune des deux versions en indiquant qu'aucune n'était utilisable. La Figure 6 illustre l'intégralité de ces résultats.

FIGURE 8. UTILISABILITE DES POST-EDITIONS – TEXTE 1



Le Tableau 13 montre le nombre exact de réponses données par les experts pour chacune des catégories avec les pourcentages correspondants.

Utilisabilité	Nbr. de réponses	Pourcentage
La PE des Musiciens est beaucoup plus utilisable que la PE des Linguistes	45	42,8 %
La PE des Musiciens est un peu plus utilisable que la PE des Linguistes	34	32,4 %
Aucune des 2 PE n'est utilisable	6	5,7 %
La PE des Linguistes est un peu plus utilisable que la PE des Musiciens	13	12,4 %
La PE des Linguistes est beaucoup plus utilisable que la PE des Musiciens	7	6,7 %
TOTAL	105	100 %

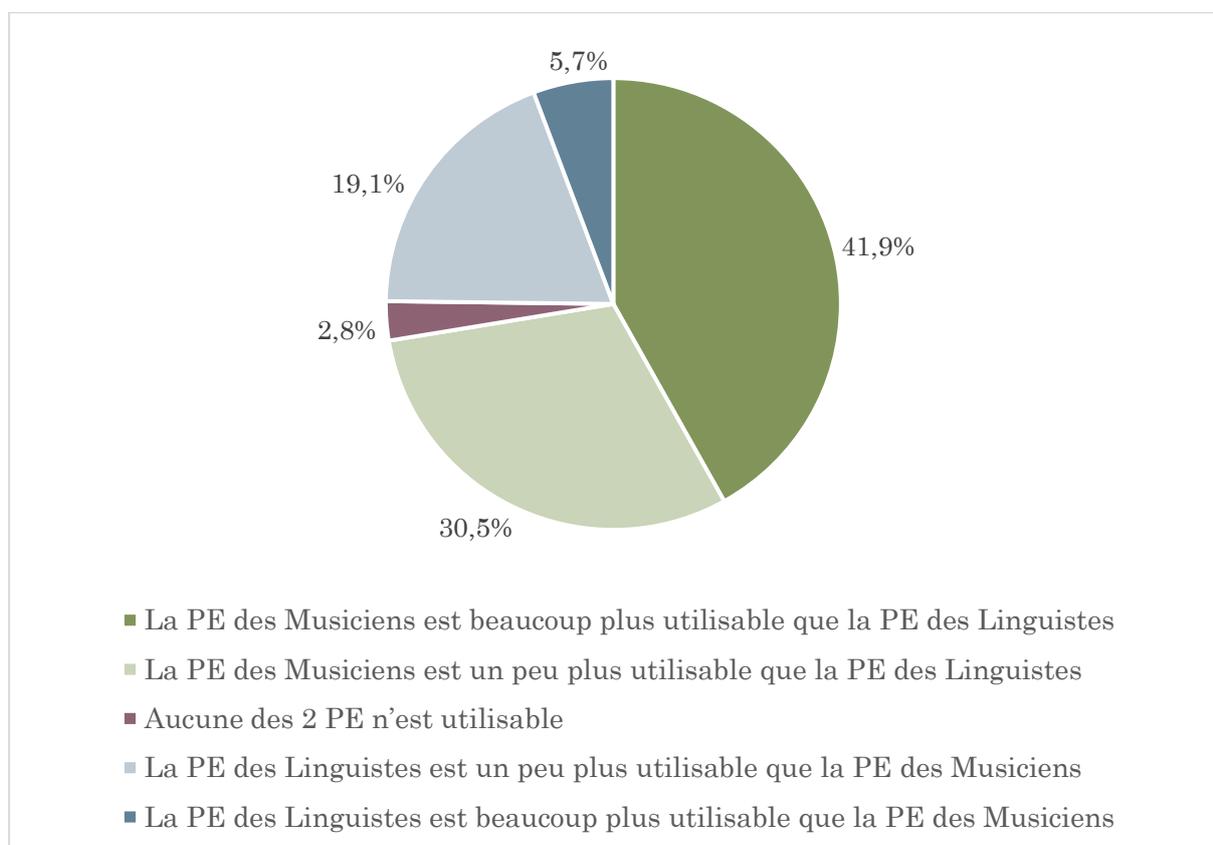
TABLEAU 13. BILAN DES REPOSES DES EXPERTS – TEXTE 1

6.3.1.2 Texte hautement spécialisé (Texte 2)

L'évaluation de l'utilisabilité des post-éditions du texte hautement spécialisé a donné des résultats tout-à-fait similaires à ceux du texte non spécialisé. D'abord, 41,9 % des réponses indiquent que les post-éditions des musiciens sont « beaucoup

plus » utilisables que celles des linguistes, et 30,5 % indiquent qu'elles sont « un peu plus » utilisables. Quant aux post-éditions effectuées par les linguistes, 19,1 % des réponses ont considéré qu'elles étaient « un peu plus » utilisables que celles des musiciens, et seulement 5,7 % « beaucoup plus » utilisables. De ce fait, les experts ont aussi déclaré une préférence pour l'utilisabilité des post-éditions effectuées par les musiciens dans le cas du texte hautement spécialisé, comme l'illustre la Figure 7. Finalement, seulement 2,8 % des réponses ont déclaré qu'aucune des deux post-éditions n'est utilisable.

FIGURE 9. UTILISABILITE DES POST-EDITIONS – TEXTE 2



Tout comme pour le texte précédent, le Tableau 14 montre le nombre exact de réponses données par les experts par catégorie et le pourcentage qu'elles représentent.

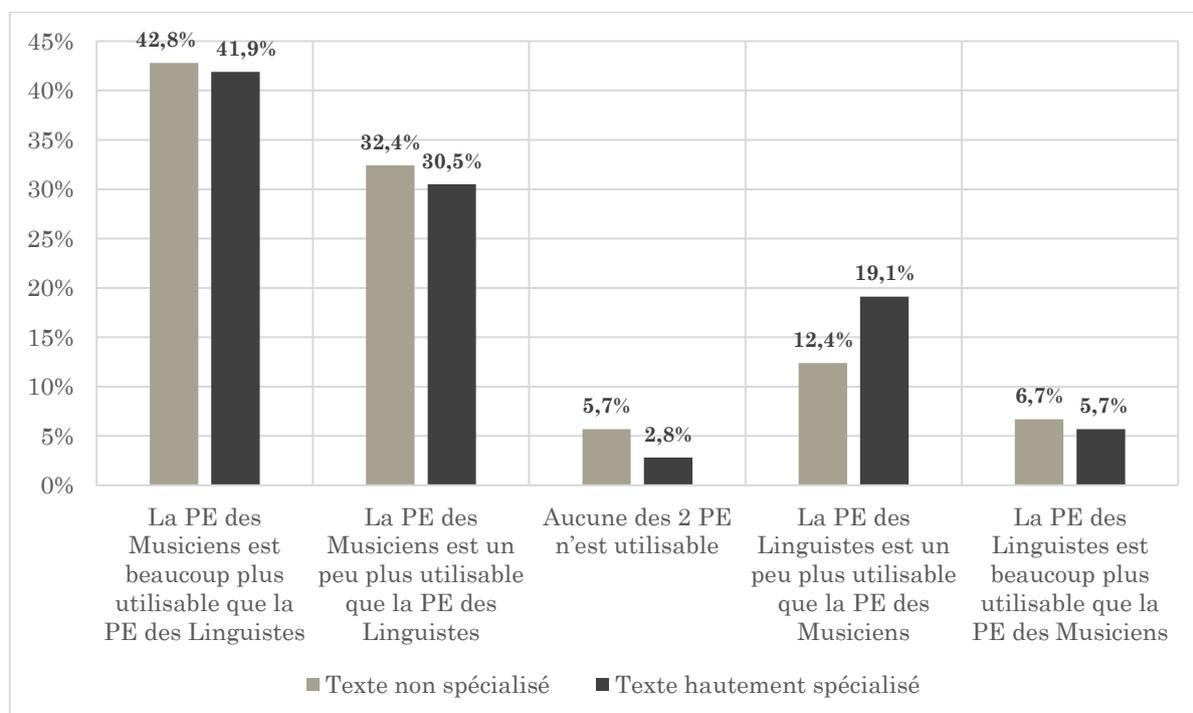
Utilisabilité	Nbr. de réponses	Pourcentage
La PE des Musiciens est beaucoup plus utilisable que la PE des Linguistes	44	41,9 %
La PE des Musiciens est un peu plus utilisable que la PE des Linguistes	32	30,5 %
Aucune des 2 PE n'est utilisable	3	2,8 %
La PE des Linguistes est un peu plus utilisable que la PE des Musiciens	20	19,1 %
La PE des Linguistes est beaucoup plus utilisable que la PE des Musiciens	6	5,7 %
TOTAL	105	100 %

TABLEAU 14. BILAN DES REPNSES DES EXPERTS – TEXTE 2

6.3.1.3 Conclusion

Dans cette section, nous avons présenté les scores d'utilisabilité des post-éditions monolingues d'après l'évaluation effectuée par les experts lors de la cinquième phase de notre expérience. Nous avons pu constater que, pour les deux textes, les experts ont considéré que les post-éditions effectuées par les musiciens sont plus utilisables que celles des linguistes. La Figure 8 synthétise les pourcentages issus des réponses données par les juges dans les questionnaires.

FIGURE 10. UTILISABILITE DES POST-EDITIONS POUR LES DEUX TEXTES



Pour résumer, nous pouvons constater que le jugement des experts n'a pas changé significativement entre le texte non spécialisé et celui hautement spécialisé. Il est néanmoins important de remarquer que la post-édition des linguistes a été jugée « un peu plus » utilisable que celle des musiciens plus fréquemment dans le texte hautement spécialisé que dans le texte non spécialisé. Ce fait pourrait s'expliquer par le type de contenu du texte et le style de post-édition des linguistes. C'est-à-dire, si l'on accepte que, naturellement, les linguistes portent plus d'attention à la forme du texte et aux aspects linguistiques que les musiciens, et que le contenu du texte hautement spécialisé est plus complexe, nous pouvons conclure que les experts apprécient davantage la lisibilité apportée par les linguistes au texte hautement spécialisé.

En conclusion, les experts ont jugé que les post-éditions des musiciens sont beaucoup plus utilisables que celles des linguistes, sans distinction du degré de spécialité du texte. Il semblerait que, en ce qui concerne le texte hautement spécialisé, les experts apprécient davantage la lisibilité d'un texte linguistiquement correct. Néanmoins, les post-éditions effectuées par les musiciens ont été considérées clairement plus aptes et utilisables dans des contextes d'enseignement et d'apprentissage de la musique.

6.3.2 Évaluations quantitative et qualitative des post-éditions monolingues à la lumière de l'utilisabilité estimée par les juges

Dans cette section, nous allons comparer les résultats des évaluations effectuées lors de la quatrième phase (section 6.2) avec les résultats de la phase d'évaluation par les juges (section 6.3.1). Comme nous l'avons indiqué précédemment, cette comparaison entre les valeurs qualitatives, appliquées aux post-éditions lors de la quatrième phase, et le jugement des experts, a pour but d'évaluer la cohérence entre les deux évaluations pour déterminer la validité de notre échelle (voir Tableau 12) à la lumière de l'avis des juges experts sur l'utilisabilité des post-éditions monolingues.

Tout d'abord, le taux de correction obtenu par les musiciens pour les deux textes, ainsi que leur taux global de correction, nous a permis d'affirmer que l'essentiel du sens des post-éditions monolingues avait été transmis grâce aux corrections

appliquées lors de la post-édition. Comme nous pouvons le constater dans la Figure 8, il semblerait que les experts aient validé cette valeur qualitative étant donné que, pour les deux textes, les post-éditions des musiciens ont été jugées « beaucoup » ou « un peu plus » utilisables que celles des linguistes, et donc clairement préférées par les experts. De ce fait, il existe une cohérence entre les deux appréciations.

D'autre part, le taux de correction obtenu par les linguistes, tant pour les deux textes que pour le taux global de correction, correspondait à des valeurs qualitatives moins encourageantes. D'après notre échelle, la post-édition des linguistes déformait le sens de manière importante (notamment pour le Texte 1) ou véhiculait très peu de sens à cause d'un taux de correction trop bas (notamment pour le Texte 2). En ce qui concerne le Texte 1, la tendance semble coïncider entre les pourcentages de la Figure 8 et les valeurs assignées aux post-éditions des linguistes.

Dans ce sens, pour le Texte 2, on pourrait attendre des niveaux d'utilisabilité très bas d'après le jugement des experts. C'est en effet le cas, mais, comme nous l'avons remarqué auparavant, les juges ont estimé que, dans 19,1 % des cas, la post-édition des linguistes est « un peu plus » utilisable que celle des musiciens, ce qui dépasse le pourcentage estimé pour le Texte 1 dans cette même catégorie. De ce fait, on pourrait affirmer que, malgré le fait que la plupart des erreurs n'ont pas été corrigées par les linguistes pour le Texte 2, les experts ont trouvé le texte hautement spécialisé des linguistes plus utilisable que le texte non spécialisé, qui a pourtant un taux de correction plus haut. Malgré cette légère déviation, la valeur qualitative de notre échelle assignée au Taux global de correction des linguistes semble aussi coïncider avec l'avis des experts : dans aucun cas les post-éditions des linguistes n'ont dépassé le score d'utilisabilité des post-éditions des musiciens.

En conséquence, nous pouvons affirmer qu'il existe une cohérence entre notre échelle d'évaluation qualitative, et l'avis des experts exprimé dans les questionnaires lors de la cinquième et dernière étape de notre étude.

6.3.3 Conclusion : incidence de l'expertise des post-éditeurs sur l'utilisabilité des post-éditions

À ce stade, nous sommes en mesure de reprendre la première hypothèse que nous avons formulé auparavant (section 5.1) :

- Plus le post-éditeur est expert dans le domaine, plus il peut rendre utilisable la traduction automatique brute d'un texte de spécialité.

Sur la base des données présentées tout au long des sections précédentes (voir Figure 8 et Tableau 12), nous sommes en mesure de confirmer notre première hypothèse. En effet, les post-éditeurs experts ont non seulement obtenu des taux de correction considérablement plus hauts que les linguistes (voir Figure 5), mais, en plus, leurs post-éditions ont été jugées par les experts nettement plus utilisables dans des contextes d'apprentissage et d'enseignement de la musique (voir Figure 6 et 7). En conséquence, notre première hypothèse est validée.

Avant de conclure ce chapitre et de répondre à nos questions de recherche principales, nous allons effectuer une analyse comparative entre les systèmes de TA employés dans notre étude : Google Translate et DeepL. Cette analyse pourrait nous apporter des nuances et des chiffres concrets qu'il serait intéressant de prendre en compte pour répondre à nos questions de recherche. Nous allons développer cette analyse dans la section suivante.

6.4 Comparaison des systèmes de TA

Dans cette section, nous allons déterminer la quantité et le type des erreurs annotées par l'annotatrice bilingue dans les TA brutes de chaque système afin de les comparer entre elles (section 6.4.1) et répondre à la première question subsidiaire que nous avons formulée :

- Lequel des deux systèmes de TA, Google Translate ou DeepL, fait le moins d'erreurs d'après les annotations de l'annotatrice bilingue ?

Nous allons ensuite analyser les préférences des participants concernant le système de TA (section 6.4.2) et répondre à la deuxième question subsidiaire que nous avons formulée :

- Lequel des deux systèmes est d'avantage choisi par les post-éditeurs pour effectuer la tâche qui leur est présentée ?

Finalement, nous allons proposer une comparaison entre les deux systèmes (section 6.4.3) qui prenne en compte la totalité des données exposées dans cette section.

6.4.1 Quantité et types d'erreurs annotées dans les TA brutes

En ce qui concerne le texte non spécialisé, le Tableau 15 montre que les erreurs sémantiques représentent 60 % des erreurs identifiées pour Google Translate, tandis qu'elles représentent moins de la moitié des erreurs pour DeepL (46,4 %). D'autre part, 50 % des erreurs annotées pour DeepL sont terminologiques. Ce type d'erreur, pour Google Translate, ne représente que 39,3 % du total des erreurs signalées dans la TA produite par ce système.

Type d'erreur	Google Translate		DeepL	
	Erreurs	Pourcentage	Erreurs	Pourcentage
Sémantique	17	60,7 %	13	46,4 %
Terminologique	11	39,3 %	14	50 %
Ajout	0	0 %	0	0 %
Omission	0	0 %	1	3,6 %
Total	28	100 %	28	100 %

TABLEAU 15. NOMBRE D'ERREURS PAR SYSTEME – TEXTE NON SPECIALISE

Pour le texte non spécialisé, Google Translate a donc enregistré plus d'erreurs sémantiques que DeepL, tandis que ce dernier a enregistré plus d'erreurs terminologiques que Google Translate. Aucun des deux systèmes n'a enregistré d'ajout, et seulement une erreur d'omission a été marquée dans la TA brute produite par DeepL. Chaque système a enregistré un total de 28 erreurs.

Pour le texte hautement spécialisé, les systèmes comptabilisent au total moins d'erreurs que celles identifiées pour le texte non spécialisé (23 pour Google Translate et 19 pour DeepL). Ce résultat est certainement surprenant mais pourrait être dû en partie au fait que le texte non spécialisé a un nombre de mots un peu plus élevé que le texte hautement spécialisé (voir Tableau 3, section 5.2.2.2 du chapitre 5), ce qui pourrait entraîner davantage d'erreurs. D'ailleurs, la

différence la plus marquée entre les deux systèmes est le nombre d'erreurs sémantiques (voir tableaux 15 et 16). En effet, le sens univoque et transparent d'un texte équivalant à la vulgarisation scientifique peut s'avérer plus vulnérable à ce type d'erreurs qu'un texte hautement spécialisé, ce dernier étant adressé à un public avec des connaissances et des outils conceptuels pour comprendre l'idée du texte malgré les défauts possibles de la TA.

Bien qu'avec des pourcentages plus bas, la tendance montrée pour le texte non spécialisé, par rapport aux erreurs sémantiques, s'est maintenue pour le texte hautement spécialisé. C'est-à-dire que Google Translate a enregistré plus d'erreurs sémantiques que DeepL : 39,2 % pour le premier et 31,6 % pour ce dernier, comme nous pouvons le constater dans le Tableau 16.

Type d'erreur	Google Translate		DeepL	
	Erreurs	Pourcentage	Erreurs	Pourcentage
Sémantique	9	39,2 %	6	31,6 %
Terminologique	13	56,5 %	11	58 %
Ajout	1	4,3 %	1	5,2 %
Omission	0	0 %	1	5,2 %
Total	23	100 %	19	100 %

TABLEAU 16. NOMBRE D'ERREURS PAR SYSTEME – TEXTE HAUTEMENT SPECIALISE

Néanmoins, dans ce cas, Google Translate a aussi enregistré plus d'erreurs terminologiques que DeepL (13 pour Google Translate contre 11 pour DeepL). Cependant, si l'on observe par système le pourcentage total d'erreurs terminologiques dans les deux textes, celui de DeepL (58 %) est légèrement plus haut que celui de Google Translate (56,5 %). Dans ce sens, concernant les erreurs terminologiques, la tendance observée dans le texte non spécialisé s'est aussi maintenue pour celui hautement spécialisé. De plus, chaque système a enregistré un seul ajout (représentant 4,3 % des erreurs annotées pour Google Translate et 5,2 % pour DeepL), et, comme dans le texte précédent, seulement DeepL a enregistré une erreur d'omission (5,2 %).

En résumé, nous sommes en mesure d'affirmer que la plupart des erreurs annotées pour les deux systèmes sont d'ordres sémantique et terminologique. Pour Google

Translate, les erreurs sémantiques représentent 51 % du total d'erreurs (26 erreurs au total), tandis que pour DeepL ils représentent 40,4 % (19 erreurs au total). À l'inverse, les erreurs terminologiques représentent 53,2 % pour DeepL (25 erreurs au total), face à 47 % enregistré pour Google Translate (24 erreurs). Chaque système a enregistré un seul ajout, et seulement DeepL a enregistré 2 omissions (4,2 % au total).

Type d'erreur	Google Translate		DeepL	
	Erreurs	Pourcentage	Erreurs	Pourcentage
Sémantique	26	51 %	19	40,4 %
Terminologique	24	47 %	25	53,2 %
Ajout	1	2 %	1	2,2 %
Omission	0	0 %	2	4,2 %
Total	51	100 %	47	100 %

TABLEAU 17. NOMBRE TOTAL D'ERREURS PAR SYSTEME DES DEUX TEXTES

Du point de vue quantitatif, Google Translate a enregistré au total plus d'erreurs que DeepL (51 contre 47), mais nous constatons que les différences entre les deux systèmes sont minimes. Ceci peut s'expliquer par le fait qu'il s'agisse de deux systèmes de TA concurrents sur le marché et avec le même type d'architecture. Nous constatons donc des performances relativement similaires qui ont toutefois tendance à favoriser le système DeepL. Dans la section suivante, nous allons analyser dans quelle mesure ceci s'est reflété dans les choix du système effectués par les post-éditeurs lors de la tâche de post-édition.

6.4.2 Choix du système de TA effectué par les post-éditeurs

Concernant le choix des systèmes, le Tableau 18 montre que les musiciens ont choisi d'avantage la TA produite par DeepL (60 %) par rapport à celle de Google Translate (40 %). Nous pouvons constater les choix effectués par chaque post-éditeur dans le Tableau 18. Il est en effet intéressant de voir que les musiciens ont choisi de post-éditer au total 9 segments issues de DeepL et 6 de Google translate, et cela pour les deux textes (indépendamment de leur degré de spécialité).

Post-éditeurs musiciens	Google Translate		DeepL	
	TEXTE 1	TEXTE 2	TEXTE 1	TEXTE 2
A	1	1	4	4
B	3	4	2	1
C	2	1	3	4
Total de segments :	6/15	6/15	9/15	9/15
% Total par texte :	40 %	40 %	60 %	60 %
Total Global :	12/30		18/30	
% Global :	40 %		60 %	

TABLEAU 18. TOTAL ET POURCENTAGE DES SEGMENTS CHOISIS PAR LES MUSICIENS

Du côté des linguistes, concernant le choix du système, la tendance observée chez les musiciens est amplifiée. Ils ont également choisi d'avantage DeepL, dans 80 % des cas, décidant de post-éditer uniquement 20 % des segments de Google Translate qui leur ont été présentés. La Tableau 19 nous permet de constater aussi que, concernant le type de système choisi pour chaque texte, les linguistes ont sélectionné DeepL 10 fois, et Google Translate 5 fois, pour le texte non spécialisé, tandis que, pour le texte hautement spécialisé, ils ont choisi 14 fois sur 15 la TA produite par DeepL.

Post-éditeurs linguistes	Google Translate		DeepL	
	TEXTE 1	TEXTE 2	TEXTE 1	TEXTE 2
A	0	1	5	4
B	3	0	2	5
C	2	0	3	5
Total de segments :	5/15	1/15	10/15	14/15
% Total par texte :	33,3 %	6,6 %	66,7 %	93,4 %
Total Global :	6/30		24/30	
% Global :	20 %		80 %	

TABLEAU 19. TOTAL ET POURCENTAGE DES SEGMENTS CHOISIS PAR LES LINGUISTES

Cette préférence marquée des linguistes pour la TA de DeepL dans le cas du texte hautement spécialisé pourrait être dû au fait que, comme ils l'ont exprimé lors des questionnaires post-tâche, le contenu des textes dépassait leurs connaissances et, donc, il était difficile pour eux de repérer des erreurs (particulièrement dans une tâche de post-édition monolingue, sans texte source). Comme nous l'avons signalé

auparavant (voir Tableau 16), la TA de DeepL pour le texte hautement spécialisé présente moins d'erreurs sémantiques. Ce type d'erreurs est possiblement plus facile à détecter par des non spécialistes que les erreurs terminologiques dans un texte spécialisé. De ce fait, un taux plus bas d'erreurs, éventuellement plus détectables, pourrait expliquer cette préférence des linguistes pour la TA produite par DeepL. Néanmoins, il faut rappeler que le taux de correction des linguistes pour le texte hautement spécialisé est très bas (voir Figure 2). Mais, ce taux a également été déclaré « un peu plus » utilisable que le texte des musiciens, plus souvent que ce qu'on pourrait attendre (voir Figure 8).

6.4.3 Conclusion de la comparaison entre les systèmes

Dans cette section, nous avons présenté les choix de systèmes de TA effectués par les musiciens et les linguistes lors de la tâche de post-édition proposée dans notre expérience. Nous pouvons donc répondre aux deux questions subsidiaires que nous avons formulées par rapport aux systèmes de TA employés.

En ce qui concerne la première question (qui interroge lequel des deux systèmes fait le moins d'erreurs d'après les annotations de l'annotatrice bilingue), nous pouvons affirmer que, au total, DeepL a fait moins d'erreurs que Google Translate (47 erreurs de DeepL contre 51 erreurs de Google Translate, voir Tableau 16). Si l'on va plus loin dans l'analyse, nous constatons que DeepL fait moins d'erreurs que Google translate dans une seule catégorie d'erreurs (Sémantiques), enregistrant 19 erreurs au total contre 26 de Google Translate. Néanmoins, DeepL fait un peu plus d'erreurs que Google Translate pour deux catégories d'erreurs : Terminologiques (25 de DeepL contre 24 de Google Translate) et erreurs d'omission (2 de DeepL, mais aucun enregistré pour Google Translate). Concernant les ajouts, chaque système a enregistré une seule erreur. En conséquence, d'après les annotations de l'experte bilingue, le léger avantage de DeepL se joue au niveau de la quantité d'erreurs sémantiques identifiées dans les TA de Google Translate.

Concernant la deuxième question cherchant à savoir lequel des deux systèmes est davantage choisi par les post-éditeurs pour effectuer la tâche qui leur est présentée, nous pouvons d'abord constater dans le Tableau 20 que les post-éditeurs (3 musiciens et 3 linguistes) ont post-édité 60 segments au total (étant donné que

chaque participant devait post-éditer 10 segments, 5 pour chaque texte, voir Tableau 6, section 5.3.3). En ce sens, les segments traduits par DeepL ont été choisis 42 fois par les post-éditeurs (représentant 70 % des segments totaux) et ceux traduits par Google Translate 18 fois (30 % des segments).

	Google Translate		DeepL	
	TEXTE 1	TEXTE 2	TEXTE 1	TEXTE 2
Total de segments :	11/30	7/30	19/30	23/30
% Total par texte :	36,6 %	23,3 %	63,4 %	76,7 %
Total Global :	18/60		42/60	
% Global :	30 %		70 %	

TABLEAU 20. TOTAL ET POURCENTAGE DES SEGMENTS CHOISIS PAR LES DEUX GROUPES DE POST-EDITEURS

Ainsi, en réponse à la deuxième question, ces résultats globaux confirment donc une préférence générale pour DeepL, tendance qui a été également observée dans les données par groupe de post-éditeurs présentées dans les tableaux 18 et 19.

Nous sommes maintenant en mesure d'aborder la conclusion de ce chapitre et de répondre à nos questions de recherche sur la base des résultats présentés tout au long de cette section.

6.5 Conclusion

Dans ce chapitre, nous avons présenté les résultats obtenus pour notre expérience. Nous allons, pour conclure, faire un bilan des réponses aux questions subsidiaires et aux hypothèses pour, finalement, répondre à nos deux questions de recherche principales.

Tout d'abord, en ce qui concerne les systèmes de TA employés, DeepL a enregistré moins d'erreurs au total que Google Translate. Concernant le texte non spécialisé, les deux systèmes ont enregistré le même nombre d'erreurs : 28 chacun (Tableau 15). Quant au texte hautement spécialisé, DeepL a enregistré un total de 19 erreurs contre 23 avec Google Translate (Tableau 16). Le bilan total montre donc que DeepL a fait 47 erreurs et Google Translate 51 (Tableau 17). Si l'on considère les types d'erreurs annotées pour chaque système, la différence la plus significative

se trouve au niveau des erreurs sémantiques annotées dans les TA de Google Translate.

DeepL a été également choisi davantage par les post-éditeurs. Au total, dans 70 % des cas, les participants ont choisi de post-éditer les TA de DeepL et dans 30 % celles de Google Translate. Nous constatons donc une cohérence entre le système qui a présenté le moins d'erreurs annotées par l'experte bilingue et le système davantage choisi par les post-éditeurs pour effectuer la tâche qui leur a été présentée.

Quant à notre première hypothèse, nous avons observé que, dans le cas des traités de musique, plus le post-éditeur est expert dans le domaine, plus il peut rendre utilisable la traduction automatique brute d'un texte de spécialité. En effet, d'après le jugement des experts, les scores d'utilisabilité des post-éditions faites par les musiciens sont au-dessus du 30 % pour les deux textes, tandis que ce même score pour les post-éditions des linguistes ne dépasse dans aucun cas 19,1 %. De ce fait, nous avons validé notre première hypothèse.

Nous avons également validé la deuxième hypothèse. Les experts voient en effet plus de fautes de fidélité que les linguistes. D'après les observations constatées dans les post-éditions de chaque groupe, les corrections effectuées par les musiciens ont un taux de correction global de 72,3 %, tandis que celui des linguistes est de 27,2 % (Tableau 12). En conséquence notre deuxième hypothèse est elle aussi validée car plus le post-éditeur est capable de voir des fautes, plus il peut les corriger et, donc, élever le taux de correction des post-éditions par rapport aux erreurs marquées par l'annotatrice.

Nous avons formulé deux questions de recherche principales auxquelles nous sommes, à ce stade, en mesure de répondre. La première question cherchait à savoir dans quelle mesure l'expertise du post-éditeur monolingue dans le domaine de spécialité est un facteur déterminant pour l'utilisabilité des post-éditions de traités de musique. Pour répondre à cette question, nous allons analyser les scores d'utilisabilité non nuancés provenant des réponses données par les experts aux questionnaires post-tâche. Pour cela, nous avons regroupé toutes les réponses des

experts qui ont considéré « plus utilisable » la post-édition des musiciens, d'une part, et, de l'autre, toutes celles qui ont préféré les linguistes, sans prendre en compte si les juges ont considéré l'une « un peu plus » ou « beaucoup plus » utilisable que l'autre.

Profil de post-éditeur	Réponses	Pourcentage
PE des musiciens plus utilisable que celle des linguistes	165	78,6 %
Aucune des 2 PE n'est utilisable	9	4,3 %
PE des linguistes plus utilisable que celle des musiciens	36	17,1 %

TABLEAU 21. BILAN TOTAL DE FAVORABILITE NON NUANCEE

Nous constatons, dans la Tableau 21, que 78,6 % des post-éditions faites par les musiciens ont été considérées plus utilisables que celles des linguistes et, inversement, 17,1 % des post-éditions des linguistes ont été considérées plus utilisables que celles effectuées par les musiciens. Par conséquent, et tout en considérant la validité de nos deux hypothèses, nous pouvons affirmer que l'expertise du post-éditeur dans le domaine de spécialité du texte détermine l'utilisabilité des post-éditions monolingues. Dans le cas de notre recherche, cela signifie que la proportion de post-éditions monolingues effectuées par les musiciens qui ont été considérées par les experts comme « plus utilisables » est de 61,5 points de pourcentage supérieure à celle des post-éditions effectuées par les linguistes.

Ces pourcentages sont cohérents avec les réponses données par les post-éditeurs dans le questionnaire post-tâche. Les linguistes ont considéré que leurs post-éditions n'étaient pas fiables à cause de leur manque de connaissances ; c'est l'opposé concernant les musiciens qui, quant à eux, ont déclaré qu'ils considèrent leurs post-éditions fiables, ou partiellement fiables. En plus, l'expérience pour ces derniers a été satisfaisante, ludique et stimulante, tandis que les linguistes ont déclaré avoir expérimenté de la frustration à cause de l'impossibilité d'assurer la qualité de la post-édition, malgré l'intérêt que l'expérience a suscité chez eux. En résumé, nous confirmons que, dans le domaine de spécialité de la musique, l'expertise des post-éditeurs joue un rôle essentiel, tant pour la perception de la

qualité finale et de la tâche de post-édition que pour l'utilisabilité des textes produits grâce à la post-édition monolingue.

La deuxième question interroge l'impact du degré de spécialisation du texte sur l'utilisabilité des post-éditions produites par les participants de notre étude. Il existe, selon nous, deux façons d'aborder la question. La première consiste à se centrer sur l'impact du degré de spécialité des textes à propos du score d'utilisabilité atteint par chaque groupe. À cet égard, nous avons indiqué auparavant que, indépendamment du degré de spécialité des textes, les post-éditions des musiciens sont toujours considérées comme « beaucoup plus » ou « un peu plus » utilisables que celles des linguistes (voir Figure 8). La deuxième consiste à prendre le pourcentage de réponses qui ont considéré les post-éditions des deux groupes comme inutilisables et les comparer afin de déterminer lequel des deux textes a le plus de réponses de ce type. Cette approche de la question assume que tant que l'un des deux groupes arrive à produire des segments utilisables, l'impact du degré de spécialité du texte sur les post-éditions produites est très bas.

	Texte non spécialisé		Texte hautement spécialisé	
	Réponses	Pourcentage	Réponses	Pourcentage
Aucune des 2 PE n'est utilisable	6/105	5,7 %	3/105	2,8 %

TABLEAU 22. BILAN DES « NON UTILISABLES » PAR TEXTE

Dans le Tableau 22, nous pouvons constater que, en général, le pourcentage de réponses considérant qu'aucune des 2 post-éditions n'est utilisable est très faible pour les deux textes. Nous observons néanmoins que, en effet, le texte hautement spécialisé obtient un pourcentage de ces réponses encore plus faible que celui du texte non spécialisé (2,8 % pour l'un et 5,7 % pour l'autre). Contrairement à ce qu'on aurait pu supposer, l'impact du texte hautement spécialisé sur l'utilisabilité des post-éditions monolingues est plus bas que celui du texte non spécialisé.

Ceci pourrait être expliqué par les différences dans la situation communicative établie par les textes. En effet, le but d'un texte non spécialisé est de communiquer aux débutants, de façon claire et précise, des notions de base sur la musique. Naturellement, les experts auront tendance à juger plus fortement des erreurs

sémantiques ou terminologiques dans ce type de texte que dans un texte hautement spécialisé, et donc d'estimer plus souvent qu'aucune des post-éditions n'est utilisable dans le texte non spécialisé. Rappelons que le texte hautement spécialisé s'adresse aux spécialistes qui ont plus d'outils techniques et conceptuels que les débutants pour comprendre le sens d'une post-édition monolingue, même si sa qualité n'est pas optimale. De ce fait, les spécialistes pourront, par exemple, élucider le sens d'un passage comprenant des erreurs des post-édition en utilisant les extraits de notation musicale, très présents dans ce type de textes (comme nous l'avons décrit et illustré dans la section 4.2.1), ce qui va au-delà des possibilités des débutants. En conséquence, plus le texte est spécialisé, plus les post-éditions monolingues peuvent être considérées utilisables grâce à la situation communicative et au lecteur auquel s'adresse chaque texte.

Dans cette section, nous avons présenté, illustré et analysé les données issues de notre expérience. Ces données ont apporté des éléments de réponse à nos questions subsidiaires et nous ont permis de valider nos hypothèses. Mais, avant tout, ces données nous ont permis de répondre aux questions principales de cette recherche. Nous pouvons donc conclure notre travail dans la section suivante.

7. Conclusion

7.1 Synthèse et résultats

L'objectif de ce mémoire était d'explorer le potentiel et les limites de la post-édition monolingue pour des textes de spécialité dans le domaine de la musique à travers une étude expérimentale menée avec des participants humains. Nous nous sommes intéressés à l'utilisabilité des post-éditions en fonction du profil du post-éditeur et du degré de spécialité des textes.

Tout d'abord, nous avons adopté une approche historique de la TA et nous avons décrit les principales architectures des systèmes de TA (chapitre 2). Ceci nous a permis de contextualiser notre étude et de mieux comprendre les enjeux propres à l'expérience que nous avons menée.

Ensuite, nous avons présenté les études les plus récentes sur la post-édition monolingue, le profil des post-éditeurs et l'utilisabilité de la TA et des post-éditions (chapitre 3). Nous avons ainsi pu montrer l'originalité de notre travail et les points communs avec d'autres études.

Nous avons également exploré les caractéristiques des traités de musique, notre typologie textuelle (chapitre 4). Ce chapitre nous a permis de porter un regard descriptif sur les traits essentiels des traités de musique et de nous interroger sur la possibilité d'utiliser des systèmes neuronaux dans ce contexte.

Par la suite, nous avons décrit en détail la méthodologie suivie pour l'expérience tout en précisant l'objectif, les questions de recherche principales et subsidiaires et les hypothèses auxquelles on souhaitait répondre (chapitre 5). Ceci nous a permis de justifier les choix méthodologiques concernant les textes, le profil des participants, les outils et les métriques que nous avons choisis pour la préparation et le déroulement de l'expérience.

Finalement, nous avons présenté les résultats de l'expérience (chapitre 6). Dans ce chapitre, nous avons répondu aux questions de recherche, et nous avons confirmé nos hypothèses, sur la base des résultats de l'expérience.

Ce travail a montré que l'expertise du post-éditeur monolingue est essentielle pour l'utilisabilité des post-éditions de traités de musique. Au total, les experts ont considéré que les post-éditions des musiciens (78,6 %) ont été davantage utilisables que celles des linguistes (17,1 %), en marquant une différence de 61,5 points de pourcentage entre les deux groupes de post-éditeurs. Ainsi, des facteurs tels que la connaissance dans le domaine, l'aisance avec la notation musicale et la familiarité avec la terminologie liée au phénomène de la musique sont déterminants pour identifier et corriger les erreurs de la TA sans recourir au texte source, c'est-à-dire, pour effectuer des post-éditions monolingues utilisables.

La notion d'utilisabilité, dans ce contexte, est liée à la fonction pédagogique de ce type de texte. Les post-éditions produites par des post-éditeurs musiciens sont donc plus utilisables que celles des post-éditeurs linguistes dans des contextes d'enseignement et d'apprentissage de la musique. Ainsi, le profil du post-éditeur est un facteur déterminant pour l'utilisabilité des post-éditions monolingues, tant pour les textes hautement spécialisés que pour ceux non spécialisés, équivalents aux textes de vulgarisation scientifique.

En définitive, nous pouvons conclure que, dans le contexte de notre expérience, la TA peut être partiellement comprise comme un outil d'aide pour « décoder » des textes de spécialité dans le domaine de la musique, écrits en langue russe, en les traduisant vers l'espagnol. Si l'objectif est de rendre utilisables les TA dans les contextes d'utilisation propres à notre type de texte, la post-édition devrait être effectuée par un musicien étant donné que l'expertise dans le domaine de la musique est un facteur essentiel dans ce contexte.

7.2 Limites et perspectives

Notre expérience présente plusieurs limites dérivées de trois choix méthodologiques que nous avons privilégiés.

La première limite, liée aux textes, est la longueur des échantillons textuels que nous avons choisis. Nos textes comptent entre 200 et 230 mots, ce qui, pour ce type de textes, représente un échantillon assez réduit. Les traités de musique sont, en général, des textes explicatifs longs, complexes et qui établissent des relations

transtextuelles qui forcément échappent à nos échantillons. Néanmoins, nous avons choisi des passages qui, dans une certaine mesure, reflètent ces caractéristiques et qui sont plutôt représentatifs de cette typologie textuelle.

La deuxième limite, liée au plan expérimental (deuxième phase, section 5.3.2), concerne le nombre d'annotateurs des fautes de la TA. Pour cette expérience, nous n'avons pu recruter qu'une seule annotatrice, russophone native et bilingue en espagnol. Même s'il s'agit d'une experte reconnue dans le domaine, avec plusieurs années d'expérience dans l'enseignement et une formation de premier niveau, nous aurions voulu pouvoir recruter plus d'annotateurs avec ce profil pour contraster et étoffer les annotations. Ceci nous aurait permis d'évaluer les post-éditions des participants sur une base plus diverse et encore plus fiable. Dans ce cas, la rareté du profil recherché pour cette tâche et le temps dont nous avons disposé ont imposé cette limite.

La troisième limite, également liée au plan expérimental (cinquième phase, section 5.3.5), est la méthode employée pour les évaluations du résultat des post-éditions par des experts du domaine. Nous avons opté pour une méthode qui compare les segments de texte provenant des deux groupes de post-éditeurs, choisis aléatoirement mais présentés aux experts en respectant l'ordre original des textes. Cela permettait aux juges de contraster les post-éditions des linguistes avec celles des musiciens sans devoir évaluer la totalité des post-éditions effectuées, ce qui aurait impliqué d'évaluer plusieurs versions de chaque texte. Cependant, cela a également impliqué que les experts n'ont pas comparé des textes post-édités par une même personne de chaque groupe, mais des fragments effectués par des personnes différentes. De ce fait, la cohésion textuelle entre les segments des textes présentés aux experts n'était pas assurée. Nous avons cependant décidé de procéder ainsi pour prioriser le jugement spontané des experts et éviter leur fatigue.

Malgré ces limites, l'objectif que nous nous étions fixé est accompli. Une des perspectives possibles de notre étude est de continuer à explorer cette typologie textuelle avec les outils et les méthodes des technologies langagières dans un projet de thèse interdisciplinaire qui s'intéresserait à une caractérisation plus

approfondie de la typologie textuelle, aux enjeux de la traduction spécialisée de ce type de texte et à la place des outils d'aide à la traduction dans ce contexte.

Un tel projet nous permettrait non seulement d'aborder des dimensions plus ambitieuses, qui dépassent les limites de ce mémoire de master, mais aussi de continuer à labourer les chemins croisés entre les technologies langagières, les humanités et les arts. Projet vaste et innovant qui, sans doute, porterait ses fruits.

Bibliographie

- Allen, Jeffrey. 2003. « Post-editing ». Dans *Computers and Translation. A translator's guide*, 297-318. Amsterdam : John Benjamins Publishing Company.
- Ayvazyan, Nune, et Anthony Pym. 2017. « West enters East: A strange case of unequal equivalences in Soviet translation theory ». Dans *Going East: Discovering New and Alternative Traditions in Translation Studies*, 221-45. Berlin : Frank & Timme.
- Barbosa de Lima Fonseca, Norma. 2019. « Analysing the impact of TAPs on temporal, technical and cognitive effort in monolingual post-editing ». *Perspectives. Studies in Translation Theory and Practice* 27 (4) : 552-88. <https://doi.org/10.1080/0907676X.2019.1597909>.
- Benoit-Otis, Marie-Hélène. 2013. « (Re)traduire les écrits de compositeurs : L'exemple de *Das Judentum in der Musik* de Wagner. » Dans *Écrits de compositeurs : une autorité en questions (XIXe et XXe siècles)*, 309-23. Paris : Vrin.
- Berman, Antoine. 1995. *L'Épreuve de l'étranger : Culture et traduction dans l'Allemagne romantique*. Paris : Gallimard.
- Bonet Heras, Josep. 2004. « La necesaria especialización del traductor técnico ». Dans *Manual de documentación y terminología para la traducción especializada*, édité par Valentín García Yebra et Consuelo Gonzalo García, 37-48. Madrid : Arco Libros.
- Castilho, Sheila, et Sharon O'Brien. 2016. « Evaluating the Impact of Light Post-Editing on Usability ». Dans *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, 310-16. Portorož, Slovénie : European Language Resources and Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2016/pdf/539_Paper.pdf.
- Castilho, Sheila, Sharon O'Brien, Fabio Alves, et Morgan O'Brien. 2014. « Does post-editing increase usability? A study with Brazilian Portuguese as Target

Language ». Dans *Proceedings of the 17th Annual conference of the European Association for Machine Translation*, 183-90. Dubrovnik, Croatie : European Association for Machine Translation.

<https://www.aclweb.org/anthology/2014.eamt-1.40>.

Chan, Sin-wai. 2015. « The development of translation technology 1967–2013 ». Dans *The Routledge encyclopedia of translation technology*, 3-31. Londres : Routledge.

Ciapuscio, Guiomar, et Inés Kuguel. 2002. « Hacia una tipología del discurso especializado: aspectos teóricos y aplicados ». Dans *Entre la terminología, el texto y la traducción*, édité par Joaquín García Palacios et María Teresa Fuentes, 37-73. Salamanca : Almar.

Doherty, Stephen, et Sharon O'Brien. 2014. « The Usability of Raw Machine Translated Output: A User-Centered Study Using Eye Tracking ». *International Journal of Human-Computer Interaction* 30 (1) : 40-51.
<https://doi.org/10.1080/10447318.2013.802199>.

Forcada, Mikel L. 2017. « Making sense of neural machine translation ». *Translation Spaces* 6:2 : 291-309. <https://doi.org/10.1075/ts.6.2.06for>.

Frolova-Walker, Marina, Jonathan Powell, Rosamund Bartlett, Izaly Zemtsovsky, Mark Slobin, Jarkko Niemi, et Yuri Sheikin. 2001. « Russian Federation (Russ. Rossiiskaya Federatsiya; Russia) ». Dans *Grove Music Online*.
<https://doi.org/10.1093/gmo/9781561592630.article.40456>.

Genette, Gérard. 1982. *Palimpsestes*. Poétique. Paris : Seuil.

Hearne, Mary, et Andy Way. 2011. « Statistical Machine Translation: A Guide for Linguists and Translators ». *Language and Linguistics Compass* 5 (5) : 205-26.

Hu, Chang, Philip Resnik, Yakov Kronrod, Vladimir Eidelman, Olivia Buzek, et Benjamin B. Bederson. 2012. « The Value of Monolingual Crowdsourcing in a Real-World Translation Scenario: Simulation using Haitian Creole Emergency SMS Messages ». Dans *Proceedings of the 6th Workshop on*

- Statistical Machine Translation*, 399-404. Edinburgh, Écosse : Association for Computational Linguistics.
- <http://www.cs.umd.edu/hcil/monotrans/publications/wmt11monotrans.pdf>.
- Hu, Ke, et Patrick Cadwell. 2016. « A Comparative Study of Post-editing Guidelines ». *Baltic Journal of Modern Computing* 4 (2) : 346-53.
- Hutchins, W. John. 2015. « Machine Translation. History of research and applications ». Dans *The Routledge encyclopedia of translation technology*, 120-36. Londres : Routledge.
- ISO 9241-11:2018. 2018. *Ergonomie de l'interaction homme-système — Partie 11 : Utilisabilité — Définitions et concepts*. 2^e éd. Vernier, Genève, Suisse : ISO 2018. <https://www.iso.org/fr/standard/63500.html>.
- ISO 18587:2017. 2017. *Services de traduction — Post-édition d'un texte résultant d'une traduction automatique — Exigences*. 1^{re} éd. Vernier, Genève, Suisse : ISO 2017. <https://www.iso.org/fr/standard/62970.html>.
- Jurafsky, Dan, et H. Martin James. 2009. *Speech and Language Processing*. Second Edition. New York : Prentice Hall.
- Koehn, Philipp. 2010. « Enabling Monolingual Translators: Post-Editing vs. Options ». Dans *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings*, 537-45. Los Angeles, Californie : Association for Computational Linguistics. <https://www.aclweb.org/anthology/N10-1078/>.
- . 2021. « The State of Neural Machine Translation (NMT) ». Omniscien Technologies. *Omniscien Blog* (blog). 2021. <https://omniscien.com/blog/state-neural-machine-translation-nmt/>.
- Koehn, Philipp, et Rebecca Knowles. 2017. « Six Challenges for Neural Machine Translation ». Dans *Proceedings of the First Workshop on Neural Machine Translation*, 28-39. Vancouver : Association for Computational Linguistics. <http://www.aclweb.org/anthology/W17-3204>.

- Koponen, Maarit, et Leena Salmi. 2015. « On the correctness of machine translation: A machine translation post-editing task ». *The Journal of Specialised Translation* 23 : 118-36.
- Läubli, Samuel, Chantal Amrhein, Patrick Düggelin, Beatriz Gonzalez, Alena Zwahlen, et Martin Volk. 2019. « Post-editing Productivity with Neural Machine Translation: An Empirical Assessment of Speed and Quality in the Banking and Finance Domain ». Dans *Proceedings of Machine Translation Summit XVII*, Volume 1: Research Track:267-72. Dublin, Irlande. <https://doi.org/arXiv:1906.01685>.
- Lee, Tong King. 2015. « Technological Interventions in Literary Meaning: A Case of Machine Translation ». Dans *Translation and Cross-Cultural Communication Studies in the Asia Pacific*, 42 : 101-14. Approaches to Translation Studies. Brill : Leiden.
- León, Jacqueline. 2008. « La traduction automatique I : les premières tentatives jusqu'au rapport ALPAC ». Dans *History of the Language Sciences*, 3 : 2767-73. Handbooks of Linguistics and Communication Science. Berlin • New York : De Gruyter Mouton.
- Leßmann, Benedikt. 2020. « Übersetzung - ein Thema der Musikforschung? » Dans *Zwischen Transfer und Transformation: Horizonte der Rezeption von Musik*, édité par Michele Calella et Benedikt Leßmann, 147-73. Band 51. Vienne : Hollitzer Verlag.
- Melby, Alan K. 2019. « Future of machine translation. Musings on Weaver's memo ». Dans *The Routledge Handbook of Translation and Technology*, 1st Edition, 419-36. Londres : Routledge.
- Mitchell, Linda, Johann Roturier, et Sharon O'Brien. 2014. « Community-based post-editing of machine-translated content: monolingual vs. bilingual ». Dans *Proceedings European Association for Machine Translation (EAMT)*, XIV : 35-44. Nice, France : European Association for Machine Translation.
- Munday, Jeremy. 2016. *Introducing Translation Studies. Theories and applications*. Fourth Edition. New York : Routledge.

- Nikolayev, Alexander Ivanovich. 2004. *Школа игры на фортепиано*. Piano. Moscow : Музыка.
- Nitzke, Jean. 2016. « Monolingual post-editing: An exploratory study on research behaviour and target text quality ». Dans *Eyetracking and Applied Linguistics*, édité par Silvia Hansen-Schirra et Grucza Sambor, 83-109. Translation and Multilingual Natural Language Processing 2. Berlin : Language science press.
- Nunes Vieira, Lucas. 2019. « Post-editing of machine translation ». Dans *The Routledge Handbook of Translation and Technology*, 1st Edition, 319-35. Londres : Routledge.
- Palisca, Claude V, et Ian D. Bent. 2001. « Theory, Theorists ». Dans *Grove Music Online*. <https://doi.org/10.1093/gmo/9781561592630.article.44944>.
- Poibeau, Thierry. 2017. *Machine Translation*. The MIT Press Essential Knowledge Series. Cambridge, Massachusetts : The MIT Press.
- . 2019. *Babel 2.0 : Où va la traduction automatique ?* Paris : Odile Jacob.
- Rigouts Terryn, Ayla, Lieve Macken, Els Lefever, Robert Vander Stichele, Koen Vanneste, et Joost Buysschaert. 2019. « Pilot study on medical translations in lay language : post-editing by language specialists, domain specialists or both? » Dans *Translating and the computer 41 (TC41) : proceedings*, 101-12. Genève : Tradulex.
- Rimsky-Korsakov, Nikolai Andreevich. 2019. *Основы оркестровки. С партитурными образцами из собственных сочинений. Учебное пособие*. 3rd Edition. Vol. I. University textbooks. Books on specialized subjects. Saint-Pétersbourg : The Planet of Music.
- Salamanca, Diego. 2018. « La traducción especializada de la tratadística musical. Corpus de armonía: un estudio de caso ». Mémoire de Master I, Toulouse, France : Université Jean Jaurès, Toulouse II.
- . 2019. « Étude et traduction inédite des « Essais sur les principes de l'harmonie » de Jean Adam Serre. Une expérience de traduction spécialisée ».

Mémoire de Master II, Toulouse, France : Université Jean Jaurès, Toulouse II.

Scarpa, Federica. 2010. *La traduction spécialisée. Une approche professionnelle à l'enseignement de la traduction*. Traduit par Marco A. Fiola. Regards sur la traduction. Ottawa : Les Presses de l'Université d'Ottawa.

Schwartz, Lane. 2014. « Monolingual Post-Editing by a Domain Expert is Highly Effective for Translation Triage ». Dans *Proceedings of the 3rd Workshop on Post-Editing Technology and Practice (WPTP)*, 34-44. Vancouver, Canada : Association for Machine Translation in the Americas.
<http://www.amtaweb.org>.

Schwartz, Lane O.B., Timothy Anderson, Jeremy Gwinnup, et Katherine M. Young. 2014. « Machine Translation and Monolingual Postediting: The AFRL WMT-14 System ». Dans *Proceedings of the Ninth Workshop on Statistical Machine Translation*, 186-94. Baltimore, Maryland USA : Association for Computational Linguistics. <https://www.aclweb.org/anthology/W14-3321.pdf>.

Screen, Ben. 2019. « What effect does post-editing have on the translation product from an end-user's perspective? » *The Journal of Specialised Translation* 31 : 133-57.

Seligman, Mark. 2019. « The Evolving Treatment of Semantics in Machine Translation ». Dans *Advances in Empirical Translation Studies. Developing Translation Resources and Technologies*, édité par Meng Ji et Michael Oakes, 53-76. Cambridge : Cambridge University Press.
<https://doi.org/10.1017/9781108525695>.

TAUS (Translation Automation User Society). 2016. *Machine Translation post-editing guidelines*. Édité par Isabella Massardo, Jaap Van der Meer, Sharon O'Brien, Fred Hollowood, Nora Aranberri, et Katrin Drescher. Amsterdam, Pays-Bas : TAUS Signature Editions.

Temizöz, Özlem. 2013. « Postediting Machine Translation Output and Its Revision. Subject-Matter Experts versus Professional Translators ». Thèse de doctorat, Espagne : Universitat Rovira i Virgili.

- Toral, Antonio, Martijn Wieling, et Andy Way. 2018. « Post-editing Effort of a Novel With Statistical and Neural Machine Translation ». *Frontiers in Digital Humanities* 5 : 9. <https://doi.org/10.3389/fdigh.2018.00009>.
- Van Egdom, Gys-Walt, et Mark Pluymaekers. 2019. « Why go the extra mile? How different degrees of post-editing affect perceptions of texts, senders and products among end users ». *The Journal of Specialised Translation* 31 : 158-76.
- Vella Ramírez, Mercedes, et José Manuel Muñoz Muñoz. 2010. « La 'Función Textual' en los textos de didáctica musical: consecuencias traductológicas ». *Entreculturas* 2 (décembre) : 133-55.
- Wiesmann, Eva. 2019. « Machine translation in the field of law: a study of the translation of italian legal texts into german ». *Comparative Legilinguistics* 37 : 117-53. <https://doi.org/10.14746/cl.2019.37.4>.

Annexes

Annexe 1 : Tableau Récapitulatif – État de l'art

Auteurs	Notions explorées	Problématique	Méthodes	Système de TA	Langues
La post-édition monolingue					
(Mitchell, Roturier, et O'Brien 2014)	PE monolingue	PE bilingue vs PE monolingue	<ul style="list-style-type: none"> ▪ Mesure du temps pour tâche donnée ▪ « Keystrokes » ▪ Évaluation (échelle de type Likert) 	Statistique	en > de en > fr
(Koehn 2010)	PE monolingue et TA	Qualité de la TA comme facteur déterminant pour la PE monolingue	<ul style="list-style-type: none"> ▪ Traduction automatique interactive ▪ Évaluation humaine avec traduction de référence 	Statistique	zh-CN > en ar > en
(Koponen et Salmi 2015)	PE monolingue – éval. de systèmes de TA	Mesurer la qualité de la TA à travers la post-édition monolingue	<ul style="list-style-type: none"> ▪ Tâche de post-édition ▪ Évaluation des post-éditions ▪ Analyse de la TA brute 	Statistique Linguistique	en > fi
(C. Hu et al. 2012)	PE monolingue – Systèmes de TA	Tester le système <i>MonoTrans2</i> avec la combinaison haïtien créole > anglais	<ul style="list-style-type: none"> ▪ Traduction et retraduction itérative des phrases annotées, évaluées ou corrigées entre post-éditeurs monolingues 	Statistique	ht > en
(Schwartz et al. 2014)	PE monolingue – Systèmes de TA	Tester le système <i>AFRL WMT-14</i> avec les combinaisons hindi > anglais russe > anglais	<ul style="list-style-type: none"> ▪ Tâche de post-édition monolingue ▪ Évaluation bilingue des post-éditions monolingues 	Statistique	hi > en ru > en
Profil des post-éditeurs					
(Nitzke 2016)	Profil du post-éditeur - PE monolingue	PE professionnels vs PE semi-professionnels	<ul style="list-style-type: none"> ▪ Évaluation de la qualité ▪ Enregistrement d'écran ▪ « Eye-tracking » ▪ Questionnaire 	Statistique	en > de

(Barbosa de Lima Fonseca 2019)	Profil du post-éditeur - PE monolingue	PE étudiants vs PE semi-professionnels vs PE professionnels	<ul style="list-style-type: none"> Mesure de l'effort temporel (e.g. mesure du temps) Mesure de l'effort technique (e.g. « keystrokes ») Mesure de l'effort cognitif (« e.g. eye-tracking ») 	Statistique	en > pt-BR
(Rigouts Terryn et al. 2019)	Profil du post-éditeur - PE complète	Spécialistes (T) + Traducteurs (R) vs Traducteurs (T) + Spécialistes (R)	<ul style="list-style-type: none"> Annotation d'erreurs dans la TA TER Comparaison : annotations vs post-éditions 	Neuronal	nl > fr
(Temizöz 2013)	Profil du post-éditeur - PE complète	Spécialistes (T) + Traducteurs (R) vs Traducteurs (T) + Spécialistes (R)	<ul style="list-style-type: none"> Enregistrement d'écran Tâche de post-édition contrôlée Questionnaires 	Statistique	en > tr
Utilisabilité de la TA					
(Forcada et al. 2018)	<i>Gisting</i> et Utilisabilité de la TA	Questionnaires de compréhension (RCQ) vs <i>Gap filling</i> (GF)	<ul style="list-style-type: none"> Tâche monolingue de compréhension de lecture Tâche du <i>Gap filling</i> avec et sans indices 	Statistique	de > en
(Doherty et O'Brien 2014)	Utilisabilité PE Monolingue	Évaluation monolingue de l'usabilité de la TA brute	<ul style="list-style-type: none"> Questionnaire (échelle de type Likert). Mesure du temps pour tâche donnée « Eye-tracking » 	Statistique	en > ja en > es en > de en > fr
(Van Egdom et Pluymaekers 2019)	Utilisabilité PE Monolingue	Usabilité des différents degrés de PE	<ul style="list-style-type: none"> Tâche de post-édition contrôlée Évaluation (échelle de type Likert) 	Statistique	en > nl
(Castilho et al. 2014)	Utilisabilité PE légère	PE légère vs TA brute	<ul style="list-style-type: none"> Enregistrement d'écran « Eye-tracking » Questionnaire 	Statistique	en > pt-BR
(Castilho et O'Brien 2016)	Utilisabilité PE légère	PE légère vs TA brute vs Texte source	<ul style="list-style-type: none"> « Eye-tracking » Questionnaire 	Statistique	en > de
(Screen 2019)	Utilisabilité PE complète	Traduction humaine vs PE complète	<ul style="list-style-type: none"> « Eye-tracking » Évaluation (échelle de type Likert) 	Neuronal	en > cy

Annexe 2 : Textes source en langue russe

Текст не специализированный (Текст 1)

Для того чтобы записать музыкальное произведение, пользуются знаками, которые называются **НОТАМИ**.

Ноты, как и звуки, бывают разной длительности. Если, например, ровно просчитать раз, два, три четыре и на каждый счёт представить себе по одному звуку, то мы получим четыре одинаковых по длительности звука. Каждый такой звук обозначается знаком 

Звук, длящийся два счёта, пишется так: 

Звук, длящийся четыре счёта, – так: 

Эти знаки имеют соответствующие названия:

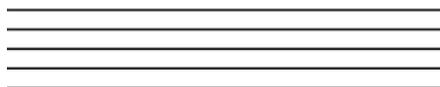
 – ЧЕТВЕРТНАЯ НОТА

 – ПОЛОВИННАЯ НОТА

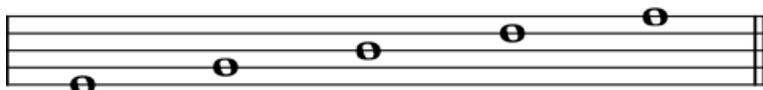
 – ЦЕЛАЯ НОТА

[...]

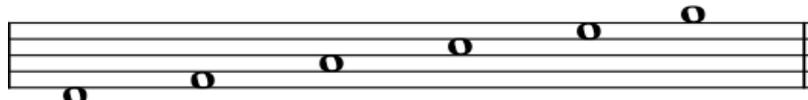
Для обозначения высоты звука пользуются нотоносцем, состоящим из пяти линеек:



Ноты пишутся на линейках:

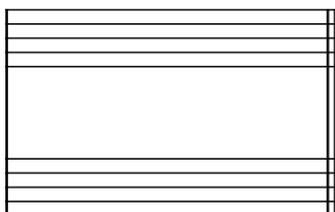


Между линейками:

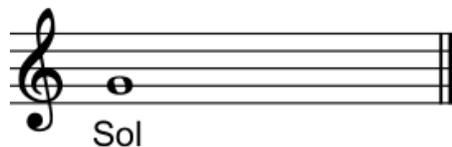


Нижняя линейка считается первой, а верхняя – пятой.

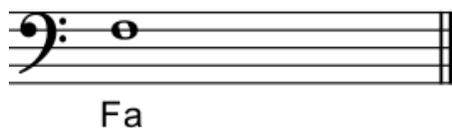
Чтобы записать большее количество разных по высоте звуков, пользуются двумя нотоносцами, соединёнными чертой:



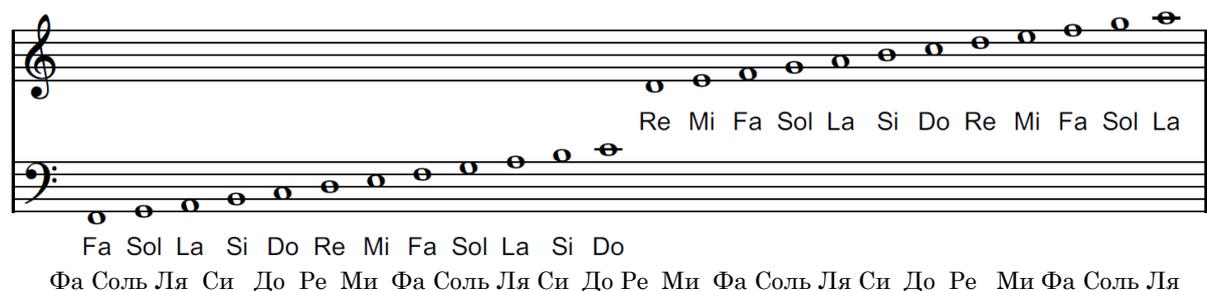
На верхнем нотном стане ставится знак , который называется СКРИПИЧНЫМ КЛЮЧОМ или КЛЮЧОМ СОЛЬ. Он указывает место ноты *соль* первой октавы на второй линейке:



На нижнем нотном стане ставится знак , который называется БАСОВЫМ КЛЮЧОМ или КЛЮЧОМ ФА. Он указывает место ноты *фа* малой октавы на четвертой линейке:



РАСПОЛОЖЕНИЕ НОТ НА НОТНОСЦАХ

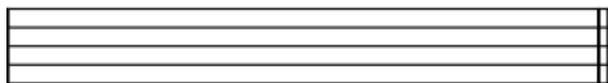


[...]

В нотной записи музыкальное произведение делится на равные части, которые называются ТАКТАМИ. Один такт отделяется от другого ТАКТОВОЙ ЧЕРТОЙ.

[...]

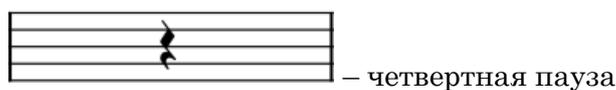
В конце пьесы ставится двойная черта:

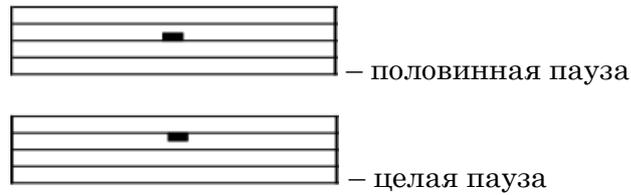


Две точки около двойной тактовой черты указывают, что вся пьеса или её часть исполняется два раза:



ПАУЗА – перерыв в звучании, знак молчания. Паузы, как и ноты, имеют разную длительность:





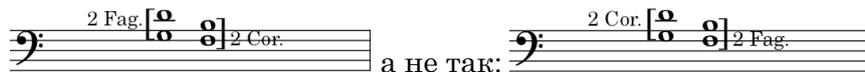
Целая пауза применяется для пустых тактов любого размера.

Texte hautement spécialisé (Texte 2)

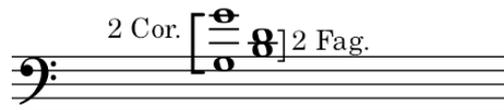
[...]

2. Наслоение, перекрещивание и окружение.

Фаготы и валторны, как уже выше было замечено, имеют тембры, наиболее сближающие между собой две группы, медную и деревянную. Четырёхголосная гармония из двух фаготов и двух валторн в особенности *piano* звучит ровно и красиво, напоминая собой гармонию из 4 валторн с оттенком несколько большей прозрачности. При *forte* валторны начинают пересиливать фаготы, и гармония из четырёх валторн оказывается предпочтительнее. Сочетание двух валторн и двух фаготов в четырёхголосном аккорде обыкновенно располагается перекрёстным способом, для большего смещения тембров. * При этом консонирующие интервалы лучше поручать валторнам, а диссонирующие – фаготам, например:



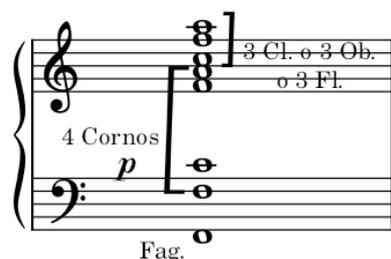
Приём окружения тоже применим, причём валторны окружают фаготы, а не наоборот:



Тот же приём окружения применим и к выдерживающим октавам трубам. В *piano* терции флейт в низком регистре, иногда с присоединением кларнетов, помещенные среди октавы труб, производят прекрасное, несколько таинственное впечатление.

В последовательном ряду аккордов медным инструментам предпочтительнее поручаются стоячие голоса, а движущиеся – деревянным.

Кларнеты своим тембром не подходят к сочетаниям с валторнами по способу окружения, но в качестве верхнего слоя гармонии в своем высоком регистре, наравне с гобоями и флейтами прекрасно дополняют аккорд из четырёх валторн *piano*; удвоение же басовой ноты можно дать фаготу:



В forte валторны пересиливают деревянные духовые, а потому верхний слой гармонии следует давать в удвоении, например:

The image shows a musical score snippet for woodwinds and brass. It consists of two staves. The upper staff is a treble clef and contains notes for 2 Flutes (2 Fl.), 2 Oboes (2 Ob.), and 2 Clarinets (2 Cl.). The lower staff is a bass clef and contains notes for 4 Horns (4 Cornos) and 1 Bassoon (Fag.). A forte dynamic marking (*f*) is placed at the beginning of the lower staff. The notes are arranged in a way that illustrates the concept of doubling the upper harmonic layer.

Annexe 3 : TA effectuées par Google Translate et DeepL

Texte non spécialisé (Texte 1)

DeepL

Para grabar una pieza musical, se utilizan unos signos llamados NOTAS.

Las notas, al igual que los sonidos, tienen distintas duraciones. Por ejemplo, si contamos exactamente uno, dos, tres cuatro y por cada cuenta imaginamos un sonido, obtendremos cuatro sonidos de la misma duración. Cada uno de estos sonidos se denota

con el signo 

Un sonido que dura dos cuentas se escribe así: 

Un sonido que dura cuatro cuentas se escribe como: 

Estas señales tienen nombres apropiados:

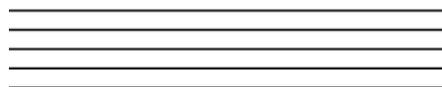
 – CUARTO DE NOTA

 – MEDIA NOTA

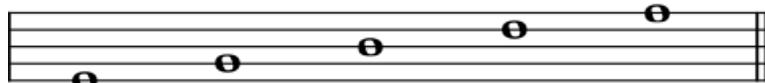
 – NOTA COMPLETA

[...]

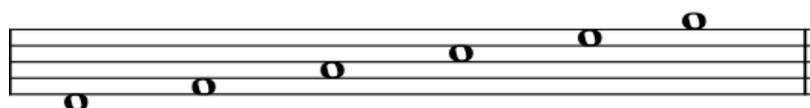
Para indicar el tono de un sonido, se utiliza un portador de notas formado por cinco reglas:



Las notas están escritas en las reglas:

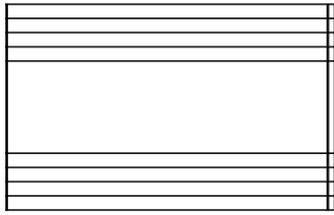


Entre los gobernantes:

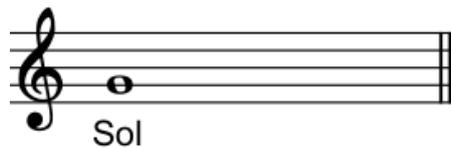


La regla inferior se considera la primera y la superior la quinta.

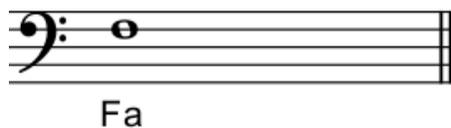
Para grabar un mayor número de tonos diferentes, se utilizan dos portadores de notas conectados por una línea:



En la regla superior se coloca el signo , que se denomina CLAVE ESCRITORA o CLAVE DE COLOR. Esto indica la posición del Sol de la primera octava en la segunda regla:



En la notocorda inferior se coloca el signo , que se denomina CLAVE DE BAJO o CLAVE DE FA. Indica la ubicación de la nota Fa de la octava baja en la cuarta regla:



LA UBICACIÓN DE LAS NOTAS EN EL NOTOCORDIO



[...]

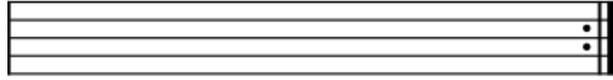
En notación musical, una pieza musical se divide en partes iguales, que se denominan TATKS. Una medida está separada de la otra por una cinta.

[...]

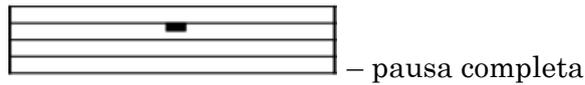
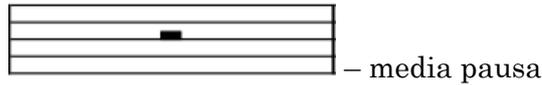
Al final de la pieza se coloca una línea doble:



Dos puntos cerca de una barra doble indican que toda la pieza, o una parte de ella, se interpreta dos veces:



PAUSA – es una pausa en el sonido, una señal de silencio. Las pausas, al igual que las notas, tienen distintas duraciones:



La pausa completa se utiliza para barras vacías de cualquier tamaño.

Texte non spécialisé (Texte 1)

Google Translate

Para grabar una pieza musical, utilizan signos llamados NOTAS.

Las notas, como los sonidos, tienen diferentes longitudes. Si, por ejemplo, calculamos exactamente uno, dos, tres, cuatro y por cada conteo imaginamos un sonido, entonces obtenemos cuatro sonidos de la misma duración. Cada uno de esos sonidos está indicado

por el signo 

Un sonido que dura dos tiempos se escribe así: 

Un sonido que dura cuatro cuentas es así: 

Estos signos tienen nombres correspondientes:

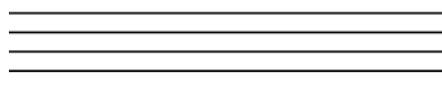
 – CUARTA NOTA

 – MEDIA NOTA

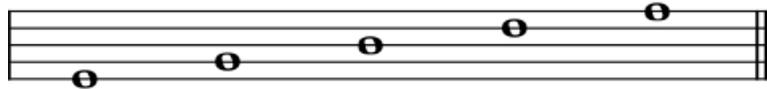
 – NOTA COMPLETA

[...]

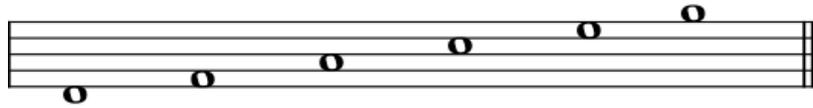
Para indicar el tono, use un pentagrama que consta de cinco reglas:



Las notas están escritas en reglas:

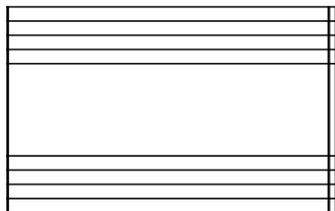


Entre los gobernantes:

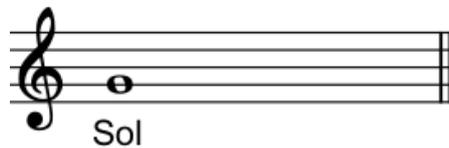


La regla de abajo se considera la primera y la de arriba es la quinta.

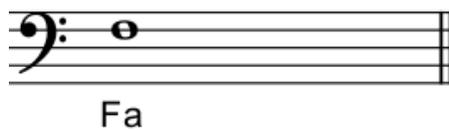
Para grabar una mayor cantidad de sonidos de diferentes alturas, use dos pentagramas conectados por una línea:



El letrero se coloca en el pentagrama superior , que se llama TREBLE CLUE o SALT KEY. Indica el lugar de la nota G de la primera octava en la segunda regla:



En el pentagrama inferior hay un signo , que se llama BASS KEY o FA KEY. Indica el lugar de la nota F de la octava menor en la cuarta regla:



POSICIÓN DE LAS NOTAS SOBRE LAS NOTAS



[...]

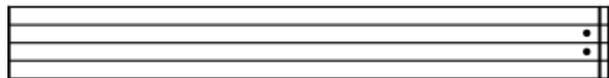
En notación musical, una pieza musical se divide en partes iguales, que se denominan TACTOS. Una barra está separada de la otra por una barra.

[...]

Al final de la jugada, se pone una doble línea:



Los dos puntos cerca de la barra doble indican que la pieza completa o parte de ella se juega dos veces:



PAUSA: una pausa en el sonido, una señal de silencio. Las pausas, como las notas, tienen diferentes longitudes:



El resto se aplica a medidas vacías de cualquier tamaño.

Texte hautement spécialisé (Texte 2)

DeepL

[...]

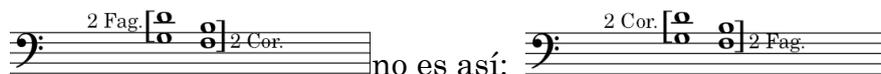
2. Acomodación, cruce y envolvimiento.

Los fagots y las trompas, como ya se ha dicho, tienen timbres que se asemejan más a los de los dos grupos, el de los metales y el de las maderas.

La armonía a cuatro voces de dos fagotes y dos trompas en particular suena suave y hermosa, recordando la armonía de 4 trompas con un toque más de transparencia. En *forte*, las trompas comienzan a sobrepasar a los fagotes y es preferible la armonía de cuatro trompas.

La combinación de dos trompas y dos fagotes en un acorde a cuatro voces suele disponerse de forma cruzada, para una mayor mezcla de tonos.

* Al hacerlo, los intervallos consonantes se asignan mejor a las trompas y los disonantes a los fagotes, por ejemplo:



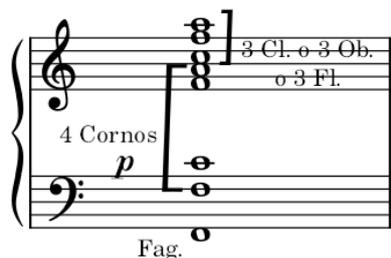
También se aplica la técnica del cerco, en la que las trompas rodean a los fagotes, y no al revés:



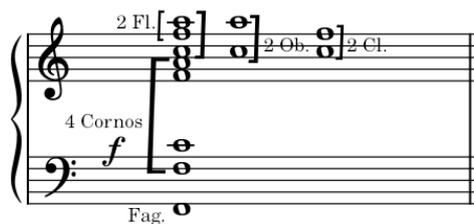
La misma técnica de rodear también se aplica a las trompetas de octava. En el *piano*, las terciaciones de la flauta en el registro bajo, a veces con la adición de clarinetes, colocados entre las octavas de las trompetas, producen una impresión hermosa y un tanto misteriosa.

En una sucesión de acordes se prefieren los instrumentos de metal a las voces de pie, y los móviles a los de madera.

Los clarinetes en su timbre no se adaptan a las combinaciones con las trompas francesas en la forma de su entorno, pero como la capa superior de la armonía en su registro alto, junto con los oboes y las flautas complementan perfectamente el acorde del *piano* de las cuatro trompas francesas; la nota baja de doblaje se puede dar al fagot:



En *forte*, las trompas se imponen a las maderas, por lo que la capa superior de la armonía debe darse en dobles, por ejemplo:



Texte hautement spécialisé (Texte 2)

Google Translate

[...]

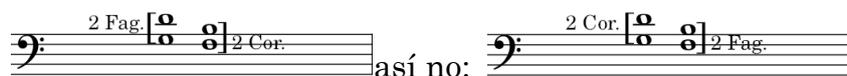
2. Colocación de capas, rebautismo y cerco.

Los fagotes y los cuernos franceses, como ya se señaló anteriormente, tienen timbres que se asemejan más entre sí a dos grupos, cobre y madera.

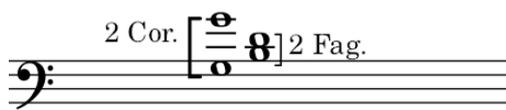
La armonía en cuatro partes de dos fagotes y dos trompas, en particular el *piano*, suena de manera uniforme y hermosa, recordando la armonía de cuatro trompas con un tinte de transparencia ligeramente mayor. Con *forte*, los cuernos franceses comienzan a dominar a los fagot, y es preferible la armonía de los cuatro cuernos franceses.

La combinación de dos trompas y dos fagotes en un acorde de cuatro partes generalmente se organiza en forma cruzada, para una mayor mezcla de timbres.

* En este caso, es mejor asignar intervalos consonantes a las trompas francesas e intervalos disonantes a los fagot, por ejemplo:



La técnica del entorno también es aplicable, y los cuernos franceses rodean los fagotes, y no al revés:

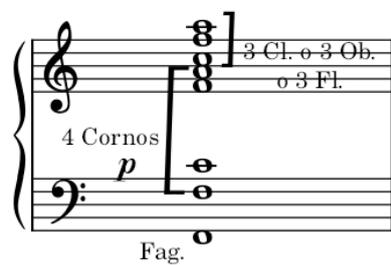


El mismo truco de ambiente se aplica a las trompetas resistentes a las octavas. En *piano*, los tercios de flautas en un registro bajo, a veces con la adición de clarinetes,

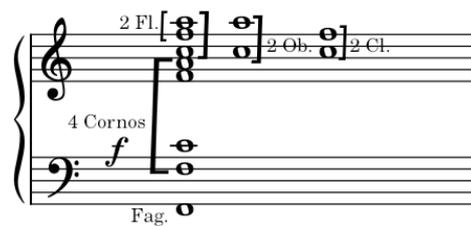
colocados entre la octava de trompetas, causan una impresión maravillosa, algo misteriosa.

En una fila secuencial de acordes, las voces de pie son preferibles para los instrumentos de metal y las de madera para los que se mueven.

Los clarinetes con su timbre no se adaptan a las combinaciones con trompas en la forma en que envuelven, pero como capa superior de armonía en su registro alto, junto con oboes y flautas, complementan perfectamente el acorde de *piano* de cuatro trompas; Doblar la nota de bajo se le puede dar al fagot:



En *forte*, los cuernos franceses dominan a los vientos de madera y, por lo tanto, la capa superior de armonía debe darse al doblar, por ejemplo:



Annexe 4 : Format d'annotation

Experiencia de posesición monolingüe

Contexto:

En esta experiencia, usted anotará dos traducciones generadas por dos sistemas de traducción automática. Sus anotaciones nos servirán para evaluar en qué medida los correctores de estas traducciones (también llamados «poseedores») serán capaces de detectar y corregir los errores que usted señale.

Las traducciones que anotará fueron efectuadas sobre dos textos. El primero, es un fragmento del «Método ruso para piano» de Alexander Nikolaev; el segundo, es un fragmento de los «Principios de orquestación» de Nikolai Rimsky-Korsakov.

Instrucciones:

Por favor, señale los errores de traducción que detecte en los textos en español traducidos por los sistemas de traducción automática. Emplee la función «Nuevo comentario» de Word² cada vez que quiera señalar un error. En cada comentario, deberá escribir uno o varios errores que apliquen a la palabra o al grupo de palabras que haya seleccionado. Adicionalmente, también podrá sugerir una mejor traducción en los comentarios.

Los tipos de error que deberá señalar son:

- **Semántico:** la traducción al español de la frase o el segmento no tiene sentido o tergiversa significativamente el sentido del texto en ruso.
- **Terminológico:** los términos no están correctamente traducidos al español.
- **Omisión:** la traducción omite información que se encuentra en el texto en ruso.
- **Adición:** la traducción presenta información que no se encuentra en el texto en ruso.

En cada página encontrará el texto original en ruso y las traducciones correspondientes al español producidas por los sistemas SIGMA y LAMBDA. También encontrará un recuadro con los tipos de error que deberá señalar.

Muchas gracias por su participación.

² Para una demostración del uso de esta función aplicada a este contexto en particular, puede dirigirse al siguiente enlace: <https://youtu.be/QvDGMIP7kHc>

Texto I

Segmento original en ruso	
<p>Для того чтобы записать музыкальное произведение, пользуются знаками, которые называются НОТАМИ.</p> <p>Ноты, как и звуки, бывают разной длительности. Если, например, ровно просчитать раз, два, три четыре и на каждый счет представить себе по одному звуку, то мы получим четыре одинаковых по длительности звука. Каждый такой звук обозначается знаком </p> <p>Звук, длящийся два счёта, пишется так: </p> <p>Звук, длящийся четыре счёта, – так: </p> <p>Эти знаки имеют соответствующие названия:</p> <p> – ЧЕТВЕРТНАЯ НОТА</p> <p> – ПОЛОВИННАЯ НОТА</p> <p> – ЦЕЛАЯ НОТА</p>	
Sistema SIGMA	Sistema LAMBDA
<p>Para grabar una pieza musical, utilizan signos llamados NOTAS.</p> <p>Las notas, como los sonidos, tienen diferentes longitudes. Si, por ejemplo, calculamos exactamente uno, dos, tres, cuatro y por cada conteo imaginamos un sonido, entonces obtenemos cuatro sonidos de la misma duración. Cada uno de esos sonidos está indicado por el signo </p> <p>Un sonido que dura dos tiempos se escribe así: </p> <p>Un sonido que dura cuatro cuentas es así: </p> <p>Estos signos tienen nombres correspondientes:</p> <p> – CUARTA NOTA</p> <p> – MEDIA NOTA</p> <p> – NOTA COMPLETA</p>	<p>Para grabar una pieza musical, se utilizan unos signos llamados NOTAS.</p> <p>Las notas, al igual que los sonidos, tienen distintas duraciones. Por ejemplo, si contamos exactamente uno, dos, tres cuatro y por cada cuenta imaginamos un sonido, obtendremos cuatro sonidos de la misma duración. Cada uno de estos sonidos se denota con el signo </p> <p>Un sonido que dura dos cuentas se escribe así: </p> <p>Un sonido que dura cuatro cuentas se escribe como: </p> <p>Estas señales tienen nombres apropiados:</p> <p> – CUARTO DE NOTA</p> <p> – MEDIA NOTA</p> <p> – NOTA COMPLETA</p>

Tipos de error:	Semántico	Terminológico	Adición	Omisión
-----------------	-----------	---------------	---------	---------

Segmento original en ruso

Для обозначения высоты звука пользуются нотоносцем, состоящим из пяти линеек:



Ноты пишутся на линейках:



Между линейками:



Нижняя линейка считается первой, а верхняя – пятой.

Sistema SIGMA

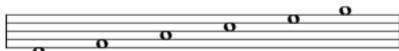
Para indicar el tono, use un pentagrama que consta de cinco reglas:



Las notas están escritas en reglas:



Entre los gobernantes:



La regla de abajo se considera la primera y la de arriba es la quinta.

Sistema LAMBDA

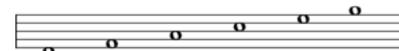
Para indicar el tono de un sonido, se utiliza un portador de notas formado por cinco reglas:



Las notas están escritas en las reglas:



Entre los gobernantes:



La regla inferior se considera la primera y la superior la quinta.

Tipos de error:	Semántico	Terminológico	Adición	Omisión
-----------------	-----------	---------------	---------	---------

Segmento original en ruso

Чтобы записать большее количество разных по высоте звуков, пользуются двумя нотоносцами, соединёнными чертой:



На верхнем нотоносце ставится знак , который называется СКРИПИЧНЫМ КЛЮЧОМ или КЛЮЧОМ СОЛЬ. Он указывает место ноты *соль* первой октавы на второй линейке:



На нижнем нотоносце ставится знак , который называется БАСОВЫМ КЛЮЧОМ или КЛЮЧОМ ФА. Он указывает место ноты *фа* малой октавы на четвёртой линейке:



Sistema SIGMA

Para grabar una mayor cantidad de sonidos de diferentes alturas, use dos pentagramas conectados por una línea:



El letrero se coloca en el pentagrama superior , que se llama TREBLE CLUE o SALT KEY. Indica el lugar de la nota G de la primera octava en la segunda regla:



En el pentagrama inferior hay un signo , que se llama BASS KEY o FA KEY. Indica el lugar de la nota F de la octava menor en la cuarta regla:



Sistema LAMBDA

Para grabar un mayor número de tonos diferentes, se utilizan dos portadores de notas conectados por una línea:



En la regla superior se coloca el signo , que se denomina CLAVE ESCRITORA o CLAVE DE COLOR. Esto indica la posición del Sol de la primera octava en la segunda regla:



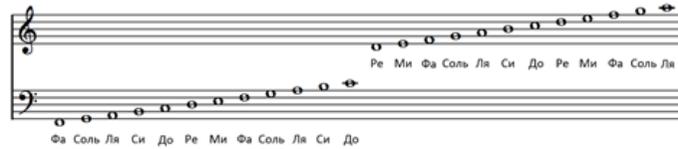
En la notocorda inferior se coloca el signo , que se denomina CLAVE DE BAJO o CLAVE DE FA. Indica la ubicación de la nota Fa de la octava baja en la cuarta regla:



Tipos de error: Semántico Terminológico Adición Omisión

Segmento original en ruso

РАСПОЛОЖЕНИЕ НОТ НА НОТНОСЦАХ

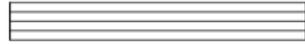


[...]

В нотной записи музыкальное произведение делится на равные части, которые называются ТАКТАМИ. Один такт отделяется от другого ТАКТОВОЙ ЧЕРТОЙ.

[...]

В конце пьесы ставится двойная черта:

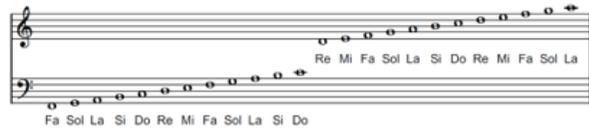


Две точки около двойной тактовой черты указывают, что вся пьеса или её часть исполняется два раза:



Sistema SIGMA

ПОЗИЦИОН DE LAS NOTAS SOBRE LAS NOTAS



[...]

En notación musical, una pieza musical se divide en partes iguales, que se denominan TACTOS. Una barra está separada de la otra por una barra.

[...]

Al final de la jugada, se pone una doble línea:



Los dos puntos cerca de la barra doble indican que la pieza completa o parte de ella se juega dos veces:



Sistema LAMBDA

LA UBICACIÓN DE LAS NOTAS EN EL NOTOCORDIO



[...]

En notación musical, una pieza musical se divide en partes iguales, que se denominan TATKS. Una medida está separada de la otra por una cinta.

[...]

Al final de la pieza se coloca una línea doble:



Dos puntos cerca de una barra doble indican que toda la pieza, o una parte de ella, se interpreta dos veces:



Tipos de error:

Semántico

Terminológico

Adición

Omisión

Segmento original en ruso

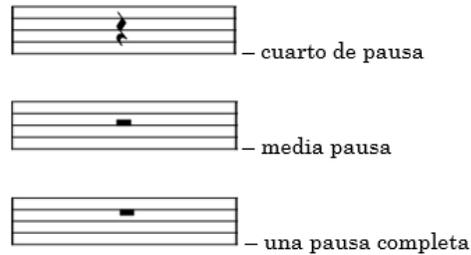
ПАУЗА – перерыв в звучании, знак молчания. Паузы, как и ноты, имеют разную длительность:



Целая пауза применяется для пустых тактов любого размера.

Sistema SIGMA

PAUSA: una pausa en el sonido, una señal de silencio. Las pausas, como las notas, tienen diferentes longitudes:



El resto se aplica a medidas vacías de cualquier tamaño.

Sistema LAMBDA

PAUSA – es una pausa en el sonido, una señal de silencio. Las pausas, al igual que las notas, tienen distintas duraciones:



La pausa completa se utiliza para barras vacías de cualquier tamaño.

Tipos de error:	Semántico	Terminológico	Adición	Omisión
------------------------	------------------	----------------------	----------------	----------------

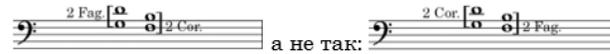
Texto II

Segmento original en ruso	
<p>[...]</p> <p style="text-align: center;"><i>2. Наслоение, перекрещивание и окружение.</i></p> <p>Фаготы и валторны, как уже выше было замечено, имеют тембры, наиболее сближающие между собой две группы, медную и деревянную. Четырёхголосная гармония из двух фаготов и двух валторн в особенности <i>piano</i> звучит ровно и красиво, напоминая собой гармонию из 4 валторн с оттенком несколько большей прозрачности. При <i>forte</i> валторны начинают пересиливать фаготы, и гармония из четырёх валторн оказывается предпочтительнее.</p>	
Sistema SIGMA	Sistema LAMBDA
<p>[...]</p> <p style="text-align: center;"><i>2. Colocación de capas, rebautismo y cerco.</i></p> <p>Los fagotes y los cuernos franceses, como ya se señaló anteriormente, tienen timbres que se asemejan más entre sí a dos grupos, cobre y madera.</p> <p>La armonía en cuatro partes de dos fagotes y dos trompas, en particular el <i>piano</i>, suena de manera uniforme y hermosa, recordando la armonía de cuatro trompas con un tinte de transparencia ligeramente mayor. Con <i>forte</i>, los cuernos franceses comienzan a dominar a los fagot, y es preferible la armonía de los cuatro cuernos franceses.</p>	<p>[...]</p> <p style="text-align: center;"><i>2. Acomodación, cruce y envolvimiento.</i></p> <p>Los fagots y las trompas, como ya se ha dicho, tienen timbres que se asemejan más a los de los dos grupos, el de los metales y el de las maderas.</p> <p>La armonía a cuatro voces de dos fagotes y dos trompas en particular suena suave y hermosa, recordando la armonía de 4 trompas con un toque más de transparencia. En <i>forte</i>, las trompas comienzan a sobrepasar a los fagotes y es preferible la armonía de cuatro trompas.</p>

Tipos de error:	Semántico	Terminológico	Adición	Omisión
-----------------	-----------	---------------	---------	---------

Segmento original en ruso

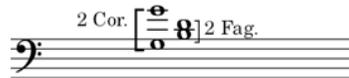
Сочетание двух валторн и двух фаготов в четырёхголосном аккорде обыкновенно располагается *перекрёстным* способом, для большего смещения тембров. * При этом консонирующие интервалы лучше поручать валторнам, а диссонирующие – фаготам, например:



а не так:



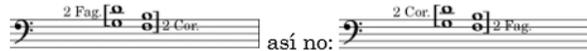
Приём *окружения* тоже применим, причём валторны окружают фаготы, а не наоборот:



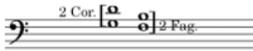
Sistema SIGMA

La combinación de dos trompas y dos fagotes en un acorde de cuatro partes generalmente se organiza en forma cruzada, para una mayor mezcla de timbres.

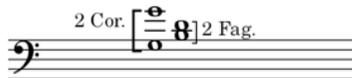
* En este caso, es mejor asignar intervalos consonantes a las trompas francesas e intervalos disonantes a los fagot, por ejemplo:



así no:



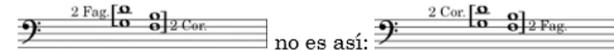
La técnica del entorno también es aplicable, y los cuernos franceses rodean los fagotes, y no al revés:



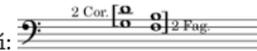
Sistema LAMBDA

La combinación de dos trompas y dos fagotes en un acorde a cuatro voces suele disponerse de forma cruzada, para una mayor mezcla de tonos.

* Al hacerlo, los intervalos consonantes se asignan mejor a las trompas y los disonantes a los fagotes, por ejemplo:



no es así:



También se aplica la técnica del cerco, en la que las trompas rodean a los fagotes, y no al revés:



Tipos de error:

Semántico

Terminológico

Adición

Omisión

Segmento original en ruso	
<p>Тот же прием окружения применим и к выдерживающим октавам трубам. В <i>piano</i> терции флейт в низком регистре, иногда с присоединением кларнетов, помещенные среди октавы труб, производят прекрасное, несколько таинственное впечатление.</p> <p>В последовательном ряду аккордов медным инструментам предпочтительнее поручаются стоячие голоса, а движущиеся – деревянным.</p>	
Sistema SIGMA	Sistema LAMBDA
<p>El mismo truco de ambiente se aplica a las trompetas resistentes a las octavas. En <i>piano</i>, los tercios de flautas en un registro bajo, a veces con la adición de clarinetes, colocados entre la octava de trompetas, causan una impresión maravillosa, algo misteriosa.</p> <p>En una fila secuencial de acordes, las voces de pie son preferibles para los instrumentos de metal y las de madera para los que se mueven.</p>	<p>La misma técnica de rodear también se aplica a las trompetas de octava. En el <i>piano</i>, las <u>terciaciones</u> de la flauta en el registro bajo, a veces con la adición de clarinetes, colocados entre las octavas de las trompetas, producen una impresión hermosa y un tanto misteriosa.</p> <p>En una sucesión de acordes se prefieren los instrumentos de metal a las voces de pie, y los móviles a los de madera.</p>

Tipos de error:	Semántico	Terminológico	Adición	Omisión
------------------------	------------------	----------------------	----------------	----------------

Segmento original en ruso

Кларнеты своим тембром не подходят к сочетаниям с валторнами по способу окружения, но в качестве верхнего слоя гармонии в своем высоком регистре, наравне с гобоями и флейтами прекрасно дополняют аккорд из четырёх валторн *piano*; удвоение же басовой ноты можно дать фаготу:

Sistema SIGMA

Los clarinetes con su timbre no se adaptan a las combinaciones con trompas en la forma en que envuelven, pero como capa superior de armonía en su registro alto, junto con oboes y flautas, complementan perfectamente el acorde de *piano* de cuatro trompas; Doblar la nota de bajo se le puede dar al fagot:

Sistema LAMBDA

Los clarinetes en su timbre no se adaptan a las combinaciones con las trompas francesas en la forma de su entorno, pero como la capa superior de la armonía en su registro alto, junto con los oboes y las flautas complementan perfectamente el acorde del *piano* de las cuatro trompas francesas; la nota baja de doblaje se puede dar al fagot:

Tipos de error:

Semántico

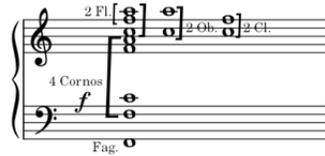
Terminológico

Adición

Omisión

Segmento original en ruso

В *forte* валторны пересиливают деревянные духовые, а потому верхний слой гармонии следует давать в удвоении, например:



Sistema SIGMA

En *forte*, los cuernos franceses dominan a los vientos de madera y, por lo tanto, la capa superior de armonía debe darse al doblar, por ejemplo:



Sistema LAMBDA

En *forte*, las trompas se imponen a las maderas, por lo que la capa superior de la armonía debe darse en dobles, por ejemplo:



Tipos de error:	Semántico	Terminológico	Adición	Omisión
------------------------	------------------	----------------------	----------------	----------------

Annexe 5 : Format de post-édition

Experiencia de posesición monolingüe

Encargo:

Una editorial trabaja en un proyecto para traducir al español dos textos especializados en música de carácter didáctico. Dado que los textos de partida están escritos en una lengua de trabajo poco convencional para los traductores hispanohablantes, no ha sido posible encontrar traductores disponibles con esta combinación lingüística.

Por lo tanto, los editores han optado por recurrir a la traducción automática y a la práctica de la posesición monolingüe³ en un intento por salvar el proyecto y, con ello, posibilitar la publicación en español del patrimonio intelectual ofrecido por los textos de partida que, de otro modo, no llegarían a traducirse.

Usted ha sido seleccionado para integrar dicho proyecto y se le pide efectuar la posesición monolingüe de los siguientes textos.

Instrucciones:

A continuación, encontrará dos textos traducidos al español divididos en cinco segmentos. Dispondrá de dos versiones por cada segmento ubicadas en las columnas **LAMBDA** y **SIGMA** respectivamente. Dichas versiones son el resultado en bruto de dos sistemas de traducción automática diferentes. Para efectuar la posesición de cada segmento, usted deberá elegir únicamente una de las dos versiones, pero siempre podrá trabajar sobre la versión que mejor le parezca.

Las imágenes y los símbolos están protegidos contra escritura. Por lo tanto, las secciones de texto representan los únicos espacios editables del documento. En la esquina inferior derecha de cada celda, encontrará un casilla (☒) que marcará para indicar la versión sobre la que ha efectuado la posesición⁴.

Para llevar a cabo esta tarea, puede recurrir a todo tipo de informaciones y fuentes externas que considere útiles, o combinar aspectos de ambas versiones; por supuesto, también podrá recurrir a sus propios conocimientos en la materia.

³ Se llama **posesición** al proceso en el que una traducción generada por un sistema de traducción automática es modificada por un humano para conseguir un producto final aceptable.

En este caso, se denomina «**posesición monolingüe**» dado que no se recurre al texto de partida para efectuar la posesición y se cuenta únicamente con la traducción automática en español.

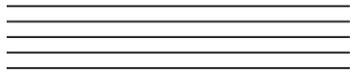
⁴ Si tiene dudas, en el siguiente enlace podrá consultar un vídeo que muestra cómo proceder para poseer en este contexto en particular: <https://youtu.be/slxnynNy6M0>.

Texto I

Seg.	LAMBDA	SIGMA
1	<p>Para grabar una pieza musical, utilizan signos llamados NOTAS. Las notas, como los sonidos, tienen diferentes longitudes. Si, por ejemplo, calculamos exactamente uno, dos, tres, cuatro y por cada conteo imaginamos un sonido, entonces obtenemos cuatro sonidos de la misma duración. Cada uno de esos sonidos está indicado por el signo </p> <p>Un sonido que dura dos tiempos se escribe así: </p> <p>Un sonido que dura cuatro cuentas es así: </p> <p>Estos signos tienen nombres correspondientes:</p> <ul style="list-style-type: none">  – CUARTA NOTA  – MEDIA NOTA  – NOTA COMPLETA 	<p>Para grabar una pieza musical, se utilizan unos signos llamados NOTAS. Las notas, al igual que los sonidos, tienen distintas duraciones. Por ejemplo, si contamos exactamente uno, dos, tres, cuatro y por cada cuenta imaginamos un sonido, obtendremos cuatro sonidos de la misma duración. Cada uno de estos sonidos se denota con el signo </p> <p>Un sonido que dura dos cuentas se escribe así: </p> <p>Un sonido que dura cuatro cuentas se escribe como: </p> <p>Estas señales tienen nombres apropiados:</p> <ul style="list-style-type: none">  – CUARTO DE NOTA  – MEDIA NOTA  – NOTA COMPLETA

2

Para indicar el tono, use un pentagrama que consta de cinco reglas:



Las notas están escritas en reglas:

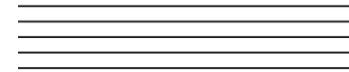


Entre los gobernantes:



La regla de abajo se considera la primera y la de arriba es la quinta. □

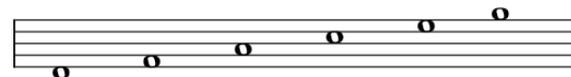
Para indicar el tono de un sonido, se utiliza un portador de notas formado por cinco reglas:



Las notas están escritas en las reglas:

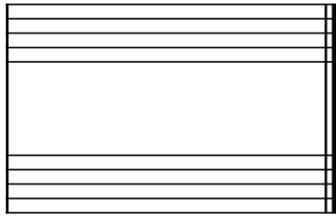


Entre los gobernantes:

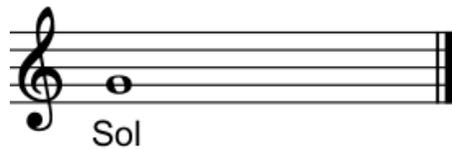


La regla inferior se considera la primera y la superior la quinta. □

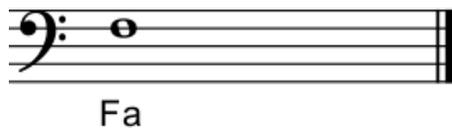
Para grabar una mayor cantidad de sonidos de diferentes alturas, use dos pentagramas conectados por una línea:



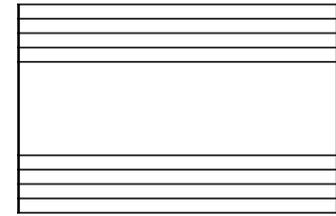
El letrero se coloca en el pentagrama superior , que se llama TREBLE CLUE o SALT KEY. Indica el lugar de la nota G de la primera octava en la segunda regla:



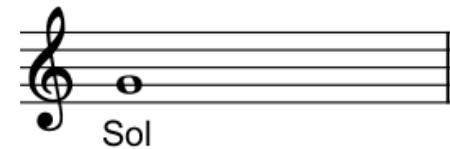
En el pentagrama inferior hay un signo , que se llama BASS KEY o FA KEY. Indica el lugar de la nota F de la octava menor en la cuarta regla:



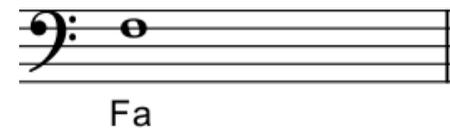
Para grabar un mayor número de tonos diferentes, se utilizan dos portadores de notas conectados por una línea:



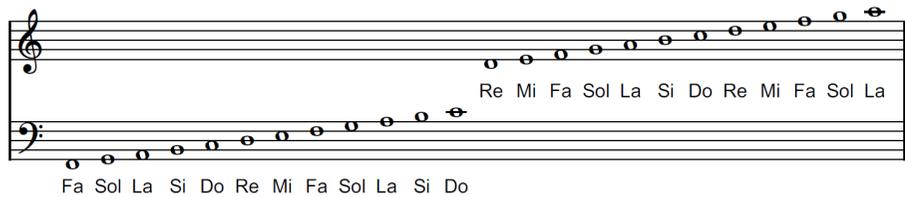
En la regla superior se coloca el signo , que se denomina CLAVE ESCRITORA o CLAVE DE COLOR. Esto indica la posición del Sol de la primera octava en la segunda regla:



En la notocorda inferior se coloca el signo , que se denomina CLAVE DE BAJO o CLAVE DE FA. Indica la ubicación de la nota Fa de la octava baja en la cuarta regla:



POSICIÓN DE LAS NOTAS SOBRE LAS NOTAS



[...]

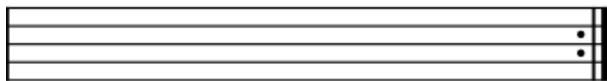
En notación musical, una pieza musical se divide en partes iguales, que se denominan TACTOS. Una barra está separada de la otra por una barra.

[...]

Al final de la jugada, se pone una doble línea:

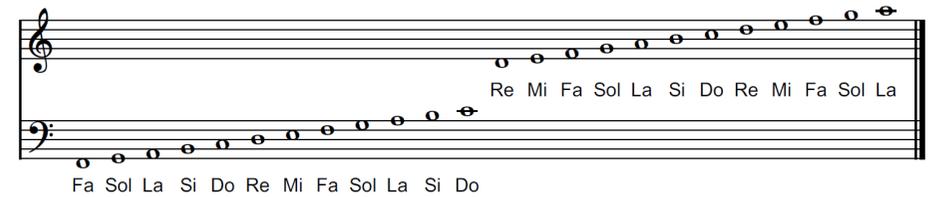


Los dos puntos cerca de la barra doble indican que la pieza completa o parte de ella se juega dos veces:



□

LA UBICACIÓN DE LAS NOTAS EN EL NOTOCORDIO

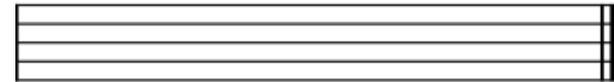


[...]

En notación musical, una pieza musical se divide en partes iguales, que se denominan TATKS. Una medida está separada de la otra por una cinta.

[...]

Al final de la pieza se coloca una línea doble:

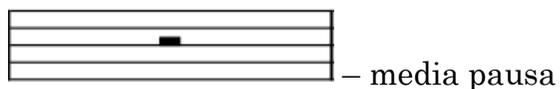


Dos puntos cerca de una barra doble indican que toda la pieza, o una parte de ella, se interpreta dos veces:



□

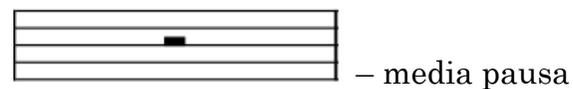
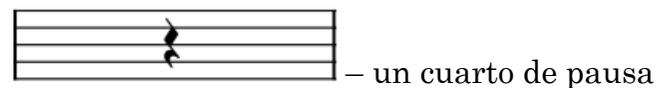
PAUSA: una pausa en el sonido, una señal de silencio. Las pausas, como las notas, tienen diferentes longitudes:



El resto se aplica a medidas vacías de cualquier tamaño.



PAUSA – es una pausa en el sonido, una señal de silencio. Las pausas, al igual que las notas, tienen distintas duraciones:

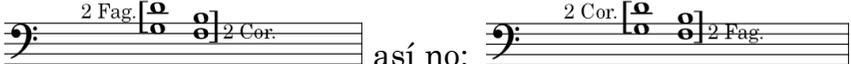
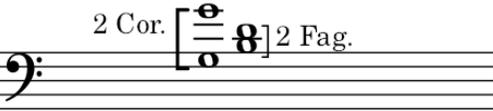
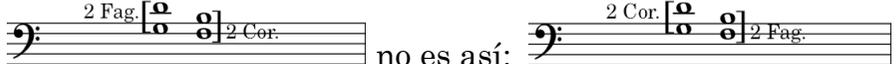
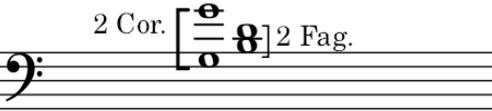


La pausa completa se utiliza para barras vacías de cualquier tamaño.



Texto II

Seg.	LAMBDA	SIGMA
1	<p>[...]</p> <p style="text-align: center;">2. Colocación de capas, rebautismo y cerco.</p> <p>Los fagotes y los cuernos franceses, como ya se señaló anteriormente, tienen timbres que se asemejan más entre sí a dos grupos, cobre y madera.</p> <p>La armonía en cuatro partes de dos fagotes y dos trompas, en particular el <i>piano</i>, suena de manera uniforme y hermosa, recordando la armonía de cuatro trompas con un tinte de transparencia ligeramente mayor. Con <i>forte</i>, los cuernos franceses comienzan a dominar a los fagot, y es preferible la armonía de los cuatro cuernos franceses.</p> <p style="text-align: right;">□</p>	<p>[...]</p> <p style="text-align: center;">2. Acomodación, cruce y envolvimiento.</p> <p>Los fagots y las trompas, como ya se ha dicho, tienen timbres que se asemejan más a los de los dos grupos, el de los metales y el de las maderas.</p> <p>La armonía a cuatro voces de dos fagotes y dos trompas en particular suena suave y hermosa, recordando la armonía de 4 trompas con un toque más de transparencia. En <i>forte</i>, las trompas comienzan a sobrepasar a los fagotes y es preferible la armonía de cuatro trompas.</p> <p style="text-align: right;">□</p>

<p>2</p>	<p>La combinación de dos trompas y dos fagotes en un acorde de cuatro partes generalmente se organiza en forma cruzada, para una mayor mezcla de timbres.</p> <p>* En este caso, es mejor asignar intervalos consonantes a las trompas francesas e intervalos disonantes a los fagot, por ejemplo:</p>  <p>La técnica del entorno también es aplicable, y los cuernos franceses rodean los fagotes, y no al revés:</p> 	<p>La combinación de dos trompas y dos fagotes en un acorde a cuatro voces suele disponerse de forma cruzada, para una mayor mezcla de tonos.</p> <p>* Al hacerlo, los intervalos consonantes se asignan mejor a las trompas y los disonantes a los fagotes, por ejemplo:</p>  <p>También se aplica la técnica del cerco, en la que las trompas rodean a los fagotes, y no al revés:</p> 
<p>3</p>	<p>El mismo truco de ambiente se aplica a las trompetas resistentes a las octavas. En <i>piano</i>, los tercios de flautas en un registro bajo, a veces con la adición de clarinetes, colocados entre la octava de trompetas, causan una impresión maravillosa, algo misteriosa.</p> <p>En una fila secuencial de acordes, las voces de pie son preferibles para los instrumentos de metal y las de madera para los que se mueven.</p>	<p>La misma técnica de rodear también se aplica a las trompetas de octava. En el <i>piano</i>, las terciaciones de la flauta en el registro bajo, a veces con la adición de clarinetes, colocados entre las octavas de las trompetas, producen una impresión hermosa y un tanto misteriosa.</p> <p>En una sucesión de acordes se prefieren los instrumentos de metal a las voces de pie, y los móviles a los de madera.</p>

4

Los clarinetes con su timbre no se adaptan a las combinaciones con trompas en la forma en que envuelven, pero como capa superior de armonía en su registro alto, junto con oboes y flautas, complementan perfectamente el acorde de *piano* de cuatro trompas; Doblar la nota de bajo se le puede dar al fagot:

Musical score for Example 4. The score is written for piano accompaniment. It features four staves: a grand staff (treble and bass clefs) and two individual staves. The grand staff contains the 4 Cornos part, marked with a piano (*p*) dynamic. The two individual staves contain the woodwind parts: 3 Cl. o 3 Ob. and 3 Fl. The bass clef staff contains the Fag. (Bassoon) part, which is marked with a piano (*p*) dynamic. The notes are arranged in a chordal structure, with the bassoon part doubling the low note of the horn section.

□

Los clarinetes en su timbre no se adaptan a las combinaciones con las trompas francesas en la forma de su entorno, pero como la capa superior de la armonía en su registro alto, junto con los oboes y las flautas complementan perfectamente el acorde del *piano* de las cuatro trompas francesas; la nota baja de doblaje se puede dar al fagot:

Musical score for Example 4. The score is written for piano accompaniment. It features four staves: a grand staff (treble and bass clefs) and two individual staves. The grand staff contains the 4 Cornos part, marked with a piano (*p*) dynamic. The two individual staves contain the woodwind parts: 3 Cl. o 3 Ob. and 3 Fl. The bass clef staff contains the Fag. (Bassoon) part, which is marked with a piano (*p*) dynamic. The notes are arranged in a chordal structure, with the bassoon part doubling the low note of the horn section.

□

5

En *forte*, los cuernos franceses dominan a los vientos de madera y, por lo tanto, la capa superior de armonía debe darse al doblar, por ejemplo:

Musical score for Example 5. The score is written for piano accompaniment. It features four staves: a grand staff (treble and bass clefs) and two individual staves. The grand staff contains the 4 Cornos part, marked with a forte (*f*) dynamic. The two individual staves contain the woodwind parts: 2 Fl. and 2 Ob. The bass clef staff contains the Fag. (Bassoon) part, which is marked with a forte (*f*) dynamic. The notes are arranged in a chordal structure, with the bassoon part doubling the low note of the horn section.

□

En *forte*, las trompas se imponen a las maderas, por lo que la capa superior de la armonía debe darse en dobles, por ejemplo:

Musical score for Example 5. The score is written for piano accompaniment. It features four staves: a grand staff (treble and bass clefs) and two individual staves. The grand staff contains the 4 Cornos part, marked with a forte (*f*) dynamic. The two individual staves contain the woodwind parts: 2 Fl. and 2 Ob. The bass clef staff contains the Fag. (Bassoon) part, which is marked with a forte (*f*) dynamic. The notes are arranged in a chordal structure, with the bassoon part doubling the low note of the horn section.

□

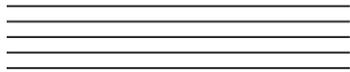
Annexe 6 : Exemple de post-édition effectué lors de la troisième phase

Texto I

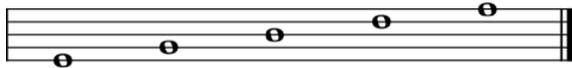
Seg.	LAMBDA	SIGMA
1	<p>Para grabar una pieza musical, utilizan signos llamados NOTAS. Las notas, como los sonidos, tienen diferentes longitudes. Si, por ejemplo, calculamos exactamente uno, dos, tres, cuatro y por cada conteo imaginamos un sonido, entonces obtenemos cuatro sonidos de la misma duración. Cada uno de esos sonidos está indicado por el signo </p> <p>Un sonido que dura dos tiempos se escribe así: </p> <p>Un sonido que dura cuatro cuentas es así: </p> <p>Estos signos tienen nombres correspondientes:</p> <ul style="list-style-type: none">  – CUARTA NOTA  – MEDIA NOTA  – NOTA COMPLETA <p style="text-align: right;">□</p>	<p>Para escribir una pieza musical, se utilizan unos signos llamados NOTAS. Las notas, al igual que los sonidos, tienen distintas duraciones. Por ejemplo, si contamos exactamente uno, dos, tres, cuatro y por cada tiempo imaginamos un sonido, obtendremos cuatro sonidos de la misma duración. Cada uno de estos sonidos se escribe con el signo </p> <p>Un sonido que dura dos tiempos se escribe así: </p> <p>Un sonido que dura cuatro tiempos se escribe así: </p> <p>Estos signos tienen nombres apropiados:</p> <ul style="list-style-type: none">  – Negra  – Blanca  – Redonda <p style="text-align: right;">☒</p>

2

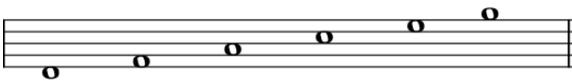
Para indicar el tono, use un pentagrama que consta de cinco reglas:



Las notas están escritas en reglas:

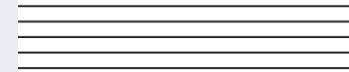


Entre los gobernantes:



La regla de abajo se considera la primera y la de arriba es la quinta. □

Para indicar el tono de un sonido, se utiliza un pentagrama formado por cinco líneas:



Las notas que están escritas en las líneas son:

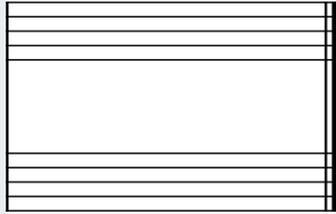


Entre los espacios son:

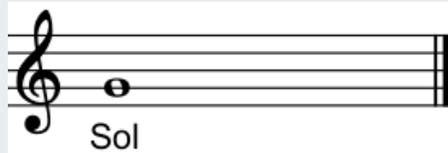


La línea inferior se considera la primera y la superior la quinta. ☒

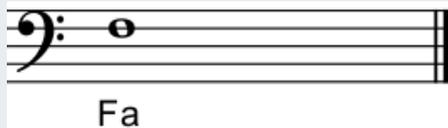
Para escribir una mayor cantidad de sonidos de diferentes alturas, se usan dos pentagramas conectados por una línea:



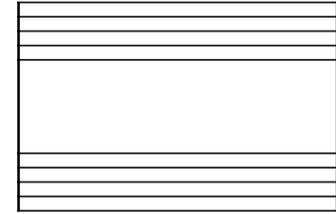
La clave que se coloca en el pentagrama superior , se llama Clave de Sol. Indica el lugar de la nota Sol de la cuarta octava que va en la segunda línea:



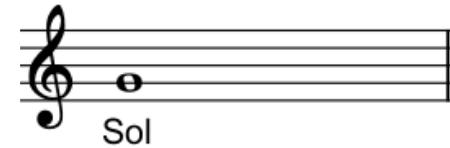
En el pentagrama inferior se coloca el signo , que se llama Clave de Fa. Indica el lugar de la nota Fa de la tercera octava que va en la cuarta línea:



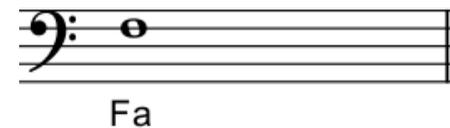
Para grabar un mayor número de tonos diferentes, se utilizan dos portadores de notas conectados por una línea:



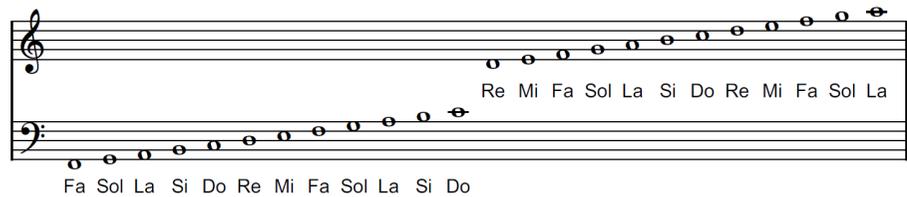
En la regla superior se coloca el signo , que se denomina CLAVE ESCRITORA o CLAVE DE COLOR. Esto indica la posición del Sol de la primera octava en la segunda regla:



En la notocorda inferior se coloca el signo , que se denomina CLAVE DE BAJO o CLAVE DE FA. Indica la ubicación de la nota Fa de la octava baja en la cuarta regla:



POSICIÓN DE LAS NOTAS SOBRE LAS NOTAS



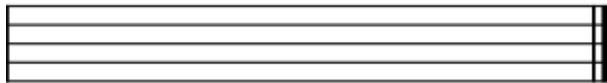
[...]

En notación musical, una pieza musical se divide en partes iguales, que se denominan TACTOS. Una barra está separada de la otra por una barra.

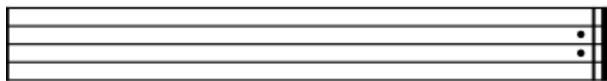
4

[...]

Al final de la jugada, se pone una doble línea:

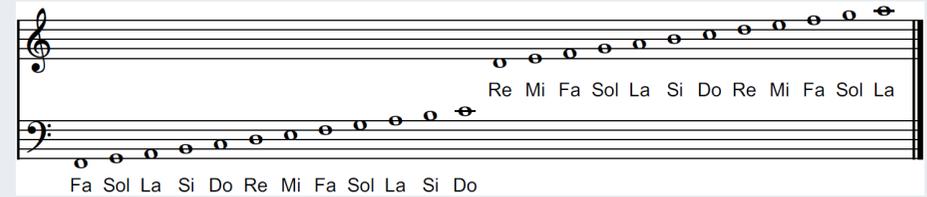


Los dos puntos cerca de la barra doble indican que la pieza completa o parte de ella se juega dos veces:



□

UBICACIÓN DE LAS NOTAS EN EL PENTAGRAMA



[...]

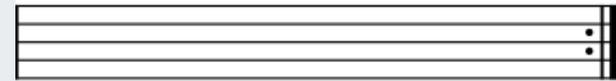
En notación musical, una pieza musical se divide en partes iguales, que se denominan COMPASES. Un compás está separado del otro por una barra.

[...]

Al final de la pieza se coloca una barra doble:



Dos puntos cerca de una barra doble indican que toda la pieza, o una parte de ella, se toca dos veces:



☒

5

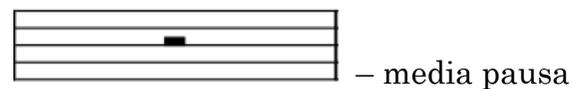
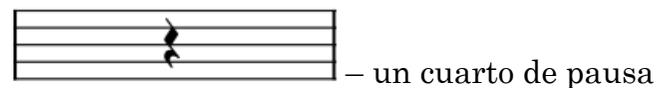
SILENCIO: es una pausa en el sonido, una señal de silencio. Los silencios, como las notas, tienen diferentes duraciones:



El silencio de redonda se aplica en compases enteros de cualquier duración.



PAUSA – es una pausa en el sonido, una señal de silencio. Las pausas, al igual que las notas, tienen distintas duraciones:

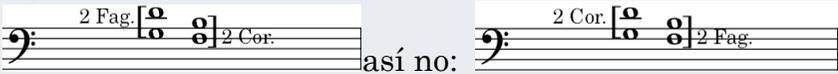
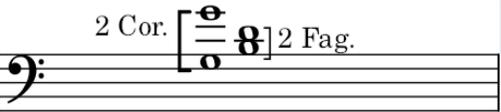
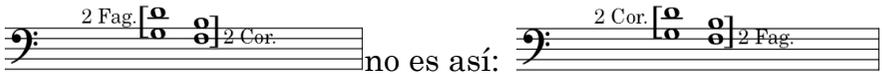
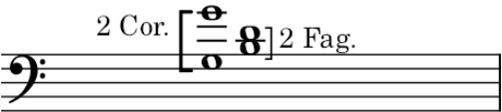


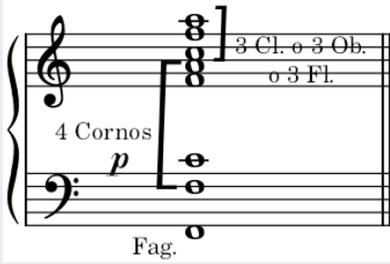
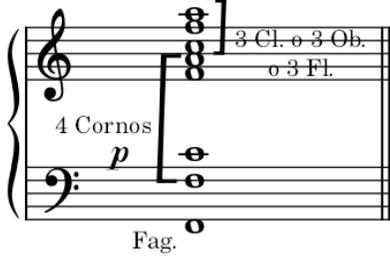
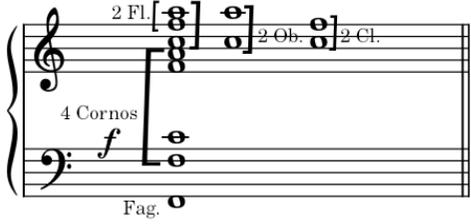
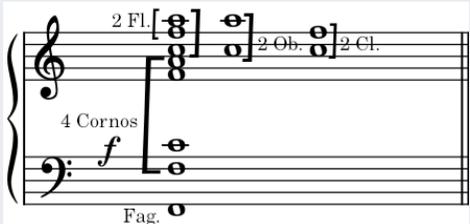
La pausa completa se utiliza para barras vacías de cualquier tamaño.



Texto II

Seg.	LAMBDA	SIGMA
1	<p>[...]</p> <p style="text-align: center;"><i>2. Colocación de acordes, cruce y envolvimiento.</i></p> <p>Los fagotes y los cuernos franceses, como ya se señaló anteriormente, tienen timbres que se asemejan más entre sí a los dos grupos, cobre y madera.</p> <p>La armonía a cuatro voces de dos fagotes y dos trompas, en particular en <i>piano</i>, suena de manera uniforme y hermosa, pareciéndose a la armonía de cuatro cuernos con un tinte de transparencia ligeramente mayor. En <i>forte</i>, los cuernos franceses comienzan a dominar a los fagotes, y es preferible la armonía de los cuatro cuernos franceses.</p> <p style="text-align: right;"><input checked="" type="checkbox"/></p>	<p>[...]</p> <p style="text-align: center;"><i>2. Acomodación, cruce y envolvimiento.</i></p> <p>Los fagots y las trompas, como ya se ha dicho, tienen timbres que se asemejan más a los de los dos grupos, el de los metales y el de las maderas.</p> <p>La armonía a cuatro voces de dos fagotes y dos trompas en particular suena suave y hermosa, recordando la armonía de 4 trompas con un toque más de transparencia. En <i>forte</i>, las trompas comienzan a sobrepasar a los fagotes y es preferible la armonía de cuatro trompas.</p> <p style="text-align: right;"><input type="checkbox"/></p>

<p>2</p>	<p>La combinación de dos cornos y dos fagotes en un acorde de cuatro voces generalmente se organiza en forma cruzada, para una mayor mezcla de timbres.</p> <p>* En este caso, es mejor asignar intervalos consonantes a los cornos franceses e intervalos disonantes a los fagotes, por ejemplo:</p>  <p>La técnica de envolvimiento también es aplicable, en la que los cornos franceses rodean a los fagotes, y no al revés:</p>  <p style="text-align: right;"><input checked="" type="checkbox"/></p>	<p>La combinación de dos trompas y dos fagotes en un acorde a cuatro voces suele disponerse de forma cruzada, para una mayor mezcla de tonos.</p> <p>* Al hacerlo, los intervalos consonantes se asignan mejor a las trompas y los disonantes a los fagotes, por ejemplo:</p>  <p>También se aplica la técnica del cerco, en la que las trompas rodean a los fagotes, y no al revés:</p>  <p style="text-align: right;"><input type="checkbox"/></p>
<p>3</p>	<p>La misma técnica de envolvimiento se aplica a las trompetas a la octava. En <i>piano</i>, las flautas en terceras en un registro bajo, a veces con la adición de clarinetes, colocados entre la octava de trompetas, causan una impresión maravillosa, algo misteriosa.</p> <p>En una secuencia de acordes, las voces de pedal son preferibles para los instrumentos de metal y las de madera para los que se mueven.</p> <p style="text-align: right;"><input checked="" type="checkbox"/></p>	<p>La misma técnica de rodear también se aplica a las trompetas de octava. En el <i>piano</i>, las terciaciones de la flauta en el registro bajo, a veces con la adición de clarinetes, colocados entre las octavas de las trompetas, producen una impresión hermosa y un tanto misteriosa.</p> <p>En una sucesión de acordes se prefieren los instrumentos de metal a las voces de pie, y los móviles a los de madera.</p> <p style="text-align: right;"><input type="checkbox"/></p>

4	<p>El timbre de los clarinetes no se adaptan a las combinaciones con cuernos en la técnica de envolvimiento, pero como capa superior de armonía en su registro alto, junto con oboes y flautas, complementan perfectamente un acorde <i>piano</i> de cuatro cuernos; El fagot puede doblar la nota del bajo:</p>  <p style="text-align: right;">☒</p>	<p>Los clarinetes en su timbre no se adaptan a las combinaciones con las trompas francesas en la forma de su entorno, pero como la capa superior de la armonía en su registro alto, junto con los oboes y las flautas complementan perfectamente el acorde del <i>piano</i> de las cuatro trompas francesas; la nota baja de doblaje se puede dar al fagot:</p>  <p style="text-align: right;">☐</p>
5	<p>En <i>forte</i>, los cuernos franceses dominan a los vientos de madera y, por lo tanto, la capa superior de armonía debe darse al doblar, por ejemplo:</p>  <p style="text-align: right;">☐</p>	<p>En <i>forte</i>, los cuernos se imponen a las maderas, por lo que la capa superior de la armonía debe doblarse, por ejemplo:</p>  <p style="text-align: right;">☒</p>

Annexe 7 : Schéma de composition des questionnaires pour les experts

Linguistes		Musiciens	
Groupes de post-éditeurs			
« Groupe A » : Musiciens		« Groupe B » : Linguistes	
Post-éditeur A		Post-éditeur D	
Post-éditeur B		Post-éditeur E	
Post-éditeur C		Post-éditeur F	
EVALUATION 1			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
B1	F1	E1	B1
C2	E2	A2	D2
B3	E3	D3	B3
F4	B4	A4	F4
F5	C5	F5	C5

EVALUATION 2			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
C1	F1	B1	F1
F2	A2	C2	E2
E3	C3	B3	E3
A4	F4	F4	B4
B5	E5	F5	C5

EVALUATION 3			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
B1	E1	C1	F1
A2	F2	F2	A2
F3	B3	E3	C3
B4	E4	A4	F4
D5	B5	B5	E5

EVALUATION 4			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
F1	C1	B1	E1
D2	B2	A2	F2
A3	E3	F3	B3
F4	B4	B4	E4
A5	F5	D5	B5

EVALUATION 5			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
A1	E1	F1	C1
E2	A2	D2	B2
F3	B3	A3	E3
B4	E4	F4	B4
E5	C5	A5	F5

EVALUATION 6			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
D1	B1	A1	E1
A2	D2	E2	A2
F3	B3	F3	B3
C4	F4	B4	E4
C5	F5	E5	C5

EVALUATION 7			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
E1	C1	D1	B1
D2	A2	A2	D2
E3	C3	F3	B3
A4	D4	C4	F4
B5	E5	C5	F5

EVALUATION 8			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
F1	B1	E1	C1
E2	B2	D2	A2
A3	D3	E3	C3
E4	B4	A4	D4
A5	E5	B5	E5

EVALUATION 9			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
C1	F1	F1	B1
C2	F2	E2	B2
D3	A3	A3	D3
E4	C4	E4	B4
A5	D5	A5	E5

EVALUATION 10			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
A1	D1	C1	F1
B2	D2	C2	F2
C3	E3	D3	A3
F4	A4	E4	C4
D5	C5	A5	D5

EVALUATION 11			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
E1	C1	A1	D1
F2	C2	B2	D2
F3	A3	C3	E3
C4	D4	F4	A4
A5	F5	D5	C5

EVALUATION 12			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
C1	D1	E1	C1
B2	E2	F2	C2
B3	D3	F3	A3
F4	B4	C4	D4
B5	F5	A5	F5

EVALUATION 13			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
A1	D1	C1	D1
A2	E2	B2	E2
F3	A3	B3	D3
B4	D4	F4	B4
E5	C5	B5	F5

EVALUATION 14			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
A1	D1	A1	D1
E2	C2	A2	E2
E3	C3	F3	A3
E4	C4	B4	D4
D5	A5	E5	C5

EVALUATION 15			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
A1	D1	A1	D1
F2	B2	E2	C2
D3	A3	E3	C3
F4	C4	E4	C4
C5	E5	D5	A5

EVALUATION 16			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
F1	A1	A1	D1
D2	B2	F2	B2
C3	D3	D3	A3
A4	D4	F4	C4
E5	B5	C5	E5

EVALUATION 17			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
F1	C1	F1	A1
F2	B2	D2	B2
E3	C3	C3	D3
D4	B4	A4	D4
A5	F5	E5	B5

EVALUATION 18			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
D1	C1	F1	C1
B2	D2	F2	B2
C3	F3	E3	C3
A4	E4	D4	B4
D5	B5	A5	F5

EVALUATION 19			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
D1	A1	D1	C1
E2	C2	B2	D2
F3	B3	C3	F3
A4	E4	A4	E4
D5	C5	D5	B5

EVALUATION 20			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
B1	E1	D1	A1
C2	F2	E2	C2
D3	B3	F3	B3
C4	E4	A4	E4
D5	A5	D5	C5

EVALUATION 21			
TEXTE 1		TEXTE 2	
LAMBDA	SIGMA	LAMBDA	SIGMA
E1	B1	B1	E1
A2	D2	C2	F2
D3	B3	D3	B3
A4	F4	C4	E4
F5	C5	D5	A5

Annexe 8 : Récapitulatif des réponses aux questionnaires post-tâche

N°	Question	Musiciens	Linguistes
1	Temps	<ul style="list-style-type: none"> • 0h50 – 2h 	<ul style="list-style-type: none"> • 1h20 – 1h30
2	Opinion de la PE monolingue	<ul style="list-style-type: none"> • Importante • Nécessaire en contexte académique • Significative : trad. Automatique ne remplace pas trad. Humaine dans la traduction de ce type de textes • Intéressante : besoin de comprendre le texte pour post-éditer 	<ul style="list-style-type: none"> • Intéressante • Peut s'avérer utile • Laborieuse • Besoin du texte source
3	Difficulté	<ul style="list-style-type: none"> • Facile : connaissance du domaine • Difficile : rigueur • Difficile : respect de « l'original » 	<ul style="list-style-type: none"> • Difficulté moyenne • Difficile • Manque de connaissances et du texte source
4	Fiabilité de la TA	<ul style="list-style-type: none"> • Fiable • Pas fiable • Partiellement fiable 	<ul style="list-style-type: none"> • Partiellement fiable • Pas fiable
5	Fiabilité des PE	<ul style="list-style-type: none"> • Fiable • Fiable • Partiellement fiable 	<ul style="list-style-type: none"> • Pas fiable : manque de connaissances • Pas fiable : manque de ressources terminologiques
6	Eval. Expérience	<ul style="list-style-type: none"> • Bien pour une première fois • Significative, satisfaisante, gratifiante • Une activité ludique, un défi intellectuel et une façon d'amplifier les connaissances. À répéter 	<ul style="list-style-type: none"> • Intéressante • Frustrante • Manque de ressources et d'outils d'aide à la traduction pour assurer la qualité
7	Utilisabilité des PE	<ul style="list-style-type: none"> • Absolument oui • Sans aucune doute • Je l'utilise déjà avec mes étudiants 	<ul style="list-style-type: none"> • Oui, pour saisir l'idée générale du texte • Non dans des contextes pédagogiques • Non
N° de réponses :		3	2