



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).

Intégration de la terminologie en traduction automatique : Le cas de la
Confédération suisse avec DeepL Pro

Baldassarre, Valentina Vincenza

How to cite

BALDASSARRE, Valentina Vincenza. Intégration de la terminologie en traduction automatique : Le cas de la Confédération suisse avec DeepL Pro. Master, 2021.

This publication URL: <https://archive-ouverte.unige.ch/unige:150596>



**UNIVERSITÉ
DE GENÈVE**

**FACULTÉ DE TRADUCTION
ET D'INTERPRÉTATION**

Valentina V. BALDASSARRE

Intégration de la terminologie en traduction automatique :

Le cas de la Confédération suisse avec DeepL Pro

Directrice de mémoire : Mme Pierrette BOUILLON

Jurée : Mme Marianne STARLANDER

Mémoire présenté à la **Faculté de traduction et d'interprétation**
(Département de traitement informatique multilingue) pour l'obtention de la
Maîtrise universitaire en traitement informatique multilingue (MATIM)

Université de Genève
Année académique 2019 – 2020
Janvier 2021

Carpe diem.

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 étudiant·e·s*, le *Règlement d'études 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 consciente 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.

Nom et prénom :

BALDASSARRE, Valentina V.

Lieu / date / signature :

Genève, le 18 janvier 2021

Valentina V. Baldassarre

Remerciements

Je tiens à remercier les personnes qui m'ont accompagnée tout au long de ce parcours académique abouti à la rédaction du présent mémoire.

J'exprime d'abord mes remerciements à ma directrice de mémoire, Madame Pierrette Bouillon, qui a immédiatement accepté de me guider dans ce projet. Merci à elle pour sa disponibilité, son expertise et ses conseils.

Je remercie également Madame Marianne Starlander, qui a accepté d'être ma jurée.

Merci à Marco Turchi¹ et Bernd Klein² pour leur disponibilité et leur expertise.

J'adresse aussi mes sincères remerciements à Martina Borghi qui a partagé avec moi les hauts et les bas de ce parcours académique et m'a apporté une aide inestimable. Je te souhaite le meilleur pour la suite de ta carrière.

Merci à Alessandro Caratelli, Laura Semeraro, Luis De Oliveira Vilaça, Sabrina Girletti, Simone Iadanza et Anette Wittwer pour leurs aides précieuses.

Enfin, je remercie de tout cœur Daniel Calderon pour son immense soutien pendant toutes les fins de semaine, soirées et vacances que j'ai dédiées à la conception et à la rédaction de ce travail. Merci pour les encouragements et le réconfort que tu as su m'apporter dans les moments de doutes.

¹ Coauteur de la métrique automatique Term Hit Rate (THR) (Farajian et al., 2018 et <http://marcoturchi.com/index.html>).

² Fondateur de Bodenseo, plateforme d'apprentissage en ligne pour ce qui concerne les langages de programmation (p.ex. Python, Perl) et Linux (<https://www.bodenseo.com/>).

Table des matières

REMERCIEMENTS	III
LISTE DES ILLUSTRATIONS	VII
LISTE DES TABLEAUX	VIII
LISTE DES ABRÉVIATIONS	X
1 INTRODUCTION	1
1.1 CONTEXTE	2
1.2 OBJECTIFS ET QUESTIONS DE RECHERCHE	3
1.3 MÉTHODOLOGIE	4
1.4 PLAN	5
2 TRADUCTION AUTOMATIQUE (TA)	7
2.1 QU'EST-CE QUE LA TRADUCTION AUTOMATIQUE ?	7
2.2 ÉVOLUTION HISTORIQUE	8
2.3 SYSTÈMES DE TRADUCTION AUTOMATIQUE	11
2.3.1 <i>Systèmes basés sur les règles</i>	11
2.3.1.1 Systèmes directs	12
2.3.1.2 Systèmes indirects	13
2.3.2 <i>Systèmes basés sur les corpus</i>	14
2.3.2.1 Systèmes basés sur les exemples	15
2.3.2.2 Systèmes statistiques	15
2.3.2.3 Systèmes neuronaux	17
2.4 ÉVALUATION DE LA TRADUCTION AUTOMATIQUE	18
2.4.1 <i>Évaluation humaine</i>	19
2.4.2 <i>Évaluation automatique</i>	21
2.5 CONCLUSIONS SUR LA TA	24
3 TRADUCTION AUTOMATIQUE ET TERMINOLOGIE	25
3.1 QU'EST-CE QUE LA TERMINOLOGIE ET POURQUOI EST-ELLE IMPORTANTE EN TA ?	25
3.2 LES SYSTÈMES DE TA ET LA TERMINOLOGIE	27
3.2.1 <i>La terminologie dans les systèmes basés sur les règles</i>	28
3.2.2 <i>La terminologie dans les systèmes statistiques</i>	29
3.2.2.1 « Static terminology integration »	30
3.2.2.2 « Dynamic terminology integration »	31
3.2.3 <i>La terminologie dans les systèmes neuronaux</i>	32
3.2.3.1 Adaptation au domaine	33
3.2.3.1.1 Méthodes « data centric »	33
3.2.3.1.2 Méthodes « model centric »	35
3.2.3.2 Spécialisation d'un système neuronal déjà entraîné	36
3.3 ÉVALUATION DE LA TERMINOLOGIE EN TA	38
3.3.1 <i>Term Hit Rate (THR)</i>	38
3.3.2 <i>TermEval</i>	40

3.4	CONCLUSION SUR LA TA ET LA TERMINOLOGIE	42
4	MÉTHODOLOGIE	43
4.1	OBJECTIFS ET QUESTIONS DE RECHERCHE	43
4.2	DÉMARCHE	44
4.2.1	<i>Assemblage du corpus</i>	45
4.2.2	<i>Constitution du corpus de test</i>	47
4.2.3	<i>Création du glossaire</i>	49
4.2.4	<i>Ajout du glossaire dans DeepL Pro</i>	53
4.2.5	<i>Traductions avec DeepL Pro</i>	54
4.2.6	<i>Évaluation humaine</i>	55
4.2.7	<i>Évaluation automatique</i>	55
4.3	CONCLUSIONS SUR LA MÉTHODOLOGIE	57
5	ANALYSE DES RÉSULTATS	58
5.1	ÉVALUATION HUMAINE	58
5.1.1	<i>TA de DeepL Pro réalisée sans la fonctionnalité « glossaire »</i>	58
5.1.1.1	Termes cibles exacts	58
5.1.1.2	Termes cibles inexacts	59
5.1.1.2.1	Recours à d'autres formes linguistiques	60
5.1.1.2.2	Emploi du langage commun au détriment d'une communication spécialisée	64
5.1.1.2.3	Erreurs de sens	64
5.1.1.2.4	Termes cibles qui ne sont pas conformes à la réalité suisse	65
5.1.1.2.5	Noms propres et acronymes en LC qui sont mal traduits	66
5.1.2	<i>TA de DeepL Pro réalisée avec la fonctionnalité « glossaire »</i>	68
5.1.2.1	Termes cibles exacts	68
5.1.2.1.1	Termes intégrés parfaitement dans la TA	69
5.1.2.1.2	Termes qui provoquent des répétitions	70
5.1.2.1.3	Erreurs de syntaxe et/ou d'orthographe dans la phrase	72
5.1.2.2	Termes cibles inexacts	73
5.1.2.2.1	Variations terminologiques qui ne sont pas détectées dans la « Réf »	73
5.1.2.2.2	Noms propres et acronymes en LC qui sont mal traduits	77
5.1.3	<i>Bilan de l'évaluation humaine</i>	78
5.2	ÉVALUATION AUTOMATIQUE	81
5.2.1	<i>BLEU</i>	81
5.2.1.1	BLEU calculé en sélectionnant l'option « Lowercase »	81
5.2.1.2	BLEU calculé en désélectionnant l'option « Lowercase »	83
5.2.2	<i>Term Hit Rate (THR)</i>	85
5.2.3	<i>Bilan de l'évaluation automatique</i>	86
5.3	CONCLUSIONS SUR L'ANALYSE DES RÉSULTATS	87
6	CONCLUSIONS	88
6.1	SYNTHÈSE ET QUESTIONS DE RECHERCHE	88
6.2	LIMITES ET PERSPECTIVES	90
7	BIBLIOGRAPHIE	92

8	WEBOGRAPHIE	99
9	ANNEXES	102
9.1	COMPOSITION DES CANDIDATS-TERMES DANS LA PARTIE ANGLAISE DU CORPUS	102
9.2	COMPOSITION DES CANDIDATS-TERMES DANS LES SEPT SOUS-ENSEMBLES DU CORPUS	103
9.3	LÉGENDE DES MATRICES SYNTAXIQUES DU GRAPHIQUE À COLONNE (SECTION 9.4) ET DE LA BOÎTE À MOUSTACHE (SECTION 9.5)	107
9.4	GRAPHIQUE À COLONNE : COMPOSITION DES TERMES DANS LES SEPT SOUS-ENSEMBLES DU CORPUS	108
9.5	BOÎTE À MOUSTACHE : COMPOSITION DES TERMES DANS LES SEPT SOUS-ENSEMBLES DU CORPUS	109
9.6	GRAPHIQUE À COLONNES REPRÉSENTANT LE SCORE BLEU	110
9.7	LISTE DES TERMES INCLUS DANS LE « GLOSSAIRE » DE DEEPL PRO	111
9.8	CORPUS DE TEST	117

Liste des illustrations

Figure 1 Classification des types de traduction (Hutchins et Somers, 1992 : 148 en Volkart, 2018 : 7).	7
Figure 2 Triangle de Vauquois (Jurafsky et Martin, 2014 : 999).	12
Figure 3 Fonctionnement des systèmes directs (Quah, 2006 : 70).	13
Figure 4 Processus traductif dans les systèmes par transfert (Quah, 2006 : 74).	13
Figure 5 Processus traductif dans les systèmes par interlingue (Quah, 2006 : 72).	14
Figure 6 Processus traductif dans les systèmes basés sur les exemples (Quah, 2006 : 81).	15
Figure 7 Processus traductif dans les systèmes statistiques (Quah, 2006 : 79).	16
Figure 8 Exemple de « plongement de mots » (Koehn, 2017 : 36).	17
Figure 9 Évaluation de la TA (Koehn, 2020 : 48).	18
Figure 10 Échelles de valeurs pour estimer la fidélité et la fluidité (Koehn, 2020 : 46).	20
Figure 11 Formule de précision et rappel (Koehn, 2010 : 223).	22
Figure 12 Formule de « f-measure » (Koehn, 2010 : 224).	22
Figure 13 Formule du Word Error Rate, WER (Koehn, 2010 : 225).	23
Figure 14 Triangle sémiotique (Baldassarre, 2016 : 11).	26
Figure 15 Extraction des équivalents en LC (L'Homme, 2008 : 251).	28
Figure 16 Création de l'interlangue (L'Homme, 2008 : 257).	29
Figure 17 Approches pour adapter un système neuronal à un domaine (Chu et Wang, 2018 : 1305).	33
Figure 18 Expérience avec le « glossaire » de DeepL Pro (LinkedIn, 2020).	37
Figure 19 Formule du Term Hit Rate (THR) (Farajian et al., 2018 : 154).	39
Figure 20 Formule de TermEval (Haque et al., 2019 : 13).	41
Figure 21 Déroulement de notre démarche méthodologique.	44
Figure 22 Résultats de l'extraction terminologique dans SynchroTerm	50
Figure 23 « Concordance » (à gauche) et « Clusters » (à droite) dans AntConc.	51
Figure 24 Création de notre glossaire dans DeepL Pro.	53
Figure 25 Extrait d'une traduction de DeepL Pro avec son « glossaire » (DeepL, 2020).	54
Figure 26 Nombre d'occurrences des termes calculées par les scripts de Python 3.	57
Figure 27 Graphique à colonnes : distribution des catégories syntaxiques (R Studio).	108
Figure 28 Boîte à moustache : distribution des catégories syntaxiques (R Studio).	109
Figure 29 Graphique à colonnes : score BLEU dans les TA de DeepL Pro.	110

Liste des tableaux

Tableau 1 Termes contenus dans le « glossaire » de DeepL Pro.	52
Tableau 2 Synonymes et variations terminologiques contenus dans le « glossaire » de DeepL Pro.	52
Tableau 3 Structure du tableau réunissant les phrases à évaluer.	54
Tableau 4 Sous-catégories des termes cibles exacts par rapport à la traduction de référence.	55
Tableau 5 Sous-catégories des termes cibles inexacts par rapport à la traduction de référence.	55
Tableau 6 « Termes exacts » et « termes inexacts » dans la TA de DeepL Pro sans « glossaire ».	58
Tableau 7 Termes cibles exacts dans DeepL Pro sans « glossaire ».	59
Tableau 8 Sous-catégories des termes cibles inexacts dans DeepL Pro sans « glossaire ».	59
Tableau 9 Recours à d'autres formes linguistiques dans DeepL Pro sans « glossaire ».	60
Tableau 10 Exemples de synonymes dans DeepL Pro sans « glossaire ».	60
Tableau 11 Exemples de variations morphosyntaxiques dans DeepL Pro sans « glossaire ».	61
Tableau 12 Exemples de variations graphiques dans DeepL Pro sans « glossaire ».	62
Tableau 13 Exemples de variations flexionnelles dans DeepL Pro sans « glossaire ».	63
Tableau 14 Exemples de variations syntaxiques faibles dans DeepL Pro sans « glossaire ».	63
Tableau 15 Emploi du langage commun vs. spécialisé dans DeepL Pro sans « glossaire ».	64
Tableau 16 Exemples d'erreurs de sens dans DeepL Pro sans « glossaire ».	65
Tableau 17 Termes qui ne sont pas conformes à la réalité suisse dans DeepL Pro sans « glossaire ».	66
Tableau 18 Noms propres et acronymes en LC qui sont mal traduits par DeepL Pro sans « glossaire ».	67
Tableau 19 « Termes exacts » et « termes inexacts » dans la TA de DeepL Pro avec « glossaire ».	68
<i>Tableau 20 Sous-catégories des termes cibles exacts dans DeepL Pro avec « glossaire ».</i>	68
Tableau 21 Termes intégrés parfaitement dans la TA de DeepL Pro avec « glossaire ».	69
Tableau 22 Mots changés après l'intégration de nos termes dans DeepL Pro avec « glossaire ».	70
Tableau 23 Répétitions causées par l'absence de synonymes dans DeepL Pro avec « glossaire ».	71
Tableau 24 Répétitions causées par l'absence de variations terminologiques DeepL Pro avec « glossaire ».	71
Tableau 25 Erreurs de syntaxe et/ou de l'orthographe causées par DeepL Pro avec « glossaire ».	73
Tableau 26 Sous-catégories des termes cibles inexacts dans DeepL Pro avec « glossaire ».	73
Tableau 27 Variations terminologiques dans DeepL Pro avec « glossaire ».	73
Tableau 28 Variations terminologiques de type morphosyntaxique dans DeepL Pro avec « glossaire ».	74
Tableau 29 Variations terminologiques de type flexionnel dans DeepL Pro avec « glossaire ».	75
Tableau 30 Forme canonique des termes dans DeepL Pro avec « glossaire ».	76
Tableau 31 Exemples de variations terminologiques de type graphique dans DeepL Pro avec « glossaire ».	77
Tableau 32 Noms propres et acronymes en LC qui sont mal traduits par DeepL Pro avec « glossaire ».	77
Tableau 33 Termes « exacts » et « inexacts » dans les TA de DeepL Pro avec et sans « glossaire ».	79
Tableau 34 Catégorisation des termes « inexacts » traduits par DeepL Pro avec et sans « glossaire ».	79
Tableau 35 Score BLEU de DeepL Pro avec et sans « glossaire ».	81
Tableau 36 Exemple n°1 de différence entre notre évaluation humaine et le BLEU avec « Lowercase ».	82
Tableau 37 Exemple n°2 de différence entre notre évaluation humaine et le BLEU avec « Lowercase ».	82
Tableau 38 Exemple n°3 de différence entre notre évaluation humaine et le BLEU avec « Lowercase ».	83

Tableau 39 Exemple n°1 de différence entre notre évaluation humaine et le BLEU sans « Lowercase ».	84
Tableau 40 Exemple n°2 de différence entre notre évaluation humaine et le BLEU sans « Lowercase ».	84
Tableau 41 Nombre d’occurrences des termes de notre glossaire dans les trois traductions.	85
Tableau 42 Résultats du Term Hit Rate (THR).	85
Tableau 43 Corrélation entre notre évaluation humaine et le score BLEU.	86
Tableau 44 Comparaison de l’ensemble des résultats.	89
Tableau 45 Impacts négatifs de DeepL Pro « avec glossaire » sur la TA et la terminologie.	89
Tableau 46 Problèmes terminologiques de DeepL Pro avec et sans « glossaire ».	90
Tableau 47 Composition des candidats-termes extraits par TermoStat dans la partie anglaise du corpus.	102
Tableau 48 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°1.	103
Tableau 49 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°2.	103
Tableau 50 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°3.	104
Tableau 51 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°4.	104
Tableau 52 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°5.	105
Tableau 53 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°6.	105
Tableau 54 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°7	106
Tableau 55 Légende des matrices syntaxiques du graphique à colonne (section 9.4) et de la boîte à moustache (section 9.5).	107
Tableau 56 Liste des termes inclus dans le « glossaire » de DeepL Pro.	116

Liste des abréviations

ALPAC : Automatic Language Processing Advisory Committee

BLEU : Bilingual Evaluation Understudy

DA : Domain adaptation

EBMT : Example-based MT

FFNN : Feed-Forward Neural Network

LC : langue cible

LS : langue source

MQM : Multidimensional Quality Metrics

MT : Machine Translation

NMT : Neural Machine Translation

OFSP : Office fédéral de la santé publique

OFSP : Office fédéral suisse de la santé publique

RBMT : Rules-Based Machine Translation

SMT : Statistical MT

TA : traduction automatique

TAAH : traduction automatique assistée par un humain

TAO : traduction assistée par ordinateur

TAS : traduction automatique statistique

TEAHQ : traduction entièrement automatique de haute qualité

TH : traduction faite par un être humain

THAO : traduction humaine assistée par un ordinateur

THR : Term Hit Rate

UT : unité de traduction

1 Introduction

Au fil des années, le volume de traduction a connu une explosion formidable dont l'intensité n'arrête pas d'augmenter et, face à la fébrilité de l'activité traductive, la traduction automatique (TA) continue à s'imposer toujours comme partie de cette solution, notamment depuis l'arrivée des systèmes neuronaux sur le marché (Schütz, 2003 : 77 et Smith, 2001 : 40). Bien que certain·e·s continuent à les percevoir comme une menace, les systèmes de TA constituent des aides précieuses. Selon les circonstances, ils peuvent effectivement donner de très bons résultats au point que certain·e·s chercheurs ou chercheuses affirment parfois que la qualité de la TA est comparable à celle d'une traduction faite par un être humain (Haque et al., 2019 : 437 et Hassan et al., 2018 : 1).

En dépit des grandes améliorations de la traduction automatique neuronale, certaines questions restent ouvertes et des problématiques de longue date ne sont pas encore résolues (Hasler et al., 2018 : 506 et Isabelle et al., 2017 : 1), en particulier : comment adapter la sortie de la traduction automatique à la terminologie d'un domaine spécialisé ? (Hasler et al., 2018 : 506) ; comment évaluer la terminologie d'une TA ? (Haque et al., 2019 : 1).

Ces problématiques bien connues n'ont suscité que peu de recherches en TA, bien que la préservation des connaissances d'un domaine donné constitue une priorité dans le processus traductif, ainsi que le besoin principal des clients de l'industrie des langues et de la traduction (Haque et al., 2019 : 2). En effet, lorsque la traduction porte sur différents domaines, les mots et leurs sens sont traduits différemment. Particulièrement dans les domaines spécialisés ou techniques, la terminologie se dessine comme le pilier de l'activité traduisante (Baldassarre, 2016 : 36 et Knowles et Koehn, 2017 : 28). De plus, l'évaluation de la terminologie est généralement confiée aux spécialistes du domaine, mais cela comporte un investissement qui n'est pas anodin sur le plan du temps et des ressources (Haque et al., 2019 : 1).

Bien que certaines études (p.ex. Chu et Wang, 2018, Farajian et al., 2018, Guerrero, 2020 et Haque et al., 2019) aient été menées dans ce domaine, ces problématiques sont encore loin d'être résolues et nous avons voulu les approfondir dans ce mémoire.

1.1 Contexte

L'évolution rapide et constante des technologies linguistiques, comme la traduction automatique, a forcément un impact sur la communication institutionnelle et sur les relations entre la population et l'État (Chancellerie fédérale, 2012 : 18). De plus, face à la place croissante du numérique, l'efficacité et l'efficacit  de l'action de l' tat sont de plus en plus un objet de d bat et un sujet de r flexion (Imelli, 2019 : 12). Il va sans dire que les cons quences pratiques sont multiples et de grande port e (Chancellerie f d rale, 2012 : 18).

Bien que certain s professionnel s continuent   critiquer fortement certains progr s, notamment en TA, puisqu'ils sont souvent sceptiques quant   l'avenir de leur m tier, les employ  s des services linguistiques doivent constamment s'adapter aux derni res technologies linguistiques,   leurs particularit s techniques et aux nouvelles m thodes pour r pondre   l'augmentation du volume des textes   traduire, tout en assurant les standards de qualit  (Chancellerie f d rale, 2012 : 18 et Gaspari, 2001 : 35). Dans ce contexte, il convient d' valuer la performance des nouvelles technologies par rapport aux prestations linguistiques et   la qualit  souhait e, voire impos e par la l gislation (Imelli, 2019 : 12).

C'est dans ces circonstances qu'  la fin du 2018 la Conf d ration suisse a charg  la Conf rence interd partementale des services linguistiques (CISL) de mener un test interd partemental avec la version payante de DeepL Pro³, un logiciel de traduction automatique g n raliste con u en 2016 sous la houlette de Jaroslaw Kutylowski⁴ et mis en ligne en 2017, qui ne demande aucune entra nement et comp tence informatique pour l'utiliser (Chancellerie f d rale, 2019 et DeepL, 2020). Le but de ce test  tait d'analyser les prestations de cette solution de TA et de tirer les enseignements n cessaires, afin de d terminer dans quelle mesure une solution de ce genre se pr te aux activit s de traduction de la Conf d ration suisse (Chancellerie f d rale, 2019). Ledit test s'est conclu avec l'achat d'une licence d'entreprise « DeepL Pro Starter⁵ » (Chancellerie f d rale, 2019). Par honn tet  intellectuelle, il importe de signaler ici que l'auteur du pr sent m moire a directement particip    cette  tude,  tant elle-m me ancienne employ e au sein des

³ Cette version permet de traduire de fa on plus efficace et rapide de longs textes et de les importer sous forme de diff rents fichiers (DeepL, 2020). De plus, le traducteur en ligne garantit la suppression des donn es apr s la traduction, ce qui assure la protection des donn es confidentielles (DeepL, 2020).

⁴ Membre de l' quipe Linguee, une soci t  qui a lanc  en 2010 le moteur de recherche de traduction du m me nom (DeepL, 2020).

⁵ Cette version permet de traduire un nombre illimit  de textes et d'ajouter depuis mai 2020 un seul glossaire sans limite d'entr es terminologiques (DeepL, 2020). Par contre, elle ne permet pas l'int gration de DeepL dans les logiciels de TAO (DeepL, 2020).

Services linguistiques de l'Office fédéral suisse de la statistique comme terminologue et traductrice.

En prenant comme exemple le cas de la Confédération suisse et en nous situant dans le même contexte d'étude que Guerrero (2020) (section 3.2.3.2), à savoir celui de l'industrie des langues et de la traduction, nous avons voulu étudier l'impact de la nouvelle fonctionnalité « glossaire »⁶ de DeepL Pro sur la TA des termes.

1.2 Objectifs et questions de recherche

Le présent mémoire vise deux objectifs principaux.

Le premier consiste à décrire le traitement de la terminologie en traduction automatique. En l'espèce, nous regarderons quels sont les problèmes terminologiques de DeepL Pro et dans quelle mesure sa nouvelle fonctionnalité « glossaire » permet de les résoudre en permettant aux utilisateurs d'effectuer des traductions, tout en garantissant une certaine terminologie.

Le deuxième objectif concerne l'évaluation de la qualité de la traduction automatique des termes avec et sans « glossaire ». En partant de la constatation que l'évaluation humaine comporte un investissement qui n'est pas anodin sur le plan du temps et des ressources, nous verrons s'il existe une corrélation entre celle-ci et les métriques automatiques.

Dans ce contexte, nous chercherons à répondre à la question de recherche principale suivante :

- Le traducteur en ligne DeepL Pro permet-il d'adapter la terminologie de ses traductions à un domaine spécialisé ?

Cette question soulève également trois sous-questions :

- Sous-question 1 : quel est l'impact de la nouvelle fonctionnalité « glossaire » sur la qualité de la traduction ?
- Sous-question 2 : le « glossaire » permet-il de prendre en compte les variations des termes ?
- Sous-question 3 : quels sont les types d'erreurs posés par la TA des termes ?

⁶ Cette fonctionnalité a été lancée en mai 2020.

Ci-après, nous donnerons un aperçu de la démarche méthodologique mise en place pour pouvoir répondre à ces questions de recherche.

1.3 Méthodologie

Pour la réalisation de notre travail, nous sommes passée par plusieurs étapes.

Premièrement, nous avons assemblé un corpus parallèle bilingue (allemand – anglais) contenant 892 unités de traduction portant sur le domaine de la pandémie de Covid-19, où chaque unité de traduction correspond à une phrase de notre corpus (section 4.2.1). Compte tenu de notre contexte, l'ensemble des textes collectés provient de l'Office fédéral suisse de la santé publique (OFSP) et ne reflète que sa terminologie.

Dans un second temps, nous avons choisi selon des critères bien définis un sous-corpus représentatif (section 4.2.2) d'où nous avons extrait des candidats-termes⁷. De cette façon, nous avons créé un glossaire (section 4.2.3), que nous avons ajouté dans l'interface web de DeepL Pro (section 4.2.4) pour l'évaluer dans ses traductions. Pour cette étude, nous avons mis en place une approche terminographique⁸, bien que plusieurs méthodes auraient pu être utilisées pour évaluer l'impact de la TA sur la terminologie, par exemple le recours à des suites de test (« *challenge set* »)⁹ (Isabelle et al., 2017) et l'annotation de termes dans des corpus de référence (« *inline annotation of the target terms in the source segment of parallel corpora* » et « *TermMarker* ») (Dinu et al., 2019 et Haque et al., 2019). Le choix d'une telle approche dépend de notre contexte, à savoir celui de l'industrie des langues et de la traduction, dans lequel : les terminographes créent des bases terminologiques¹⁰ servant de référence pour traduire des textes spécialisés et les traducteurs ou les traductrices recourent à ces outils informatiques d'aide à la traduction pour travailler de façon plus efficace et rapide.

⁷ Il s'agit de mots ou suites de mots qui sont susceptibles de correspondre à des unités terminologiques (L'Homme, 2020 : 201).

⁸ Une collection terminologique devrait généralement être créée par un ou une terminographe, qui s'intéresse à la collecte, à la consignation, au traitement et à la présentation des données terminologiques (Baldassarre, 2016 : 17, ISO 1087-2 : 1989(F) : 13).

⁹ Cette méthode consiste à produire à la main un set de phrases pour évaluer le traitement de certaines divergences en TA (p.ex. accord entre sujet et verbe, recours à certains temps verbaux, etc.) (Isabelle et al., 2017 : 1).

¹⁰ Pour ce faire, le terminographe effectue sept tâches : mettre en forme un corpus s'il n'existe pas déjà, repérer et collecter de candidats-termes, les analyser et les synthétiser pour finalement encoder et organiser les données terminologiques retenues (L'Homme, 2020 : 51 - 52).

L'étape suivante a consisté à traduire en anglais avec DeepL Pro notre sous-corpus avec et sans « glossaire » (section 4.2.5). Ce dernier a représenté donc notre variable indépendante, alors que la qualité des traductions notre variable dépendante.

Puis nous avons procédé à l'évaluation humaine de la terminologie (section 4.2.6). Bien que dans ce contexte également, plusieurs approches existent, comme le « *jugement intuitif* » et l'« *évaluation comparative* » (section 2.4.1), nous avons effectué nous-même une évaluation manuelle basée sur la comparaison de la terminologie entre traduction de référence et la TA. En nous inspirant de l'évaluation humaine effectuée par Haque et al. (2019) visant à classer la TA des termes en deux catégories, à savoir « *error terms* » et « *incorrect terms* » (section 3.3.2), nous avons analysé les « termes exacts » et les « termes inexacts », tout en tenant compte de leurs sous-catégories le cas échéant (section 4.2.6).

En partant de la constatation que l'évaluation humaine comporte un investissement qui n'est pas anodin sur le plan du temps et des ressources, nous avons également mesuré la performance de DeepL Pro de façon automatique en recourant au score BLEU et au Term Hit Rate (THR) (section 4.2.7). De ce fait, nous avons vu s'il existe une correspondance entre évaluation humaine et évaluation automatique.

Enfin, les résultats obtenus nous ont permis de répondre à notre question de recherche ainsi qu'à la série de sous-questions qui en découlent (section 6.1).

1.4 Plan

Après avoir exposé le contexte dans lequel s'inscrit ce mémoire et détaillé nos objectifs et notre démarche, nous présenterons la traduction automatique (TA) au chapitre 2. Afin de poser les bases nécessaires à la compréhension de notre travail, nous définirons celle-ci (section 2.1) et nous retracerons son historique (section 2.2) ; puis nous nous intéresserons aux différents types de systèmes (section 2.3) ; finalement, nous regarderons les méthodes permettant de juger la qualité d'un système de TA (section 2.4).

Le chapitre 3 portera sur le traitement de la terminologie en traduction automatique. Nous commencerons par définir la terminologie en montrant son importance en TA (section 3.1). Dans un second temps, nous approfondirons comment elle est traitée dans les différents

systèmes de TA, selon l'architecture (section 3.2), et nous aborderons les métriques automatiques permettant d'évaluer la terminologie en TA (section 3.3).

Nous exposerons la méthodologie de notre travail au chapitre 4. D'abord, nous approfondirons les objectifs et les questions de recherche (section 4.1), puis nous passerons en revue la démarche qui a été entreprise pour mener à bien le projet de ce mémoire, en éclaircissant chaque étape nécessaire à son bon déroulement (section 4.2).

Le chapitre 5 sera consacré à l'analyse des résultats. En premier lieu, nous aborderons l'évaluation humaine effectuée pour comparer la traduction de DeepL avec et sans la fonctionnalité « glossaire » (sections 5.1.1 et 5.1.2). Puis nous exposerons le résultat de l'évaluation automatique (section 5.2), c'est-à-dire le score BLEU (section 2.4.2) et le Term Hit Rate (THR) (section 3.3.1).

Dans notre conclusion au chapitre 6, nous récapitulerons d'abord notre expérience et nous répondrons à nos questions de recherches (section 6.1). Après cela, nous évoquerons les limites de notre travail et nous proposerons de différentes pistes de recherches qui pourront être conduites à l'avenir (section 6.2).

2 Traduction automatique (TA)

Notre mémoire a pour objectif de décrire et d'évaluer le traitement de la terminologie en traduction automatique. Afin de poser les bases nécessaires à la compréhension de notre travail, ce chapitre porte sur la présentation de la traduction automatique dans son ensemble. En premier lieu, nous définirons celle-ci (section 2.1) et nous retracerons son historique (section 2.2) ; puis nous présenterons les différents types de systèmes (section 2.3) et, finalement, nous exposerons les méthodes permettant de juger la qualité d'un système de TA (section 2.4). Enfin, nous ferons un résumé du chapitre (section 2.5).

2.1 Qu'est-ce que la traduction automatique ?

La traduction automatique (TA) constitue un champ d'étude interdisciplinaire réunissant des linguistes, lexicologues, informaticiens et traducteurs qui poursuivent un objectif commun : la conception et le développement de ressources et logiciels qui permettent de traduire automatiquement un texte d'une langue naturelle de départ (ou langue source LS) dans une langue d'arrivée (ou langue cible LC) (Bouillon et Clas, 1993 : 15 et Yvon, 2014). Contrairement à la traduction faite par un être humain (TH), la TA comprend différents degrés d'automatisation (Figure 1) (Hutchins et Somers, 1992 : 148).

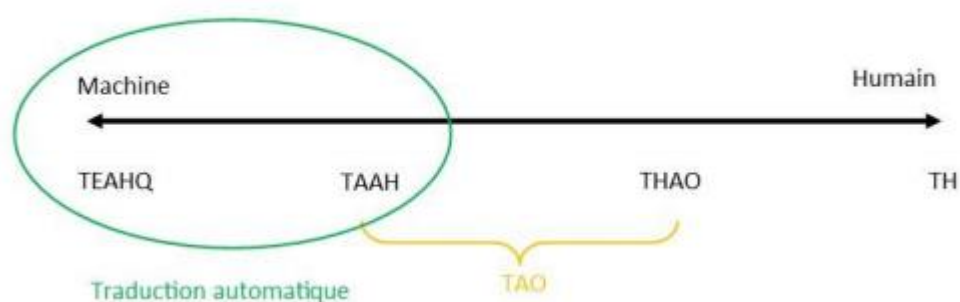


Figure 1 Classification des types de traduction (Hutchins et Somers, 1992 : 148 en Volkart, 2018 : 7).

Dans le contexte de la traduction automatique, nous trouvons aux deux extrêmes la traduction entièrement automatique de haute qualité (TEAHQ) et la traduction automatique assistée par

un être humain (TAAH) (Hutchins et Somers, 1992 : 148). Si dans la TEAHQ le texte présente une très bonne qualité et aucune contribution humaine n'est nécessaire, dans le deuxième cas, il est important que le traducteur ou la traductrice humain·e intervienne avant, pendant ou après le processus traductif pour améliorer le résultat (Quah, 2006 : 7 - 8). Plus précisément, l'intervention humaine avant la traduction est connue sous le nom de « préédition », tandis que celle après s'appelle « post-édition » (Hutchins et Somers, 1992 : 148). Hutchins et Somers (1992 : 148) désignent la contribution humaine pendant le processus traductif comme une « intervention interactive ». Le terme TA n'inclut pas les outils informatiques d'aide à la traduction, tels que les mémoires de traduction et les bases terminologiques, qui s'inscrivent dans le cadre de la traduction assistée par ordinateur (TAO) et qui sont repris sous la dénomination générique de « *traductive* » (Bouillon et Clas, 1993 : 15).

Lorsqu'on parle de traduction automatique, il est toujours fondamental d'en étudier l'évolution afin de comprendre ce que le marché offre actuellement et quels en seront les développements (Schütz, 2003 : 77). Dans cette optique, nous donnerons ci-après un aperçu historique de la TA en parcourant les événements les plus marquants.

2.2 Évolution historique

Certain·e·s retracent l'origine de la traduction automatique au XVII^e siècle avec les idées de Leibniz, Descartes et autres philosophes qui parlent d'une langue universelle, philosophique et logique (Hutchins, 2004 : 1). Bien qu'il faille se garder de voir ces travaux comme des systèmes pilotes de TA, leurs propositions ont été source d'inspiration pour nombreuses études (Hutchins, 2004 : 1 et Poibeau, 2019 : 32).

Dans les années 1930, deux chercheurs imaginent des systèmes mécaniques visant à la gestion de dictionnaires multilingues et à la traduction semi-automatique (Poibeau, 2019 : 35). En 1933, Artsrouni dépose une demande de brevet pour un « cerveau mécanique » capable de conserver et de retrouver différents types d'informations, dont des données linguistiques dans plusieurs langues (Poibeau, 2019 : 35 - 26). Contrairement à Artsrouni, l'invention de Trojanskij porte sur un processus de traduction assistée se déroulant en trois étapes : un ou une analyste cherche premièrement des éléments de traduction au niveau des mots ; puis un dispositif qui tient compte de la fonction du mot dans la phrase source génère « la traduction » dans la langue cible ; en fin de course, un correcteur ou une correctrice vérifie que le texte soit

exact du point de vue stylistique (Poibeau, 2019 : 37). Bien qu'à l'époque les inventions des deux chercheurs aient été largement ignorées, ils peuvent être considérés comme les précurseurs des technologies apparues dans les années 1950 (Hutchins, 2004 : 12 - 13 et Poibeau, 2019 : 37 - 38).

L'après-guerre voit l'apparition des premiers ordinateurs, ce qui marque le début d'une phase d'effervescence. Comme Steedman (2007) le dit, « *human knowledge is expressed in language. So computation linguistics is very important* ».

Les premières recherches en TA sont attribuées à Booth qui, à la fin des années 1940, s'intéresse au traitement de la langue par un système automatique dans le sillage de Artsrouni et Trojanskij (Poibeau, 2019 : 39).

En 1943, McCulloch et Pitts développent le premier modèle neuronal (Koehn, 2020 : 31). Selon Jurafsky et Martin (2014 : 9) il s'agit de « *a kind of computing element that could be described in terms of propositional logic* ». Ce modèle est repris par Rosenblatt en 1957 avec le réseau des neurones artificiels « perceptrons »¹¹ (Koehn, 2020 : 31). Toutefois, la publication du livre « Perceptrons » (Minsky and Papert, 1969) montrant leurs limites, ainsi que les limitations au niveau informatique pendant l'après-guerre et autres opinions négatives sur les réseaux neuronaux artificiels causent l'arrêt des travaux de recherche dans le domaine pour plus d'une décennie (Koehn, 2020 : 31).

En 1949, le mathématicien Weaver publie un mémorandum, où il propose d'utiliser les techniques de cryptage pour traduire automatiquement (Quah, 2006 : 58 - 59). En pleine guerre froide, où il est très important de comprendre les communications chiffrées des ennemis, cet ouvrage marque le point de départ de la traduction automatique et contribue à lancer des recherches dans le domaine (Poibeau, 2019 : 44 - 45 et Quah, 2006 : 59).

Pendant les années 1950, le chercheur Bar-Hillel s'intéresse à la TA et dans ses travaux, il affirme que la traduction entièrement automatique n'est pas en mesure de produire des résultats de haute qualité, puisque la machine ne peut pas éliminer les ambiguïtés sémantiques (Bar-Hillel, 1951 : 229). En conséquence, il faut se tourner vers une TA mixte dans laquelle l'être humain intervient dans le processus (Bar-Hillel, 1951 : 230 - 231). Dans la foulée de la

¹¹ Les « perceptrons » constituent une sous-catégorie des « réseaux de neurones acycliques » (« *Feed-Forward Neural Network, FFNN* ») (Gelly, 2017 : 7 - 16). Il ne se composent que d'une seule couche, qui correspond à une liste de neurones où chacun a le même nombre d'entrées binaires et seulement une sortie binaire (Koehn, 2020 : 31).

première conférence du domaine organisée par Bar-Hillel, en 1954 la première démonstration de traduction automatique a lieu : le système mis au point conjointement par l'équipe de l'Université de Georgetown et par IBM montre que la traduction automatique peut donner de bons résultats et encourage la recherche (Poibeau, 2019 : 46).

Le climat d'effervescence caractérisant les recherches en traduction automatique dans les années 1950 change drastiquement en 1966, quand le comité indépendant « Automatic Language Processing Advisory Committee (ALPAC) » rédige un rapport d'évaluation sur la TA (Koehn, 2020 : 34 et Poibeau, 2019 : 55 - 58). Selon ce rapport, la traduction automatique de haute qualité ne donne pas de bons résultats et, en matière de coûts, un traducteur ou une traductrice humain·e revient moins cher qu'une TA (Poibeau, 2019 : 58). Pour cette raison, il faut plutôt investir dans la recherche en linguistique et le développement de méthodes qui améliorent la traduction humaine (Koehn, 2020 : 35). Malgré les conclusions négatives tirées après le rapport d'évaluation de l'ALPAC et l'arrêt brutal des recherches dans le domaine, certain·e·s chercheurs et chercheuses poursuivent leurs travaux et différents systèmes de traduction automatique sont mis au point, tels que METEO (Koehn, 2020 : 35).

Jurafsky et Martin (2014 : 12) parlent de la fin des années 1980 comme de la période du « *return of empiricism* ». Le nombre croissant de textes sur support numérique incluant aussi des textes bilingues « alignés » relance la recherche en traduction automatique en révolutionnant les approches (Poibeau, 2019 : 69). Dans ce contexte, deux courants voient le jour : la traduction automatique à base d'exemples (« Example-based MT, EBMT ») et la traduction automatique statistique (« Statistical MT, SMT ») (Hearne et Way, 2011 : 205 et Koehn, 2020 : 37). En même temps, certain·e·s chercheurs et chercheuses étudient les réseaux des neurones dans le sillage des premières recherches dans ce domaine (Koehn, 2020 : 39). Toutefois, les modèles développés ne donnent pas de résultats intéressants à cause du nombre restreint de données disponibles pour les entraîner, et la recherche dans le domaine s'interrompt à nouveau en favorisant la diffusion des modèles statistiques, qui conçoivent la traduction comme un problème d'optimisation statistique (Koehn, 2020 : 36 - 39). Les années 1980 et 1990 sont également caractérisées par la recherche sur la traduction de la parole, ce qui implique l'intégration de la reconnaissance vocale, de la synthèse vocale et des modules de traduction (Hutchins, 2005 : 4). À cette époque, les outils de traduction assistée par ordinateur (TAO) suscitent un grand succès (Rochard, 2017 212 - 227).

La traduction statistique a connu véritable essor depuis son apparition jusqu'à 2016, quand elle a été supplantée par la traduction dite « neuronale », qui actuellement progresse rapidement (Koehn, 2020 : 39 et Poibeau, 2019 : 117). Depuis les dernières décennies, aussi le traitement de la parole fait l'objet de recherches intensives et la traduction automatique directement à partir de la parole est un champ de recherche très actif (Poibeau, 2019 : 195). Au cours des prochaines années, les études prendront différentes directions pour améliorer les systèmes de traduction automatique et, à la lumière des pas de géant faits dans le domaine, on peut encore s'attendre à de nouveaux progrès (Poibeau, 2019 : 199).

2.3 Systèmes de traduction automatique

Les systèmes de traduction automatique peuvent être catégorisés différemment selon les critères utilisés, tels que paires de langues (p.ex. systèmes bilingues ou multilingues), direction (p.ex. systèmes unidirectionnels ou bidirectionnels) ou types de données utilisées (systèmes basés sur des informations linguistiques représentées sous forme de règles ou systèmes fondés sur des données non linguistiques comme des corpus) (Bouillon et Clas, 1993 : 15 - 16). Nous choisirons de décrire les systèmes selon leur mode de fonctionnement, soit le type de données qu'ils utilisent pour traduire automatiquement un texte. Dans cette perspective, nous aborderons premièrement les systèmes basés sur les règles (section 2.3.1), puis nous passerons en revue les systèmes basés sur les corpus (section 2.3.2). Bien que plusieurs systèmes hybrides existent, nous ne les analyserons pas en détail. En général, ces systèmes combinent les éléments axés sur la connaissance et sur les données et reprennent les approches précédentes en évitant leurs erreurs et en tirant profit des solutions développées (Thurmair, 2009 : 1).

2.3.1 Systèmes basés sur les règles

Les systèmes basés sur les règles (« *Rules-Based Machine Translation, RBMT* ») disposent de connaissances linguistiques qui leur permettent d'identifier les mots, leur associer une catégorie grammaticale et une forme fléchie et d'analyser la structure de la phrase (L'Homme, 2008 : 248 - 250). Pour cette raison, ils sont aussi connus sous le nom de « systèmes linguistiques » (Bouillon et Clas, 1993 : 47). Leur fonctionnement est bien illustré dans le triangle de Vauquois (**Figure 2**) (Jurafsky et Martin, 2014 : 999).

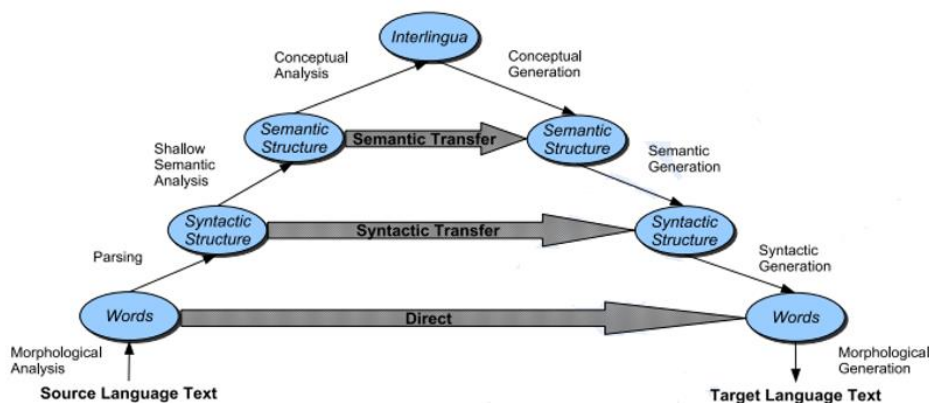


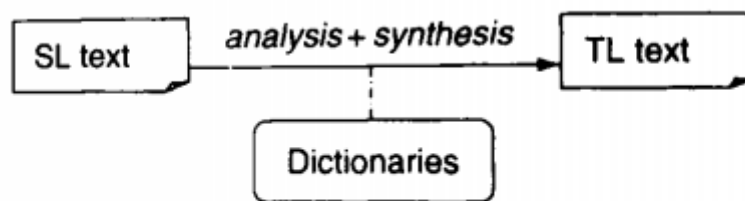
Figure 2 Triangle de Vauquois (Jurafsky et Martin, 2014 : 999).

Comme L'Homme (2008 : 258) affirme « [p]lus on analyse le texte source, plus la représentation intermédiaire est épurée et peut être réutilisée directement pour la génération du texte cible sans passer par un module de transfert proprement dit. En revanche, lorsque l'analyse du texte source est sommaire [...] un grand nombre de correspondances interlinguistiques doivent être formulées. » Le triangle permet donc de distinguer les systèmes restant au plus proche du texte et procédant mot à mot (système direct) de ceux visant un niveau de représentation plus abstraite, plus ou moins indépendante des langues (systèmes indirects) (Poibeau, 2019 : 21 - 22).

Les systèmes basés sur les règles se classent en deux grandes catégories que nous détaillerons ci-après : systèmes directs et systèmes indirects (Poibeau, 2019 : 21).

2.3.1.1 Systèmes directs

Les systèmes directs, dits de première génération ou minimalistes, se situent en bas du triangle et consistent à fournir la traduction en langue cible en deux étapes (Figure 3): après une analyse morphologique superficielle du texte source visant à attribuer à chaque mot sa catégorie grammaticale, la consultation du dictionnaire fournit les équivalences traductives mot à mot ; puis l'application de règles locales de réarrangement de l'ordre des mots permet de générer la traduction en langue cible (Bouillon et Clas, 1993 : 46 et Jurafsky et Martin, 2014 : 1000).



SL = source language; TL = target language

Figure 3 Fonctionnement des systèmes directs (Quah, 2006 : 70).

2.3.1.2 Systèmes indirects

Les systèmes indirects, également appelés de deuxième génération ou maximalistes, prévoient une étape d'analyse syntaxique complète qui leur permet de désambiguïser la phrase source en lui attribuant un sens (Bouillon, 2018b : 3). Ils ne mettent pas en relation des mots entre langue source et cible, mais des représentations des phrases (Bouillon et Clas, 1993 : 46 - 47). Cela peut être fait à travers deux approches : par « transfert » ou par « interlingue ».

Dans l'approche par transfert, la traduction est générée en trois étapes (Figure 4).

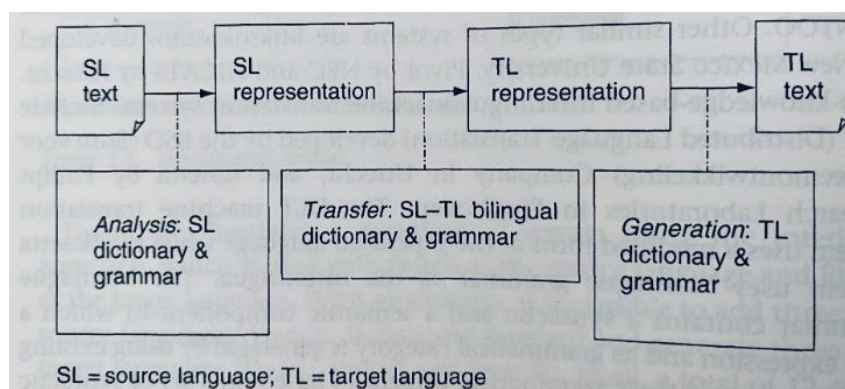


Figure 4 Processus traductif dans les systèmes par transfert (Quah, 2006 : 74).

L'analyse (« parsing »), permet d'extraire une représentation qui indique les relations entre les mots ; puis le « transfert », où le système associe les informations qu'il a extraites du texte source à celles dont il a besoin pour produire la représentation cible; enfin la synthèse (« generation »), qui consiste à fournir la traduction en langue cible (Arnold et al., 1994 : 50 et

L'Homme, 2008 : 256). Dans ce contexte, les représentations intermédiaires, qui sont de deux types selon le moment où elles ont lieu (entre l'analyse et le transfert et entre le transfert et la synthèse), dépendent alors des langues en question (Hutchins, 1986 : 73 - 75).

Dans l'approche par interlingue, le processus traductif présente deux modules (Figure 5) : analyse linguistique du texte source qui aboutit à une représentation intermédiaire indépendante de langue source et de la langue cible ; puis génération en langue cible en partant de la représentation intermédiaire (Bouillon et Clas, 1993 : 47 et Jurafsky et Martin, 2014 : 1005). Cette représentation, qui est élaborée en utilisant une interlangue basée sur des primitives sémantiques, logiques et/ou un modèle du domaine, constitue le principal avantage de ces approches, puisque « *no bilingual knowledge is needed to add a new language to an existing system* » (Bouillon et Clas, 1993 : 47 et Forcada, 2010 : 215 - 223).

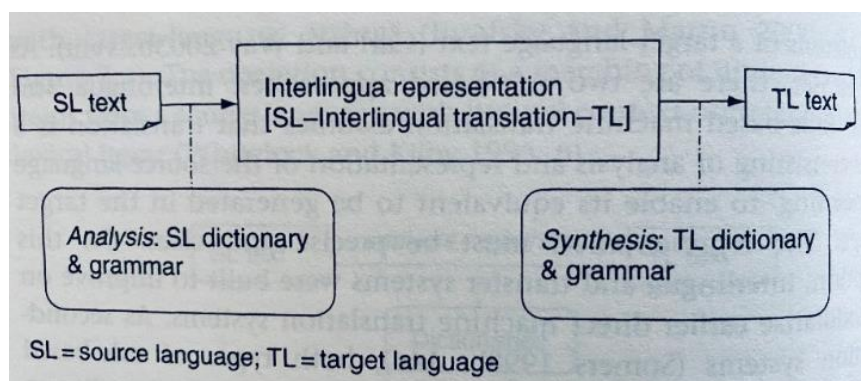


Figure 5 Processus traductif dans les systèmes par interlingue (Quah, 2006 : 72).

2.3.2 Systèmes basés sur les corpus

Le fonctionnement des systèmes basés sur les corpus repose sur l'exploitation d'une grande quantité de textes connus sous le nom de « corpus » (Jurafsky et Martin, 2014 : 1017). Ceux-ci sont généralement multilingues et de deux types : parallèles¹² ou comparables¹³ (Prochasson, 2009 : 1 - 106). Les systèmes basés sur les corpus peuvent être classés en trois catégories : les systèmes basés sur les exemples, les systèmes statistiques et les systèmes neuronaux.

¹² Les corpus sont parallèles lorsqu'ils sont composés de traductions mutuelles (Koehn, 2017 : 23). Ils sont aussi appelés « *bitextes* » (Jurafsky et Martin, 2014 : 1017).

¹³ Les corpus sont comparables lorsqu'ils sont formés par de textes ayant des traits communs (Prochasson, 2009 : 3).

2.3.2.1 Systèmes basés sur les exemples

Les systèmes basés sur les exemples ou par analogie, qui reposent sur la notion de « similarité » dans le contexte de corpus parallèles bilingues, fournissent une traduction en passant par trois étapes (Figure 6) (Hearne et Way, 2011 : 205 et Quah, 2006 : 81). En premier lieu, ils cherchent parmi des exemples préexistants celui qui est le plus proche de la phrase à traduire (Quah, 2006 : 81). Pratiquement, ces « exemples » correspondent à des unités de traduction emmagasinées dans des « mémoires de traduction¹⁴ », qui ont été créées à partir d'un corpus bilingue (Koehn, 2020 : 36 et Quah, 2006 : 81). Ensuite, le système compare à l'aide d'un algorithme l'exemple le plus proche avec le correspondant en langue cible, qui constituera le point de départ pour la traduction (Quah, 2006 : 81). Enfin, il recombine ces segments pour générer le texte en langue cible (Quah, 2006 : 81).

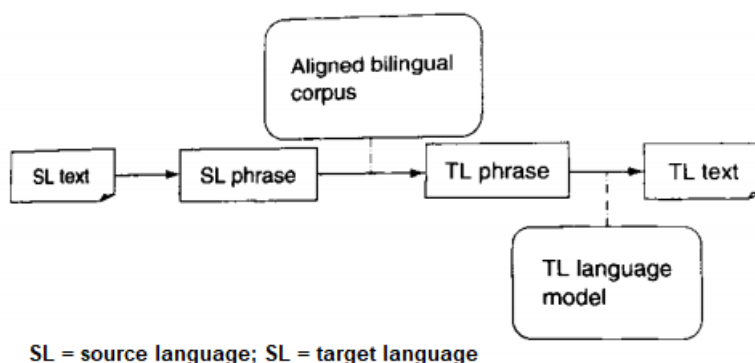


Figure 6 Processus traductif dans les systèmes basés sur les exemples (Quah, 2006 : 81).

2.3.2.2 Systèmes statistiques

Les systèmes statistiques qui permettent d'obtenir un équivalent cible en décomposant la phrase en mots ou en séquences de mots (« *unigrams*, *bigrams*, *n-grams* »), partent du principe qu'indépendamment du couple des langues considérées, il existe plusieurs traductions

¹⁴ Dans le contexte des mémoires de traduction, qui sont des bases de données où on enregistre les phrases, paragraphes ou segments de textes qui ont déjà été traduits auparavant, chaque entrée, ou segment, comprend le texte en langue d'origine, connu sous le nom de « texte source » et sa traduction, connue sous le nom de « texte cible ». Ces paires sont appelées « unités de traduction » ou « UT » (SDL Trados, 2020).

possibles d'une phrase donnée et le choix parmi ces possibilités dépend d'une question de sensibilité et de précision de la traduction visée (Hearne et Way, 2011 : 209 - 211, Koehn, 2017 : 81 - 127 et Poibeau, 2019 : 95). Pour le calcul de ces probabilités, deux modèles mathématiques peuvent être utilisés : le « *noisy-channel model* » et le « *log-linear model* » (Hearne et Way, 2011 : 206).

Indépendamment de la formule mathématique utilisée, le fonctionnement des systèmes statistiques repose sur deux étapes (Figure 7) : l'entraînement et le décodage (Hearne et Way, 2011 : 205).

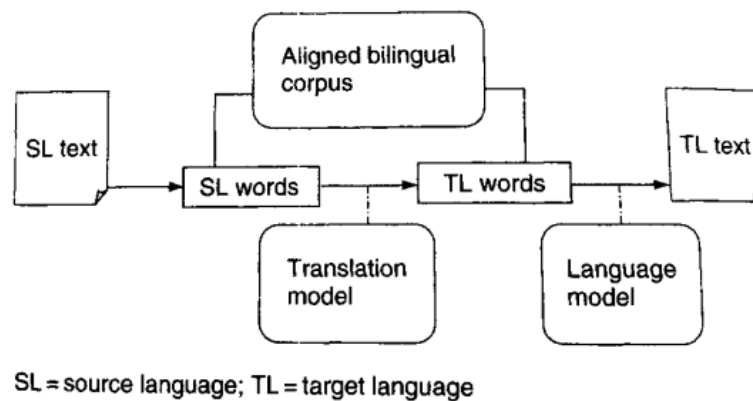


Figure 7 Processus traductif dans les systèmes statistiques (Quah, 2006 : 79).

L'entraînement consiste à extraire un modèle statistique de traduction d'un corpus parallèle et un modèle statistique de la langue cible d'un corpus monolingue qui est généralement beaucoup plus grand (Hearne et Way, 2011 : 205). En d'autres termes, lorsqu'une nouvelle phrase est soumise au système, celui-ci découpe la phrase en segments et recherche les traductions les plus probables pour chaque segment parmi le modèle de traduction (Poibeau, 2019 : 106). Ensuite, le modèle de langage permet de calculer la probabilité des séquences ainsi obtenues dans la langue cible (Poibeau, 2019 : 106). Comme le précisent Hearne et Way (2011 : 205), le modèle de traduction identifie la traduction la plus fidèle en parcourant la grande quantité de textes alignés, tandis que le modèle de langage détermine la fluidité en analysant un corpus monolingue. Enfin, le « décodeur » fait un tri parmi un éventail de choix possibles et maximise la probabilité que la phrase cible générée par le système soit correcte dans la langue d'arrivée en tenant compte du modèle de traduction et du modèle du langage (Poibeau, 2019 : 107).

2.3.2.3 Systèmes neuronaux

Les systèmes neuronaux considèrent la phrase comme un tout et pour représenter le sens des mots, ils procèdent par une analyse contextuelle qui tient compte des mots dans le contexte (Poibeau, 2019 : 121 -122). Comme le dit Koehn en reprenant la citation de Firth, « *you shall know a word by the company it keeps* » (Koehn, 2020 : 108). En d'autres termes, la traduction neuronale reprend les avantages des celle statistique, mais elle se sert de représentations de mots plus riches et analyse les phrases entières, en évitant l'assemblage de fragments de différentes traductions (Poibeau, 2019 : 121).

Le fonctionnement des systèmes neuronaux repose sur deux éléments principaux : le réseau de neurones et le « plongement de mots » (« *word embeddings* » ; **Figure 8**) (Poibeau, 2019 : 121).

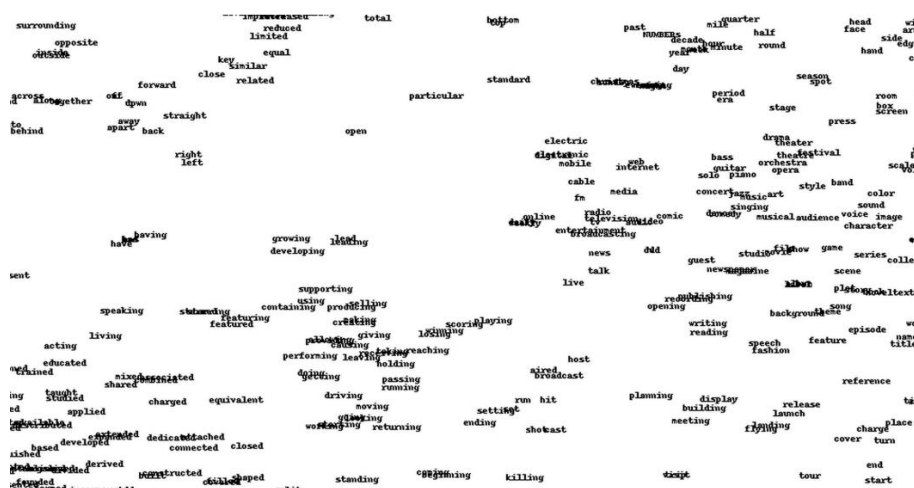


Figure 8 Exemple de « plongement de mots » (Koehn, 2017 : 36).

Plus précisément, la traduction en langue cible est produite par une architecture faisant appel à un réseau de neurones qui ne se compose à la base que « d'un encodeur et d'un décodeur » (Poibeau, 2019 : 133 - 135). L'encodeur analyse le segment source en procédant mot par mot et cette analyse permet d'obtenir une matrice, où chaque ligne et chaque colonne correspondent à un mot donné et l'intersection d'une ligne et d'une colonne donne le nombre des fois où les deux mots apparaissent ensemble dans le corpus (co-occurrences) (Poibeau, 2019 : 123). Ensuite, la matrice est normalisée et simplifiée en rassemblant graduellement les contextes les

plus similaires (Poibeau, 2019 : 124). Ce processus, où l'encodeur comprime toutes les informations d'une langue source dans un vecteur d'une longueur n formé par des « cases » contenant chacune une valeur numérique, prend le nom de « plongement des mots » (Bahdanau et al., 2014 : 1 et Poibeau, 2019 : 124). Enfin, le décodeur se sert du vecteur produit par l'encodeur pour générer la traduction en langue cible (Bahdanau et al., 2014 : 1).

Au cours des années, l'architecture de la traduction neuronale s'est perfectionnée et actuellement, les systèmes neuronaux se fondent dans la plupart des cas sur des paramètres globaux dits « *hyperparamètres* » qui sont déterminés empiriquement (Poibeau, 2019 : 137). Cela contribue à percevoir ces systèmes comme des boîtes noires, dont la manque de lisibilité constitue leur principal point faible et un des majeurs thèmes de recherche (Poibeau, 2019 : 137).

2.4 Évaluation de la traduction automatique

À présent, les questions les plus fréquemment posées lorsqu'on parle de TA sont : « *How good is [...] machine translation today ?* » et « *When will we reach fully automated high quality machine translation?* » (Koehn, 2010 : 217 et 2020 : 29). Aujourd'hui comme hier, personne ne pense que la traduction automatique de haute qualité des textes complexes représente un objectif qui peut être atteint à très court terme (Smith, 2001 : 39).



Facebook's AI Just Set A New Record In Translation And Why It Matters

Linguists, update your resumes because Baidu thinks it has cracked fast AI translation

Microsoft AI translates news as well as humans, takes on Google Translate

SDL Cracks Russian to English Neural Machine Translation

Figure 9 Évaluation de la TA (Koehn, 2020 : 48).

En effet, la traduction automatique ne donne pas encore des résultats qui sont comparables à ceux d'une traduction faite par un être humain (Figure 9). Toutefois, elle n'a pas besoin d'être

impeccable pour en pouvoir profiter (Koehn, 2010 : 20). Dans cette optique, on distingue trois champs d'application de la TA, où le niveau de qualité et rapidité augmente progressivement : « assimilation » ou, en anglais, « *gisting* » (qualité acceptable pour assurer la compréhension immédiate du contenu, p.ex. consultation d'un site internet étranger), « communication » (bonne qualité pour permettre les échanges, tels que ceux par e-mail) et « dissémination » (très bonne qualité pour garantir la diffusion des informations, p.ex. publications en plusieurs langues) (Koehn, 2010 : 20). Selon Koehn (2010 : 21) « *[t]he most widespread use of machine translation is *gisting** », puisque « *[t]ranslation does not have to be perfect to bring across the meaning* ».

Indépendamment du champ d'application de la TA, il existe plusieurs méthodes pour évaluer la qualité d'une traduction automatique. Avant de les détailler, il importe de dire que peu importe la métrique utilisée, il faut que celle-ci soit : peu coûteuse et rapide (« *low cost* ») ; elle doit pouvoir déterminer la qualité réelle de la traduction automatique (« *meaningful* » et « *correct* ») ; elle doit enfin être systématique (« *consistent* »), afin que la traduction et tous les systèmes soient mesurés de la même façon par tous ou toutes les juges (« métrique stable ») et que le résultat de leur évaluation ne diffère pas (« accord entre les juges » ou, en anglais, « *inter-annotator agreement* ») (Koehn, 2010 : 220 - 221). Dans le cadre de l'évaluation humaine, l'accord entre les juges peuvent être mesuré en utilisant le « *score Kappa* », égal au nombre de fois où tous ou toutes les juges ont fait le même choix divisé par le nombre total de segments à évaluer (Koehn, 2020 :48). Compte tenu de ces conditions, il est de possible classer les méthodes d'évaluation en deux grands groupes : évaluation humaine (section 2.4.1) et évaluation automatique (section 2.4.2). Le choix des méthodes passe quatre étapes : identification des besoins à satisfaire et des attentes envers le système ; élaboration d'une liste de critères ordonnés en fonction de leur importance et choix d'une technique permettant de réunir les données relatives à ces derniers ; collecte et analyse des données ; formulation d'un jugement sur la base des résultats de l'analyse des données (Bouillon et Clas, 1993 : 266).

2.4.1 Évaluation humaine

Il existe plusieurs méthodes d'évaluation humaine, dont celles « intrinsèques » estimant la qualité de la traduction automatique (évaluation manuelle, jugement intuitif (fluidité et fidélité), évaluation comparative, analyse des erreurs) et les approches « extrinsèques »

déterminant l'utilité que celle-ci a dans le contexte où elle est utilisée (évaluation fondée sur la tâche) (Kit et Wong, 2015 : 251 - 274). Ci-après, nous les passerons brièvement en revue.

L'évaluation manuelle consiste à déterminer si chaque phrase traduite automatiquement est correcte ou non (Koehn, 2010 : 218). Cette tâche est généralement confiée à des juges bilingues, capables de comprendre aussi bien le texte source que celui en langue cible (Koehn, 2010 : 218). Cependant, les juges ne sont pas toujours disponibles à accomplir ce travail, en conséquence on fait souvent appel à des évaluateurs ou des évaluateurs monolingues qui ne comprennent que la langue cible, mais qui peuvent estimer la qualité de la traduction automatique en utilisant une traduction de référence (« *reference translation* ») (Koehn, 2010 : 218).

Pour arriver à des résultats plus précis, les juges se basent souvent sur deux critères d'évaluation : fluidité (« *fluency* ») et fidélité (« *adequacy* ») (Koehn, 2020 : 45). La fluidité concerne la langue cible, plus précisément la justesse grammaticale et le choix idiomatique des mots, tandis que la fidélité vise à garantir l'exactitude du texte de départ (Koehn, 2020 : 46). Afin d'obtenir une évaluation détaillée basée sur ces deux critères, Koehn (2020 : 46) propose une échelle de valeurs (Figure 10).

Adequacy		Fluency	
5	all meaning	5	flawless English
4	most meaning	4	good English
3	much meaning	3	non-native English
2	little meaning	2	disfluent English
1	none	1	incomprehensible

Figure 10 Échelles de valeurs pour estimer la fidélité et la fluidité (Koehn, 2020 : 46).

Cette méthode, qui est connue sous le nom de « jugement intuitif », peut causer des difficultés si les juges n'arrivent pas à discerner ces deux critères et à estimer la qualité toujours de la même façon et avec la même constance (Koehn, 2020 : 46 - 47).

Pour résoudre ces difficultés, il est possible dans le contexte d'évaluation de plusieurs systèmes de TA d'effectuer une évaluation comparative, où les juges attribuent un score à la meilleure

traduction et ils n'assignent aucun point s'ils jugent les traductions équivalentes (Koehn, 2020 : 47). Toutefois, l'évaluation comparative ne donne pas des renseignements sur le type d'erreurs.

Celui-ci peut être identifié en faisant une analyse des erreurs. Concrètement, les juges identifient les erreurs suivant des standards (« *Multidimensional Quality Metrics, MQM* ») et attribuent un coefficient à chaque erreur en fonction des conséquences qu'elle a sur la fidélité et la fluidité, de la difficulté de sa correction et du temps nécessaire à la correction (Hutchins et Somers, 1992 : 164 et Volkart, 2018 : 24). Bien que l'analyse des erreurs permette d'identifier les inexactitudes, ces dernières ne sont pas perçues de la même façon par tous ou toutes les juges (Kit et Wong, 2014 : 223). En conséquence, les catégories d'erreurs ne sont pas toujours claires et certaines erreurs peuvent appartenir à plusieurs catégories différentes (Kit et Wong, 2014 : 223 en Volkart, 2018 : 25).

L'évaluation fondée sur la tâche tient compte de la fonction de la traduction automatique dans un certain contexte. Dans cette optique, elle peut par exemple servir à mesurer l'effort de la post-édition en quantifiant le temps et les modifications nécessaires pour obtenir une traduction de bonne qualité (Koehn, 2020 : 44). En faisant cela, il est également possible d'estimer la productivité des traducteurs et la qualité (intrinsèque) de la traduction automatique, puisque l'effort de post-édition est moins important lorsque la TA est de bonne qualité (Kit et Wong, 2014 : 213 - 236 en Volkart, 2018 : 25 et Koehn, 2020 : 44).

2.4.2 Évaluation automatique

Les méthodes d'évaluation automatique permettent de mesurer la qualité d'une traduction automatique, en gagnant du temps et en obtenant des résultats qui sont objectifs (Koehn, 2020 : 52 - 53). De façon similaire à l'évaluation humaine, celle automatique regroupe nombreuses méthodes, telles que précision, rappel et f-measure, Word Error Rate (WER) (dit aussi « Edit distance »), Translation Edit Rate (TER), Human Translation Edit Rate (HTER), BLEU, METEOR et NIST. Par la suite, nous les présenterons en faisant un panoramique.

Dans le cadre des mesures de « précision et rappel » (Figure 11), la précision calcule le nombre des mots corrects dans la sortie du système automatique par rapport au nombre total que celle-ci contient (Koehn, 2010 : 223). En revanche, le rappel mesure le nombre des mots corrects dans le texte traduit automatiquement sur le nombre total des mots figurant dans la traduction de référence (Koehn, 2010 : 223).

$$\text{precision} = \frac{\text{correct}}{\text{output-length}}$$

$$\text{recall} = \frac{\text{correct}}{\text{reference-length}}$$

Figure 11 Formule de précision et rappel (Koehn, 2010 : 223).

L'utilisation d'une seule de ces deux mesures risque de donner des résultats incorrects (Koehn, 2010 : 223). En effet, une traduction courte ne contenant que des mots corrects peut avoir une précision élevée, mais, en même temps, ne pas inclure plusieurs mots de la traduction de référence (Koehn, 2010 : 223). Par contre, un texte qui se compose de nombreux mots corrects ainsi que de mots ne pas figurant dans la traduction de référence, peut présenter un rappel important (Koehn, 2010 : 223). Puisque toutes les deux mesures sont importantes en traduction automatique, il existe la possibilité de les combiner parmi la « *f-measure* » (Koehn, 2010 : 224). Ce score est obtenu en divisant le produit de la précision et du rappel par la moitié de la somme de la précision et du rappel (Figure 12) (Koehn, 2010 : 224).

$$\text{f-measure} = \frac{\text{precision} \times \text{recall}}{(\text{precision} + \text{recall})/2}$$

Figure 12 Formule de « f-measure » (Koehn, 2010 : 224).

La qualité d'une traduction peut également être mesurée à travers le « Word Error Rate (WER) », égal au nombre minimum d'édicions nécessaires (insertion, suppression et remplacement), afin qu'une séquence soit identique à celle dans la traduction de référence (« distance de Levenshtein ») divisé par le nombre de mots figurant dans cette dernière (Figure 13) (Koehn, 2010 : 224). Cette métrique est aussi connue sous le nom de « *Edit distance* » (Przybocki et al., 2006 : 2039). Toutefois, ce score tend à sous-estimer des traductions correctes, mais dont la formulation diffère sensiblement de la traduction de référence (Gerlach, 2015 : 113 en Volkart, 2018 : 27).

$$\text{WER} = \frac{\text{substitutions} + \text{insertions} + \text{deletions}}{\text{reference-length}}$$

Figure 13 Formule du Word Error Rate, WER (Koehn, 2010 : 225).

Le Translation Edit Rate (TER) repose sur le même principe du Word Error Rate (WER), cependant il tient compte également compte des déplacements (« *shifts* ») de séquences de mots (Snover et Dorr, 2006 : 225).

Le Human Translation Edit Rate (HTER) reprend le calcul du Translation Edit Rate (TER), mais il se base sur une traduction de référence qui correspond à la traduction automatique post-éditée par un traducteur humain (Snover et al., 2009 : 259).

Plusieurs études évaluant la qualité d'un texte traduit automatiquement font également appel à la méthode automatique BLEU, qui repose sur la notion de précision (Scansani et al., 2019 : 83 et Farajian et al., 2018 :154). Contrairement à la « précision » qui tient compte uniquement des mots en reflétant le concept de « fidélité », BLEU considère la longueur des séquences des mots et alors reproduit la « fluidité » (Papineni et al., 2002 : 313). Ce score, qui se fonde sur le principe de comparer une traduction de référence (« T_{ref} ») avec une traduction produite automatiquement (« T_{Auto} »), est calculé en découpant les deux traductions en segments de longueur 1 à n et en comparant le nombre de segments communs (Poibeau, 2019 : 158). Normalement, le résultat plus fiable est obtenu avec une valeur de $n = 4$ (Poibeau, 2019 : 158). Dans le contexte du calcul de BLEU, la précision est modifiée (« *modified n-gram precision* ») (Papineni et al., 2002 : 312). En clair, les n -grammes corrects figurant dans le texte traduit automatiquement ne sont comptés que selon le nombre de leurs occurrences dans la traduction de référence (Papineni et al., 2002 : 312). Cela veut dire qu'un mot donné, comme « terminologie », apparaissant respectivement dans la traduction de référence et dans le texte traduit automatiquement trois fois et sept fois ne sera compté que trois fois. Toutefois, la précision modifiée peut nuire aux phrases longues et favoriser les phrases courtes (Papineni et al., 2002 : 313 - 314). Pour résoudre ce problème, une « pénalité de brièveté » est utilisée pour réduire le score des phrases plus courtes que la référence (Papineni et al., 2002 : 313 - 314).

Bien que BLEU soit largement répandu et offre plusieurs avantages, tels que la possibilité d'évaluer la sortie du système automatique par rapport à plusieurs traductions de référence, cette méthode d'évaluation présente des points faibles (Koehn, 2010 : 229). Dans le cadre de

notre mémoire, nous nous limiterons à mentionner sa tendance à sous-estimer la qualité du texte traduit automatiquement ainsi que celle des systèmes neuronaux (Shterionov et al., 2017 : 74). Plus précisément, BLEU ne calcule pas les mots en fonction de leur importance (Koehn, 2010 : 229). Cet aspect est partiellement résolu avec deux autres métriques automatiques dérivées de BLEU : « *NIST* », qui donne plus de poids aux n-grammes dont la valeur informative est supérieure et la fréquence est plus faible ; et « *METEOR* », qui contrairement à BLEU n'identifie pas seulement les n-grammes identiques, mais prend aussi en compte les variantes morphologiques ainsi que les synonymes (Banerjee et Lavie, 2005 : 65 - 72 et Doddington, 2002 : 138 - 145).

2.5 Conclusions sur la TA

Ce chapitre porte sur les fondements de la traduction automatique. Premièrement, nous avons jugé primordial de définir la notion de TA (section 2.1), puis nous avons parcouru les évènements les plus marquants dans le domaine (section 2.2). D'une part, nous avons évoqué certains précurseurs. De l'autre, nous avons surtout mentionné les recherches qui ont eu un véritable impact sur la traduction automatique. Après cela, nous avons décrit les systèmes de traduction automatique selon le type de données utilisées pour le développement du système (2.3). Enfin, nous avons présenté les différentes méthodes actuellement utilisées pour évaluer une traduction automatique (section 2.4). Dans le prochain chapitre (3), nous mettrons l'accent sur le traitement de la terminologie dans la traduction automatique.

3 Traduction automatique et terminologie

Lorsque la traduction porte sur différents domaines, les mots sont traduits différemment. En conséquence, il est important que les systèmes de traduction automatique fournissent des traductions qui préservent les connaissances du domaine. Dans cette optique, nous définirons la terminologie et nous éclaircirons la raison de son importance en TA (section 3.1). Ensuite, nous nous intéresserons à la façon dans laquelle la terminologie est traitée dans les différents systèmes de TA en faisant référence à leur classement dans le chapitre précédent (section 2.3). Puis nous exposerons les métriques automatiques permettant d'évaluer la terminologie en TA (section 3.3). Enfin, nous résumerons le présent chapitre (section 3.4).

3.1 Qu'est-ce que la terminologie et pourquoi est-elle importante en TA ?

Depuis que se manifeste et se développe le langage, nous sommes témoins de la présence de langues spécialisées (Baldassarre, 2016 :11). La science, la recherche et la technique, comme de nombreux autres domaines de spécialité tels la politique, la culture, l'économie et le commerce évoluent et le rythme accéléré de ces changements se reflète dans la diversification généralisée des savoirs, ainsi que dans la forte augmentation de la communication spécialisée (Baldassarre, 2016 : 11 et CST, 2014 :10). C'est là que la terminologie, que nous définirons ci-après, joue un rôle essentiel pour garantir la qualité de la communication spécialisée, en permettant de diffuser les informations d'une façon exacte, claire et cohérente (CST, 2014 : 10 - 11).

La « terminologie » présente plusieurs acceptions (Baldassarre, 2016 : 16). En premier lieu, elle désigne l'ensemble des unités lexicales rattachées à un domaine de spécialité, qui peut être vu comme un « *paramètre classificateur des sens* » (Baldassarre, 2016 : 11, L'Homme, 2020 : 59 - 60). Elle définit également la science qui étudie, d'une part, les concepts et leurs formes linguistiques dans le cadre des vocabulaires spécialisés (étude théorique) et, d'autre part, les méthodes propres au travail terminologique (terminographie¹⁵) (CST, 2014 : 16). Plus précisément, la terminologie repose sur trois éléments fondamentaux : concept, objet et terme,

¹⁵ La terminographie s'intéresse à la collecte, à la consignation, au traitement et à la présentation des données terminologiques, alors que la terminologie traite des aspects théoriques portant sur les concepts et les termes, (Baldassarre, 2016 : 17 et ISO 1087-2: 1989(F) : 13).

dont les liens existants peuvent être représentés par le triangle sémiotique (Figure 14) (Baldassarre, 2016 : 12 et CST, 2014 :17).

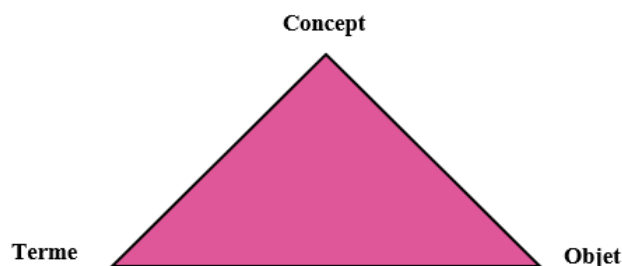


Figure 14 Triangle sémiotique (Baldassarre, 2016 : 11).

Selon l'ISO, le concept est la représentation abstraite de l'ensemble des caractéristiques communes essentielles à un groupe d'entités, qui déterminent « l'objet » (Baldassarre, 2016 : 12 et ISO 1087-1 : 2000(F) : 1). Il est également considéré comme une unité de connaissance, puisque la terminologie est liée aux langues de spécialités et, en conséquence, à un domaine restreint (Baldassarre, 2016 : 12 et ISO 704 : 2009 (F) : 3). L'objet constitue une entité abstraite ou concrète de la réalité qui « *peut être perçue ou conçue* » sans aucune ambiguïté (Baldassarre, 2016 : 12 et ISO 704 : 2009(F) : 2). Enfin, le terme, qui peut être défini comme « *une manière succincte de référencer le concept* », représente la forme linguistique du concept (Baldassarre, 2016 : 13 et ISO 704 : 2009(F) : 22). Idéalement, dans une langue de spécialité un terme ne désigne qu'un seul concept (« monosémie ») afin d'assurer une communication précise et ad hoc (Baldassarre, 2016 : 15 et ISO 704 : 2009(F) : 34). C'est ce qu'on appelle le « principe de biunivocité » (L'Homme, 2020 : 29). La pratique nous oblige néanmoins à admettre qu'un même terme renvoie souvent à plusieurs concepts (« polysémie ») et que divers termes ayant des formes écrite et phonétique variées sont attribués au même concept (« synonymie »), pour cette raison, ils peuvent être interchangeables dans des contextes particuliers (Baldassarre, 2016 : 15, ISO 704 : 2009(F) : 34 - 35 et L'Homme, 2020 : 29). Un seul terme peut également correspondre à plusieurs concepts différents en fonction du domaine (« homonymie ») (Baldassarre, 2016 : 15 et ISO 704 : 2009(F) : 34 - 35). De plus, des formes linguistiques différentes liées à un même concept (« variations terminologiques ») peuvent cohabiter dans un même texte, raison pour laquelle il est important de les définir clairement, afin d'assurer la qualité de la communication spécialisée (L'Homme, 2020 : 81).

Car la terminologie joue un rôle fondamental en communication spécialisée, il est nécessaire de choisir de façon appropriée les termes et leurs différentes formes linguistiques qui doivent être employées dans un texte. Cela s'applique notamment en traduction, où il faut préserver les connaissances du domaine de la langue originale (Haque et al., 2019 : 1 et Poibeau, 2019 : 11). Nous pouvons constater cette nécessité à plusieurs niveaux. Dans le cadre des organismes publics, par exemple, il faut garantir la cohérence conceptuelle dans la législation ainsi que l'harmonisation des dialogues dans la coopération internationale (CST, 2014 : 10 - 11). En ce qui concerne les entreprises, leur volonté d'être concurrentielle sur la scène tant nationale qu'internationale est étroitement liée à une utilisation correcte, cohérente et uniforme de leurs informations et de leur documentation (Baldassarre, 2016 : 36 et CST, 2014 : 11 - 12). En conséquence, la terminologie utilisée représente une vitrine linguistique qui va « accrocher » les usagers pour communiquer d'une façon efficace et incontournable (Baldassarre, 2016 : 37). Ces contextes apportent à la terminologie un caractère politique, économique, de productivité et de commercialisation, tout comme une exigence de qualité, de sécurité et de rentabilité qui ne sont pas négligeables (Baldassarre, 2016 : 36 et CST, 2014 : 9).

Dans ces circonstances, il est alors fondamental que les systèmes de traduction automatique fournissent également des traductions qui garantissent l'emploi exact et cohérent de la terminologie (Farajian et al., 2018 : 149). Toutefois, cela constitue un grand défi pour les systèmes de TA, qui emploient souvent des mots du langage commun au détriment de la communication spécialisée. Par exemple, DeepL Pro a traduit « Präsenzunterricht verzichten », qui en français veut dire « renoncer à l'enseignement présentiel », par « do without attendance classes » au lieu d'avoir utilisé la formulation employée en communication spécialisée « forgo face-to-face teaching ». Aussi, certains noms propres ou acronymes sont traduits avec des fautes d'orthographe (p.ex. DeepL Pro a écrit « The Federal Office of Public Health is thus following the recommendations of the European *Centre* for Disease Prevention and Control (ECDC) ») ou alors quelques variations terminologiques produisent des incohérences avec la répétition de certains concepts (p.ex. DeepL Pro a donné plusieurs solutions en ce qui concerne « Covid-19 vaccin », « COVID-19 vaccin », « Covid 19 vaccin » et « COVID 19 »).

3.2 Les systèmes de TA et la terminologie

Les systèmes de traduction automatique mettent en place différentes techniques pour générer une traduction, où la terminologie correspond au domaine spécialisé traité (Vasconcellos et al.,

2001 : 698). Par la suite, nous verrons comment les différents systèmes de TA, que nous avons passés en revue dans le chapitre précédent (section 2.3), permettent de traiter la terminologie d'un domaine spécialisé.

3.2.1 La terminologie dans les systèmes basés sur les règles

Les systèmes basés sur les règles, qui « sont fondés sur des listes d'équivalence terme à terme », permettent aux utilisateurs d'effectuer des traductions, tout en garantissant une certaine terminologie (Bouillon et Clas, 1993 : 46). En l'espèce, ils permettent de définir la façon dans laquelle traduire certains mots ou expressions en créant un nouveau dictionnaire bilingue qui a la priorité sur les dictionnaires généraux, intégrés par défaut aux systèmes (Jurafsky et Martin, 2014 : 998 et Vasconcellos et al., 2001 : 702).

Les systèmes directs reposent sur la consultation d'un seul dictionnaire, qui inclut toutes les informations monolingues et bilingues (Bouillon, 2018a : 4 - 19). Les informations monolingues, qui sont utilisées pour les étapes de compréhension et génération, incluent les aspects morphologiques (catégorie grammaticale, type de flexions du mot, etc.), syntaxiques (rection, type d'auxiliaires, etc.) et sémantiques des termes (Bouillon, 2018a : 4 - 5). En revanche, les informations bilingues, qui trouvent leur application dans l'étape de la traduction, servent à remplacer les unités de la langue source par l'équivalent en langue cible, tout en appliquant un nombre limité de tests et d'actions¹⁶ (Figure 15) (Bouillon, 2018a : 4 - 19).

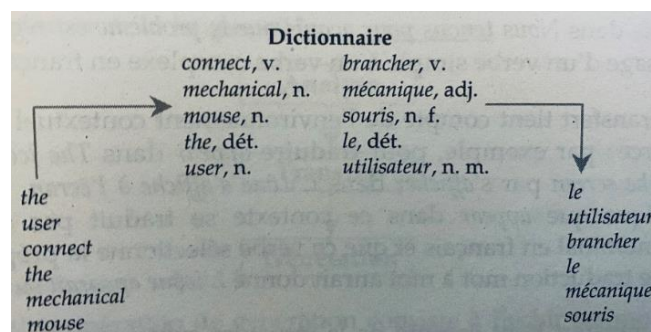


Figure 15 Extraction des équivalents en LC (L'Homme, 2008 : 251).

¹⁶ Les tests permettent de choisir la bonne traduction dans le système linguistique, s'il y en a plusieurs (« indicateurs de sens »), alors que les actions permettent de traiter certaines divergences, c'est-à-dire de changer la syntaxe sous certaines conditions (« X misses Y » en anglais devient « Y manque à X » en français) (Bouillon, 2018a : 8 - 15).

En revanche, les systèmes indirects par transfert comprennent deux dictionnaires différents, chacun pour la langue en question, plus des règles d'équivalences lexicales servant à fournir les équivalents en langue cible (Hutchins et Somers, 1992 : 79). Le dictionnaire de la langue source comprend les informations nécessaires à effectuer l'analyse syntaxique et à procéder à la désambiguïsation, telles que les catégories grammaticales, les traits sémantiques, etc. (Hutchins et Somers, 1992 : 79). Le dictionnaire de la langue cible est moins détaillé, puisqu'il ne contient que des données morphologiques (Hutchins et Somers 1992 : 79).

En ce qui concerne les systèmes indirects par interlangue, ils détectent les équivalents traductiques en représentant ensemble toutes les phrases qui expriment le même concept indépendamment de leur langue (Figure 16) (Jurafsky et Martin, 2004 : 1005).

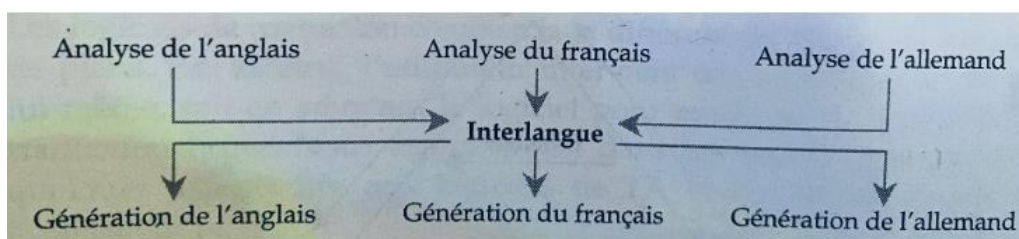


Figure 16 Création de l'interlangue (L'Homme, 2008 : 257).

3.2.2 La terminologie dans les systèmes statistiques

Les systèmes statistiques, permettant d'obtenir un équivalent cible en décomposant la phrase en mots ou en séquences de mots (« *unigrams*, *bigrams*, *n-grams* »), reposent sur le principe que, peu importe du couple des langues considérées, il existe plusieurs traductions possibles d'une phrase donnée et le choix parmi ces possibilités dépend d'une question de sensibilité et de précision de la traduction visée (section 2.3.2.2). Dans cette optique, l'emploi d'une terminologie exacte et cohérente constitue un défi qui n'est pas anodin (Pinnis, 2015 : 11 - 12). Il est donc nécessaire d'entraîner et de paramétrer les systèmes statistiques, afin qu'ils choisissent les bons termes à utiliser dans une traduction en excluant les mots ou groupe de mots qui ne correspondent pas à cette terminologie, bien qu'ils aient une probabilité élevée d'apparaître dans le même contexte. Cet objectif peut être atteint de deux façons (Pinnis, 2015 : 26 - 95) :

- Pendant l'entraînement du système statistique (« *static terminology integration* »), l'on peut modifier les paramètres du modèle de traduction et celui du langage ;
- Durant l'étape de la traduction (« *dynamic terminology integration* »), l'on peut utiliser un système déjà entraîné et prétraiter le texte source (annotation des termes cibles dans le texte source en leur assignant des scores de probabilités).

3.2.2.1 « *Static terminology integration* »

En ce qui concerne la première approche (« *static terminology integration* »), il est possible de faire une distinction selon la façon d'inclure la terminologie dans les systèmes statistiques (Pinnis, 2015 : 76) :

- « *Intégration directe de la terminologie* » (ajout de collections terminologiques provenant d'un domaine spécialisé, appelées en anglais « *in-domain term collections* », aux données d'entraînement)
- « *Intégration indirecte de la terminologie* » (les données d'entraînement ne se composent que de corpus parallèles et/ou monolingues contenant la « *in-domain terminology* » ; cette approche est connue sous le nom d'adaptation au domaine, « *domain-adaptation* »).

Parmi celles qui intègrent « directement » la terminologie dans un système de TAS, nous citerons à titre d'exemple les approches proposées par Bouamor et al. (2012) et Arcan et al. (2014a). Bouamor et al. (2012 : 676) appliquent la méthode « *terminology as a corpus* » qui se compose de deux phases. En premier lieu, ils intègrent une collection terminologique bilingue dans le corpus parallèle qui est employé pour entraîner le modèle de traduction du système statistique (Bouamor et al., 2012 : 676 et Pinnis, 2015 : 81). Dans un second temps, ils ajoutent la liste d'équivalents cibles de cette collection terminologique au corpus monolingue qui est utilisé pour entraîner le modèle de langage (Bouamor et al., 2012 : 676 et Pinnis, 2015 : 81). Bien qu'elle permette d'obtenir de bons résultats, cette méthode n'aide cependant pas vraiment le système à associer la forme canonique des termes contenue dans les corpus avec leurs

variations flexionnelles dans le texte à traduire (Pinnis, 2015 : 81 - 82). Arcan et al. (2014a : 65) spécialisent un système statistique par rapport à un domaine donné, en passant par deux étapes. Premièrement, ils extraient une liste de termes bilingues de phrases parallèles fournies par des traducteurs professionnels (Arcan et al., 2014a : 54 - 64). Puis ils l'utilisent pour entraîner le système statistique (Arcan et al., 2014a : 54 - 64). Cela leur permet de constater une amélioration de la qualité de la terminologie des textes traduits, ainsi qu'une augmentation de la productivité des traducteurs (Arcan et al., 2014a : 54). En effet, ces derniers ne doivent plus consulter des bases terminologiques pendant l'étape de la traduction, puisque le système est déjà entraîné avec la liste de termes figurant dans toutes leurs différentes variations flexionnelles (Arcan et al., 2014a : 54).

En ce qui concerne l'intégration « indirecte » de la terminologie, nous évoquerons les études de Koehn et Schroeder (2007) et Bertoldi et Federico (2009). Koehn et Schroeder (2007 : 224 - 225) expérimentent plusieurs méthodes pour spécialiser un système de TAS et, finalement, ils identifient comme meilleure celle qui consiste à entraîner le modèle de traduction avec des données « *out-of-domain* » et « *in-domain* » et le modèle de langage seulement avec des données « *in-domain* ». Dans ce contexte, il est important de signaler que les données « *out-of-domain* » sont beaucoup plus nombreuses que celles « *in-domain* » (quatre millions contre un million de mots) (Koehn et Schroeder, 2007 : 224 - 225). En revanche, Bertoldi et Federico (2009) créent un corpus bilingue qui doit être utilisé pour spécialiser un système de TAS. Pour ce faire, ils laissent traduire automatiquement de données monolingues spécifiques au domaine traité, sans cependant les post-éditer (Bertoldi et Federico, 2009 : 182). Cela leur permet d'obtenir une amélioration du BLEU score de 22.60% au 28.10% (Bertoldi et Federico, 2009 : 182).

3.2.2.2 « *Dynamic terminology integration* »

L'intégration de la terminologie dans l'étape de la traduction (« *dynamic terminology integration* ») consiste à utiliser un système déjà entraîné et à prétraiter le texte source pour atteindre deux objectifs (Pinnis, 2015 : 95). En premier lieu, on vise à identifier les termes dans le texte source à l'aide d'une collection terminologique bilingue (Pinnis, 2015 : 95). Puis on cherche à générer la traduction en langue cible en annotant les termes cibles dans le texte source et en leur assignant des scores de probabilités (Pinnis, 2015 : 95). Dans ce contexte, nous ferons

référence à Carl et Langlais (2002). Ils conçoivent l’outil « *Intelligent Terminology Database (ITDB)* » qui permet de retrouver des termes et leurs variations terminologiques dans les textes ainsi que de fournir la traduction en faisant référence à la terminologie bilingue stockée (Carl et Langlais, 2002 : 1 - 7). Aussi, Babych et Hartley (2003) font usage d’une liste spéciale « *do not translate* » qui contient les dénominations de certaines organisations qui ne doivent pas être traduites, et d’une méthode de prétraitement qui restreint la traduction aux phrases identifiées (Pinnis, 2015 : 96). Ainsi, ils réussissent à améliorer la qualité de la traduction pour la paire de langues anglais - russe (Pinnis, 2015 : 96). Il est important de signaler que les principes de certaines études menées dans le contexte de l’intégration de la terminologie dans les systèmes statistiques ont été ensuite repris pour la traduction automatique neuronale (Chen et al., 2017 : 40, Chu et Wang, 2018 : 1305). En particulier, Chu et Wang (2018 : 1304 - 1305) mentionnent les méthodes qui se focalisent sur l’exploitation des données (« *data centric* ») et celles qui trouvent leur application dans la spécialisation du système automatique (« *model centric* »). Nous approfondirons ceux-ci dans la section suivante (3.2.3).

3.2.3 La terminologie dans les systèmes neuronaux

Pendant les dernières années, les systèmes neuronaux (section 2.3.2.3) se sont imposés comme les approches les plus prometteuses en traduction automatique (Bojar et al., 2016 : 131 – 198). Cependant, lorsqu’ils sont employés pour traduire des textes ou des groupes de mots dont le domaine et, en conséquence, la terminologie ne correspondent pas à ceux des données utilisées pour leur entraînement (« *out of domain texts* » ou « *out-of-vocabulary words, OOVs* »), la performance de ces systèmes est réduite (Knowles et Koehn, 2017 : 28 et Huck et al., 2019 : 5809). En effet, celle-ci dépend largement du degré de correspondance du domaine entre le corpus exploité pour l’entraînement du système neuronal et les textes à traduire (Chen et al., 2017 : 40). Bien qu’aujourd’hui il soit fondamental qu’une traduction de haute qualité reflète la terminologie d’un domaine spécialisé, il n’est pas toujours possible de repérer des corpus parallèles du domaine cible (Chu et Wang, 2018 : 1304). Les corpus varient considérablement d’un domaine à l’autre et les traductions faites à partir de corpus portant sur différents domaines sont peu fiables (Chen et al., 2017 : 40). Comme Knowles et Koehn (2017 : 28) le précisent, « *NMT systems have lower quality out of domain, to the point that they completely sacrifice adequacy for the sake of fluency* ». Cet aspect, qui en traduction automatique est aussi connu sous le nom de « *domain mismatch* », constitue un des défis majeurs en traduction neuronale,

car dans différents domaines les mots et leurs sens sont traduits différemment (Knowles et Koehn, 2017 : 28). De façon similaire aux systèmes statistiques, ce défi peut être résolu par des approches, que nous pouvons classer en deux catégories selon l'étape dans laquelle on adapte le système neuronal à la terminologie d'un domaine spécialisé : pendant l'entraînement (« *adaptation au domaine* » ou, en anglais, « *domain adaptation* ») ou durant l'étape de la traduction en utilisant un système déjà entraîné.

3.2.3.1 *Adaptation au domaine*

En ce qui concerne l'adaptation d'un système pendant son entraînement, de nombreuses méthodes ont été développées et elles peuvent être classées de la façon suivante (Figure 17).

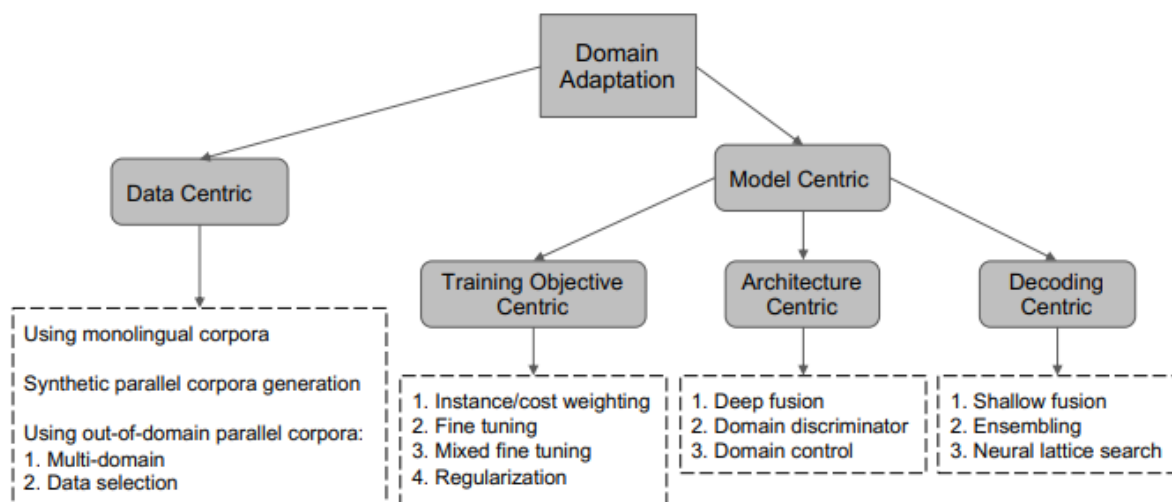


Figure 17 Approches pour adapter un système neuronal à un domaine (Chu et Wang, 2018 : 1305).

Ci-après, nous dresserons un panorama des celles-ci en faisant référence à des études, où elles ont été appliquées.

3.2.3.1.1 Méthodes « *data centric* »

Les méthodes du premier groupe (« *data centric* ») permettent d'adapter les systèmes neuronaux à la terminologie d'un domaine spécialisé en utilisant des données d'entraînement, qui peuvent être extraites à partir de corpus monolingues ou parallèles ou, s'il n'existe pas de

grands corpus à disposition, elles sont générées sous forme de « *synthetic parallel data* », (Chu et Wang, 2018 : 1304 - 1307). Nous pouvons rappeler ici l'étude de Currey et al. (2017 : 148 - 154) qui, pour améliorer la performance d'un système neuronal utilisé pour traduire paires de langues pour lesquelles il existe un nombre restreint de corpus parallèles, augmentent les données d'entraînement du système. Pour ce faire, ils créent un bitexte à partir d'un corpus monolingue en langue cible, où chaque phrase source est une TA des phrases en LC (« *back-translation* ») (Currey et al., 2017 : 148 - 154). Ensuite, ils mélangent ce bitexte avec le corpus parallèle qui est déjà utilisé pour entraîner le système (Currey et al., 2017 : 148 - 154). Dans le but de spécialiser un système de TA, Chinea-Rios et al. (2017) génèrent aussi des « *synthetic parallel data* ». En particulier, ils extraient à partir d'un corpus monolingue contenant des phrases dans la langue source celles similaires au domaine spécialisé (Chinea-Rios et al., 2017 : 138 - 147). Ensuite, ils les font traduire automatiquement et les intègrent aux données d'entraînement du système (Chinea-Rios et al., 2017 : 138 - 147).

Dans le cadre du recours à des corpus parallèles « *in-domain* » et « *out-of-domain* » permettant d'entraîner un système portant sur différents domaines, Chu et Wang (2018 : 1309 - 1310) suggèrent d'utiliser les méthodes « *multi-domain* » et « *data selection* ». En appliquant la première méthode (« *multi-domain* »), on prépare le système neuronal à atteindre deux objectifs : générer les phrases rattachées au domaine spécialisé en ajoutant des balises dans la langue source de chaque corpus qui permettent de classifier les phrases selon leur domaine d'appartenance ; accorder aux différents corpus une attention égale indépendamment de leur taille en utilisant des suréchantillons (Chu et Wang, 2018 : 1310). Dans la deuxième méthode (« *data selection* »), plusieurs travaux ont déjà prouvé que la performance des systèmes neuronaux est extrêmement sensible à la taille des données (Chen et al., 2017 : 41 - 45). Dans cette optique, les chercheurs et les chercheuses explorent les différentes possibilités pour choisir de données d'entraînement visant à spécialiser le système par rapport à un domaine et proposent différentes techniques. À titre d'exemple, Van der Wees et al. (2017 : 1408) mettent au point la technique « *gradual fine-tuning* » qui, par rapport à la sélection conventionnelle des données, prouve une amélioration du BLEU score en réduisant graduellement la taille des données utilisées pour l'entraînement du système.

3.2.3.1.2 Méthodes « *model centric* »

En ce qui concerne les méthodes de la catégorie « *model centric* », celles-ci permettent de spécialiser le système automatique en fonction du domaine et de la terminologie, tout en apportant des modifications à ses paramètres d'entraînement, à l'algorithme de son décodeur ou à son architecture (Chu et Wang, 2018 : 1307).

Quant à l'entraînement du système, nous nous limiterons à approfondir la technique conventionnelle la plus populaire, c'est-à-dire « *fine tuning* », en rappelant l'étude de Luong et Manning (2015). Cette étude montre dans quelle mesure l'entraînement d'un système avec des « *out-domain training data* » et l'ajout progressif de « *in-domain training data* » améliore l'adaptation du système à un domaine spécialisé (Chu et Wang, 2018 : 1313 et Moslem, 2020 : 9 - 10). D'une façon similaire, Farajian et al. (2018 : 156) spécialisent le système pour l'adapter à une terminologie et un domaine donnés (« *fine tuning* ») mais, au lieu d'utiliser toutes les « *in-domain training data* », extraient seulement des phrases (« *instance-based adaptation* »).

Parmi les méthodes qui permettent de modifier l'algorithme du décodeur (« *guided decoding* »), nous n'évoquerons que la technique « *ensembling* » (appelée aussi « *ensemble decoding* »). Cette technique tire son origine du travail de Freitag et Al-Onaizan (2016) reposant sur l'usage simultané de deux modèles différents (« *base model* » et « *fine-tuned model* », correspondant respectivement au « *out-of-domain model* » et au « *in-domain model* ») (Hasler et al., 2018 : 506 et Moslem, 2020 : 11). Dans le but d'adapter le système à la terminologie d'un domaine spécialisé, Freitag et Al-Onaizan (2016 : 10 - 11) entraînent premièrement le système neuronal avec les « *in-domain data* », mais ils initialisent tous les paramètres du « *fine-tuned model* » avec ceux du « *base model* » développé avec les « *out-of-domain data* ». Ensuite, ils continuent l'entraînement du système en ajoutant au modèle initial le « *base model* » (Freitag et Al-Onaizan, 2016 : 11). Dans une optique différente, Dinu et al. (2019) modifient l'algorithme du décodeur pour enseigner au système à comment utiliser les entrées terminologiques, lorsqu'elles sont fournies sous forme d'input supplémentaire à la phrase source. En particulier, ils extraient des bases terminologiques existantes et en faisant référence à leur contenu, ils marquent (« *term annotation* ») les bi-termes dans la langue source (Dinu et al., 2019 : 3064 - 3054).

Dans le cadre des méthodes modifiant l'architecture du système neuronal, nous ne mentionnerons que la méthode « *domain control* » en faisant référence à l'étude de Kobus et

al. (2017). Cette méthode permet d'utiliser un seul réseau neuronal couvrant plusieurs domaines parmi des informations (p.ex. l'indication du domaine d'appartenance) stockées sous forme de balises au niveau des mots (Chu et Wang, 2018 : 1312).

3.2.3.2 Spécialisation d'un système neuronal déjà entraîné

Dans le cadre de l'adaptation d'un système déjà entraîné à une certaine terminologie durant l'étape de la traduction, nous avons identifié trois possibilités.

- Vérification de la terminologie à l'aide des bases terminologiques éventuellement disponibles dans l'environnement d'un logiciel de traduction assistée par ordinateur (TAO).
- Utilisation d'un traducteur en ligne (p.ex. Microsoft Custom Translator et Amazon Translate), qui permet d'importer la mémoire de traduction ou un glossaire sous forme de segments alignés contenant la terminologie du domaine.
- Recours à certains traducteurs en ligne (DeepL Pro, Google Translate Advanced, Customer Translate), qui depuis 2020¹⁷ offrent la possibilité aux usagers d'intégrer leur propre terminologie sous forme de base de données terminologiques, afin que le logiciel en tienne compte dans la traduction.

Dans le cadre de notre mémoire, nous ne nous focaliserons que sur la dernière possibilité. Dans ce contexte, nous ne constatons à l'heure actuelle que l'existence d'un seul travail documenté, à savoir celui de Guerrero (2020), ainsi qu'une expérience partagée seulement sous forme de post dans le milieu du réseau professionnel LinkedIn¹⁸ (Figure 18).

¹⁷ D'après son site internet, DeepL Pro a été le premier traducteur en ligne à proposer la fonctionnalité « glossaire », lancée en mai 2020 (DeepL, 2020).

¹⁸ Dans le but de le décrire dans le présent mémoire, nous avons demandé à l'auteur du post de consulter le travail en question. Nous n'avons pas encore reçu une réponse (www.linkedin.com).

Interesting results testing the [DeepL](#) custom vocabulary feature on medical device translations. We observed a reduction of the terminology error rate from 56% to 2% for the language pair EN-DE. This customization approach appears much more promising to improve terminology consistency compared to re-training machine translation models with hand-crafted or automatically generated bilingual training data. What do you think? Any experiences with other language combinations?

...
[DeepL](#) Now we would only need this feature accessible via API! Is it too late to add this to our Christmas wishlist? [#machinetranslation](#) [#machinelearning](#) [#artificialintelligence](#) [#aihealthcare](#)

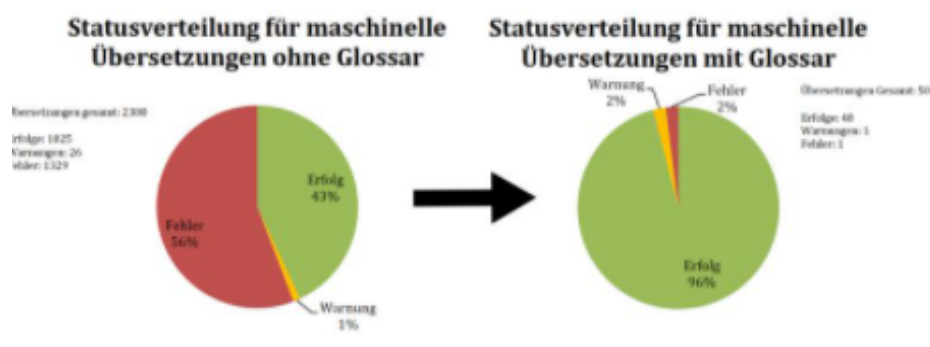


Figure 18 Expérience avec le « glossaire » de DeepL Pro (LinkedIn, 2020).

L'étude de Guerrero (2020) vise à comparer la qualité de la TA fournie respectivement par Google Translate Basic et Google Translate Advanced dans le contexte du processus traductif de CPSL – Language Services & Digital Solutions, fournisseur de solutions multilingues (traduction, interprétation, localisation et multimédia) (CPSL, 2020). Compte tenu de la possibilité de définir la façon dans laquelle traduire certains termes mise à disposition par Google Translate Advanced, qui permet d'intégrer un maximum de 1000 glossaires distincts par projet pour un total de 10,4 millions (10 485 760) d'octets UTF-8 par glossaire, Guerrero (2020 : 8) ajoute un glossaire portant sur le domaine médical avant de lancer la traduction automatique avec Google Translate Advanced (Google Cloud, 2020). Ensuite, elle compare la traduction automatique de Google Translate Basic (sans glossaire) avec celle de Google Translate Advanced (avec glossaire), en effectuant une évaluation humaine (« jugement intuitif » et « analyse des erreurs » ; section 2.4.1) et une évaluation automatique (« Edit distance » ; section 2.4.2) (Guerrero, 2020 : 8). À la lumière des scores obtenus, Guerrero (2020 : 18) ne remarque que des améliorations mineures, puisque Google Translate Advanced fait des erreurs terminologiques et syntaxiques (p.ex. terminologie n'est pas prise en compte lorsque les termes figurent au pluriel dans le texte source, alors que dans le glossaire ils sont

enregistrés au singulier ; acronymes bien traduits, mais ordre des mots différents et incorrects, etc.), bien qu'il utilise un glossaire comme aide supplémentaire à la traduction. À l'avenir, il sera intéressant de voir si d'autres études seront menées dans ce sens.

3.3 Évaluation de la terminologie en TA

Bien que la préservation des connaissances d'un domaine spécialisé constitue une priorité dans le processus traductif, et bien que le besoin principal des clients de l'industrie des langues et de la traduction et que les systèmes de TA soient régulièrement mis à jour pour des raisons telles que la disponibilité de nouvelles données pour l'entraînement du système ou la mise au point des dernières technologies pour améliorer les performances des outils, l'évaluation de la terminologie d'une traduction automatique représente un domaine peu traité en TA (Haque et al., 2019 : 1 - 2). Généralement, cette tâche qui implique un investissement qui n'est pas anodin sur le plan du temps et des ressources est confiée aux spécialistes du domaine (Haque et al., 2019 : 1). En l'espèce, ils comparent la terminologie d'une traduction avec celle utilisée dans le domaine spécialisé (Haque et al., 2019 : 1). Il s'avère dès lors nécessaire de trouver une solution moins onéreuse que l'évaluation humaine. D'après nos recherches, nous constatons qu'actuellement seulement deux métriques automatiques, dont seulement une a été appliquée à quelques études (p.ex. Term Hit Rate (THT) dans le cadre de Scansasi et al., 2019), ont été mises au point pour estimer la qualité de la terminologie d'une traduction automatique. Il s'agit du Term Hit Rate (THR) et de TermEval, que nous approfondirons ci-après (Farajian et al., 2018 et Haque et al., 2019). Toutefois, il est fondamental de signaler que certaines méthodes d'évaluation automatiques, telles que BLEU, METEOR, précision et rappel et TER pourraient être employées pour mesurer, entre autre, la qualité de la terminologie issue d'un système de traduction automatique (section 2.4.2).

3.3.1 Term Hit Rate (THR)

Farajian et al. (2018 : 154) proposent d'utiliser le « *Terme Hit Rate (THR)* » pour mesurer la capacité d'un système de TA à apprendre la terminologie d'un domaine spécialisé. En partant du principe qu'il ne faut pas tenir compte de tous les mots qui apparaissent dans un texte, mais seulement des « termes » y contenus, ce score est obtenu en divisant le nombre total des occurrences de termes correctement traduits par le système de TA par le nombre total des

occurrences de termes figurant dans la traduction de référence (Figure 19) (Farajian et al., 2018 : 154). Il s'agit alors d'un score de « rappel » au niveau des termes. Afin d'éviter d'attribuer des scores plus élevés aux systèmes automatiques générant des traductions où la fréquence d'un certain terme est supérieure à celle du même terme figurant dans la traduction de référence (surestimation d'un système de TA), lorsqu'on calcule le Term Hit Rate (THR), il faut compter les occurrences des termes de la traduction automatique (« $count_{clip}(term)$ ») en fonction du nombre de fois maximum qu'ils apparaissent dans la traduction de référence (« $count_{ref}(term)$ ») (Farajian et al., 2018 : 154). Étant donné par exemple le set de termes « X, Y et Z », figurant dans la traduction automatique respectivement cinq, quatre et sept fois et dans la traduction de référence respectivement trois, cinq et sept fois, on ne comptera le terme « X » que trois fois pour calculer le Term Hit Rate (THR). Dans le cadre de cet exemple, ce score sera égal à 0,93 (nombre total des occurrences de termes correctement traduits par le système de TA soit $5 \underline{3} + 4 + 7 = 14$ divisé par le nombre total des occurrences des termes figurant dans la traduction de référence soit $\underline{3} + 5 + 7 = 15$; score = $14 / 15 = 0,93$), donc la terminologie sera correcte dans le 93% des cas.

$$THR = \frac{\sum_{term \in ref} count_{clip}(term)}{\sum_{term \in ref} count_{ref}(term)}$$

Figure 19 Formule du Term Hit Rate (THR) (Farajian et al., 2018 : 154).

Selon Haque et al. (2019 : 4), le Term Hit Rate (THR) présente deux limites. Ils notent que cette métrique pourrait inclure un terme de la traduction de référence qui n'est pas correctement traduit dans la traduction automatique, mais qui figure erronément dans une autre partie de la TA (Haque et al., 2019 : 4). Étant donné par exemple le set de termes sources « X et Y » où chaque terme figure de façon individuelle dans respectivement les phrases deux et cinq d'un texte source composé de sept phrases, il est important de vérifier que leurs équivalents cibles « X_1 et Y_1 » sont situés dans les mêmes phrases correspondantes de la TA et dans la même position des termes sources. Si un des termes cibles, par exemple « Y_1 », figure dans la phrase six de la TA, alors que le terme source correspondant « Y » est situé dans la phrase source cinq, cela causerait des ambiguïtés au niveau du calcul du Term Hit Rate (THR). De plus, selon eux,

le Term Hit Rate (THR) ne permet pas de traiter les termes ayant plus d'un équivalent cible dans un certain domaine, leurs variations lexicales et flexions, ainsi que les cas d'homonymie (Haque et al., 2019 : 4). En conséquence, le calcul de ce score pourrait être susceptible d'être incorrect (Haque et al., 2019 : 4). Toutefois, nous pensons que le premier risque pourrait être largement réduit en mettant en place deux approches : calculer le Term Hit Rate (THR) seulement dans le contexte de phrases qui contiennent de la terminologie et, donc, écarter celles qui n'incluent aucun terme ; effectuer à titre de comparaison une évaluation humaine ou intégrer une autre métrique automatique. En ce qui concerne les variations terminologiques, nous retenons qu'elles pourraient être incluses dans le calcul du score, soit en appliquant la technique de « racinisation »¹⁹ (en anglais « *stemming* »), soit en les ajoutant directement au glossaire faisant l'objet d'une étude. Ces possibilités représentent effectivement le point de départ pour la mise au point de la métrique TermEval (section 3.3.2).

3.3.2 TermEval

Sur la base des limites du Term Hit Rate (THR) qu'ils ont identifié, Haque et al. (2019 : 13) proposent d'utiliser alors le score « *TermEval* », qui tient également compte de la position des termes, ainsi que des variations lexicales et flexionnelles des termes les plus fréquentes (« *lexical and inflectional variations, LIVs* ») (Haque et al., 2019 : 13). Plus précisément, ils identifient la « bonne position des termes » dans les phrases du corpus de test et ils marquent les termes sources et cibles du glossaire à évaluer dans la TA, en utilisant un outil dit « *TermMarker* » qu'ils ont eux-mêmes développé (Haque et al., 2019 : 9). Concrètement, TermMarker affiche chaque paire de phrases du corpus de test et une personne appelée « *annotator* » marque manuellement les termes appartenant au domaine spécialisé et les ajoute avec les relatives variations terminologiques plus fréquentes et significatives dans une base de données terminologiques (Haque et al., 2019 : 9). L'identification des termes peut être supportée aussi par une extraction terminologique automatique qui est directement intégrée dans TermMarker (Haque et al., 2019 : 9 - 10). Une fois les termes marqués, les phrases du corpus qui ne contiennent aucun terme sont écartées du corpus de test et celles contenant de la

¹⁹ Procédure qui permet de transformer les flexions en leur racine en supprimant les suffixes (Boukhari, 2018 : 42).

terminologie sont regroupées dans un « *gold-testset* » qui fera l'objet de l'évaluation exprimée en termes de TermEval (Haque et al., 2019 : 9).

En ce qui concerne sa formule, le score TermEval demande de compter pour chaque phrase du texte source avec indice « n » et pour chaque terme source avec indice « s » (où l'indice « s » dépend de la phrase « n » où le terme source est situé) le nombre d'équivalents cibles (« Hyp_n ») correspondant au terme source (« R_v »), à condition qu'il existe au moins une occurrence d'un terme source (« R_v ») dans la TA sous forme d'équivalent cible (« Hyp_n »). Comme Haque et al. (2019 : 13) le précisent: « *[t]he proposed evaluation metric starts evaluation process with forming a tuple with each source sentence from the test set, its translation (i.e. hypothesis), the list of source terms appearing in the source sentence, their reference translations (i.e. terms), and LIVs of the reference terms* ». En conséquence, s'il est possible de détecter l'équivalent cible d'un terme d'une phrase source dans la phrase correspondante de la TA (« *hypothesis (Hyp)* »), cela veut dire que le système de TA a bien traduit la terminologie (Haque et al., 2019 : 13). Dans cette optique, il faut donc diviser la somme du nombre des équivalents cibles correctement traduits en TA (« Hyp_n ») par le nombre total des paires de termes source et cible contenus dans le glossaire (« NT ») (Figure 20) (Haque et al., 2019 : 13).

$$\text{TermEval} = \frac{\sum_{n=1}^N \sum_{s=1}^S \sum_{v=1}^V \begin{cases} 1 & \text{if } R_v \in \text{Hyp}_n; \text{ break;} \\ 0 & \text{otherwise} \end{cases}}{NT} \quad (1)$$

where: N : number of sentences in the test set
 S : number of source terms in the n^{th} source sentence
 V : number of reference translations (including LIVs) for s^{th} the source term
 R_v : v^{th} reference term for s^{th} the source term
 Hyp_n : translation of n^{th} source input sentence
 NT : total number of terms in the test set

Figure 20 Formule de TermEval (Haque et al., 2019 : 13).

Par exemple, si l'on considère le set de termes sources « X, Y et Z » où chaque terme figure de façon individuelle dans respectivement les phrases deux, trois et cinq d'un texte source composé de sept phrases, il faudra compter leurs équivalents cibles « X_1 , Y_1 et Z_1 » seulement

s'ils sont situés dans les mêmes phrases correspondantes de la TA et dans la même position des termes sources. Si un des termes cibles, par exemple « Z_1 », figure dans la phrase six de la TA, alors que le terme source correspondant « Z » est situé dans la phrase source cinq, il ne faudra pas l'inclure dans le calcul. Dans le cadre de cet exemple, le TermEval sera égal à 0,66 (nombre des équivalents cibles corrects soit 2 divisé par le nombre des paires de termes source et cible contenus dans le glossaire soit 3 ; score $2 / 3 = 0,66$), donc la terminologie de la TA sera correcte dans le 66% des cas.

Pour vérifier la fiabilité du TermEval, Haque et al. (2019 : 14 - 15) effectuent une évaluation humaine dont le but est de classer la terminologie contenue dans les différentes phrases en deux catégories (« *error* » et « *correct* »). En faisant cela, ils tiennent compte uniquement des caractéristiques morphologiques et syntaxiques des termes à analyser sans le besoin de juger toute la TA (Haque et al., 2019 : 14 - 15). Finalement, ils voient que les résultats de TermEval sont assez similaires à ceux de l'évaluation humaine, ce qui veut dire qu'il y a une très bonne corrélation entre l'évaluation humaine et l'automatique (Haque et al., 2019 : 17).

3.4 Conclusion sur la TA et la terminologie

Ce chapitre porte sur traitement de la terminologie en traduction automatique. En premier lieu, nous avons défini la notion de « terminologie » et nous avons éclairci la raison de son importance en TA (section 3.1). Puis nous avons approfondi la façon dans laquelle la terminologie est traitée dans les différents systèmes de TA (section 3.2) en faisant référence à leur catégorisation dans le chapitre précédent (2.3). Enfin, nous avons présenté les métriques automatiques permettant d'évaluer la terminologie en TA (section 3.3). Plus précisément, nous avons approfondi le Term Hit Rate (THR) (section 3.3.1) et TermEval (section 3.3.2). Dans le prochain chapitre (4), nous poserons notre attention sur la méthodologie développée pour répondre à notre question de recherche principale, ainsi qu'à la série de sous-questions qui en découlent.

4 Méthodologie

À la lumière des concepts théoriques préalablement évoqués, dans ce chapitre, nous passerons en revue la démarche qui a été entreprise pour mener à bien le projet de ce mémoire en éclaircissant chaque étape nécessaire à son bon déroulement. D'abord, nous décrirons les objectifs et les questions de recherche (section 4.1), puis nous retracerons les différentes tâches accomplies (section 4.2). Nous résumerons pour finir ce quatrième chapitre (section 4.3).

4.1 Objectifs et questions de recherche

Notre mémoire vise à décrire et évaluer le traitement de la terminologie en TA. Bien que la recherche sur la TA ait fait des pas de géant et les possibilités pour améliorer les systèmes de traduction automatique soient nombreuses, l'emploi d'une terminologie exacte et cohérente reste une problématique loin d'être résolue (chapitre 3). En prenant comme exemple le cas de la Confédération suisse et en nous situant dans le même contexte d'étude de Guerrero (2020) (section 3.2.3.2), nous avons voulu étudier l'impact de la TA sur la traduction des termes. Dans ce contexte, nous chercherons à répondre à la question de recherche principale suivante :

- Le traducteur en ligne DeepL Pro permet-il d'adapter la terminologie de ses traductions à un domaine spécialisé ?

Cette question soulève également trois sous-questions :

- Sous-question 1 : quel est l'impact de la nouvelle fonctionnalité « glossaire » sur la qualité de la traduction ?
- Sous-question 2 : le « glossaire » permet-il de prendre en compte les variations des termes ?
- Sous-question 3 : quels sont les types d'erreurs posés par la TA des termes ?

Pour pouvoir répondre à ces questions de recherche, nous sommes passée par plusieurs étapes que nous présenterons ci-après (section 4.2).

4.2 Démarche

Dans cette section, nous approfondirons l'ensemble des étapes nécessaires au bon déroulement de notre démarche (Figure 21). Afin de mener à bien celle-ci, nous avons eu recours à différents outils que nous approfondirons dans le cadre des étapes où ils ont été utilisés.

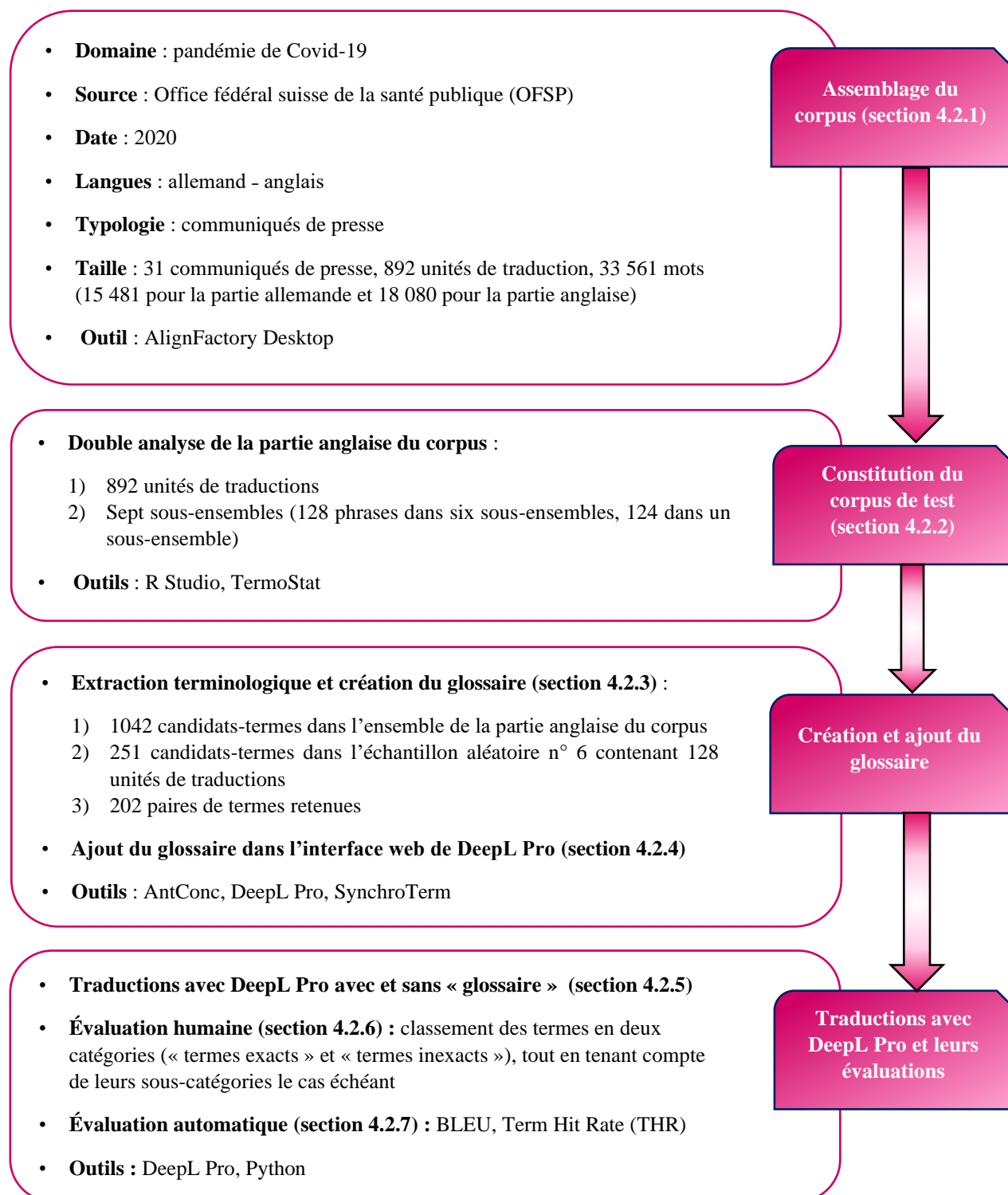


Figure 21 Déroulement de notre démarche méthodologique.

4.2.1 Assemblage du corpus

Pour assembler notre corpus, constituant le point de départ de notre étude, nous avons premièrement dû définir ses critères de sélection, à savoir : domaine, source, date, langues, taille, typologie et degré de spécialisation des textes collectés. Ci-après, nous rentrerons dans le détail de ceux-ci.

Nous avons choisi de nous focaliser sur un domaine apparu récemment qui a connu une véritable explosion dès 2020 : la pandémie de Covid-19. Comme *The Economist* (2020) le dit, depuis janvier 2020, le nombre de publications portant sur ce domaine a doublé tous les 14 jours et déjà pendant la première semaine du mois de mai 2020, il a atteint 1 363 unités (*The Economist*, 2020). Riedeger (2020) parle de l'explosion de contenus liée au Covid-19 comme d'une épidémie d'information qui, comme le virus, se propage dans toutes les langues par les canaux les plus divers, tels que les sites internet institutionnels et privés, les médias sociaux, les journaux en ligne, les vidéos, les blogs, etc. Compte tenu de cette explosion de contenus soudaine, nous avons pensé que la description et l'évaluation de la terminologie issue de ce domaine très récent pouvaient donner des résultats intéressants dans le cadre de notre mémoire.

Afin d'explorer un corpus contenant la terminologie suisse issue de la pandémie de Covid-19, nous avons considéré seulement les textes produits par l'Office fédéral suisse de la santé publique (OFSP). Cet Office, qui est responsable de la santé de la population, est le centre de compétences pour les questions de santé en Suisse (Office fédéral suisse de la santé publique, 2020). Il édite donc régulièrement des publications dans le domaine de la santé et aux vues des évènements récents, portant sur la pandémie de Covid-19 (Office fédéral suisse de la santé publique, 2020).

Compte tenu des évènements historiques²⁰ et de notre choix de vouloir assembler un corpus contenant la terminologie suisse liée à la pandémie de Covid-19, nous avons tenu compte de tous les textes publiés entre le 1^{er} janvier et le 31 octobre 2020. Pour des questions d'accessibilité, nous n'avons sélectionné que le matériel interactif disponible sur le site web de l'Office fédéral suisse de la santé publique qui se présente seulement en format écrit. Pratiquement, nous avons défini un intervalle temporaire (1^{er} janvier - 31 octobre 2020) de

²⁰ Le 8 janvier, l'Organisation mondiale de la Santé (OMS) annonçait l'apparition d'une nouvelle forme de coronavirus en Chine (Le Temps, 2020).

publication dans la section dédiée aux communiqués de presse disponible sur le site internet de l'Office. Nous avons créé alors un corpus synchronique et clos²¹.

Puisque ce corpus devait servir de référence pour créer le glossaire bilingue et évaluer la qualité de la TA des termes, nous avons assemblé un corpus parallèle bilingue où la terminologie était parfaitement équivalente dans les deux langues. Pour une consultation plus immédiate, nous avons aligné les textes collectés et sauvegardés sous formes de fichiers TXT avec l'aide de AlignFactory Desktop²², qui nous a permis d'obtenir des « bitextes ». Pour chacune des deux langues, chaque segment correspondait à une phrase du corpus. La paire de langues de notre corpus était allemand - anglais. Le choix d'une telle combinaison linguistique repose sur deux raisons. La première est qu'actuellement, la fonctionnalité « glossaire » de DeepL Pro n'est disponible que pour les paires de langues suivantes : allemand - anglais, français - anglais, anglais - allemand et anglais - français (DeepL, 2020). La deuxième raison porte sur notre ancienne expérience professionnelle au sein des Services linguistiques de l'Office fédéral de la statistique, où nous avons pu constater que l'allemand est la langue de rédaction la plus fréquente. Dans la perspective des évaluations des traductions automatiques de DeepL Pro, nous sommes partie alors du principe que les communiqués de presse sont rédigés en allemand. En conséquence, la partie allemande du corpus a constitué le texte source, tandis que la partie anglaise correspondante a représenté la traduction de référence.

En ce qui concerne la typologie des textes collectés, il importe de dire que l'Office fédéral suisse de la santé publique fournit les informations sous différentes formes (p.ex. publications, communiqué de presse, etc.). Dans le cadre de notre mémoire, nous n'avons considéré que les communiqués de presse, qui sont généralement destinés à informer les usagers sur un événement précis. Cette décision repose sur deux motivations principales. Nous croyons que le recours à un système de TA peut représenter une aide précieuse pour traduire et diffuser rapidement ce type de contenus, à condition que la qualité des traductions automatiques soit jugée satisfaisante. La deuxième raison est liée à l'usage d'un système automatique généraliste faisant l'objet de notre évaluation, qui n'est pas entraîné avec notre corpus.

²¹ Un corpus « synchronique » donne « une photo » de l'usage d'une langue pendant un temps limité (Bowker et Pearson, 2002 : 11 - 13). Un corpus dit « clos » veut dire que aucun texte sera ajouté avant la fin de cette recherche (Bowker et Pearson, 2002 : 11 - 13).

²² AlignFactory Desktop est une application qui permet d'aligner des documents afin d'alimenter un outil d'aide à la traduction, un moteur de traduction automatique neuronale ou un concordancier bilingue (Terminotix, 2020a). Cela peut être fait soit en se servant de marqueurs de langues présents dans le nom des fichiers et définis par l'utilisateur ou soit en utilisant des algorithmes de détection de la langue et d'empreinte électronique sur le type et le contenu des fichiers (Terminotix, 2020a).

À la lumière de nos critères de sélection, notre corpus parallèle bilingue (allemand - anglais) portant sur le domaine de la pandémie de Covid-19 s'est finalement composé de 31 communiqués de presse, dont l'auteur était l'Office fédéral suisse de la santé publique. Dans l'ensemble, le corpus a compté 892 unités de traduction et 33 561 mots (15 481 pour la partie allemande et 18 080 pour la partie anglaise).

Dans le cadre des évaluations que nous avons effectuées pour mesurer la performance de DeepL Pro en matière de terminologie, nous n'avons retenu qu'un échantillon de phrases (section 9.8) et de termes (section 9.7) à évaluer. Ci-après, nous regarderons en détail le processus de leur sélection.

4.2.2 Constitution du corpus de test

Pour des raisons de volume de données à traiter, nous avons choisi de nous baser sur un échantillon représentatif de notre corpus en vue de l'évaluation du traitement de la terminologie dans DeepL Pro. Pour le choisir, nous sommes passés par trois étapes, où nous nous sommes intéressés principalement à la partie anglaise de notre corpus (traduction de référence).

En premier lieu, nous avons cherché à savoir combien de candidats-termes étaient contenus dans notre corpus cible et quelle était leur composition. Pour obtenir cette information, nous avons employé TermoStat²³ (2020) qui, à l'aide d'un logiciel d'étiquetage externe (TreeTagger) et d'expressions régulières, a désambiguïsé les mots pouvant recevoir plus d'une catégorie syntaxique et les a extraits selon de différentes catégories syntaxiques prédéfinies (annexe 9.1). En total, TermoStat a détecté 1042 candidats-termes.

Dans un second temps, nous avons décidé d'extraire aléatoirement des phrases et de les regrouper de façon arbitraire en sept parties. Cela nous a permis d'obtenir des sous-ensembles du corpus où nous avons pu vérifier si les candidats-termes se distribuaient selon leur catégorie syntaxique de manière similaire à l'ensemble du corpus. Compte tenu de l'importance du concept de « *statistical significance* »²⁴ et, en conséquence, de la nécessité d'obtenir un échantillon représentatif dans le contexte de notre étude, où les phrases sont choisies à partir

²³ TermoStat est un outil d'acquisition automatique de termes qui exploite une méthode de mise en opposition de corpus spécialisés et non-spécialisés afin d'identifier les termes (TermoStat, 2020).

²⁴ Koehn (2004 : 391 - 395) souligne l'importance d'effectuer une évaluation sur la base d'un échantillon représentatif (« *representative sample* »), afin que les résultats obtenus reflètent réellement la qualité globale de la traduction.

de différentes parties du corpus, pour éviter d’avoir des résultats biaisés, nous avons accompli cette tâche automatiquement en utilisant l’environnement de développement de R Studio²⁵ (Koehn, 2004 : 391). Le code élaboré²⁶ nous a premièrement permis de diviser le corpus en sept sous-ensembles, en calculant le nombre de phrases nécessaires à les constituer (six sous-ensembles contenant 128 phrases et un contenant 124). Dans un second temps, une variable vectorielle contenant indices compris entre 1 et 892 (total des phrases contenues dans notre corpus) a été créée à l’intérieur d’une boucle « while » pour choisir aléatoirement des phrases à inclure dans les différents sous-ensembles. Lorsque la variable vectorielle n’a plus détecté de phrases à inclure dans l’un des sous-ensembles, la boucle « while » s’est terminée et les sept sous-ensembles du corpus ont été automatiquement sauvegardés dans différents fichiers TXT.

Une fois les 7 fichiers obtenus, nous avons pu voir en utilisant TermoStat si les candidats-termes se distribuaient selon leur catégorie syntaxique de manière similaire dans les sept sous-ensembles du corpus (annexe 9.2) et si leur composition reflétait celle du corpus entier (annexe 9.1). En particulier, nous avons vérifié qu’une condition soit satisfaite : au moins toutes les catégories syntaxiques des candidats-termes extraits de l’ensemble de la partie anglaise du corpus, dont la fréquence était égale ou supérieure au 1%, devaient être contenues dans les différents sous-ensembles indépendamment de leurs valeurs de fréquence dans ceux-ci. Compte tenu des valeurs généralement tolérées en statistique²⁷, nous nous sommes basée sur un seuil de précision de 99%. Pour visualiser et comparer les données des sept sous-ensembles générées par TermoStat (annexe 9.2), nous avons utilisé deux représentations différentes : un graphique à colonnes (annexe 9.4), et une boîte à moustache (annexe 9.5), aussi connue sous le nom anglais de « box plot ». Pour une meilleure compréhension de ceux-ci, où les catégories syntaxiques sont indiquées sous forme d’acronymes en faisant référence aux premières lettres de leurs noms, nous recommanderons d’utiliser la légende des matrices syntaxiques (annexe 9.3). Comme les résultats obtenus²⁸ nous ont suggéré qu’au moins toutes les catégories syntaxiques des candidats-termes extraits de l’ensemble de la partie anglaise du corpus, dont la fréquence est égale ou supérieure au 1%, étaient contenues dans 99% des cas dans les sept

²⁵ R Studio est un logiciel proposant un environnement de développement et des outils adaptés au langage de programmation de R, qui est destiné aux statistiques et à la science des données (01net, 2020 et Wikipedia, 2020).

²⁶ Pour programmer en R la façon de sélectionner aléatoirement des phrases, nous avons fait appel à l’expertise du Dr. De Oliveira Vilaça (<https://genev.unige.ch/research/people/Luis-Miguel-De-Oliveira-Vilaca>).

²⁷ En statistique, le seuil de précision généralement toléré et plus répandu est égal au 95% (Koehn, 2004 : 391). Le niveau de 99 % est le plus prudent et le niveau de 90 % est rarement utilisé (Wikipedia, 2020).

²⁸ Notre condition sera intégralement satisfaite dans trois cas (échantillon n° 1, 4 et 6) sur sept. Dans quatre cas (échantillon n° 2, 3, 5 et 7), le seuil de précision sera égal au 99%.

sous-ensembles du corpus, nous avons donc pu choisir un échantillon aléatoire (n° 6) qui, dans le cadre de nos évaluations avec DeepL Pro, a constitué la traduction de référence.

Afin d'extraire les candidats-termes dans les deux langues pour pouvoir créer notre glossaire bilingue ainsi que le texte source à faire traduire par DeepL Pro, nous avons repéré à partir de l'alignement préalable de notre corpus les segments correspondants en langue allemande et nous les avons sauvegardés dans un nouveau fichier TXT. Finalement, sur la totalité des unités de traduction de notre corpus (892), nous en avons choisi alors 128 (14,28%), servant de base à la sélection de nos termes.

4.2.3 Création du glossaire

Pour adapter la terminologie d'une TA à un domaine spécialisé, il est nécessaire d'intégrer au système de TA une collection terminologique déjà existante, également dite « glossaire », ou de la créer si elle n'est pas disponible. Car aucun glossaire n'existait dans notre cas, nous avons dû l'élaborer. Compte tenu du but et du contexte de notre étude, visant à regarder l'impact d'un glossaire sur la TA, nous avons mis en place une approche terminographique se déroulant en deux étapes.

- En premier lieu, nous avons effectué une extraction terminologique à l'aide de SynchroTerm²⁹ qui, à l'aide d'algorithmes statistiques, syntaxiques et morphologiques, nous a permis de retrouver automatiquement les candidats-termes de notre échantillon d'unités de traductions ainsi que leurs contextes³⁰ (Figure 22) (Terminotix, 2020b).
- Puis nous avons épuré la liste des candidats-termes en utilisant AntConc³¹, afin de collecter les informations qui nous ont servies à l'analyse et à la synthèse des données terminologiques (L'Homme, 2020 : 54). Les fonctionnalités de cet outil, en particulier

²⁹ SynchroTerm est un extracteur terminologique permettant d'accélérer l'extraction et la création de fiches terminologiques qui prend en charge 30 langues différentes (Terminotix, 2020b).

³⁰ Pour ce faire, nous avons dû définir les paramètres de l'outil. Dans la section « Options d'extraction » de SynchroTerm, nous saisissons notre combinaison linguistique (allemand - anglais), le nombre minimum (2) et maximum (8, pour ne pas inclure seulement les termes simples et complexes, mais aussi des autres syntagmes) des mots à extraire, ainsi que le nombre minimum d'occurrences (3).

³¹ Antconc et un concordancier gratuit qui traite de l'identification automatique de la structure de textes et qui peut être appliquée à la recherche en générale (EduTech Wiki, 2020a).

« Concordance »³² et « Clusters »³³, nous ont permis de retrouver les occurrences et les cooccurrents des candidats-termes, ainsi que leurs variations terminologiques (Figure 23). Nous avons également eu recours aux expressions régulières. À titre d'exemple, nous avons utilisé les opérateurs de quantification (« ? » pour repérer 0 ou 1 occurrence du caractère précédent) et de disjonction (« [Aa] » pour inclure les majuscules et/ou minuscules), ainsi que la troncature pour retrouver les formes variables d'un terme (« ease, easing, approved, approve») et les frontières pour délimiter notre suggestion (p.ex. \b pour limiter la recherche à la chaîne de caractères indiqués) (Baldassarre et Borghi, 019 : 3).. Nous avons employé les expressions régulières individuellement ou en combinaison (p.ex. « \b[Lt]ockern?\b » ou « \b[Mm]asures?\b ») (Baldassarre et Borghi, 2019 : 3).

The screenshot displays the SynchroTerm 2019 interface. It features two main tables: 'Source Extraction' and 'Target Extraction'. The 'Source Extraction' table lists various expressions and their occurrences, with a 'PNG' column. The 'Target Extraction' table lists corresponding terms and their percentages and occurrences. Below these tables, there are sections for 'Source Selection' and 'Target Selection', each showing a list of translation units with checkboxes and context snippets.

Annotations in red boxes point to specific elements:

- A box labeled 'Liste des possibles correspondances dans la langue cible' points to the 'Target Extraction' table.
- A box labeled 'Dans la colonne PNG « 1 » veut dire qu'il y a la probabilité d'avoir un groupe nominal' points to the 'PNG' column in the 'Source Extraction' table.
- A box labeled 'Pourcentage de spécificité' points to the '%' column in the 'Target Extraction' table.
- A box labeled 'Contexte' points to the translation unit snippets at the bottom of the interface.

Figure 22 Résultats de l'extraction terminologique dans SynchroTerm

³² « Concordance » nous a permis d'explorer le contexte d'usage des termes.

³³ « Clusters » nous a servi à repérer les voisins directs des termes analysés.

Rank	Freq	Range	Cluster
1	17	1	the covid
2	7	1	national covid
3	6	1	against covid
4	5	1	of covid
5	3	1	for covid
6	3	1	situation covid
7	2	1	a covid
8	2	1	in covid
9	1	1	combat covid
10	1	1	combating covid
11	1	1	coronavirus (covid)
12	1	1	coronavirus covid
13	1	1	current covid

Figure 23 « Concordance » (à gauche) et « Clusters » (à droite) dans AntConc.

Pour choisir les termes à inclure dans notre glossaire, nous avons tenu compte de plusieurs aspects, tels que :

- La spécificité, servant à identifier les candidats-termes à retenir et à laisser de côté ceux qui ne sont pas spécifiques à notre domaine (L’Homme, 2020 : 187). Dans cette perspective, nous avons inclus dans notre glossaire plusieurs termes, tels que « coronavirus, quarantaine, epidemic », alors que nous n’avons pas retenu « also, bar, people ».
- La fréquence, en partant du principe qu’une forme linguistique figurant un certain nombre de fois dans des textes représentatifs est fort probablement un terme (L’Homme, 2020 : 64). Pour considérer la fréquence « significative », nous avons considéré à la fois le nombre d’occurrences des candidats-termes et leur répartition (L’Homme, 2020 : 64). En effet, un terme ne doit pas apparaître dans quelques phrases, mais dans plusieurs (L’Homme, 2020 : 64). Dans notre étude, nous avons constaté une corrélation entre la spécificité et la fréquence. En effet, les candidats-termes plus fréquents avaient un score de spécificité très élevé, tels que « coronavirus, mask, rapid test ».
- La prédominance de candidats-termes de nature nominale qui, en terminologie représentent la « partie du discours privilégiée » (L’Homme, 2020 : 187). Dans notre glossaire, ceux-ci ont représenté la plupart des termes (146 sur 202).
- La complexité (« Federal Council, task force, patient safety, isolation unit ») et le nombre de séquences pouvant correspondre à un terme complexe, dont par exemple la

séquence « adjectif + nom » (« new coronavirus, close contact, rapid test, confirmed case ») ou celle composée par un syntagme prépositionnel (« number of cases, level of contact ») (L’Homme, 2020 : 187).

Finalement, sur un total de 251 candidats-termes, nous n’en avons retenu que 202 pour les enregistrer dans le « glossaire » de DeepL Pro (Tableau 1).

Termes contenus dans le « glossaire » de DeepL Pro			
Catégorie	Exemples	Nombre	%
Termes complexes	infected person, hygienic mask, high-risk country	116	57%
Termes simples	coronavirus, mask, SN-STF, FOPH	31	15%
Collocations	relax the measures, measures to combat the coronavirus	38	19%
Verbes	test, ease, control, apply	17	9%
Total		202	100

Tableau 1 Termes contenus dans le « glossaire » de DeepL Pro.

Parmi ceux-ci, nous avons inclus des synonymes et des variations terminologiques (Tableau 2).

Synonymes et variations terminologiques contenus dans le « glossaire » de DeepL Pro			
Catégorie	Exemples	Nombre	%
Variations terminologiques de type graphique ³⁴ , flexionnel ³⁵ et morphosyntaxique ³⁶	Covid-19 vaccine, COVID-19 vaccine, precautionary measure, precautionary measures, easing the measures, ease measures	33	77%
Synonymes	Swiss Confederation, Swiss federal government	10	23%
Total		43	100%

Tableau 2 Synonymes et variations terminologiques contenus dans le « glossaire » de DeepL Pro.

En vue de l’ajout du glossaire dans DeepL Pro et de nos évaluations, nous avons listé nos termes dans un fichier TXT (section 9.7). Conformément à la pratique terminographique, nous les avons enregistrés dans leur forme canonique.

³⁴ Celles-ci se résument à des distinctions graphiques minimales, telles que l’ajout d’un signe diacritique ou d’un trait d’union (L’Homme, 2020 : 83).

³⁵ Elles regroupent les formes différentes fléchies d’un terme (p.ex. singulier et pluriel « set of precautionary measures, sets of precautionary measure »), comme aussi les différentes formes verbales (« test, testing, tested ») (L’Homme, 2020 : 83).

³⁶ Les variations morphosyntaxiques font alterner des parties du discours et entraînent des transformations dans les phrases (L’Homme, 2020 : 83).

4.2.4 Ajout du glossaire dans DeepL Pro

La nouvelle fonctionnalité « glossaire » de DeepL Pro est actuellement disponible seulement dans son interface web (DeepL, 2020). Concrètement, nous avons dû cliquer sur « Ajouter un nouveau glossaire » (Figure 24), sélectionner la paire de langues de notre choix et ajouter manuellement notre liste de termes préalablement enregistrés dans un fichier TXT (section 9.7). En effet, ce traducteur en ligne ne permet pas d'importer une collection terminologique déjà existante et cela constitue un investissement de temps et une limite du traducteur en ligne qui ne sont pas anodins. De plus, nous avons constaté l'impossibilité d'ajouter plusieurs équivalents cibles pour un seul terme source (p.ex. les deux synonymes « Swiss Confederation » et « Swiss federal government » pour se référer à « Bund » et les deux variations terminologiques « rapid test » et « rapid testing » pour traduire « Schnelltest ») et d'inclure dans le glossaire des termes contenant de symboles spéciaux, tels que les parenthèses et les virgules (p.ex. « Krisenstab des Bundesrates zur Bewältigung der Corona Krise (KSBC) »). À cause de cela, DeepL Pro présente deux autres limites : il ne permet pas d'employer des synonymes et des variations terminologiques pour éviter des répétitions ou pour préciser le sens d'un terme (section 5.1.2.1.2) ; il peut mal traduire des noms propres ou des acronymes contenant de symboles spéciaux (5.1.2.2.2).

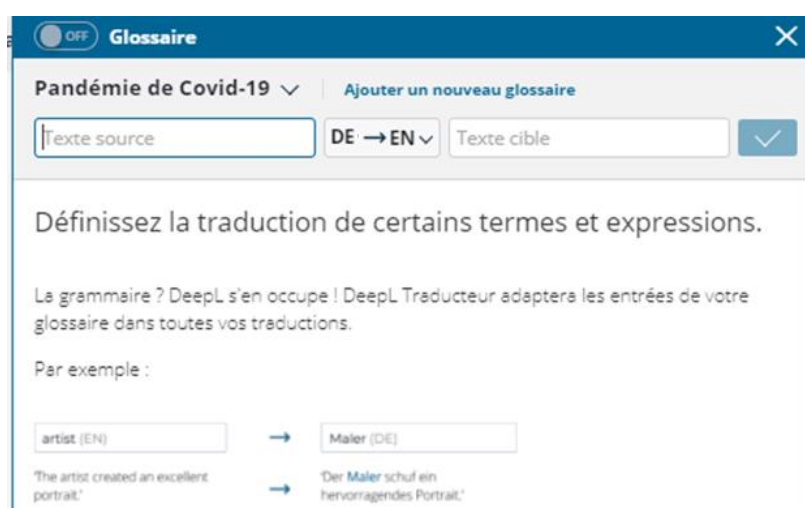


Figure 24 Création de notre glossaire dans DeepL Pro.

4.2.5 Traductions avec DeepL Pro

Une fois les 202 termes ajoutés au « glossaire » de DeepL Pro, nous lui avons fait traduire en anglais³⁷ l'ensemble de notre échantillon de segments en allemand avec et sans « glossaire ». Car la nouvelle fonctionnalité « glossaire » est actuellement disponible seulement en ligne, la traduction a été faite dans l'interface web de DeepL Pro. Il faut noter que DeepL Pro a mis en évidence les termes repris de notre liste lorsque notre terminologie a été intégrée (Figure 25).



Figure 25 Extrait d'une traduction de DeepL Pro avec son « glossaire » (DeepL, 2020).

Afin de marquer les termes à évaluer dans le texte source, dans la traduction de référence ainsi que dans les traductions automatiques de DeepL Pro, nous avons créé un tableau réunissant tous les segments ordonnés numériquement et les 202 termes de notre glossaire, signalés en gras. Ce tableau (section 2.7) se compose de cinq colonnes (Tableau 3), contenant chacune le numéro de la phrase à évaluer et les segments avec respectivement : le texte source (« Src »), la traduction de référence (« Réf ») et les traductions automatiques de DeepL Pro réalisées sans la fonctionnalité « glossaire » (« Syst sans glos ») ou avec (« Syst avec glos »).

Phrase n°	Src	Réf	Syst sans glos	Syst avec glos
1)	Ab dem 11. Mai können nicht nur Läden, Märkte und obligatorische Schulen , sondern auch Museen und Bibliotheken sowie Gastrobetriebe wieder öffnen.	From 11 May, not only will shops, markets and primary and lower secondary schools be allowed to reopen, but also museums, libraries and restaurants.	From 11 May, not only shops, markets and compulsory schools , but also museums, libraries and catering establishments will be able to reopen.	From 11 May, not only shops, markets and primary and lower secondary schools will be able to reopen, but also museums, libraries and catering establishments.

Tableau 3 Structure du tableau réunissant les phrases à évaluer.

³⁷ DeepL Pro permet de saisir deux variétés d'anglais : l'anglais britannique, qui constituera l'objet de notre choix, et l'anglais américain.

4.2.6 Évaluation humaine

En ce qui concerne l'évaluation humaine des traductions de DeepL Pro réalisées avec (section 5.1.2) et sans (section 5.1.1) la fonctionnalité « glossaire », nous avons vérifié s'il existait une correspondance terminologique entre la traduction de référence et les TA. Pour ce faire, nous nous sommes inspirée de l'évaluation humaine effectuée par Haque et al. (2019) visant à classer la TA des termes en deux catégories, à savoir « *error terms* » et « *incorrect terms* » (section 3.3.2). Nous avons donc classé les 202 termes de notre glossaire chaque fois que nous les avons rencontrés (260 occurrences) en deux catégories (« termes exacts » et « termes inexacts »), tout en tenant compte de leurs sous-catégories le cas échéant (Tableau 4 et Tableau 5).

Sous-catégories des termes cibles exacts par rapport à la traduction de référence dans DeepL Pro avec « glossaire »
Termes intégrés parfaitement dans la TA
Termes qui provoquent des répétitions
Erreurs de syntaxe et/ou d'orthographe dans la phrase

Tableau 4 Sous-catégories des termes cibles exacts par rapport à la traduction de référence.

Sous-catégories des termes cibles inexacts par rapport à la traduction de référence	
DeepL Pro sans « glossaire »	DeepL Pro avec « glossaire »
Recours à d'autres formes linguistiques ³⁸	Variations terminologiques qui ne sont pas détectées dans la traduction de référence
Emploi en LC du langage commun au détriment de la communication spécialisée	/
Erreurs de sens	/
Termes cibles qui ne sont pas conformes à la réalité suisse	/
Noms propres et acronymes en LC qui sont mal traduits	Noms propres et acronymes en LC qui sont mal traduits

Tableau 5 Sous-catégories des termes cibles inexacts par rapport à la traduction de référence.

4.2.7 Évaluation automatique

En ce qui concerne l'évaluation automatique des traductions de DeepL Pro réalisées avec et sans « glossaire », nous avons choisi de calculer le score BLEU (section 2.4.2) et le Term Hit Rate (THR) (section 3.3.1). Le BLEU permet d'estimer la qualité d'une TA en comparant le nombre de ses chaînes de mots et de n-grammes, qui se révèlent identiques à la traduction de

³⁸ Les autres formes linguistiques incluent les synonymes et les variations terminologiques.

référence (Hearne et Way, 2011 : 219). En revanche, le Term Hit Rate (THR) calcule le nombre total des occurrences de termes correctement traduits par le système de traduction automatique par rapport au nombre total des occurrences de termes figurant dans la traduction de référence (Farajian et al., 2018 : 154). En utilisant ces métriques, nous avons donc estimé la qualité globale de la TA et celle des termes utilisés.

Pour le calcul du score BLEU, nous avons utilisé la plateforme Tilde MT³⁹, où nous avons ajouté les fichiers de notre texte source, de la traduction de référence et des outputs de DeepL Pro. Car nous avons voulu voir l'impact de la différence de casse (p.ex. COVID-19, Covid-19, federal government, Federal Government) sur le BLEU, nous avons calculé ce score sur tout notre corpus de test deux fois : en sélectionnant et en désélectionnant l'option « Lowercase ».

Pour obtenir le Term Hit Rate (THR), nous sommes passée par quatre étapes où nous n'avons considéré que la traduction de référence et les TA.

Premièrement, nous avons ajouté à la partie anglaise de notre glossaire les différents types de variations terminologiques des termes, afin de les inclure dans le calcul du score. Nous avons par exemple intégré « COVID-19 vaccine, COVID-19 vaccines, covid-19 vaccine, covid-19 vaccines, etc. » ainsi que « working from home, work from home, etc. ».

Dans un second temps, nous avons créé des scripts dans Python 3⁴⁰ qui, en tenant compte de la traduction de référence et des TA, ainsi que de notre nouvelle liste de terme ajoutées sous forme de fichier TXT, nous ont permis de compter le nombre de fois que chaque terme de notre glossaire, y compris ses variations terminologiques, figurait dans les trois traductions (Figure 26). Il est important de signaler que le nombre d'occurrences obtenu était supérieur à celui calculé à l'issue de notre évaluation humaine, puisque les scripts de Python 3 ont compté certains termes plusieurs fois. Étant donné par exemple deux termes (p.ex. « mask » et « hygienic mask »), dont un figurant à la fois comme terme simple (« mask ») et aussi comme élément d'un terme complexe (« hygienic mask »), Python 3 a compté deux fois le mot (« mask ») qui objectivement est apparu dans deux dénominations, alors que dans le cadre de notre évaluation humaine nous l'avons considéré de façon indépendante (une occurrence pour « mask » et une occurrence pour « hygienic mask »). Bien que ce décompte n'ait pas reflété celui que nous avons fait dans le cadre de l'évaluation humaine, cette différence n'a causé

³⁹ <https://www.letsmt.eu/>.

⁴⁰ <https://swan.cern.ch>.

aucun problème. En effet, les scripts de Python 3 ont compté les occurrences dans les trois traductions (« Réf, Syst sans glos, Syst avec glos ») selon le même critère.

```
extraordinary situation -> 1
Extraordinary situation -> 1
Face masks must also be worn -> 1
face-to-face teaching -> 2
FDFA -> 1
FDHA -> 2
FDJP -> 1
Federal Council -> 24
Federal Council Coronavirus Crisis Unit -> 1
federal crisis management of the COVID-19 pandemic -> 1
```

Figure 26 Nombre d'occurrences des termes calculées par les scripts de Python 3.

Afin d'éviter d'attribuer des scores plus élevés aux traductions de DeepL Pro où la fréquence d'un certain terme pouvait être supérieure à celle du même terme figurant dans la traduction de référence (surestimation d'un système de TA), nous avons ensuite dû adapter le nombre des occurrences des termes dans les traductions automatiques selon le nombre de fois maximum qu'ils sont figurés dans la traduction de référence. En l'espèce, nous avons adapté le nombre d'occurrences de 30 termes et/ou variations terminologiques contenus dans les traductions automatiques de DeepL Pro.

Enfin, nous avons pu appliquer la formule du Term Hit Rate (THR) (Figure 19).

À la lumière de l'ensemble des résultats obtenus, y compris ceux dérivant de notre évaluation humaine, nous avons finalement pu répondre à notre question de recherche, ainsi qu'à la série de sous-questions qui en découlent. Dans le chapitre suivant (5), nous présenterons et discuterons les résultats, afin de répondre à nos questions de recherches et tirer des conclusions sur notre travail (chapitre 6).

4.3 Conclusions sur la méthodologie

Dans ce chapitre, nous avons retracé la démarche entreprise pour mener à bien le projet de ce mémoire en détaillant chaque étape nécessaire à son bon déroulement. D'abord, nous avons défini les objectifs et les questions de recherche (section 4.1), puis nous avons passé en revue les différentes tâches effectuées pour réaliser notre travail (section 4.2).

5 Analyse des résultats

Dans ce chapitre, nous nous pencherons sur l'analyse des résultats de nos évaluations. Nous commencerons par décrire l'évaluation humaine (sections 5.1), puis nous reprendrons les résultats de l'évaluation automatique (section 5.2). Enfin, nous récapitulerons le présent chapitre (section 5.3).

5.1 Évaluation humaine

Par la suite, nous détaillerons les résultats de l'évaluation humaine, réalisée par nous-mêmes pour comparer les traductions de DeepL Pro réalisées avec (section 5.1.2) et sans (section 5.1.1) la fonctionnalité « glossaire ». Concrètement, nous commenterons les termes rencontrés en fonction de leur classement (« termes exacts » et « termes inexacts ») tout en tenant compte de leurs sous-catégories le cas échéant (section 4.2.6).

5.1.1 TA de DeepL Pro réalisée sans la fonctionnalité « glossaire »

Notre évaluation humaine de la traduction de DeepL Pro réalisée sans la fonctionnalité « glossaire » a donné les résultats suivants (Tableau 6).

Classification des termes traduits par DeepL Pro sans « glossaire »	Nombre d'occurrences des termes	%
Termes cibles exacts par rapport à la traduction de référence (section 5.1.1.1)	136	52%
Termes cibles inexacts par rapport à la traduction de référence (section 5.1.1.2)	124	48%
Total	260	100%

Tableau 6 « Termes exacts » et « termes inexacts » dans la TA de DeepL Pro sans « glossaire ».

5.1.1.1 Termes cibles exacts

Dans 52% des cas (N=136/260), DeepL Pro a bien traduit nos termes. Dans ce contexte, nous avons noté un cas⁴¹ (Tableau 7), où le terme choisi par DeepL Pro était tout à fait approprié au

⁴¹ Phrase n° 9.

contexte, bien qu'il soit différent de celui contenu dans la traduction de référence. En particulier, le traducteur ou la traductrice humain-e ont traduit « Impfung » (en français « vaccination ») avec « vaccine⁴² » (en allemand « Impstoff »), alors qu'il ne s'agit pas de l'équivalent correct (« vaccination⁴³ »).

Src	Réf	Syst sans glos
Aufgrund ihres Mandats und ihrer Erfahrung könne GAVI wesentlich dazu beitragen, dass eine künftige Impfung gegen COVID-19 auch die Menschen in den ärmsten Ländern erreiche.	Due to its mandate and experience, Gavi is well placed to ensure that any future vaccine against COVID-19 will also reach people in the poorest countries	With its mandate and experience, GAVI can make a significant contribution to ensuring that future vaccination against COVID-19 reaches people in the poorest countries.

Tableau 7 Termes cibles exacts dans DeepL Pro sans « glossaire ».

5.1.1.2 Termes cibles inexacts

Dans 48% des cas (N=124/260), DeepL Pro n'a pas bien traduit nos termes. Pour identifier les problèmes terminologiques du traducteur en ligne, nous avons classé les termes inexacts en cinq sous-catégories (Tableau 8).

Sous-catégories des termes cibles inexacts par rapport à la traduction de référence dans DeepL Pro sans « glossaire »		
Sous-catégories	Nombre d'occurrences	%
Recours à d'autres formes linguistiques	72	28%
Emploi du langage commun au détriment d'une communication spécialisée	28	11%
Erreurs de sens	15	6%
Termes non conformes à la réalité suisse	6	2%
Noms propres et acronymes en LC mal traduits	3	1%
Total	124	48%

Tableau 8 Sous-catégories des termes cibles inexacts dans DeepL Pro sans « glossaire ».

À l'aide d'exemples concrets, nous commenterons ci-après les résultats obtenus par sous-catégorie.

⁴² Le terme « vaccine » est défini par Merriam-Webster (2020) comme « *a preparation of killed microorganisms, living attenuated organisms, or living fully virulent organisms that is administered to produce or artificially increase immunity to a particular disease* ».

⁴³ Le terme « vaccination » est défini par Merriam-Webster (2020) comme « *the act of vaccinating* ».

5.1.1.2.1 Recours à d'autres formes linguistiques

Dans 28% des cas (N=72/260), DeepL Pro a fait recours à d'autres formes linguistiques que nous avons catégorisées de la façon suivante (Tableau 9).

Recours à d'autres formes linguistiques dans DeepL Pro sans « glossaire »		
Types de formes linguistiques	Nombre d'occurrences	%
Synonymes	33	13%
Variations terminologiques de type morphosyntaxique ³⁶	18	7%
Variations terminologiques de type graphique ³⁴	15	6%
Variations terminologiques de type flexionnel ³⁵	3	1%
Variations terminologiques de type syntaxique faible ⁴⁴	3	1%
Total	72	28

Tableau 9 Recours à d'autres formes linguistiques dans DeepL Pro sans « glossaire ».

Les cas de **synonymie** (Tableau 10) et de **variation morphosyntaxique** (Tableau 11) sont les plus nombreux.

Src	Réf	Syst sans glos
Coronavirus: Bundesrat lockert weitere Massnahmen ab dem 11. Mai 2020	Coronavirus: Federal Council to ease further measures from 11 May	Coronavirus: Federal Council relaxes further measures from 11 May 2020
Derzeit ist es jedoch entscheidend, sich auf das Bewältigen dieser Krise zu konzentrieren	At the moment however, the priority is on managing this crisis	However, at this stage it is crucial to focus on tackling this crisis
Er ruft insbesondere auch die ältere Bevölkerung dazu auf, zu Hause zu bleiben .	It is calling on the elderly in particular to remain at home .	It also calls in particular on the elderly to stay at home .
Den Test auf das SARS-Coronavirus-2 vergütet der Bund mit einer Pauschale von 169, jenen für den Antikörpernachweis mit 113 Franken.	It will reimburse the cost of tests for SARS-Coronavirus-2 at a flat rate of 169 francs, and of antibody tests at a rate of 113 francs.	The Swiss Confederation pays a lump sum of CHF 169 for the test for SARS coronavirus-2 and CHF 113 for the detection of antibodies .

Tableau 10 Exemples de synonymes dans DeepL Pro sans « glossaire ».

⁴⁴ Elles regroupent les termes où la préposition (p.ex. « of » au lieu de « for », « with » au lieu de « by », « with + nom » au lieu d'un terme complexe) servant à rattacher les éléments variés, nous avons noté trois exemples (L'Homme, 2020 : 83).

Src	Réf	Syst sans glos
Identifizierung von Spezialmassnahmen im Bereich der Innovationsförderung mit dem Ziel, aus Schweizer Wissenschafts-Know-how rasch Produkte oder Dienstleistungen hervorbringen und damit wesentlich zur Bekämpfung von Covid-19 beitragen zu können.	To identify special measures that can be introduced to encourage innovation, with the aim of exploiting Swiss scientific know-how to rapidly generate products and services in order to effectively combat COVID-19.	Identification of special measures in the field of innovation promotion with the aim of rapidly generating products or services from Swiss scientific know-how and thus making a significant contribution to the fight against Covid-19.
Die Covid-19-Pandemie zeigt deutlich, wie wichtig die Prävention und die Kontrolle von Infektionen sind.	The COVID-19 pandemic clearly shows the importance of infection prevention and control .	The Covid 19 pandemic clearly shows the importance of preventing and controlling infections .
Parallel zu diesen Öffnungsschritten werden die Einreisebeschränkungen gelockert.	In parallel to this easing of measures, restrictions on entering Switzerland will also be relaxed.	. In parallel with these opening steps, entry restrictions will be relaxed.
Coronavirus: Keine Quarantäne bei der Einreise aus den Grenzregionen	Coronavirus: No quarantine for persons entering from border regions	Coronavirus: No quarantine on entry from border regions
Die Menschen sollen weiterhin die Hygiene- und Abstandsregeln einhalten.	people should continue to follow the rules on hygiene and social distancing .	people should continue to comply with hygiene and distance regulations
Die Armeepothek hat bis heute 21 Millionen Hygienemasken an die Kantone verteilt.	To date, the army pharmacy has distributed 21 million hygienic masks to the cantons.	To date, the army pharmacy has distributed 21 million hygiene masks to the cantons.

Tableau 11 Exemples de variations morphosyntaxiques dans DeepL Pro sans « glossaire ».

Dans le contexte des **variations graphiques**, qui concernent essentiellement des distinctions graphiques mineures, comme l'ajout ou la suppression d'un trait d'union (p.ex. Covid 19, Covid-19, distance-learning, distance learning) ou des différences de casse (p.ex. COVID-19, Covid-19, federal government, Federal Government), nous avons vu que dans la plupart des cas celles-ci ne sont pas justifiées par un effet miroir avec la langue source (Tableau 12).

Dans le cadre de la TA, ces variations graphiques ont causé dans 10 cas des incohérences lorsque certains concepts ont été répétés, bien que les termes en langue source aient toujours été écrits de la même façon. À titre d'exemple, nous avons vu que « Covid-19 » a toujours été écrit tel quel, alors que le concept « Covid-19-Impfstoff » (en français « vaccin contre le Covid-19 ») a été traduit de quatre manières différentes (une⁴⁵ occurrence pour « Covid-19 vaccin », deux⁴⁶ pour « COVID-19 vaccin », une⁴⁷ pour « Covid 19 vaccin » et deux⁴⁸ pour « COVID

⁴⁵ Phrase n° 12.

⁴⁶ Phrases n°40 et 102.

⁴⁷ Phrase n° 66.

⁴⁸ Phrases n° 88 et 116.

19 »). Cela s'est vérifié aussi pour « SwissCovid App » et « SwissCovid app », respectivement apparus une fois⁴⁹. En ce qui concerne les variations graphiques « Covid-19 vaccin », « COVID-19 vaccin », « SwissCovid App » et « SwissCovid app », il a été intéressant de remarquer que ces incohérences existaient aussi dans la traduction de référence.

Src	Réf	Syst sans glos
Damit erhält die Schweizer Bevölkerung Zugang zu einem weiteren Covid-19-Impfstoff – sofern die klinische Testphase und die Zulassung durch Swissmedic erfolgreich verlaufen.	The Swiss public will therefore gain access to a further COVID-19 vaccine - provided it successfully passes the clinical trial phase and is approved by Swissmedic.	This will give the Swiss population access to a further Covid 19 vaccine - provided that the clinical trial phase and Swissmedic's approval are successful.
COVID-19-Impfstoff: Bund unterzeichnet Vertrag mit AstraZeneca	COVID-19 vaccine: Swiss federal government signs agreement with AstraZeneca	COVID 19 vaccine: Federal government signs contract with AstraZeneca
Der Bundesrat verabschiedete zudem die Verordnung über das Proximity-Tracing-System	The Federal Council also adopted the Ordinance on the Proximity Tracing System	The Federal Council also approved the ordinance on the proximity tracing system
Damit ermöglicht er den Start der SwissCovid App.	thereby making it possible to start using the SwissCovid app	thus enabling the launch of the SwissCovid App.
Diese können den Unterricht flexibel gestalten und die Möglichkeiten von Fernunterricht weiter nutzen.	Teaching can be organised in a flexible manner and they and the may continue to make use of distance-learning channels.	They can organise their lessons flexibly and continue to use the possibilities of distance learning.
Damit ermöglicht er den Start der SwissCovid App.	thereby making it possible to start using the SwissCovid app	thus enabling the launch of the SwissCovid App.
Das NCSC nimmt auch nach der öffentlichen Lancierung der SwissCovid App Meldungen von Testergebnissen entgegen, um den Datenschutz und die Sicherheit der App laufend zu gewährleisten.	The NCSC will continue to accept reports of test results following the public launch of the SwissCovid app so that data protection and the security of the app can be assured.	Even after the public launch of the SwissCovid app , the NCSC will continue to receive reports of test results in order to guarantee the privacy and security of the app on an ongoing basis.

Tableau 12 Exemples de variations graphiques dans DeepL Pro sans « glossaire ».

Dans le cadre des **variations flexionnelles**, qui incluent les différentes formes fléchies d'un terme, nous n'avons observé que trois⁵⁰ cas où la phrase de la TA était un calque de la langue source (**Tableau 13**). Pour cette raison, les termes de DeepL Pro différaient de ceux utilisés dans la traduction de référence (p.ex. emploi du singulier au lieu du pluriel, usage du passé composé au lieu du présent).

⁴⁹ Phrases n° 53 pour « SwissCovid App » et 54 pour « SwissCovid app ».

⁵⁰ Phrases n° 26, 48 et 93.

Src	Réf	Syst sans glos
Der Abstand kann weiterhin unterschritten werden, wenn eine Maske getragen wird oder Trennwände vorhanden sind.	If masks are worn or partitions are in place, the distance may be further reduced.	The distance can still be reduced if a mask is worn or if there are partitions.
Alle Personen, die mittels eines Schnelltests positiv getestet wurden , sollten sich umgehend in Isolation begeben.	However, anyone who tests positive following a rapid test should go into isolation immediately.	All persons who have tested positive using a rapid test should be placed in isolation immediately.
Die Pflicht zum Tragen einer Gesichtsmaske gilt in obligatorischen Schulen, Schulen der Sekundarstufe II und der Tertiärstufe, in Institutionen der familienergänzenden Kinderbetreuung sowie in den Trainingsbereichen von Sport- und Fitnessseinrichtungen nur dann, wenn sie im betreffenden Schutzkonzept vorgesehen ist.	Face masks must also be worn in all schools and higher education establishments, in child-care facilities and in the training areas of sport and fitness centres if part of the centre's precautionary measures.	The obligation to wear a face mask applies in compulsory schools, upper secondary and tertiary schools, in institutions providing supplementary childcare and in the training areas of sports and fitness facilities only if it is provided for in the relevant protection concept.

Tableau 13 Exemples de variations flexionnelles dans DeepL Pro sans « glossaire ».

Enfin, nous avons noté trois⁵¹ exemples de **variations syntaxiques faibles** (Tableau 14), qui regroupent les termes où la préposition servant à rattacher les éléments varie (p.ex. « of » au lieu de « for », « with » au lieu de « by » et « with + nom » au lieu d'un terme complexe « nom + nom ») (L'Homme, 2020 : 83).

Src	Réf	Syst sans glos
Schutz von besonders gefährdeten Arbeitnehmenden	Protection for employees at especially high risk	Protection of particularly exposed workers
An privaten Veranstaltungen mit über 15 Personen darf künftig nur sitzend konsumiert werden. Dieses Verbot ist eine der zentralen Massnahmen, mit der die Zahl der Infektionen mit dem Coronavirus reduziert werden konnte.	At private events attended by over 15 people , food and drink may only be consumed while seated. The ban is one of the key measures that has helped to reduce the number of coronavirus infections in Switzerland.	At private events with more than 15 people , consumption will in future only be allowed seated. This ban is one of the key measures to reduce the number of infections with coronavirus .

Tableau 14 Exemples de variations syntaxiques faibles dans DeepL Pro sans « glossaire ».

⁵¹ Phrases n° 32, 98 et 107.

5.1.1.2.2 Emploi du langage commun au détriment d'une communication spécialisée

Dans 11% des cas (N=28/260), DeepL Pro a utilisé des mots du langage commun au détriment d'une communication spécialisée (Tableau 15). À titre d'exemple, l'expression « Präsenzunterricht verzichten » (en français « renoncer à l'enseignement présentiel ») a été traduite par « do without attendance classes ». Bien que l'on comprenne le concept, il ne s'agit pas d'une formulation utilisée en communication spécialisée. De plus, le terme « eidgenössischen Volksbegehren » (en français « requête populaire ») a été traduit par DeepL Pro avec « federal referendum », alors que le terme en allemand recouvre à la fois les concepts d'initiative populaire (en anglais « popular initiative ») et de demande de référendum (en anglais « optional referendum ») (TERMDAT, 2020).

Src	Réf	Syst sans glos
Ab Montag, 2. November, müssen Hochschulen auf Präsenzunterricht verzichten .	From Monday, 2 November, higher education institutions will have to forgo face-to-face teaching .	As of Monday, 2 November, universities will have to do without attendance classes .
Die Versorgung der gesamten Bevölkerung mit Lebensmitteln, Medikamenten und Waren des täglichen Gebrauchs ist sichergestellt, es sind genügend Vorräte angelegt.	Provision of the entire population with foodstuffs, medicinal products and essential everyday items is assured; there are sufficient stocks.	The supply of the entire population with food, medicines and goods for daily use is assured and sufficient stocks have been built up.
Wer sich auf Anordnung eines Arztes oder einer Behörde in Quarantäne begeben muss, hat Anrecht auf den Corona-Erwerbsersatz .	Anyone instructed by a doctor or the authorities to go into quarantine is entitled to compensation for loss of earnings related to the corona pandemic .	Anyone who has to go into quarantine by order of a doctor or an authority is entitled to the Corona compensation .
Bis Ende Mai gilt noch der Fristenstillstand bei eidgenössischen Volksbegehren .	Deadlines for submitting federal popular requests remain suspended until the end of May.	Until the end of May, the deadline for federal referendums will be suspended.
Deshalb sicherte sich der Bund im Sommer 2020 ein Immunotherapeutikum von Molecular Partners.	In summer 2020, the Swiss federal government therefore secured access to an immunotherapeutic agent from Molecular Partners.	This is why the German government secured an immunotherapeutic drug from Molecular Partners in the summer of 2020.

Tableau 15 Emploi du langage commun vs. spécialisé dans DeepL Pro sans « glossaire ».

5.1.1.2.3 Erreurs de sens

Dans 6% des cas (N=15/260), DeepL Pro a mal traduits nos termes et a causé des erreurs de sens (Tableau 16). Il n'a par exemple pas été en mesure de transposer en langue cible le nom propre « Moderna », lorsqu'il était lié à une forme possessive (Src « Modernas mRNA-

Impfstoff » ; Réf « Moderna's mRNA vaccine ») et l'a finalement traduit avec un adjectif (Syst sans glos « Modern's mRNA vaccine »). De plus, le traducteur en ligne a confondu dans deux cas⁵² le gouvernement suisse (Src « Bund » ; Réf « federal government ») avec celui allemand (Syst sans glos « German government »).

Src	Réf	Syst sans glos
Covid-19-Impfstoff: Bund unterzeichnet Vertrag mit Biotech-Unternehmen Moderna	Covid-19 vaccine: federal government signs agreement with biotechfirm Moderna	Covid-19 vaccine: German government signs contract with biotech company Moderna
Neu gelten dieselben Vorgaben für alle Musterschutzkonzepte gibt es keine mehr.	The same requirements will apply for all sets of precautionary measures; there are to be no more model sets of precautionary measures .	Now the same requirements apply to all concepts; design protection concepts no longer exist.
Die Änderung tritt am 29.10.2020 in Kraft.	The changes come into force on 29 October.	The amendment will enter into force on 29.10.2020.
Der Bundesrat setzt nach den erfolgten Lockerungsschritten noch verstärkt auf eigenverantwortliches Handeln Modernas mRNA-Impfstoff setzt auf eine neuartige Technologie:	Following the steps to ease measures, the Federal Council is relying even more heavily on the public to behave responsibly Moderna's mRNA vaccine is based on an innovative technology:	Following the relaxation steps that have been taken, the Federal Council is placing even greater emphasis on independent action Modern's mRNA vaccine is based on a novel technology:
Diese Veranstaltungen sollen wenn möglich vermieden werden	Gatherings should be avoided were possible.	These events should be avoided if possible
Der Bundesrat wird das System ausser Betrieb nehmen, sobald es für die Bekämpfung des Coronavirus nicht mehr erforderlich ist.	As soon as the system is no longer required to combat the coronavirus , the Federal Council will take it out of service.	The Federal Council will take the system out of service as soon as it is no longer necessary for the control of the coronavirus .
Am 8. Juni sollen in einem dritten Schritt die Mittel-, Berufs- und Hochschulen wieder Präsenzveranstaltungen abhalten dürfen.	In phase three, set to commence on 8 June, upper secondary schools, vocational schools and higher education institutions will be allowed to resume face-to-face teaching .	In a third step, secondary schools, vocational colleges and universities will be allowed to hold face-to-face events again on 8 June.

Tableau 16 Exemples d'erreurs de sens dans DeepL Pro sans « glossaire ».

5.1.1.2.4 Termes cibles qui ne sont pas conformes à la réalité suisse

Dans 2% des cas (N= 6/260), DeepL Pro a employé des termes qui ne sont pas conformes à la réalité suisse (Tableau 17). Il a par exemple traduit le terme allemand « Mittelschule » (en français « école du degré secondaire supérieur ») avec « secondary school », alors que la terminologie correcte « upper secondary school » spécifie qu'il s'agit bien d'une école secondaire, mais de degré supérieur.

⁵² Phrases n° 12 et 33.

Src	Réf	Syst sans glos
Die Pflicht zum Tragen einer Gesichtsmaske gilt in obligatorischen Schulen, Schulen der Sekundarstufe II und der Tertiärstufe, in Institutionen der familienergänzenden Kinderbetreuung sowie in den Trainingsbereichen von Sport- und Fitnessseinrichtungen nur dann, wenn sie im betreffenden Schutzkonzept vorgesehen ist.	Face masks must also be worn in all schools and higher education establishments, in child-care facilities and in the training areas of sport and fitness centres if part of the centre's precautionary measures.	The obligation to wear a face mask applies in compulsory schools, upper secondary and tertiary schools, in institutions providing supplementary childcare and in the training areas of sports and fitness facilities only if it is provided for in the relevant protection concept.
Ab Montag, 2. November, müssen Hochschulen auf Präsenzunterricht verzichten.	From Monday, 2 November, higher education institutions will have to forgo face-to-face teaching.	As of Monday, 2 November, universities will have to do without attendance classes.
Die kantonalen Gymnasien können dieses Jahr auch auf die schriftlichen Maturitätsprüfungen verzichten, nachdem die Erziehungsdirektorenkonferenz EDK bereits beschlossen hatte, die mündlichen Prüfungen nicht durchzuführen.	Cantonal baccalaureate schools will be permitted to dispense with written baccalaureate examinations this year, following a decision of the Swiss Conference of Cantonal Directors of Education (EDK) not to hold oral examinations (see separate press release).	This year, the cantonal Gymnasiums will also be able to dispense with the written Matura examinations, after the EDK Education Directors' Conference had already decided not to hold the oral examinations.
Am 8. Juni sollen in einem dritten Schritt die Mittel-, Berufs- und Hochschulen wieder Präsenzveranstaltungen abhalten dürfen.	In phase three, set to commence on 8 June, upper secondary schools, vocational schools and higher education institutions will be allowed to resume face-to-face teaching.	In a third step, secondary schools, vocational colleges and universities will be allowed to hold face-to-face events again on 8 June.

Tableau 17 Termes qui ne sont pas conformes à la réalité suisse dans DeepL Pro sans « glossaire ».

5.1.1.2.5 Noms propres et acronymes en LC qui sont mal traduits

Dans 1%⁵³ des cas (N=3/260), DeepL Pro a mal traduit des noms propres et des acronymes (Tableau 18). Par exemple, l'acronyme allemand « KSBC » n'a pas d'équivalent cible et, pour cette raison, a été omis dans la traduction de référence. En revanche, DeepL Pro l'a transposé en langue cible en gardant sa forme linguistique en allemand. Concernant la traduction en anglais du terme « Erziehungsdirektorenkonferenz EDK » (en français « Conférence des directeurs cantonaux de l'instruction publique, CDIP »), DeepL Pro n'a pas inclus dans sa traduction l'adjectif « cantonal », puisqu'il n'apparaissait pas dans le texte source, bien qu'en principe il doit faire partie de la dénomination officielle. De plus, le traducteur en ligne a utilisé la forme possessive, alors que le terme officiel en anglais contient la préposition « of ». Enfin, dans un cas nous avons trouvé une faute d'orthographe (« The Federal Office of Public

⁵³ Plus celui de « Moderna » classé dans une catégorie différente, car il a causé une erreur de sens.

Health is thus following the recommendations of the European *Centre* for Disease Prevention and Control (ECDC) »).

Src	Réf	Syst sans glos
Vor diesem Hintergrund haben der Krisenstab des Bundesrates zur Bewältigung der Corona Krise (KSBC) , das Staatssekretariat für Bildung, Forschung und Innovation (SBFI) und das Bundesamt für Gesundheit (BAG) eine wissenschaftliche Science Task Force einberufen.	The Swiss National COVID-19 Science Task Force has therefore been set up by the Federal Council Coronavirus Crisis Unit , the State Secretariat for Education, Research and Innovation SERI and the Federal Office of Public Health FOPH.	Against this background, the Federal Council's Crisis Management Team for the Corona Crisis (KSBC) , the State Secretariat for Education, Research and Innovation (SERI) and the Federal Office of Public Health (FOPH) have convened a scientific science task force.
Die kantonalen Gymnasien können dieses Jahr auch auf die schriftlichen Maturitätsprüfungen verzichten, nachdem die Erziehungsdirektorenkonferenz EDK bereits beschlossen hatte, die mündlichen Prüfungen nicht durchzuführen.	Cantonal baccalaureate schools will be permitted to dispense with written baccalaureate examinations this year, following a decision of the Swiss Conference of Cantonal Directors of Education (EDK) not to hold oral examinations (see separate press release).	This year, the cantonal Gymnasiums will also be able to dispense with the written Matura examinations, after the EDK Education Directors' Conference had already decided not to hold the oral examinations.
Damit folgt das Bundesamt für Gesundheit den Empfehlungen des Europäischen Zentrums für die Prävention und die Kontrolle von Krankheiten (ECDC) .	The Federal Office of Public Health is thus following the recommendations of the European Center for Disease Prevention and Control (ECDC) .	The Federal Office of Public Health is thus following the recommendations of the European Centre for Disease Prevention and Control (ECDC) .

Tableau 18 Noms propres et acronymes en LC qui sont mal traduits par DeepL Pro sans « glossaire ».

À l'issue de notre analyse, nous retenons que la terminologie contenue dans l'ensemble de la traduction de DeepL Pro réalisée sans la fonctionnalité « glossaire » n'est pas de très bonne qualité, puisqu'elle ne reflète pas intégralement celle utilisée par l'Office fédéral suisse de la santé publique. En effet, celle-ci est pleinement correcte seulement dans 52% des cas (N=136/260). Dans le 48% des cas restants (N=124/260), nous avons constaté différentes erreurs de terminologie (**Tableau 8**) qui, dans le contexte d'un organisme public, tel que la Confédération suisse, où la terminologie constitue une exigence de qualité, empêchent de garantir la cohérence conceptuelle de la législation, l'harmonisation des dialogues dans la coopération internationale ainsi que la communication efficace et incontournable avec la population.

Dans la section suivante (5.1.2), nous verrons si la fonctionnalité « glossaire » de DeepL Pro nous a permis d'apporter des améliorations sur le plan terminologique.

5.1.2 TA de DeepL Pro réalisée avec la fonctionnalité « glossaire »

Selon l'évaluation humaine, DeepL Pro avec « glossaire » a bien traduit les termes sources en langue cible dans 86% des cas (Tableau 19). Il s'agit donc d'une amélioration de 34% (86% contre 52%) par rapport à notre évaluation précédente (section 5.1.1).

Classification des termes traduits par DeepL Pro avec « glossaire »	Nombre d'occurrences des termes	%
Termes cibles exacts par rapport à la traduction de référence (section 5.1.2.1)	223	86%
Termes cibles inexacts par rapport à la traduction de référence (section 5.1.2.2)	37	14%
Total	260	100%

Tableau 19 « Termes exacts » et « termes inexacts » dans la TA de DeepL Pro avec « glossaire ».

5.1.2.1 Termes cibles exacts

En regardant cependant plus en détail les termes exacts par rapport à la référence, nous avons constaté que l'amélioration est apparente (Tableau 20). En effet, dans 8% des cas (N=32/260), DeepL Pro n'a pas été capable d'utiliser des synonymes ou des variations terminologiques pour exprimer un même concept (N=21/223), ce qui a causé des répétitions (section 5.1.2.1.2). De plus, dans 4% des cas (N=11/260), l'intégration de la terminologie a eu un impact négatif sur les phrases, alors que ces dernières étaient correctes avant l'ajout du « glossaire » (section 5.1.2.1.3).

En conséquence, l'amélioration de la TA des termes par rapport à notre évaluation précédente (section 5.1.1) correspond effectivement à 21% (73% dans DeepL Pro avec « glossaire » contre 52% dans DeepL Pro sans « glossaire »).

Sous-catégories des termes cibles exacts par rapport à la traduction de référence dans DeepL Pro sans « glossaire »		
Sous-catégories	Nombre d'occurrences	%
Termes intégrés parfaitement dans la TA	191	73%
Termes qui provoquent des répétitions	21	9%
Erreurs de syntaxe et/ou d'orthographe dans la phrase	11	4%
Total	223	86%

Tableau 20 Sous-catégories des termes cibles exacts dans DeepL Pro avec « glossaire ».

À l'aide d'exemples concrets, nous commenterons ci-après les « termes exacts » par sous-catégorie.

5.1.2.1.1 Termes intégrés parfaitement dans la TA

Dans 73% des cas (N=191/260), DeepL Pro a pleinement tenu compte des termes contenus dans le glossaire, sans provoquer d'impacts négatifs sur la phrase (Tableau 21).

Src	Réf	Syst sans glos	Syst avec glos
Wer sich auf Anordnung eines Arztes oder einer Behörde in Quarantäne begeben muss, hat Anrecht auf den Corona-Erwerbsersatz .	Anyone instructed by a doctor or the authorities to go into quarantine is entitled to compensation for loss of earnings related to the corona pandemic .	Anyone who has to go into quarantine by order of a doctor or an authority is entitled to the Corona compensation .	Anyone who has to go into quarantine by order of a doctor or an authority is entitled to the compensation for loss of earnings related to the corona pandemic .
Schweizer Diplomaten und Diplomaten sind neu ebenfalls für notwendige dienstliche Tätigkeiten von der Quarantäne befreit .	Swiss diplomats will now also be exempt from the requirement to go into quarantine when returning from official business trips.	Swiss diplomats are now also exempt from quarantine for necessary official activities.	Swiss diplomats are now also exempt from the requirement to go into quarantine for necessary official activities.
Die Versorgung der gesamten Bevölkerung mit Lebensmitteln, Medikamenten und Waren des täglichen Gebrauchs ist sichergestellt, es sind genügend Vorräte angelegt.	Provision of the entire population with foodstuffs, medicinal products and essential everyday items is assured; there are sufficient stocks.	The supply of the entire population with food, medicines and goods for daily use is assured and sufficient stocks have been built up.	The provision of the entire population with food, medicines and goods for daily use is guaranteed, and sufficient stocks have been built up.

Tableau 21 Termes intégrés parfaitement dans la TA de DeepL Pro avec « glossaire ».

Parmi ces cas, nous avons remarqué une phrase⁵⁴, qui dans le cadre de la traduction de DeepL Pro sans « glossaire » correspondait parfaitement à la traduction de référence, alors que ce n'était plus le cas après l'intégration de notre terminologie (Tableau 22). Plus précisément, nous avons noté le remplacement du mot « further » par « other » (synonymie).

⁵⁴ Phrase n°3.

Src	Réf	Syst sans glos	Syst avec glos
Weitere Impfstoffprojekte Evaluation	Further vaccine projects in under evaluation	Further vaccine projects under evaluation	Other vaccine projects under evaluation

Tableau 22 Mots changés après l'intégration de nos termes dans DeepL Pro avec « glossaire ».

5.1.2.1.2 Termes qui provoquent des répétitions

Dans 9% des cas (N=21/260), DeepL Pro a provoqué des répétitions de termes, puisqu'il n'a pas permis d'alterner les différentes formes linguistiques pour exprimer un même concept, soit sous forme de synonyme (4%, soit N=10/260 ; [Tableau 23](#)), soit sous forme de variation terminologique (5%, soit N=11/260 ; [Tableau 24](#)). DeepL Pro ne tient compte de ces formes linguistiques que si pour chaque synonyme ou variation terminologique en langue cible (p.ex. « combat COVID-19 » et « combat the coronavirus »), il existe des équivalents correspondants en langue source (p.ex. « Bekämpfung von Covid-19 » et « Bekämpfung des Coronavirus »).

Src	Réf	Syst sans glos	Syst avec glos
Bund beteiligt sich an COVAX- Initiative zur Beschaffung von COVID-19-Impfstoffen	The Swiss Confederation joins the COVAX initiative for the procurement of COVID-19 vaccines	Federal government participates in COVAX initiative for the procurement of COVID-19 vaccines	Swiss federal government participates in COVAX initiative for the procurement of COVID-19 vaccines
Bern, 16.10.2020 - Der Bund will der Schweizer Bevölkerung einen raschen Zugang zu COVID-19-Impfstoffen gewährleisten. Den Test auf das SARS-Coronavirus-2 vergütet der Bund mit einer Pauschale von 169, jenen für den Antikörpernachweis mit 113 Franken.	Berne, 16.10.2020 - The Swiss federal government wants to assure members of the Swiss public rapid access to COVID-19 vaccines. It will reimburse the cost of tests for SARS-Coronavirus-2 at a flat rate of 169 francs, and of antibody tests at a rate of 113 francs.	Bern, 16.10.2020 - The Swiss Confederation wants to guarantee the Swiss population rapid access to COVID-19 vaccines. The Swiss Confederation pays a lump sum of CHF 169 for the test for SARS coronavirus-2 and CHF 113 for the detection of antibodies.	Bern, 16.10.2020 - The Swiss federal government wants to guarantee the Swiss population rapid access to Covid-19 vaccines. The Swiss federal government will pay a flat-rate fee of CHF 169 for the test for SARS-Coronavirus-2 and CHF 113 for the detection of antibodies.
Alle Betriebe und Organisatoren von Veranstaltungen müssen ein Schutzkonzept haben, das sich entweder auf ein Branchenkonzept oder auf die Vorgaben des BAG und des Staatssekretariats für	All companies and event organizers must have a protection concept that is based either on an industry concept or on the requirements of the FOPH and the State Secretariat for Economic Affairs (Seco).	All companies and organisers of events must have a protection concept based either on a sector concept or on the specifications of the FOPH and the State Secretariat for Economic Affairs (Seco).	All businesses and organisers of gatherings must have a precautionary measure based either on a sectoral concept or on the specifications of the FOPH and the State Secretariat for

Wirtschaft (Seco) abstützt.			Economic Affairs (Seco).
Voraussetzung ist, dass der Aufenthalt im Ausland nicht mehr als fünf Tage dauert und ein Schutzkonzept erstellt und umgesetzt wird.	The persons concerned must not remain abroad for more than five days, and precautionary measures must be planned and taken.	The prerequisite is that the stay abroad does not exceed five days and that a protection concept is drawn up and implemented.	The prerequisite is that the stay abroad does not exceed five days and that a precautionary measure is drawn up and implemented.

Tableau 23 Répétitions causées par l'absence de synonymes dans DeepL Pro avec « glossaire ».

Src	Réf	Syst sans glos	Syst avec glos
Ziel ist auch, den Anstieg der Fallzahlen so stark zu bremsen, dass die Kantone das Contact Tracing weiterhin konsequent und umfassend sicherstellen können.	The intention is also to considerably curb the number of new cases so that the cantons can continue to contract trace effectively.	The aim is also to slow down the increase in the number of cases to such an extent that the cantons can continue to ensure contact tracing consistently and comprehensively.	The aim is also to slow down the increase in the number of cases of people infected with the virus to such an extent that the cantons can continue to ensure contact tracing consistently and comprehensively.
Der starke Anstieg der Fallzahlen in den letzten Tagen ist besorgniserregend.	The rapid rise in coronavirus cases in the last few days is a cause for great concern.	The sharp increase in the number of cases in recent days is worrying.	The sharp increase in the number of cases of people infected with the virus in recent days is a cause for concern.
Angesichts der anhaltend tiefen Fallzahlen hat der Bundesrat am 19. Juni 2020 die ausserordentliche Lage aufgehoben und weitere Lockerungen beschlossen.	In view of the continuing low number of coronavirus cases , the Federal Council downgraded the situation from 'extraordinary' on 19 June and eased further measures.	In view of the persistently low number of cases , the Federal Council lifted the extraordinary situation on 19 June 2020 and decided on further easing.	In view of the continuing low number of cases of people infected with the virus , the Federal Council no longer applied the extraordinary situation on 19 June 2020 and decided on further ease measures.
Coronavirus: weitere Massnahmen zur Eindämmung der Epidemie, Einführung von Schnelltests , Reisequarantäne neu geregelt.	Coronavirus: further measures to contain the epidemic, introduction of rapid testing , new rules on travel quarantine.	Coronavirus: further measures to contain the epidemic, introduction of rapid tests , new rules on travel quarantine.	Coronavirus: further measures to contain the epidemic, introduction of rapid tests , travel quarantine reorganised
Alle Personen, die mittels eines Schnelltests positiv getestet wurden, sollten sich umgehend in Isolation begeben.	However, anyone who tests positive following a rapid test should go into isolation immediately.	All persons who have tested positive using a rapid test should be placed in isolation immediately.	All persons who have tested positive using a rapid test should go into isolation immediately.

Tableau 24 Répétitions causées par l'absence de variations terminologiques DeepL Pro avec « glossaire ».

5.1.2.1.3 Erreurs de syntaxe et/ou d'orthographe dans la phrase

Dans 4% des cas (N=11/260), notre terminologie a bien été incluse dans la traduction, mais au détriment de la syntaxe, de la morphosyntaxe et/ou de l'orthographe de la phrase (Tableau 25). DeepL Pro a par exemple ajouté dans trois⁵⁵ phrases des mots qui n'étaient pas nécessaires. Dans d'autres circonstances, le traducteur en ligne a fait une hypertraduction causant des erreurs de sens (phrase n° 9 « Gavi says it can play a key role »). En effet, il n'a pas détecté qu'en allemand le subjonctif I du verbe pouvoir « könne » est utilisé pour rapporter ce qu'une personne, qui ne correspond pas au sujet de la phrase, dit. Dans un cas, le terme intégré n'a pas été accordé avec l'adjectif (phrase n° 78 « these easing of measures »).

Src	Réf	Syst sans glos	Syst avec glos
Neu muss auch in den Aussenbereichen von Einrichtungen und Betrieben eine Maske getragen werden , wie zum Beispiel in Aussenbereichen von Läden und an Veranstaltungsorten, auf Terrassen in Restaurants und Bars und auf Wochen- und Weihnachtsmärkten.	Masks must now also be worn outside establishments and facilities, for example in outdoor areas of shops and at event venues, on terraces of restaurants and bars, and at farmers' and Christmas markets.	Now, masks must also be worn in the outdoor areas of institutions and businesses, such as in the outdoor areas of shops and at event locations, on terraces in restaurants and bars and at weekly and Christmas markets.	Newly, <i>wear masks must also be worn</i> in the outdoor areas of institutions and businesses, such as outside shops and event venues, on terraces in restaurants and bars and at weekly and Christmas markets.
Aufgrund ihres Mandats und ihrer Erfahrung könne GAVI wesentlich dazu beitragen , dass eine künftige Impfung gegen COVID-19 auch die Menschen in den ärmsten Ländern erreiche.	Due to its mandate and experience, Gavi is well placed to ensure that any future vaccine against COVID-19 will also reach people in the poorest countries	With its mandate and experience, GAVI can make a significant contribution to ensuring that future vaccination against COVID-19 reaches people in the poorest countries.	With its mandate and experience, <i>Gavi says it can play a key role</i> in ensuring that future vaccines against COVID-19 reach people in the poorest countries.
Parallel zu diesen Öffnungsschritten werden die Einreisebeschränkungen gelockert.	In parallel to this easing of measures , restrictions on entering Switzerland will also be relaxed.	. In parallel with these opening steps , entry restrictions will be relaxed.	In parallel with <i>these easing of measures</i> , restrictions on entering Switzerland will be eased.
Für die Reihenfolge der Lockerungen hat der Bundesrat mehrere Risikofaktoren berücksichtigt.	The Federal Council has taken into account a range of risk factors in determining the phases for easing the measures .	The Federal Council has taken several risk factors into account for the order of loosening .	The Federal Council has taken several risk factors into account for the order of ease measures .
Am 8. Juni sollen in einem dritten Schritt die Mittel-, Berufs- und Hochschulen wieder	In phase three , set to commence on 8 June, upper secondary schools, vocational schools and	In a third step , secondary schools, vocational colleges and universities will be allowed to hold	In a third phase three , on 8 June, secondary schools, vocational schools and higher

⁵⁵ Phrases n° 8, 15 et 48.

Präsenzveranstaltungen abhalten dürfen.	higher education institutions will be allowed to resume face-to-face teaching.	education face-to-face events again on 8 June.	education institutions will again be allowed to hold face-to-face teaching.
---	--	--	---

Tableau 25 Erreurs de syntaxe et/ou de l'orthographe causées par DeepL Pro avec « glossaire ».

5.1.2.2 Termes cibles inexacts

Pour identifier les problèmes terminologiques de DeepL Pro avec « glossaire » causant des termes « inexacts », nous avons classé ceux-ci en deux sous-catégories (Tableau 26).

Sous-catégories des termes cibles inexacts par rapport à la traduction de référence dans DeepL Pro avec « glossaire »		
Sous-catégorie	Nombre d'occurrences	%
Variations terminologiques qui ne sont pas détectées dans la « Réf »	24	9%
Noms propres et acronymes en LC qui sont mal traduits	13	5%
Total	37	14%

Tableau 26 Sous-catégories des termes cibles inexacts dans DeepL Pro avec « glossaire ».

Tout comme dans la présentation des résultats de l'évaluation humaine de la traduction de DeepL Pro sans la fonctionnalité « glossaire » (section 5.1.1), nous commenterons ci-après les différentes sous-catégories avec des exemples concrets.

5.1.2.2.1 Variations terminologiques qui ne sont pas détectées dans la « Réf »

Dans 9% des cas (N=24/260), DeepL Pro a inclus nos termes dans la traduction, mais sous forme de variations terminologiques de différents types qui ne sont pas détectées dans la traduction de référence (Tableau 27).

DeepL Pro avec « glossaire » : Variations terminologiques qui ne sont pas détectées dans la « Réf »		
Types	Nombre d'occurrences	%
Variations morphosyntaxiques ³⁶ (Tableau 28)	11	4%
Variations flexionnelles ³⁵ (Tableau 29)	9	3%
Variations graphiques ³⁴ (Tableau 2)	4	2%
Total	24	9%

Tableau 27 Variations terminologiques dans DeepL Pro avec « glossaire ».

Src	Réf	Syst sans glos	Syst avec glos
Der Bundesrat empfiehlt weiterhin, wo möglich im Home-Office zu arbeiten , auch um Spitzenauslastungen im öffentlichen	The Federal Council continues to recommend that people continue to work from home , not least to avoid overcrowding on public transport.	The Federal Council continues to recommend working in the home office wherever possible, also in order to avoid peak workloads in the public	The Federal Council continues to recommend working from home wherever possible, also in order to avoid peak workloads in the public
Die Covid-19-Pandemie zeigt deutlich, wie wichtig die Prävention und die Kontrolle von Infektionen sind.	The COVID-19 pandemic clearly shows the importance of infection prevention and control .	The Covid 19 pandemic clearly shows the importance of preventing and controlling infections .	The COVID-19 pandemic clearly shows how important the prevention and control of infections is.
Bern, 30.04.2020 - Das Bundesamt für Gesundheit passt seine Kampagne zum Schutz vor dem neuen Coronavirus an und schaltet in der Grundfarbe auf Pink.	Berne, 30.04.2020 - The Federal Office of Public Health is adapting its public awareness campaign to protect against the new coronavirus and is switching the base colour to pink.	Bern, 30.04.2020 - The Federal Office of Public Health is adapting its campaign to protect against the new coronavirus and switching to pink in the basic colour.	Bern, 30.04.2020 - The Federal Office of Public Health is adapting its campaign for protection against the new Coronavirus and switching to pink in the basic colour.
Aufgrund der epidemischen Entwicklung und gestützt auf Empfehlungen der Wissenschaft, lockert der Bundesrat ab dem 27. April die Massnahmen zum Schutz der Bevölkerung vor dem neuen Coronavirus.	In view of the epidemiological trend and based on expert recommendations, from 27 April the Federal Council will begin to ease the measures put in place to protect the public from the new coronavirus.	In view of the epidemic development and based on scientific recommendations, the Federal Council will relax measures to protect the population from the new coronavirus from 27 April.	In view of the epidemiological trend and based on scientific recommendations, the Federal Council is putting in place from 27 April onwards an easing of measures to protect the public from the new Coronavirus.
Deshalb kommt die Maske nur ergänzend zu den Distanz- und Hygieneregeln zum Einsatz.	That is why the mask is only useful in combination with physical distancing and hygienic rules .	Therefore, the mask is only used in addition to the distance and hygiene rules .	Therefore, the mask is only used in addition to the distance and hygiene rules .

Tableau 28 Variations terminologiques de type morphosyntaxique dans DeepL Pro avec « glossaire ».

Src	Réf	Syst sans glos	Syst avec glos
Der Bundesrat wird die Auswirkungen der Lockerungen auf die Entwicklung der Epidemie mit einem Monitoring genau beobachten.	The Federal Council will closely monitor the impact of easing measures on the development of the epidemic.	The Federal Council will closely monitor the impact of the relaxation of the rules on the development of the epidemic.	The Federal Council will closely monitor the impact of the ease measures on the development of the epidemic.
Alle Personen , die mittels eines Schnelltests positiv getestet wurden , sollten sich umgehend in Isolation begeben.	However, anyone who tests positive following a rapid test should go into isolation immediately.	All persons who have tested positive using a rapid test should be placed in isolation immediately.	All persons who have tested positive using a rapid test should go into isolation immediately.
Für die Reihenfolge der Lockerungen hat der Bundesrat mehrere Risikofaktoren berücksichtigt.	The Federal Council has taken into account a range of risk factors in determining the phases for easing the measures .	The Federal Council has taken several risk factors into account for the order of loosening .	The Federal Council has taken several risk factors into account for the order of ease measures .

infizierte Personen sollen frühzeitig entdeckt, behandelt und isoliert , die Übertragungsketten eruiert und weitere Übertragungen verhindert werden.	infected persons should be detected, treated and placed in isolation at an early stage, transmission chains should be identified and further transmissions prevented.	infected persons are to be detected, treated and isolated at an early stage, the transmission chains are to be determined and further transmissions prevented.	The aim is to detect, treat and place in isolation infected persons at an early stage, identify transmission chains and prevent further transmission.
---	--	---	--

Tableau 29 Variations terminologiques de type flexionnel dans DeepL Pro avec « glossaire ».

En qui concerne les **variations flexionnelles**, DeepL Pro a repris la forme canonique des termes enregistrés dans le glossaire dans 5 occurrences de termes sur un total de 9 (Tableau 30).

Src	Réf	Syst sans glos	Syst avec glos
Der Trainingsbetrieb ist für alle Sportarten ab dem 6. Juni ohne Einschränkung der Gruppengröße wieder erlaubt.	Training activities for all sports may resume without any restrictions on group sizes from 6 June.	Training is permitted again for all sports from 6 June onwards without any restrictions on group size.	Training will be allowed again for all sports from 6 June onwards without any restriction on group size.
Die Änderung tritt am 29.10.2020 in Kraft.	The changes come into force on 29 October.	The amendment will enter into force on 29.10.2020.	The change will come into force on 29.10.2020.
Der Zivilschutz übernimmt Aufgaben in verschiedenen Bereichen: Insbesondere kann er das Gesundheitswesen und die Institutionen im Pflegebereich unterstützen, indem er beispielsweise bei der ambulanten Betreuung von Betagten und Pflegebedürftigen mithilft, Mahlzeiten verteilt oder Fahrdienste übernimmt.	Civil defence units will carry out tasks in various areas: in particular, they can support the health service and social care institutions, for example by assisting in the outpatient care of the elderly and those in need of care, distributing meals or providing transport services.	Civil protection takes on tasks in various areas: In particular, it can support the health care system and institutions in the care sector, for example by assisting in the outpatient care of the elderly and people in need of care, distributing meals or providing transport services.	The civile defence unit performs tasks in various areas: In particular, it can support the healthcare and social care institutions, for example by helping with outpatient care for the elderly and those in need of care, distributing meals or providing transport services.
Voraussetzung ist, dass der Aufenthalt im Ausland nicht mehr als fünf Tage dauert und ein Schutzkonzept erstellt und umgesetzt wird.	The persons concerned must not remain abroad for more than five days, and precautionary measures must be planned and taken.	The prerequisite is that the stay abroad does not exceed five days and that a protection concept is drawn up and implemented.	The prerequisite is that the stay abroad does not exceed five days and that a precautionary measure is drawn up and implemented.
Die Pflicht zum Tragen einer Gesichtsmaske gilt in obligatorischen Schulen, Schulen der Sekundarstufe II und der Tertiärstufe, in Institutionen der familienergänzenden Kinderbetreuung sowie in den Trainingsbereichen	Face masks must also be worn in all schools and higher education establishments, in child-care facilities and in the training areas of sport and fitness centres if part of the centre's precautionary measures .	The obligation to wear a face mask applies in compulsory schools, upper secondary and tertiary schools, in institutions providing supplementary childcare and in the training areas of sports and fitness facilities only if it is provided for in	Face masks must be worn in compulsory schools, upper secondary level schools and higher education establishments, in child-care facilities and in the training areas of sports and fitness facilities only if this is provided

von Sport- und Fitnessseinrichtungen nur dann, wenn sie im betreffenden Schutzkonzept vorgesehen ist.	the relevant protection concept .	for in the precautionary measure concerned.
--	--	--

Tableau 30 Forme canonique des termes dans DeepL Pro avec « glossaire.

Dans le cadre des **variations graphiques** (Tableau 31) par rapport à la traduction de référence, nous avons constaté une nette amélioration (2%, soit N=4/260 dans DeepL Pro avec « glossaire » contre 6%, soit N=15/260 dans DeepL Pro sans « glossaire ») par rapport à notre évaluation précédente (section 5.1.1.2.1).

Bien que l'enregistrement des termes dans le « glossaire » de DeepL Pro réduise ce problème (8 variations graphiques dans la TA de DeepL Pro avec « glossaire » contre 10 variations graphiques dans la TA de DeepL Pro sans « glossaire »), nous avons quand même constaté en TA des cas d'incohérence dans la répétition de deux concepts (p.ex. « Covid-19-Impfstoff » a été traduit de deux manières différentes, à savoir « Covid-19 vaccin » et « COVID-19 vaccin », qui sont respectivement apparus trois⁵⁶ fois).

Src	Réf	Syst sans glos	Syst avec glos
Er hat das EJPD damit beauftragt, gemeinsam mit dem EDI und dem EDA die Ausdehnung der Grenzkontrollen und Einreiseverbote auf Reisende aus weiteren, stark von der Ausbreitung des Coronavirus betroffenen Länder zu prüfen.	It has tasked the FDJP to work with the FDHA and the FDFA to consider extending border checks and entry bans to travellers from other countries severely affected by the spread of the coronavirus .	It instructed the FDJP to examine, together with the FDHA and the FDFA, the extension of border controls and entry bans to travellers from other countries severely affected by the spread of the coronavirus .	It instructed the FDJP, together with the FDHA and the FDFA, to examine the extension of border controls and entry bans to travellers from other countries severely affected by the spread of the Coronavirus .
Dieses Verbot ist eine der zentralen Massnahmen, mit der die Zahl der Infektionen mit dem Coronavirus reduziert werden konnte.	The ban is one of the key measures that has helped to reduce the number of coronavirus infections in Switzerland.	This ban is one of the key measures to reduce the number of infections with coronavirus .	This ban is one of the key measures which has made it possible to reduce the number of coronavirus infections .
Bern, 16.10.2020 - Der Bund will der Schweizer Bevölkerung einen raschen Zugang zu COVID-19-Impfstoffen gewährleisten.	Berne, 16.10.2020 - The Swiss federal government wants to assure members of the Swiss public rapid access to COVID-19 vaccines .	Bern, 16.10.2020 - The Swiss Confederation wants to guarantee the Swiss population rapid access to COVID-19 vaccines .	Bern, 16.10.2020 - The Swiss federal government wants to guarantee the Swiss population rapid access to Covid-19 vaccines .

⁵⁶ Phrases n° 12, 66 et 102 pour « Covid-19 vaccine » et phrases n° 40, 88 et 116 pour « COVID-19 vaccine ».

Bund beteiligt sich an COVAX- Initiative zur Beschaffung von COVID-19-Impfstoffen	The Swiss Confederation joins the COVAX initiative for the procurement of COVID-19 vaccines	Federal government participates in COVAX initiative for the procurement of COVID 19 vaccines	Swiss federal government participates in COVAX initiative for the procurement of COVID-19 vaccines
---	---	--	--

Tableau 31 Exemples de variations terminologiques de type graphique dans DeepL Pro avec « glossaire ».

5.1.2.2.2 Noms propres et acronymes en LC qui sont mal traduits

Les erreurs de noms propres et d'acronymes sont plus nombreuses dans la version avec glossaire que sans (5%, soit N=13/260 dans DeepL Pro avec « glossaire » contre 1%, soit N=3/260 dans DeepL Pro sans « glossaire ») (Tableau 32).

Src	Réf	Syst sans glos	Syst avec glos
Den Test auf das SARS-Coronavirus-2 vergütet der Bund mit einer Pauschale von 169, jenen für den Antikörpernachweis mit 113 Franken.	It will reimburse the cost of tests for SARS-Coronavirus-2 at a flat rate of 169 francs, and of antibody tests at a rate of 113 francs.	The Swiss Confederation pays a lump sum of CHF 169 for the test for SARS coronavirus-2 and CHF 113 for the detection of antibodies .	The Swiss federal government will pay a flat-rate fee of CHF 169 for the test for SARS-Coronavirus-2 and CHF 113 for the detection of antibodies .
Identifizierung von Spezialmassnahmen im Bereich der Innovationsförderung mit dem Ziel, aus Schweizer Wissenschafts-Know-how rasch Produkte oder Dienstleistungen hervorbringen und damit wesentlich zur Bekämpfung von Covid-19 beitragen zu können.	To identify special measures that can be introduced to encourage innovation, with the aim of exploiting Swiss scientific know-how to rapidly generate products and services in order to effectively combat COVID-19 .	Identification of special measures in the field of innovation promotion with the aim of rapidly generating products or services from Swiss scientific know-how and thus making a significant contribution to the fight against Covid-19 .	Identify special measures to promote innovation with the aim of rapidly producing products or services from Swiss scientific know-how and thus making a significant contribution to the deal with combat COVID-19 .
Vor diesem Hintergrund haben der Krisenstab des Bundesrates zur Bewältigung der Corona Krise (KSBC) , das Staatssekretariat für Bildung, Forschung und Innovation (SBFI) und das Bundesamt für Gesundheit (BAG) eine wissenschaftliche Science Task Force einberufen.	The Swiss National COVID-19 Science Task Force has therefore been set up by the Federal Council Coronavirus Crisis Unit , the State Secretariat for Education, Research and Innovation SERI and the Federal Office of Public Health FOPH .	Against this background, the Federal Council's Crisis Management Team for the Corona Crisis (KSBC) , the State Secretariat for Education, Research and Innovation (SERI) and the Federal Office of Public Health (FOPH) have convened a scientific science task force.	Against this background, the Federal Council Coronavirus Crisis Unit (KSBC) , the State Secretariat for Education, Research and Innovation (SERI) and the Federal Office of Public Health (FOPH) have convened a Swiss National COVID-19 Science Task Force to manage the crisis.

Tableau 32 Noms propres et acronymes en LC qui sont mal traduits par DeepL Pro avec « glossaire »

Ceci s'explique par différentes raisons, notamment l'impossibilité d'inclure dans le glossaire des termes contenant de symboles spéciaux, tels que les parenthèses et les virgules. Dans cette optique, nous n'avons pas pu signaler à DeepL Pro que l'acronyme allemand « KSBC » n'avait pas d'équivalent cible et, que, pour ne pas être transposé en langue cible, il devait figurer dans une entrée terminologique contenant « Krisenstab des Bundesrates zur Bewältigung der Corona Krise (KSBC) ». Pour cette raison, DeepL Pro l'a traduit tel quel. Nous avons rencontré la même difficulté, lorsque nous avons essayé d'exclure les parenthèses ajoutées par DeepL Pro sans « glossaire » dans les acronymes. En effet, notre but était d'obtenir des résultats équivalents à ceux de la traduction de référence (p.ex. « Federal Office of Public Health FOPH »). Cependant, le traducteur en ligne a intégré ces symboles par effet miroir avec le texte source (p.ex. « Bundesamt für Gesundheit (BAG) »).

Par contre, l'impossibilité du « glossaire » de DeepL Pro d'inclure de termes contenant de symboles spéciaux n'a pas causé de problèmes pour la traduction en langue cible du terme « State Secretariat for Education, Research and Innovation (SERI) », puisque DeepL Pro a parfaitement intégré la virgule en faisant référence au texte source (« Staatssekretariat für Bildung, Forschung und Innovation (SBFI) »).

Concernant les hypothèses faites sur les possibles raisons empêchant une intégration de nos termes, nous retenons que DeepL a parfois confondu certaines dénominations à cause de leur catégorie syntaxique (p.ex. un nom, tel que « Bekämpfung », avec sa forme verbale « bekämpfen »). Pour cette raison, il n'a pas donné la bonne traduction (p.ex. « deal with Covid-19 » au lieu de « combat Covid-19 »). Dans d'autres cas, notre terminologie n'a pas simplement été détectée par DeepL Pro (p.ex. « detection of antibodies »), bien que nous l'ayons incluse dans le « glossaire ».

5.1.3 Bilan de l'évaluation humaine

Après l'intégration de notre collection terminologique, la TA des termes semble avoir enregistré une amélioration de 34% (Tableau 34).

Catégorisation des termes traduits par DeepL Pro avec et sans « glossaire » en deux catégories : termes « exacts » et « inexacts »			
Termes cibles exacts par rapport à la traduction de référence		Termes cibles inexacts par rapport à la traduction de référence	
DeepL Pro sans « glossaire »	DeepL Pro avec « glossaire »	DeepL Pro sans « glossaire »	DeepL Pro avec « glossaire »
52%, soit N=136/260	86%, soit N=223/260	48%, soit N=124/260	14%, soit N=37/260

Tableau 33 Termes « exacts » et « inexacts » dans les TA de DeepL Pro avec et sans « glossaire ».

Cependant, nous avons montré que seulement dans 73% des cas (N=191/260), DeepL Pro avec « glossaire » a intégré parfaitement notre terminologie en TA et cela sans aucune répétition et aucun impact négatif sur les phrases (Tableau 20). En conséquence, l'amélioration effective de la TA des termes après l'ajout du glossaire correspond à 21% (73% dans DeepL Pro avec « glossaire » contre 52% dans DeepL Pro sans « glossaire »).

En ce qui concerne les erreurs terminologiques de DeepL Pro qui déterminent les « termes inexacts », ils varient en fonction de l'intégration du « glossaire » (Tableau 34).

Termes cibles inexacts par rapport à la traduction de référence	
Sous-catégories dans DeepL Pro sans « glossaire »	Sous-catégories dans DeepL Pro avec « glossaire »
Recours en LC à d'autres formes linguistiques (28% soit N=72/260), à savoir : Synonymes (13%, soit N=33/260) Variations terminologiques (15%, soit N=39/260) de type : Morphosyntaxique (7%, soit N=18/260) Graphique (6%, soit N=15/260) Flexionnel (1%, soit N=3/260) Syntaxique faible (1%, soit N=3/260)	Variations terminologiques (9%, soit N=24/260), qui ne sont pas détectées dans la traduction de référence, dont le type est : Morphosyntaxique (4%, soit N=11/260) Graphique (2%, soit N=4/260) Flexionnel (3%, soit N=9/260)
Emploi du langage commun au détriment d'une communication spécialisée (11%, soit N=28/260)	/
Erreurs de sens (5%, soit N=15/260)	/
Termes cibles qui ne sont pas conformes à la réalité suisse (2%, soit N=6/260)	/
Noms propres et acronymes en LC qui sont mal traduits (1%, soit N=3/260)	Noms propres et acronymes en LC qui sont mal traduits (5%, soit N=13/260)

Tableau 34 Catégorisation des termes « inexacts » traduits par DeepL Pro avec et sans « glossaire ».

Après l'intégration de la nouvelle fonctionnalité, nous n'avons plus constaté trois problématiques terminologiques de DeepL Pro sans « glossaire » (emploi du langage commun

vs. spécialisé, erreurs de sens et présence d'une terminologie qui n'était pas conforme à la réalité suisse).

De plus, le recours à d'autres formes linguistiques différentes de celles utilisées dans la traduction de référence est largement diminué (28%, soit N=72/260 dans DeepL Pro sans « glossaire » vs. 9%, soit N=24/260 dans DeepL Pro avec « glossaire »). En particulier, nous n'avons plus détecté des synonymes.

Cependant, nous avons encore remarqué que les termes cibles repris du glossaire sont présents en TA sous forme de variations terminologiques (9%, soit N=24/260), alors que ces formes linguistiques n'apparaissent pas dans la traduction de référence. Bien qu'une phrase puisse être traduite de nombreuses façons et, en conséquence, les variations terminologiques puissent varier, ce phénomène peut représenter une source de dispersion et incohérence terminologique et, alors, nuire à la qualité de la traduction. Par contre, on peut rappeler ici Koehn (2020 : 4), qui considère la traduction comme un processus approximatif, car les traducteurs doivent faire des choix et plusieurs traducteurs font des choix différents. Et cela s'applique également en TA, bien que les différentes traductions soient générées différemment. En conséquence, il est possible qu'une variation d'un terme cible soit pertinente dans le contexte de la traduction de référence, mais lorsqu'elle remplace l'équivalent cible dans la TA peut causer des problèmes morphologiques et syntaxiques (Haque et al., 2019 : 20).

Cette baisse a également concerné les variations terminologiques de type graphique (6%, soit N=15/260 dans DeepL Pro sans « glossaire » contre 2%, soit N=4/260 dans DeepL Pro avec « glossaire ») et, en conséquence, les cas d'incohérence dans la répétition de certains concepts.

En ce qui concerne les noms propres et les acronymes en LC qui ont mal été traduits, nous avons constaté une augmentation de ceux-ci après l'intégration de la nouvelle fonctionnalité du traducteur en ligne (1%, soit N=3/260 dans DeepL Pro sans « glossaire » contre 5%, soit N=13/260 dans DeepL Pro avec « glossaire »).

Pour toutes ces raisons, nous retenons que la problématique de l'emploi d'une terminologie exacte et cohérente dans DeepL Pro n'est pas encore résolue avec l'intégration d'une collection terminologique.

Dans la section suivante (5.2), nous verrons si nos évaluations humaines ont été confirmées par le score BLEU (section 5.2.1) et le Term Hit Rate (section 5.2.2).

5.2 Évaluation automatique

En ce qui concerne l'évaluation automatique des traductions de DeepL Pro réalisées avec et sans « glossaire », nous avons choisi de calculer le score BLEU (section 5.2.1) et le Term Hit Rate (THR) (section 5.2.2). De cette façon, nous avons pu estimer respectivement la qualité globale de la TA et celle des termes y contenus par rapport à la traduction de référence.

5.2.1 BLEU

Le score BLEU calculé sur tout notre sous-corpus en sélectionnant et en désélectionnant l'option « Lowercase » a confirmé que DeepL Pro a été plus performant (+9.83 et +9.67 points respectivement avec et sans l'option « Lowercase »), lorsqu'on a intégré sa nouvelle fonctionnalité « glossaire » (Tableau 35).

Score BLEU de DeepL Pro avec et sans « glossaire » en sélectionnant et en désélectionnant l'option « Lowercase »			
Sélection de l'option « Lowercase »		Désélection de l'option « Lowercase »	
DeepL Pro sans « glossaire »	DeepL Pro avec « glossaire »	DeepL Pro sans « glossaire »	DeepL Pro avec « glossaire »
39.68	49.51	38.41	48.08

Tableau 35 Score BLEU de DeepL Pro avec et sans « glossaire ».

Pour vérifier la fiabilité dudit score, nous avons approfondi ses valeurs au niveau des segments. Par la suite, nous regarderons en détail ceux-ci dans le cadre du calcul effectué en sélectionnant (section 5.2.1.1) et en désélectionnant (section 5.2.1.2) l'option « Lowercase ».

5.2.1.1 BLEU calculé en sélectionnant l'option « Lowercase »

Dans le cadre du score BLEU calculé en sélectionnant l'option « Lowercase », nous avons pu voir que DeepL Pro avec « glossaire » a donné de meilleurs résultats que sans « glossaire » dans 89 phrases sur 128. Sur ces 89 phrases, nous en avons cependant remarqué 6⁵⁷ que, contrairement au score BLEU obtenu, nous avons jugées incorrectes dans le cadre de notre

⁵⁷ Phrases n° 9, 15, 34, 48, 78, 92.

évaluation humaine (section 5.1.2.1.3). Dans l'exemple suivant (Tableau 36), BLEU a privilégié la traduction de DeepL Pro avec « glossaire » contenant notre terminologie, bien que celle-ci n'ait pas été parfaitement intégrée dans la phrase et, en conséquence, a nui à la qualité de la TA.

	BLEU	Phrase 34
Src	-	Der Arbeitgeber kann ein ärztliches Attest verlangen, das aufzeigt, weshalb eine angestellte Person zu einer besonders gefährdeten Personengruppe gehört.
Réf	100.00	The employer may request a medical certificate stating why the employee is considered to be at especially high risk.
Syst sans glos	13.31	The employer may require a medical certificate showing why an employed person belongs to a particularly vulnerable group.
Syst avec glos	19.11	The employer can demand a medical certificate that shows why an employee belongs to a staff at especially high risk employee.

Tableau 36 Exemple n°1 de différence entre notre évaluation humaine et le BLEU avec « Lowercase ».

Dans 13⁵⁸ phrases, les traductions de DeepL Pro réalisées avec et sans « glossaire » ont présenté les mêmes scores. Cependant, deux⁵⁹ phrases traduites par DeepL Pro avec « glossaire » étaient à notre avis de meilleure qualité. À titre d'exemple (Tableau 37), BLEU a attribué le même score aux deux TA, alors que nous avons jugé la traduction réalisée sans l'ajout de notre terminologie de mauvaise qualité, car elle ne contient pas la bonne terminologie.

	BLEU	Phrase 39
Src	-	Coronavirus: Bundesrat lockert weitere Massnahmen ab dem 11. Mai 2020
Réf	100.00	Coronavirus : Federal Council to ease further measures from 11 May
Syst sans glos	58.77	Coronavirus : Federal Council relaxes further measures from 11 May 2020
Syst avec glos	58.77	Coronavirus : Federal Council eases further measures from 11 May 2020

Tableau 37 Exemple n°2 de différence entre notre évaluation humaine et le BLEU avec « Lowercase ».

Enfin, le score BLEU était plus élevé dans la version sans glossaire qu'avec seulement dans 26 phrases. Sur ces 26 phrases, nous en avons identifié 10⁶⁰ que nous avons jugées de qualité inférieure ou égale à celle des phrases fournies par DeepL Pro avec « glossaire ». Dans l'exemple suivant (Tableau 38), BLEU a privilégié la traduction de DeepL Pro sans

⁵⁸ Phrases n° 19, 24, 39, 45, 53, 54, 81, 84, 90, 96, 113, 119 et 127.

⁵⁹ Phrases n° 39 et 53.

⁶⁰ Phrases n° 4, 21, 23, 31, 42, 66, 82, 105, 111 et 112.

« glossaire », alors que nous avons considéré celle réalisée après l'intégration de notre terminologie comme meilleure.

	BLEU	Phrase 105
Src	-	Ziel ist auch, den Anstieg der Fallzahlen so stark zu bremsen, dass die Kantone das Contact Tracing weiterhin konsequent und umfassend sicherstellen können.
Réf	100.00	The intention is also to considerably curb the number of new cases so that the cantons can continue to contract trace effectively.
Syst sans glos	23.82	The aim is also to slow down the increase in the number of cases to such an extent that the cantons can continue to ensure contact tracing consistently and comprehensively.
Syst avec glos	19.79	The aim is also to slow down the increase in the number of cases of people infected with the virus to such an extent that the cantons can continue to ensure contact tracing consistently and comprehensively.

Tableau 38 Exemple n°3 de différence entre notre évaluation humaine et le BLEU avec « Lowercase ».

En conclusion, notre évaluation humaine a été équivalente au score BLEU calculé en sélectionnant l'option « Lowercase » dans 110 phrases sur 128 phrases. Nous avons donc constaté une corrélation entre notre évaluation humaine et le score BLEU de 86%.

5.2.1.2 BLEU calculé en désélectionnant l'option « Lowercase »

Dans le cadre du score BLEU calculé en désélectionnant l'option « Lowercase », nous avons pu voir que DeepL Pro avec « glossaire » a obtenu de meilleurs résultats que sans le « glossaire » dans 90 phrases sur 128. Sur ces 90 phrases, nous avons retrouvé les 6⁶¹ que, dans le cadre du BLEU calculé en sélectionnant l'option « Lowercase » (section 5.2.1.1), nous avons jugées incorrectes. De plus, nous avons constaté qu'en désélectionnant l'option « Lowercase » il existe une corrélation plus élevée entre notre évaluation humaine et l'évaluation automatique. Plus précisément, une phrase traduite après l'ajout du « glossaire », que nous avons considérée comme meilleure pendant l'évaluation humaine, a obtenu un score plus élevé (Tableau 39), alors que ce n'était pas le cas dans notre calcul précédent (section 5.2.1.1).

⁶¹ Phrases n° 9, 15, 34, 48, 78, 92.

	BLEU	Phrase 31
Src	-	Die Covid-19-Pandemie zeigt deutlich, wie wichtig die Prävention und die Kontrolle von Infektionen sind.
Réf	100.00	The COVID - 19 pandemic clearly shows the importance of infection prevention and control .
Syst sans glos	44.02	The Covid 19 pandemic clearly shows the importance of preventing and controlling infections .
Syst avec glos	45.72	The COVID - 19 pandemic clearly shows how important the prevention and control of infections is.

Tableau 39 Exemple n°1 de différence entre notre évaluation humaine et le BLEU sans « Lowercase ».

Dans le contexte de 14⁶² phrases, DeepL Pro avec et sans « glossaire » ont présenté les mêmes résultats.

Les phrases traduites par DeepL Pro sans « glossaire » ont obtenu de meilleurs résultats qu'avec le « glossaire » dans 24 cas. Sur ces 24 phrases, nous en avons cependant identifié 10⁶³ que nous avons jugées incorrectes dans le cadre de notre évaluation humaine (section 5.1.2.1.3). Bien que le nombre des phrases dont le score diffère de notre évaluation humaine coïncide à la valeur correspondante de notre calcul précédent (section 5.2.1.1), les phrases en question ne sont pas les mêmes. Dans l'exemple suivant (Tableau 40), BLEU a privilégié la traduction de DeepL Pro sans « glossaire », alors que nous l'avons jugée inexacte, car elle ne contient pas la bonne terminologie.

	BLEU	Phrase 25
Src	-	Die Versorgung der gesamten Bevölkerung mit Lebensmitteln, Medikamenten und Waren des täglichen Gebrauchs ist sichergestellt, es sind genügend Vorräte angelegt.
Réf	100.00	Provision of the entire population with foodstuffs, medicinal products and essential everyday items is assured ; there are sufficient stocks.
Syst sans glos	19.42	The supply of the entire population with food, medicines and goods for daily use is assured and sufficient stocks have been built up.
Syst avec glos	17.41	The provision of the entire population with food, medicines and goods for daily use is guaranteed , and sufficient stocks have been built up.

Tableau 40 Exemple n°2 de différence entre notre évaluation humaine et le BLEU sans « Lowercase ».

⁶² Phrases n° 113, 19, 24, 39, 45, 54, 58, 72, 81, 84, 90, 96, 119 et 127.

⁶³ Phrases n° 4, 23, 25, 42, 66, 82, 99, 105, 111 et 112.

En conclusion, notre évaluation humaine a été équivalente au score BLEU calculé en désélectionnant l'option « Lowercase » dans 112 phrases sur 128 phrases. Nous avons donc constaté une corrélation entre notre évaluation humaine et le score BLEU de 87,5%.

5.2.2 Term Hit Rate (THR)

Pour calculer le Term Hit Rate (THR), nous avons considéré le nombre de fois que les termes de notre glossaire sont apparus dans la traduction de référence (« Réf »), ainsi que dans les traductions de DeepL Pro réalisées avec (« Syst avec glos ») et sans (« Syst sans glos ») la fonctionnalité « glossaire » (Tableau 41). Il est toutefois important de rappeler que, pour éviter d'attribuer des scores plus élevés aux traductions de DeepL Pro où la fréquence d'un certain terme pouvait être supérieure à celle du même terme figurant dans la traduction de référence (surestimation d'un système de TA), nous avons compté les occurrences des termes de la TA selon le nombre de fois maximum qu'ils sont apparus dans la traduction de référence (section 4.3).

Nombre d'occurrences des termes de notre glossaire dans les trois traductions		
Réf	Syst sans glos	Syst avec glos
342	182	260

Tableau 41 Nombre d'occurrences des termes de notre glossaire dans les trois traductions.

Les résultats obtenus nous ont montré que la traduction de DeepL Pro avec « glossaire » (« Syst avec glos ») contient plus de termes de notre glossaire par rapport à ceux figurant dans la TA faite sans l'intégration de notre terminologie (« Syst sans glos »).

Après cela, nous avons enfin appliqué la formule du Term Hit Rate (THR) (Figure 19) qui nous a donné les résultats suivants (Tableau 42).

Term Hit Rate (THR)			
Syst sans glos		Syst avec glos	
(0,532163743)	53%	(0,760233918)	76%

Tableau 42 Résultats du Term Hit Rate (THR).

En clair, la terminologie de la traduction de DeepL Pro réalisée avec « glossaire » est correcte dans 76% des cas contre 53% de la TA faite sans intégrer notre terminologie. Ces résultats nous ont alors confirmé que la nouvelle fonctionnalité « glossaire » de DeepL Pro a eu un impact positif sur la qualité de la terminologie. De plus, ils nous ont montré qu’il existe une très bonne corrélation entre le Term Hit Rate (THR) et l’évaluation humaine. En effet, dans notre évaluation humaine (section 5.1.3) nous avons estimé que la qualité de la terminologie fournie par DeepL Pro avec et sans « glossaire » était correcte dans respectivement le 73% et 52% des cas.

5.2.3 Bilan de l’évaluation automatique

À la lumière des scores obtenus, nous avons pu constater une corrélation entre nos évaluations humaines et les évaluations automatiques.

En ce qui concerne le BLEU, ce score nous a confirmé que DeepL Pro a été plus performant (+9.83 points avec l’option « Lowercase » et +9.67 points sans l’option « Lowercase »), lorsqu’on a intégré sa nouvelle fonctionnalité « glossaire ».

De plus, nous avons vu qu’il existe une corrélation légèrement plus élevée entre notre évaluation humaine et le score BLEU calculé en désélectionnant l’option « Lowercase » (87,5%, soit N=112/128 contre 86%, soit N=110/128 en sélectionnant l’option « Lowercase ») (Tableau 43).

Corrélation entre évaluation humaine et score BLEU en sélectionnant et en désélectionnant l’option « Lowercase »			
Sélection de l’option « Lowercase »		Désélection de l’option « Lowercase »	
Nombre de phrases évaluées de façon équivalente	%	Nombre de phrases évaluées de façon équivalente	%
110 sur 128	(0,859375) 86%	112 sur 128	(0,8671875) 87,5%

Tableau 43 Corrélation entre notre évaluation humaine et le score BLEU.

À cause de la corrélation imparfaite entre nos évaluations humaines et les évaluations automatiques du score BLEU nous pensons que notre travail mériterait d’être poursuivi, en

intégrant d'autres métriques automatiques, telles que : METEOR (section 2.4.2) qui, contrairement à BLEU, n'identifie pas seulement les n-grammes identiques, mais prend aussi en compte les variantes morphologiques ainsi que les synonymes (Denkowski et Lavie, 2011 : 86 - 87) ; et Translation Edit Rate (TER) (section 2.4.2), qui calcule le nombre minimum d'édérations nécessaires (insertion, suppression, remplacement et déplacement), afin qu'une séquence soit identique à celle dans la traduction de référence (Snover et Dorr, 2006 : 225).

En revanche, les résultats fournis par le calcul du Term Hit Rate (THR) semblent correspondre pleinement à nos évaluations humaines. Plus précisément, nous avons remarqué une très bonne corrélation en ce qui concerne les résultats de DeepL Pro avec (73% à l'issue de notre évaluation humaine contre 76% selon le THR) et sans « glossaire » (52% à l'issue de notre évaluation humaine contre 53% selon le THR).

Toutefois, il serait intéressant de tester cette métrique dans le cadre d'un corpus plus grand, afin de vérifier qu'elle n'inclut pas un terme de la traduction de référence qui n'est pas correctement traduit dans la traduction automatique, mais qui figure erronément dans une autre partie de la TA.

5.3 Conclusions sur l'analyse des résultats

Dans ce chapitre, nous avons approfondi les résultats de nos évaluations. En particulier, nous avons commencé par décrire ceux de l'évaluation humaine effectuée par nous-même (sections 5.1), soit les données obtenues à l'issue de l'analyse des traductions de DeepL Pro réalisées avec (section 5.1.2) et sans « glossaire » (section 5.1.1). Après cela, nous avons fait un bilan de nos évaluations humaines (section 5.1.3). Ensuite, nous nous sommes penchée sur les données objectives (section 5.2), c'est-à-dire le score BLEU (section 5.2.1) et le Term Hit Rate (THR) (section 5.2.2). Tout comme pour les évaluations humaines, nous avons également résumé les résultats de ces métriques automatiques (section 5.2.3). Dans le prochain chapitre (6), nous récapitulerons notre travail et les résultats de notre étude, afin d'apporter les réponses à notre question de recherche principale, ainsi qu'à la série de sous-questions qui en découlent.

6 Conclusions

Dans cette dernière section, nous récapitulerons notre travail et les résultats obtenus afin de répondre à notre question de recherche principale ainsi qu'à la série de sous-questions qui en découlent (section 6.1). Puis nous exposerons les limites de notre étude et proposerons des perspectives pour de futurs travaux (section 6.2).

6.1 Synthèse et questions de recherche

Le but du présent mémoire était de décrire et d'évaluer le traitement de la terminologie en TA. Pour mener à bien notre étude, nous sommes passée par plusieurs étapes nécessaires au bon déroulement de notre démarche.

D'abord, nous avons assemblé un corpus parallèle bilingue (section 4.2.1). Dans un second temps, nous avons choisi, selon des critères bien définis, un sous-corpus représentatif (section 4.2.2) d'où nous avons extrait des candidats-termes. De cette façon, nous avons pu créer un glossaire (section 4.2.3), que nous avons ajouté dans l'interface web du traducteur en ligne (section 4.2.4), pour l'évaluer dans les traductions de DeepL Pro. L'étape suivante a consisté à traduire en anglais avec DeepL Pro notre sous-corpus avec et sans « glossaire » (section 4.2.5). Puis, nous avons procédé à l'évaluation humaine (section 4.2.6) et automatique (section 4.2.7) de la terminologie.

Enfin, les résultats obtenus (chapitre 5) nous permettront de répondre à notre question de recherche principale ainsi qu'à la série de sous-questions qui en découlent.

- Le traducteur en ligne DeepL Pro permet-il d'adapter la terminologie de ses traductions à un domaine spécialisé ?
 - Sous-question 1 : quel est l'impact de la nouvelle fonctionnalité « glossaire » sur la qualité de la traduction ?
 - Sous-question 2 : le « glossaire » permet-il de prendre en compte les variations des termes ?
 - Sous-question 3 : quels sont les types d'erreurs posés par la TA des termes ?

En qui concerne notre question de recherche principale, nous constatons que DeepL Pro ne permet pas encore d'adapter pleinement la terminologie de ses traductions à un domaine spécialisé. Bien que sa nouvelle fonctionnalité « glossaire » nous ait permis d'obtenir des améliorations en matière de terminologie, l'emploi d'une terminologie intégralement exacte et cohérente reste encore une question ouverte (Tableau 44).

Comparaison de l'ensemble des résultats					
Évaluation humaine		Score BLEU		Term Hit Rate (THR)	
Syst sans glos	Syst avec glos	Syst sans glos	Syst avec glos	Syst sans glos	Syst avec glos
136 occurrences de termes exacts sur 260, soit 52%	191 occurrences de termes exacts sur 260, soit 73%	39.68 avec « Lowercase » 38.41 sans « Lowercase »	49.51 avec « Lowercase » 48.08 sans « Lowercase »	(0,532163743) soit 53% de termes exacts	(0,760233918) soit 76% de termes exacts

Tableau 44 Comparaison de l'ensemble des résultats.

Concernant nos premières deux sous-questions, nous avons vu que la nouvelle fonctionnalité « glossaire » améliore la qualité des termes ainsi que la qualité globale la TA (Tableau 44), toutefois elle présente des limites (Tableau 45).

Impacts négatifs de DeepL Pro « avec glossaire » sur la TA et la terminologie		
Catégorie	Nombre d'occurrences	%
Termes cibles exacts, mais qui provoquent des répétitions	21	9%
Termes cibles exacts, mais erreurs de syntaxe et/ou d'orthographe de la phrase	11	4%
Total	32	13%

Tableau 45 Impacts négatifs de DeepL Pro « avec glossaire » sur la TA et la terminologie.

Plus précisément, le « glossaire » a eu dans 4% des cas des impacts négatifs sur la phrase (section 5.1.2.1.3) et, lorsque nous avons donné plusieurs traductions pour un terme source, soit sous forme de synonyme, soit sous forme de variation terminologique, il ne nous a pas permis dans 8% des cas d'alterner les différentes formes linguistiques pour exprimer un même concept (section 5.1.2.1.2). En effet, le traducteur en ligne ne tient compte de ces formes linguistiques que si pour chaque synonyme ou variation terminologique en langue cible, il existe des équivalents correspondants en langue source. Cela représente une limite du

« glossaire » de DeepL Pro, puisqu'en traduction on fait souvent appel aux synonymes et aux variations terminologiques pour éviter des répétitions ou pour préciser le sens d'un terme.

Dans le cadre de notre troisième sous-question, nous avons vu (section 5.1.3) que les problèmes terminologiques de DeepL Pro varient en fonction de l'intégration du « glossaire » (Tableau 46).

Problèmes terminologiques de DeepL Pro avec et sans « glossaire »	
Types d'erreurs dans DeepL Pro sans « glossaire »	Types d'erreurs dans DeepL Pro avec « glossaire »
<p>Recours en LC à d'autres formes linguistiques (28% soit N=72/260), à savoir :</p> <p>Synonymes (13%, soit N=33/260)</p> <p>Variations terminologiques (15%, soit N=39/260) de type :</p> <p>Morphosyntaxique (7%, soit N=18/260)</p> <p>Graphique (6%, soit N=15/260)</p> <p>Flexionnel (1%, soit N=3/260)</p> <p>Syntaxique faible (1%, soit N=3/260)</p>	<p>Variations terminologiques (9%, soit N=24/260), qui ne sont pas détectées dans la traduction de référence, dont le type est :</p> <p>Morphosyntaxique (4%, soit N=11/260)</p> <p>Graphique (2%, soit N=4/260)</p> <p>Flexionnel (3%, soit N=9/260)</p>
Emploi du langage commun au détriment d'une communication spécialisée (11%, soit N=28/260)	/
Erreurs de sens (5%, soit N=15/260)	/
Termes cibles qui ne sont pas conformes à la réalité suisse (2%, soit N=6/260)	/
Noms propres et acronymes en LC qui sont mal traduits (1%, soit N=3/260)	Noms propres et acronymes en LC qui sont mal traduits (5%, soit N=13/260)

Tableau 46 Problèmes terminologiques de DeepL Pro avec et sans « glossaire ».

6.2 Limites et perspectives

Nous devons souligner que notre travail comporte certaines limites.

Nous avons choisi de nous focaliser sur une typologie textuelle et d'en regarder en détail un sous-corpus représentatif uniquement pour des raisons de volume de données à traiter. Il serait judicieux d'élargir les données d'analyse pour voir si les observations tirées de notre expérience s'appliquent aussi à d'autres textes et à différents domaines.

Nous avons évalué l'impact du « glossaire » de DeepL Pro sur la TA des termes en mettant en place une approche terminographique dans la création de la collection terminologique à intégrer dans le traducteur en ligne. Il serait intéressant de tester d'autres méthodes, telles que le recours à des suites de test (Isabelle et al., 2017), l'extraction d'une seule catégorie de termes

(p.ex. tous les termes nominaux) ou l'intégration dans le « glossaire » de DeepL Pro de tous les candidats-termes identifiés par un logiciel d'extraction terminologique.

L'évaluation humaine a été effectuée seulement par une personne, à savoir nous-même. Pour réduire le degré de subjectivité et obtenir des résultats plus représentatifs, il serait pertinent d'inclure un plus grand nombre de juges. De façon similaire, il conviendrait de faire appel aussi à autres métriques automatiques, telles que METEOR, précision et rappel ainsi que Translation Edit Rate (TER) (section 2.4.2). À l'avenir, il sera aussi important de voir si d'autres métriques automatiques seront mises au point pour mesurer la qualité de la TA des termes.

Dans le cadre de notre évaluation humaine, nous avons analysé uniquement les termes. Il serait possible d'évaluer également la qualité globale de la TA, en effectuant par exemple une évaluation fondée sur la tâche qui mesure l'effort de la post-édition pour obtenir une traduction de très bonne qualité, où la terminologie reflète pleinement le domaine spécialisé traité. En faisant cela, on pourrait aussi estimer la productivité des traducteurs et la qualité intrinsèque de la traduction automatique.

Compte tenu de la récente implémentation de la fonctionnalité visant à intégrer la terminologie, nous avons évalué celle-ci seulement dans le contexte de DeepL Pro. Nous pensons donc que cette expérience pourrait être répétée en comparant différents systèmes de TA du même type, tels que DeepL Pro, Google Translate Advanced et Customer Translate. Il serait également intéressant de tester paires de langues différentes ou d'évaluer la même combinaison linguistique en inversant la langue source et la langue cible.

Bien que notre travail n'ait pas donné des résultats généralisés, il mériterait d'être poursuivi en raison des limites que nous venons d'évoquer.

En menant cette étude, nous avons voulu nous pencher sur le traitement de la terminologie en TA, soit une problématique bien connue dans le domaine, mais peu traitée. Nous espérons que notre travail permettra de se familiariser avec cette thématique et qu'elle encouragera la recherche en TA.

7 Bibliographie

ARCAN, M., GIULIANO, C., TURCHI, M., & BUITELAAR, P. (2014a). Identification of Bilingual Terms from Monolingual Documents for Statistical Machine Translation. Dans : Proceedings of the 4th International Workshop on Computational Terminology (Computerm), pp. 22 - 31. <http://dx.doi.org/10.3115/v1/W14-4803>.

ARCAN, M., TURCHI, M., TONELLI, S. & BUITELAAR, P. (2014b). Enhancing Statistical Machine Translation with Bilingual Terminology in a CAT Environment. Dans : Proceedings of the 11th Biennial Conference of the Association for Machine Translation in the Americas (AMTA 2014), pp. 54 - 68. <http://dx.doi.org/10.13140/2.1.1019.8404>.

ARNOLD, DJ. (1994). Machine Translation: An Introductory Guide. Manchester: NCC Blackwell.

BABYCH, B. & Hartley, A. (2003). Improving Machine Translation Quality with Automatic Named Entity Recognition. Dans : Proceedings of the 7th International EAMT workshop on MT and other Language Technology Tools, Improving MT through other Language Technology Tools: Resources and Tools for Building MT. pp. 1 - 8. <https://dl.acm.org/doi/10.5555/1609822.1609823>.

BAHDANAU, D., CHO, K. & BENGIO, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate. Dans : Proceedings of 3rd International Conference on Learning Representations, ICLR 2015. <https://arxiv.org/abs/1409.0473>.

BALDASSARRE, V. & BORGHI, M. (2019). Familiarisation avec la terminologie du domaine de « mobilité et transport » en vue d'accomplir un mandat de traduction fictif de l'OFS (FR/IT). Travail de fin de cours : Univ. Genève.

BALDASSARRE, V. (2016). Terminologie et terminographie : entre théorie et pratique. Master : Univ. Genève. <https://archive-ouverte.unige.ch/unige:88739>.

BANERJEE, S. & LAVIE, A. (2005). METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. Dans : Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization, pp. 65 - 72. <https://www.aclweb.org/anthology/W05-0909>.

BERTOLDI, N., & FEDERICO, M. (2009). Domain Adaptation for Statistical Machine Translation with Monolingual Resources. Dans : Proceedings of the Fourth Workshop on Statistical Machine Translation, pp. 182 - 189. <https://dl.acm.org/doi/10.5555/1626431.1626468>.

BOUAMOR, D., SEMMAR, N. & ZWEIGENBAUM, P. (2012). Identifying bilingual Multi-Word Expressions for Statistical Machine Translation. Dans : Proceedings of the Seventh International Conference on Language Resources and Evaluation, pp. 674 - 679. http://www.lrec-conf.org/proceedings/lrec2012/pdf/886_Paper.pdf.

- BOUILLON, P. & CLAS, A. (1993). La traductique : études et recherches de traduction par ordinateur. Montréal : Presses de l'Université de Montréal.
- BOUILLON, P. (2018a). Les systèmes directs – Cours de traduction automatique 1. Matériel de cours : Univ. Genève.
- BOUILLON, P. (2018b). Les systèmes indirects – Cours de traduction automatique 1. Matériel de cours : Univ. Genève.
- BOUILLON, P., VANDOOREN, F. & LEHMANN, S. (1999). Traitement automatique des langues naturelles. Bruxelles : Duculot.
- BOUKHARI, K. (2018). Indexation Non Supervisée de Documents Textuels : Application aux documents Biomédicaux. Thèse de Doctorat : Univ. Sousse.
- BOWKER, L. & PEARSON, J. (2002). Working with Specialized Language : A Practical Guide to Using Corpora. London [etc: Routledge].
- CABRÉ CASTELLVÍ, M. T. & CORMIER, M. C. (1998). La terminologie : théorie, méthode et applications. Paris : A. Colin ; Ottawa.
- CARL, M., & LANGLAIS, P. (2002). An Intelligent Terminology Database as a Pre-processor for Statistical Machine Translation. Dans : COLING-02 on COMPUTERM 2002: second international workshop on computational terminology-Volume 14, pp. 1 – 7. <https://www.aclweb.org/anthology/W02-1402>.
- CHANCELLERIE FÉDÉRALE. (2012). Les services linguistiques de l'administration fédérale. https://www.bk.admin.ch/dam/bk/fr/dokumente/sprachdienste/flyer_sprachdienste/erv.pdf.download.pdf/brochure_serviceslinguistiquesadmfed.pdf
- CHEN, B., CHERRY, C., FOSTER, G. LARKIN, S. (2017). Cost Weighting for Neural Machine Translation Domain Adaptation. Dans : Proceedings of the First Workshop on Neural Machine Translation, pp. 40 – 46. <https://www.aclweb.org/anthology/W17-3205/>.
- CHINEA-RÍOS, M., PERIS, A. & CASACUBERTA, F. (2017). Adapting Neural Machine Translation with Parallel Synthetic Data. Dans : Proceedings of the Second Conference on Machine Translation, pp. 138 – 147. <https://www.aclweb.org/anthology/W17-4714>.
- CHU, C. & DABRE, R. (2019). Multilingual Multi-Domain Adaptation Approaches for Neural Machine Translation. <http://arxiv.org/abs/1906.07978>.
- CHU, C. & WANG, R. (2018). A Survey of Domain Adaptation for Neural Machine Translation. Dans : Proceedings of the 27th International Conference on Computational Linguistics, pp. 1304 - 1319. <https://www.aclweb.org/anthology/C18-1111>.
- CURREY, A., MICELI BARONE A. V., HEAFIELD, K. (2017). Copied Monolingual Data Improves Low-Resource Neural Machine Translation. <https://www.aclweb.org/anthology/W17-4715>.
- DENKOWSKI, M. & LAVIE, A. (2011). Meteor 1.3: Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems. Dans : Proceedings of the Sixth

Workshop on Statistical Machine Translation, pp. 85 - 91.
<https://www.aclweb.org/anthology/W11-2107>.

DINU, G., MATHUR, P., FEDERICO, M. & AL-ONAIKAN, Y. (2019). Training Neural Machine Translation To Apply Terminology Constraints. Dans : Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 3063 - 3068.
<https://www.aclweb.org/anthology/P19-1294>.

DODDINGTON, G. (2002). Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrence Statistics. Dans : Proceedings of the second international conference on Human Language Technology Research, pp. 138 - 145.
<https://dl.acm.org/doi/pdf/10.5555/1289189.1289273>

FARAJIAN, M., A., BERTOLDI, N., NEGRI, M., TURCHI, M. & FEDERICO, M. (2018). Evaluation of Terminology Translation in Instance-Based Neural MT Adaptation. Dans : Proceedings of the 21st Annual Conference of the European Association for Machine Translation, pp. 149 - 158. https://rua.ua.es/dspace/bitstream/10045/76037/1/EAMT2018-Proceedings_17.pdf.

FORCADA, M. (2010). Machine translation today. Dans : GAMBIER, Y., VAN DOORSLAER, L. Handbook of Translation Studies, 1, pp. 215-223. John Benjamins Publishing Company.

FREITAG, M. and AL-ONAIKAN, Y. (2016). Fast domain adaptation for neural machine translation. <https://arxiv.org/abs/1612.06897>.

GASPARI, F. (2001). Teaching Machine Translation to Trainee Translators: a Survey of Their Knowledge and Opinions. Dans : MT Summit VIII Workshop on Teaching Machine Translation, pp. 35 - 44. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.581.2260>.

GELLY, G. (2017). Réseaux de neurones récurrents pour le traitement automatique de la parole. Thèse de Doctorat : Univ. Paris-Saclay. <https://tel.archives-ouvertes.fr/tel-01615475/document>.

GERLACH, J. 2015. Improving statistical machine translation of informal language: a rule-based preediting approach for French Forums. Thèse de Doctorat : Univ. Genève. <https://archive-ouverte.unige.ch/unige:73226>.

GOODFELLOW, I. I., BENGIO, Y. & COURVILLE, A. (2016). Deep Learning. London, England: The MIT Press.

HAQUE, R., HASANUZZAMAN, M. & WAY, A. (2019). TermEval: An automatic metric for Evaluating Terminology Translation in MT. Dans : Proceedings of 20th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2019). http://doras.dcu.ie/24170/1/CICLing_Term.pdf.

HASLER, E., DEGISPERT, A., IGLESIAS, G. & BYRNE, B. (2018). Neural Machine Translation Decoding with Terminology Constraints. Dans : Proceedings of the 2018

Conference of the North American Chapter of the Association for Computational Linguistics, pp. 506 - 512. <https://arxiv.org/abs/1805.03750>.

HASSAN, H., AUE, A., CHEN, C., CHOWDHARY, V., CLARK, J., FEDERMANN, C., HUANG, X., JUNCZYS-DOWMUNT, M., LEWIS, W., LI, M., LIU, S., LIU, T., LUO, R., MENEZES, A., QIN, T., SEIDE, F., TAN, X., TIAN, F., WU, L., WU, S., XIA, Y., ZHANG, D., ZHANG, Z., and ZHOU, M. (2018). Achieving human parity on automatic Chinese to English news translation. <https://arxiv.org/abs/1803.05567>.

HEARNE, M. & WAY, A. (2011). Statistical Machine Translation: A Guide for Linguists and Translators. *Language and linguistics compass* 5.5, pp. 205–226.

HUCK, M., HANGYA, V., FRASER, A. (2019). Better OOV Translation with Bilingual Terminology Mining. Dans : Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 5809 - 5815. <https://www.aclweb.org/anthology/P19-1581>.

HUTCHINS, W. J. & SOMERS, H. L. (1992). An introduction to machine translation. London: Academic Press.

HUTCHINS, W. J. (1986). Machine Translation: Past, Present, Future. Chichester: E. Horwood.

IMELLI, M. L. (2019). Prestations linguistiques de la Confédération : faire ou faire faire ? Cahier de l'IDHEAP 309/2019. Travail de mémoire : Institut de hautes études en administration publique (IDHEAP), Univ. Lausanne. urn:nbn:ch:serval-BIB_37066E6CF8C84.

ISABELLE, P., CHERRY, C. & FOSTER, G. (2017). A Challenge Set Approach to Evaluating Machine Translation. <https://www.aclweb.org/anthology/D17-1263>.

ISO. (1989). Terminology : vocabulary = Terminologie : vocabulaire. Genève: ISO.

ISO. (2000). Terminology work : vocabulary. Part 1 Partie 1. Theory and application = Travaux terminologiques : vocabulaire. Théorie et application. 1re éd. 2000-10-15. Genève: Organisation internationale de normalisation.

ISO. (2009). Terminology work : principles and methods = Travail terminologique : principes et méthodes. 3rd ed. 2009-11-01. Geneva: ISO.

JURAFSKY, D. & MARTIN, J.H. (2014). Speech and Language Processing : an Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Second edition, Indian subcontinent edition. Harlow: Pearson Education.

KAPLAN, A. (1951). Yehoshua Bar-Hillel. The present state of research on mechanical translation. *Journal of the Association for Information Science and Technology, American Documentation* (Cambridge, Mass.), Vol. 2 (4), pp. 229 - 237.

KHAYRALLAH, H., KUMAR, G., DUH, K., POST, M., & KOEHN, P. (2017). Neural Lattice Search for Domain Adaptation in Machine Translation. Dans : Proceedings of the Eighth

International Joint Conference on Natural Language Processing (Volume 2: Short Papers, pp. 20 - 25. <https://www.aclweb.org/anthology/I17-2004>.

KIT, C. & WONG, T.M. (2015). Evaluation in machine translation and computer-aided translation. Dans : CHAN, S.W. (2015). The Routledge encyclopedia of translation technology pp. 251 - 274. London : Routledge.

KOBUS, C., CREGO, J. & SENELLART, J. Domain Control for Neural Machine Translation. Dans : Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017, pp. 372 - 378. <https://arxiv.org/pdf/1612.06140>

KOEHN, P. & KNOWLES, R. (2017). Six Challenges for Neural Machine Translation. Dans : Proceedings of the First Workshop on Neural Machine Translation, pp. 28 - 39. www.aclweb.org/anthology/W17-3204/.

KOEHN, P. & SCHROEDER, J. (2007). Experiments in Domain Adaptation for Statistical Machine Translation. Dans : Proceedings of the Second Workshop on Statistical Machine Translation, pp. 224 - 227. <https://www.aclweb.org/anthology/W07-0733>.

KOEHN, P. (2004). Statistical Significance Tests for Machine Translation Evaluation. Dans : Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, pp. 388 - 395. <https://www.aclweb.org/anthology/W04-3250>.

KOEHN, P. (2010). Statistical machine translation. Cambridge: Cambridge University Press.

KOEHN, P. (2017). Neural Machine Translation. <https://arxiv.org/abs/1709.07809>

KOEHN, P. (2020). Neural Machine Translation. Cambridge: Cambridge University Press

L'HOMME, M. (2008). Initiation à la traductique. Montréal : Lingatech.

L'HOMME, M. (2020). La terminologie : principes et techniques. Montréal : Presses de l'Université de Montréal.

LUONG, MT. & MANNING, CD. (2015). Stanford neural machine translation systems for spoken language domains. Dans : Proceedings of the International Workshop on Spoken Language Translation, pp. 76 - 79. <https://nlp.stanford.edu/pubs/luong-manning-iwslt15.pdf>.

MARIEN, B. & BEUD, J. (2003). Guide pratique pour l'utilisation de la statistique en recherche : le cas des petits échantillons. Agence Universitaire de la Francophonie.

MINSKY, M. L. (1969). Perceptrons : an introduction to computational geometry. Cambridge Mass etc : The MIT Press.

PALMER, F. R. (1990). Modality and the English Modals. London: Longman.

PAPINENI, K., ROUKOS, S., WARD, T. & ZHU, W. J. (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. <https://dl.acm.org/doi/10.3115/1073083.1073135>.

PINNIS, M. (2015). Terminology Integration in Statistical Machine Translation. Thèse de Doctorat : Univ. Latvia. <http://dx.doi.org/10.13140/RG.2.1.3385.7687>.

- POIBEAU, T. (2019). Babel 2.0 : où va la traduction automatique ? Paris: Odile Jacob, 2019.
- PROCHASSON, E. (2009). Alignement multilingue en corpus comparables spécialisés. Thèse de doctorat : Univ. Nantes. <https://tel.archives-ouvertes.fr/tel-00462248>.
- PRZYBOCKI, M., SANDERS, G. LE, E. (2006). Edit Distance: A Metric for Machine Translation Evaluation. <http://mt-archive.info/LREC-2006-Przybocki.pdf>.
- QUAH, C. K. (2006). Translation and Technology. Basingstoke: Palgrave.
- ROCHARD, M. (2017). Du papier au numérique, traduction et mutations technologiques. Dans FORUM. Revue internationale d'interprétation et de traduction / International Journal of Interpretation and Translation, vol. 15 (2), pp. 212 - 227.
- SCANSANI, R., BENTIVOGLI, L., BERNARDINI, S. & FERRARESI, A. (2019). MAGMATic: A Multi-domain Academic Gold Standard with Manual Annotation of Terminology for Machine Translation Evaluation. Dans : Proceedings of Machine Translation Summit XVII Volume 1: Research Track. <https://www.aclweb.org/anthology/W19-6608>.
- SCHÜTZ, J. (2003). HUTCHINS, W. J. (ed.), Early Years in Machine Translation: Memoirs and Biographies of Pioneers, Vol. 97 Amsterdam Studies in the Theory and History of Linguistic Science. Machine Translation, vol. 18 (1), pp. 77 - 79.
- SEHELLART, J., DIENES, P., VÁRADI, T. (2001). New Generation Systran Translation System. Dans MT Summit VIII. <http://www.mt-archive.info/MTS-2001-Senellart-1.pdf>
- SENNRICH, R., HADDOW, B. & BIRCH, A. (2016). Improving Neural Machine Translation Models with Monolingual Data. Dans : Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 86 - 96. <https://www.aclweb.org/anthology/P16-1009>.
- SHTERIONOV, D., NAGLE, P., CASANELLAS, L., et al. (2017). Empirical evaluation of NMT and PBSMT quality for large-scale translation production. Dans : Proceedings of the 20th Annual Conference of the European Association for Machine Translation.), pp. 74 - 79
- SMITH, R. 2001. Machine translation: potential for progress. English today 17.4, pp. 38 - 44
- SNOVER, M. & DORR, B. (2006). A Study of Translation Edit Rate with Targeted Human Annotation. Dans : Proceedings of Association for Machine Translation in the Americas, pp. 223 - 231. https://www.cs.umd.edu/~snover/pub/amta06/ter_amta.pdf
- SNOVER, M., MADNANI, N., DORR, B. & SCHWARTZ, R. (2009). Fluency, adequacy, or HTER?: exploring different human judgments with a tunable MT metric. Dans : Proceedings of the Fourth Workshop on Statistical Machine Translation. Association for Computational Linguistics, pp. 259 -268. <https://www.aclweb.org/anthology/W09-0441>.
- SUNDETOVA, A., KARIBAYEYA, A & TUKEYEV, U. (2014). Structural Transfer Rules for Kazakh-to-English Machine Translation. <https://dergipark.org.tr/en/download/article-file/395223>

THURMAIR, G. (2009). Comparing different architectures of hybrid Machine Translation systems. Dans : Proceedings of the MT Summit XII, pp. 340 - 347. <http://www.mt-archive.info/MTS-2009-Thurmair.pdf>.

VAN DER WEES, M., BISAZZA, A., MONZ, C. (2017). Dynamic data selection for neural machine translation. Dans : Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp. 1400 – 1410. <https://arxiv.org/pdf/1708.00712.pdf>.

, M., AVEY, B., GDANIEC, C., GERBER, L. LEÓN, M., MITAMURA, T. (2001). TERMINOLOGY AND MACHINE TRANSLATION. In WRIGHT, S.E., & BUDIN, G. (2001). Handbook for Terminology Management, Amsterdam/Philadelphia : John Benjamins, Vol. 2: Application-Oriented Terminology Management, pp. 697 - 723.

VAUQUOIS, B. (2003). Automatic Translation - A Survey of Different Approaches. In NIRENBURG, S., SOMERS, H. & WILKS, Y. Readings in Machine Translation. Cambridge Mass etc : The MIT Press.

VOLKART, L. (2018). Traduction automatique statistique vs. neuronale : Comparaison de MTH et DeepL à La Poste Suisse. Master: Univ. Genève. <https://archive-ouverte.unige.ch/unige:113749>.

WANG, R., ZHAO, H., LU, B., UTIYAMA, M. & SUMITA, E. (2016). Connecting Phrase based Statistical Machine Translation Adaptation. Dans : Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers, pp. 3135 - 3145. <https://www.aclweb.org/anthology/C16-1295>.

YVON, F. (2014). Machine Translation. Chapter in Oxford Bibliographies in Linguistics (Mark Aronoff, ed.), New York: Oxford University Press, 2014.

8 Webographie

- 01net.https://www.01net.com/telecharger/windows/Programmation/base_de_donne/fiches/131618.html. Date de la dernière consultation : 27 novembre 2020.
- ANTCONC. <https://www.laurenceanthony.net/software.html>. Date de la dernière consultation : 8 novembre 2020.
- BODENSEO. <https://www.bodenseo.com/>. Date de la dernière consultation : 8 novembre 2020.
- CHANCELLERIE FÉDÉRALE (2019). L'administration fédérale achète un logiciel de traduction automatique. <https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-77610.html>. Date de la dernière consultation : 8 novembre 2020.
- CHANCELLERIE FÉDÉRALE. (2020). Conférence interdépartementale des services linguistiques (CISL). [https://www.bk.admin.ch/bk/fr/home/chancellerie-federale/organisation-de-la-chancellerie-federale/organisations-interdepartementales/Konferenz-Sprachdienste-Bundesverwaltung%20\(KOSD\).html](https://www.bk.admin.ch/bk/fr/home/chancellerie-federale/organisation-de-la-chancellerie-federale/organisations-interdepartementales/Konferenz-Sprachdienste-Bundesverwaltung%20(KOSD).html). Date de la dernière consultation : 7 novembre 2020.
- CONFÉRATION SUISSE. (2019). Bericht DeepL-Test Auswertung der Testergebnisse und Empfehlungen der Arbeitsgruppe "Maschinelle Übersetzung". <https://www.newsd.admin.ch/newsd/message/attachments/59735.pdf>. Date de la dernière consultation : 7 novembre 2020.
- CPSL. Entreprise de traduction et solutions multilingues - CPSL. <https://www.cpsl.com/fr/>. Date de la dernière consultation : 8 décembre 2020.
- DEEPL. (2020). Personnalisez DeepL Traducteur avec le glossaire. <https://www.deepl.com/blog/20200506.html>. Date de la dernière consultation : 7 novembre 2020.
- DEEPL. <https://www.deepl.com/press.html> Date de la dernière consultation : 7 novembre 2020.
- EDUTECH WIKI. (2020a). AntConc. <http://edutechwiki.unige.ch/fr/AntConc>. Date de la dernière consultation : 7 novembre 2020.
- EDUTECH WIKI. (2020b). Les boîtes à moustache. https://edutechwiki.unige.ch/fr/Les_bo%C3%AEtes_%C3%A0_moustache#:~:text=Afin%20d%27interpr%C3%A9ter%20une%20bo%C3%AEte,asym%C3%A9trie%20et%20les%20valeurs%20atypiques. Date de la dernière consultation : 28 novembre 2020.
- GOOGLE CLOUD. (2020). Créer et utiliser des glossaires (version Advanced). <https://cloud.google.com/translate/docs/advanced/glossary>. Date de la dernière consultation : 4 décembre 2020.
- GUERRERO, L. (2020). NMT plus a Bilingual Glossary: Does this Really Improve Terminology Accuracy and Consistency?. <https://asling.org/tc42/wp-content/uploads/TC42->

[20Nov-NMT-plus-a-bilingual-glossary.pdf](#). Date de la dernière consultation : 4 décembre 2020.

HUTCHINS, W. J. (2004). Two precursors of machine translation: Artsrouni and Trojanskij. <http://www.hutchinsweb.me.uk/IJT-2004.pdf>. Date de la dernière consultation : 7 novembre 2020.

HUTCHINS, W. J. (2005). The history of machine translation in a nutshell, <http://www.hutchinsweb.me.uk/Nutshell-2005.pdf>. Date de la dernière consultation : 7 novembre 2020.

LE TEMPS. (2020). Les étapes de la propagation du coronavirus dans le monde. <https://www.letemps.ch/monde/etapes-propagation-coronavirus-monde>. Date de la dernière consultation : 8 novembre 2020.

LinkedIn. <https://www.linkedin.com/>. Date de la dernière consultation : 22 décembre 2020.

MERRIAM-WEBSTER. Vaccination. Vaccination | Definition of Vaccination by Merriam-Webster. <https://www.merriam-webster.com/dictionary/vaccination>. Date de la dernière consultation : 17 décembre 2020.

MERRIAM-WEBSTER. Vaccine. Vaccine | Definition of Vaccine by Merriam-Webster. <https://www.merriam-webster.com/dictionary/vaccine>. Date de la dernière consultation : 17 décembre 2020.

MOSLEM, Y. (2020). NMT Domain Adaption Techniques. <https://blog.machinetranslation.io/wp-content/uploads/2020/10/DomainAdaptation.pdf>. Date de la dernière consultation : 8 novembre 2020.

OFFICE FÉDÉRAL DE LA SANTÉ PUBLIQUE. (2020). <https://www.bag.admin.ch/bag/fr/home/das-bag/publikationen.html>. Date de la dernière consultation : 8 novembre 2020.

OFFICE FÉDÉRAL DE LA SANTÉ PUBLIQUE. <https://www.bag.admin.ch/bag/fr/home/das-bag/auftrag-ziele.html>. Date de la dernière consultation : 8 novembre 2020.

RIDIEGER, H. (2020). Coronavirus e traduzione: persi tra terminologie, lessici e aspetti culturali. <https://blog.zhaw.ch/language-matters/2020/06/17/coronavirus-e-traduzione/>. Date de la dernière consultation : 7 novembre 2020.

SDL TRADOS. Qu'est-ce qu'une mémoire de traduction. <https://www.sdltrados.com/fr/solutions/translation-memory/>. Date de la dernière consultation : 7 novembre 2020.

STEEDMAN, M. (2007). Computational Linguistics. <https://plato.stanford.edu/entries/computational-linguistics/>. Date de la dernière consultation : 7 novembre 2020.

TAUS. (2017). Josep Bonet: The Constant Factor is Change. <https://blog.taus.net/josep-bonet-constant-factor-is-change>. Date de la dernière consultation : 19 novembre 2020.

TERMDAT. Banque de données terminologiques de l'administration fédérale. www.termdat.bk.admin.ch. Date de la dernière consultation : 17 décembre 2020.

TERMINOTIX. (2020a). AlignFactory Desktop, <https://terminotix.com/index.asp?content=item&item=49&lang=fr>. Date de la dernière consultation : 8 novembre 2020.

TERMINOTIX. (2020b). SynchroTerm : logiciel de dépouillement terminologique bilingue, <https://terminotix.com/index.asp?content=item&item=7&lang=fr#youtube>. Date de la dernière consultation : 8 novembre 2020.

TERMOSTAT. termostat.ling.umontreal.ca. Date de la dernière consultation : 8 décembre 2020.

THE ECONOMIST. (2020). Speeding up science during the pandemic. <https://www.economist.com/leaders/2020/05/09/speeding-up-science-during-the-pandemic>. Date de la dernière consultation : 8 novembre 2020.

TILDE MT. <https://www.letsmt.eu/>. Date de la dernière consultation : 8 décembre 2020.

TURCHI, M. <http://marcoturchi.com/index.html>. Date de la dernière consultation : 8 décembre 2020.

UNIVERSITÉ DE GEVÈVE. (2019). Entretien : la parole à...Pierrette Bouillon. <https://www.unige.ch/fti/ebulletin/archives/avril2019/entretien/entretien-1>. Date de la dernière consultation : 19 novembre 2020.

VIGNAUX, B. (2020) https://www.liglab.fr/files/qa_traduction_auto_bd.pdf. Date de la dernière consultation : 30 octobre 2020.

WIKIPEDIA. Expression régulière. https://fr.wikipedia.org/wiki/Expression_r%C3%A9guli%C3%A8re. Date de la dernière consultation : 14 novembre 2020.

WIKIPEDIA. Marge d'erreur. https://fr.wikipedia.org/wiki/Marge_d%27erreur. Date de la dernière consultation : 14 novembre 2020.

WIKIPEDIA. R (langage). [https://fr.wikipedia.org/wiki/R_\(langage\)](https://fr.wikipedia.org/wiki/R_(langage)). Date de la dernière consultation : 27 novembre 2020.

9 Annexes

9.1 Composition des candidats-termes dans la partie anglaise du corpus

Nombre de termes : 1042			
Matrice	Fréquence	%	Exemples
Adjectif Nom	248	24 %	precautionary measure, high risk
Nom Nom	182	17 %	contact tracing
Nom	181	17 %	measure, vaccine
Verbe	82	8 %	allow, wear, decide, ease
Adjectif	62	6 %	high, medical
Adjectif Nom Nom	56	5 %	exposed hospital staff
Nom Préposition Nom	45	4 %	risk of infection
Adjectif Adjectif Nom	32	3 %	local epidemiological situation
Adjectif Coord_Conjunction Adjectif Nom	27	3 %	private and public institution
Adjectif Nom Préposition Nom	18	2 %	high risk of infection
Nom Préposition Adjectif Nom	17	2 %	introduction of rapid testing
Adverbe	16	2 %	also, as, now, therefore
Nom Nom Nom	16	2 %	task force expert
Nom Préposition Nom Nom	14	1 %	cost of coronavirus test
Adjectif Nom Préposition Adjectif Nom	7	1 %	viral penetration of human cell
Nom Préposition Adjectif Nom Nom	5	0.5 %	basis of continuous assessment
Nom Adjectif Nom	5	0.5 %	outpatient medical practice
Adjectif Nom Préposition Nom Nom	4	0.4 %	consistent tracing of infection chain
Adjectif Adjectif Nom Nom	4	0.4 %	scheduled public transport service
Nom Nom Nom Nom	3	0.3 %	food home delivery service
Nom Préposition Nom Préposition Nom	2	0.2 %	number of case of infection
Adjectif Nom Préposition Adjectif Nom Nom	2	0.2 %	exact number of armed force personnel
Adjectif Nom Nom Nom	2	0.2 %	federal government sign agreement
Nom Nom Nom Préposition Nom	1	0.1 %	science task force of researcher
Nom Nom Préposition Nom	1	0.1 %	training area of sport
Adjectif Adjectif Adjectif Nom	1	0.1 %	swiss public rapid access
Adjectif Adjectif Nom Nom Nom	1	0.1 %	swiss federal government sign agreement
Nom Nom Préposition Nom Nom	1	0.1 %	sale area of grocery product
Adjectif Nom Adjectif Nom Nom	1	0.1 %	continuous assessment cantonal baccalaureate school
Nom Préposition Nom Nom Nom	1	0.1 %	question of child day-care facility
Adjectif Adjectif Nom Préposition Nom Nom	1	0.1 %	continuing low number of coronavirus case
Nom Nom Préposition Adjectif Nom	1	0.1 %	model set of precautionary measure
Nom Préposition Adjectif Adjectif Nom	1	0.1 %	supply of vital medical product
Nom Nom Adjectif Nom	1	0.1 %	health minister take place
Adjectif Nom Préposition Adjectif Adjectif Nom	1	0.1 %	sufficient supply of important medical product

Tableau 47 Composition des candidats-termes extraits par TermoStat dans la partie anglaise du corpus.

9.2 Composition des candidats-termes dans les sept sous-ensembles du corpus

Nombre de termes du sous-ensemble n° 1 : 231			
Matrice	Fréquence	%	Example
Adjectif Nom	62	27 %	large-scale event
Nom	55	24 %	measure, people, event,
Nom Nom	33	14 %	classroom teaching
Verbe	17	7 %	take, ease, continue, protect
Adjectif	15	6 %	public, medical, private
Nom Préposition Nom	11	5 %	number of people
Adjectif Nom Nom	8	3 %	public awareness campaign
Adjectif Adjectif Nom	5	2 %	accessible indoor space
Adjectif Nom Préposition Nom	5	2 %	significant movement of people
Adjectif Coord_Conjunction Adjectif Nom	4	2 %	public and private event
Nom Préposition Adjectif Nom	4	2 %	series of national measure
Nom Nom Nom	2	1 %	mountain transport system
Nom Adjectif Nom	2	1 %	inpatient medical sector
Adverbe	2	1 %	publicly, gradually
Adjectif Nom Préposition Adjectif Nom Nom	1	0.4 %	exact number of armed force personnel
Nom Préposition Adjectif Nom Nom	1	0.4 %	manufacturer of promising vaccine candidate
Nom Préposition Adjectif Adjectif Nom	1	0.4 %	supply of vital medical product
Adjectif Nom Préposition Adjectif Nom	1	0.4 %	low level of direct contact
Adjectif Adjectif Nom Nom	1	0.4 %	accessible public transport area
Nom Préposition Nom Nom	1	0.4 %	list of risk area

Tableau 48 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°1.

Nombre de termes du sous-ensemble n° 2 : 227			
Matrice	Fréquence	%	Exemples
Nom	53	23 %	people, measure, coronavirus
Nom Nom	46	20 %	border region, task force
Adjectif Nom	45	20 %	high risk, rapid test
Verbe	21	9 %	take, reopen, allow, permit
Nom Préposition Nom	13	6 %	number of case
Adjectif Nom Nom	9	4 %	robust healthcare system
Adjectif	7	3 %	precautionary, medical, strict
Adjectif Coord_Conjunction Adjectif Nom	6	3 %	professional or medicalreason,
Nom Préposition Adjectif Nom	5	2 %	set of precautionary measure
Adjectif Adjectif Nom	5	2 %	local epidemiological situation
Adverbe	4	2 %	also, again, gradually
Adjectif Nom Préposition Nom	4	2 %	maximum number of people
Nom Nom Nom	3	1 %	task force expert
Nom Préposition Nom Nom	3	1 %	cost of coronavirus test
Adjectif Nom Préposition Nom Nom	2	1 %	serious case of coronavirus disease
Nom Préposition Adjectif Nom Nom	1	0.4 %	basis of continuous assessment grade

Tableau 49 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°2.

Nombre de termes du sous-ensemble n° 3 : 223			
Matrice	Fréquence	%	Exemples
Adjectif Nom	65	29 %	social distancing
Nom	46	21 %	measure, canton, coronavirus
Nom Nom	41	18 %	contact tracing
Verbe	17	8 %	continue, reopen, apply
Adjectif	13	6 %	able, medical
Adjectif Nom Nom	11	5 %	first procurement programme
Adjectif Adjectif Nom	8	4 %	important protective measure
Nom Préposition Nom	5	2 %	dose of vaccine, result of people, list of country
Nom Préposition Adjectif Nom	4	2 %	effort of countless expert, introduction of rapid testing
Adjectif Coord_Conjunction Adjectif Nom	3	1 %	secondary and vocational school
Nom Nom Nom	3	1 %	contact tracing capacity
Nom Préposition Nom Nom	2	1 %	mean of attendance list
Adverbe	2	1 %	also, responsibly
Adjectif Nom Nom Nom	1	0.4 %	existing product range restriction
Nom Nom Nom Préposition Nom	1	0.4 %	science task force of researcher
Nom Adjectif Nom	1	0.4 %	outpatient medical practice

Tableau 50 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°3.

Nombre de termes du sous-ensemble n° 4 : 211			
Matrice	Fréquence	%	Exemples
Adjectif Nom	51	24 %	federal government, high risk
Nom	40	19 %	measure, contact, event
Nom Nom	37	18 %	protection concept
Verbe	16	8 %	decide, protect, ensure
Adjectif Nom Nom	15	7 %	swiss health service
Nom Préposition Nom	13	6 %	number of infection
Adjectif	12	6 %	public, federal, swiss
Nom Nom Nom	6	3 %	innovation funding instrument
Adjectif Coord_Conjunction Adjectif Nom	5	2 %	rapid and secure access
Adjectif Adjectif Nom	4	2 %	scientific advisory body
Nom Préposition Adjectif Nom	3	1 %	list of high-risk country
Adverbe	2	1 %	also, currently
Adjectif Nom Préposition Nom	2	1 %	reciprocal right of entry
Nom Nom Nom Nom	1	0.5 %	year baccalaureate examination grade
Nom Préposition Nom Nom	1	0.5 %	report of test result
Adjectif Adjectif Nom Nom	1	0.5 %	scheduled public transport service
Adjectif Nom Préposition Adjectif Nom	1	0.5 %	accessible area of administrative building
Adjectif Nom Nom Nom	1	0.5 %	federal government sign agreement

Tableau 51 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°4.

Nombre de termes du sous-ensemble n° 5 : 211			
Matrice	Fréquence	%	Exemples
Adjectif Nom	55	26 %	precautionary measure
Nom Nom	48	23 %	vaccine dose, face mask
Nom	38	18 %	measure, vaccine, country
Verbe	18	9 %	include, wear, enter, infect
Nom Préposition Nom	13	6 %	number of case
Adjectif	12	6 %	swiss, federal, necessary
Adjectif Nom Nom	9	4 %	rapid antigen test
Adjectif Coord_Conjunction Adjectif Nom	4	2 %	safe and effective vaccine
Nom Préposition Adjectif Nom	3	1 %	set of precautionary measure
Adjectif Adjectif Nom	2	1 %	swiss scientific community
Nom Adjectif Nom	2	1 %	contact take place
Adjectif Nom Préposition Adjectif Nom	2	1 %	standard set of precautionary measure
Adverbe	2	1 %	also, medically
Nom Préposition Nom Nom	1	0.5 %	mixture of antibody therapy
Adjectif Nom Préposition Nom	1	0.5 %	gradual relaxation of measure

Tableau 52 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°5.

Nombre de termes du sous-ensemble n° 6 : 251			
Matrice	Fréquence	%	Exemple
Adjectif Nom	55	22 %	precautionary measure
Nom Nom	52	21 %	coronavirus case
Nom	47	19 %	measure, vaccine, number,
Verbe	20	8 %	remain, protect, follow
Adjectif Nom Nom	12	5 %	civil society demonstration
Adjectif	12	5 %	swiss, federal, scientific
Nom Préposition Nom	11	4 %	number of coronavirus
Adjectif Adjectif Nom	10	4 %	swiss federal government
Nom Préposition Nom Nom	6	2 %	range of risk factor
Adjectif Coord_Conjunction Adjectif Nom	4	2 %	political and civil society
Adverbe	3	1 %	also, therefore, effectively
Adjectif Nom Préposition Nom	3	1 %	first phase of measure
Nom Préposition Adjectif Nom Nom	2	1 %	child of compulsory school age
Nom Nom Nom	2	1 %	hospital occupancy rate
Nom Nom Préposition Adjectif Nom	1	0.4 %	model set of precautionary measure
Adjectif Nom Préposition Adjectif Nom	1	0.4 %	unequal treatment of large retailer
Nom Adjectif Nom Nom	1	0.4 %	evaluation civil defence unit
Adjectif Adjectif Nom Préposition Nom Nom	1	0.4 %	continuing low number of coronavirus case
Nom Nom Préposition Nom	1	0.4 %	training area of sport
Nom Préposition Nom Préposition Nom	1	0.4 %	number of case of people
Adjectif Adjectif Nom Nom Nom	1	0.4 %	swiss federal government sign agreement
Nom Adjectif Nom	1	0.4 %	corona swiss diplomat
Adjectif Nom Préposition Adjectif Adjectif Nom	1	0.4 %	sufficient supply of important medical product
Nom Préposition Adjectif Nom	1	0.4 %	set of precautionary measure
Adjectif Nom Préposition Nom Nom	1	0.4 %	consistent tracing of infection chain
Adjectif Adjectif Adjectif Nom	1	0.4 %	swiss public rapid access

Tableau 53 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°6.

Nombre de termes du sous-ensemble n° 7 : 229			
Matrice	Fréquence	%	Exemple
Adjectif Nom	76	33 %	federal government
Nom	45	20 %	measure, people, vaccine
Nom Nom	30	13 %	vaccine dose
Verbe	20	9 %	take, wear, apply, decide
Nom Préposition Nom	8	3 %	number of case
Adjectif	8	3 %	public, swiss, medical
Adjectif Nom Nom	8	3 %	actual vaccination strategy
Adjectif Adjectif Nom	7	3 %	continuing low level, cantonal
Nom Nom Nom	4	2 %	child day-care facility
Adjectif Nom Préposition Nom	3	1 %	high risk of infection
Adjectif Coord_Conjunction Adjectif Nom	3	1 %	public or private event
Adverbe	3	1 %	also, particularly, accordingly
Nom Préposition Adjectif Nom	3	1 %	set of precautionary measure
Adjectif Nom Préposition Adjectif Nom	2	1 %	corresponding set of precautionary measure
Adjectif Adjectif Nom Nom	2	1 %	classic online ad material
Nom Préposition Nom Préposition Nom	1	0.4 %	number of case of infection
Adjectif Nom Préposition Nom Nom	1	0.4 %	sound system of contract tracing
Nom Préposition Adjectif Nom Nom	1	0.4 %	conduct of civil defence assignment
Nom Préposition Nom Nom	1	0.4 %	member of conscript formation
Nom Nom Adjectif Nom	1	0.4 %	health minister take place
Nom Nom Nom Nom	1	0.4 %	food home delivery service
Adjectif Nom Préposition Adjectif Nom Nom	1	0.4 %	specific task of civil defence unit

Tableau 54 Composition des candidats-termes extraits par TermoStat dans le sous-ensemble n°7

9.3 Légende des matrices syntaxiques du graphique à colonne (section 9.4) et de la boîte à moustache (section 9.5)

Acronyme	Matrice
A	Adjectif
A*	Adverbe
AAAN	Adjectif Adjectif Adjectif Nom
AAN	Adjectif Adjectif Nom
AANN	Adjectif Adjectif Nom Nom
AANNN	Adjectif Adjectif Nom Nom Nom
AANPNN	Adjectif Adjectif Nom Préposition Nom Nom
ACAN	Adjectif Coord_Conjunction Adjectif Nom
ADN	Nom Adjectif Nom
AN	Adjectif Nom
ANANN	Adjectif Nom Adjectif Nom Nom
ANN	Adjectif Nom Nom
ANNN	Adjectif Nom Nom Nom
ANPAAN	Nom Préposition Adjectif Adjectif Nom
ANPAAN	Adjectif Nom Préposition Adjectif Adjectif Nom
ANPAN	Adjectif Nom Préposition Adjectif Nom
ANPANN	Adjectif Nom Préposition Adjectif Nom Nom
ANPN	Adjectif Nom Préposition Nom
ANPNN	Adjectif Nom Préposition Nom Nom
N	Nom
NN	Nom Nom
NNAN	Nom Nom Adjectif Nom
NNN	Nom Nom Nom
NNNN	Nom Nom Nom Nom
NNNP	Nom Nom Nom Préposition Nom
NNPAN	Nom Nom Préposition Adjectif Nom
NNPN	Nom Nom Préposition Nom
NNPNN	Nom Nom Préposition Nom Nom
NPAN	Nom Préposition Adjectif Nom
NPANN	Nom Préposition Adjectif Nom Nom
NPN	Nom Préposition Nom
NPNN	Nom Préposition Nom Nom
NPNN	Nom Préposition Nom Nom Nom
NPNPN	Nom Préposition Nom Préposition Nom
V	Verbe

Tableau 55 Légende des matrices syntaxiques du graphique à colonne (section 9.4) et de la boîte à moustache (section 9.5).

9.6 Graphique à colonnes représentant le score BLEU

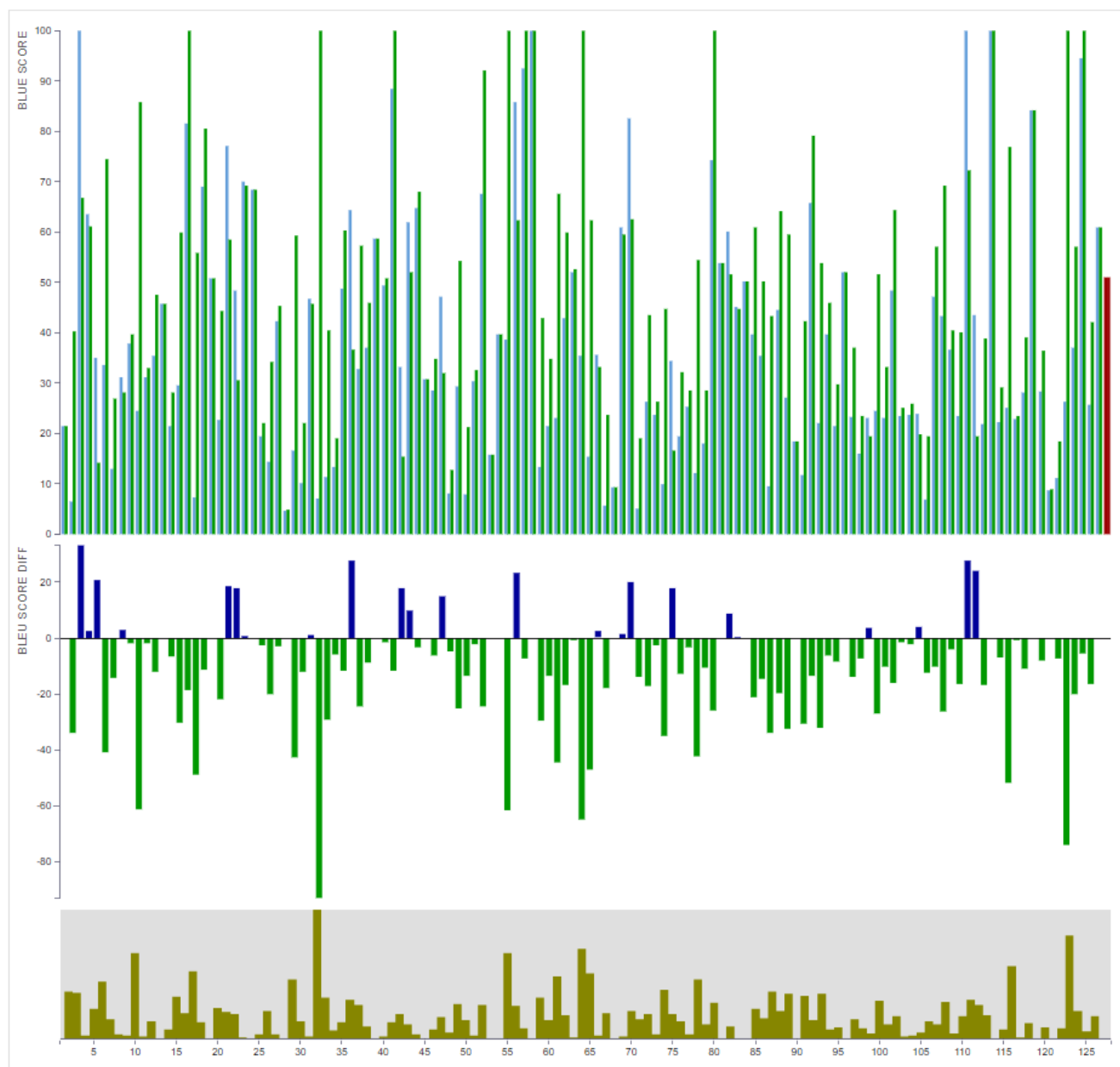


Figure 29 Graphique à colonnes : score BLEU dans les TA de DeepL Pro.

9.7 Liste des termes inclus dans le « glossaire » de DeepL Pro

Termes en allemand	Termes en anglais
verabschieden	adopt
Antikörpernachweis	antibody test
Antigen	antigen
gelten	apply
Armee	armed force
Armeeapotheke	army pharmacy
AstraZeneca	AstraZeneca
Gymnasium	baccalaureate school
verbieten	ban
Veranstaltungsverbot	ban on events
Biotech-Unternehmen	biotechfirm
Betrieb	business
Institution der familienergänzenden Kinderbetreuung	child-care facility
Zivilschutz	civile defence unit
klinische Testphase	clinical trial phase
schliessen	close
Bekämpfung von Covid-19	combat COVID-19
Bekämpfung des Coronavirus	combat the coronavirus
in Kraft treten	come into force
obligatorische Schule	compulsory school
Nachverfolgen von Kontakten	contact tracing
Coronavirus	Coronavirus
Fallzahl	coronavirus case
COVAX-Initiative	COVAX initiative
COVAX- Initiative zur Beschaffung von COVID-19-Impfstoffen	COVAX initiative for the procurement of COVID-19 vaccines
COVID-Bestimmung	COVID provision
Covid-19-Gesetz	COVID-19 Act
Covid-19-Pandemie	COVID-19 pandemic
COVID-19-Patientin und Patient	COVID-19 patient
Covid-19-Impfstoff	Covid-19 vaccine
Covid-19-Impfstoff	COVID-19 vaccine
COVID-19 Vaccine Global Access Facility	COVID-19 Vaccine Global Access Facility
bekämpfen	deal with
beschliessen	decide

künstliche Proteine	designed protein
entdecken	detect
Entwicklung der Epidemie	development of the epidemic
Abstand	distance
Fernunterricht	distance-learning
Distanz- und Hygieneregeln	distancing and hygienic rule
Verteilung von COVID-19-Impfstoff	distribution of COVID-19 vaccine
lockern	ease
Lockerung	ease measures
Öffnungsschritt	easing of measures
ECDC	ECDC
wirtschaftlicher Schäden	economic impact
EDK	EDK
Einreisebeschränkung	entry restriction
epidemische Entwicklung	epidemiological trend
besonders gefährdete Personengruppe	especially high risk employee
Europäisches Zentrum für die Prävention und die Kontrolle von Krankheiten	European Center for Disease Prevention and Control
Veranstaltung mit bis zu 1000 Personen	event for up to 1000 people
von der Quarantäne befreien	exempt from the requirement to go into quarantine
ausserordentliche Lage	extraordinary situation
Pflicht zum Tragen einer Gesichtsmaske	Face mask must be worn
Präsenzunterricht	face-to-face teaching
Präsenzveranstaltung	face-to-face teaching
EDA	FDFA
EDI	FDHA
EJPD	FDJP
Bundesrat	Federal Council
Krisenstab des Bundesrates zur Bewältigung der Corona Krise (KSBC)	Federal Council Coronavirus Crisis Unit
bundesrätlicher Krisenstab im Kontext von COVID-19	federal crisis management of the COVID-19 pandemic
Bund	federal government
Bundesamt für Gesundheit	Federal Office of Public Health
eidgenössisches Volksbegehren	federal popular request
Empfehlung des Bundes betreffend Hygiene und Abstandhalten einzuhalten	federal recommendation on hygiene and social distancing
Ordnungsbuss	fine
BAG	FOPH

verzichen	forgo
Veranstaltung	gathering
GAVI	Gavi
Isolation begeben	go into isolation
Quarantäne begeben	go into quarantine
Quarantäne einhalten	go into quarantine
Menschenansammlung von mehr als fünf Personen	group of more than five people
Gesundheitszustand	health condition
Gesundheitswesen	health service
Gesundheitswesen	healthcare
Schule der Tertiärstufe	higher education establishment
Hochschule	higher education institution
Risikogebiet	high-risk country
Hygienemaske	hygienic mask
Hygiene und Abstandhalten	hygiene and social distancing
Hygienemask	hygienic mask
Identifizierung von Spezialmassnahme	identify special measure
Immunotherapeutikum	immunotherapeutic agent
infizierte Person	infected person
Prävention und Kontrol von Infektion	infection prevention and control
Massnahmen verschärfen	introduce more stringent measures
Abstand halten	keep your distance
grosse Veranstaltung	large-scale event
Grossveranstaltung	large-scale event
Corona-Erwerbsersatz	loss of earnings related to the corona pandemic
Entscheidungsfindung im Kontext von COVID-19	make decisions about action to be taken in the COVID-19 crisis
Bewältigen der Krise	manage the crisis
Maske	mask
Maske schützt die anderen	mask protects others
Massnahme	measure
Massnahme zur Bekämpfung des Coronavirus	measure introduced to combat the coronavirus
Massnahme zur Bekämpfung des Virus	measure put in place to combat the virus
Massnahme zum Schutz der Bevölkerung vor dem neuen Coronavirus	measure put in place to protect the public from the new coronavirus
Massnahme zur Eindämmung der Epidemie	measure to contain the epidemic
Schutz der Bevölkerung vor dem neuen Coronavirus	measures to protect the public being eased
Botenmolekül	messenger molecule
gering halten	minimise

Musterschutzkonzept	model set of precautionary measure
Moderna	Moderna
mRNA	mRNA
mRNA-Impfstoff	mRNA vaccine
Nationale Zentrum für Cybersicherheit	National Cybersecurity Centre
NCSC	NCSC
in die Armbeuge niesen	neeze into the crook of your arm
Reisequarantäne neu geregelt	new rules on travel quarantine
aufheben	no longer apply
verbieten	not permit
Fallzahl	number of cases of people infected with the virus
Zahl der Infektionen mit dem Coronavirus	number of coronavirus infections
Fallzahl	number of new cases
Zahl der Neuinfektionen	number of new cases of infection
Anzahl Neuinfektion	number of new infection
Anzahl Hospitalisationen	number of people requiring hospital treatment
Verordnung über das Proximity-Tracing-System	Ordinance on the Proximity Tracing System
ambulante Betreuung	outpatient care
Pandemie	pandemic
Trennwand	partition
ausstehender Test	pending test result
Einreise aus den Grenzregionen	persons entering from border regions
dritter Schritt	phase three
Körperkontakt	physical contact
Pilotphase	pilot phase
isolieren	place in isolation
positiver Test mit Covid-19	positive test result
Schutzkonzept	precautionary measure
Schutzmassnahme	precautionary measure
vorsorglich	precautionary measure
Bundespräsidentin	President
private Veranstaltung mit über 15 Personen	private event attended by over 15 people
Beschaffung eines Impfstoffes	procurement of a vaccine
schützen	protect
Schutz vor dem neuen Coronavirus	protect against the new coronavirus
Gesundheitsschutz	protect our health
Schutz von besonders gefährdeten Arbeitnehmender	Protection for employee at especially high risk
Versorgung	provision
eigenverantwortliches Handeln	public to behave responsibly

Quarantänepflicht	quarantine requirement
Schnelltest	rapid test
Schnelltest	rapid testing
reduzieren	reduce
verringern	relax
Haus bleiben	remain at home
untersagt sein	remain prohibited
Maskentragpflicht	requirement to wear a mask
Einschränkung	restriction
Einschränkung der Grundrechte	restriction on basic rights
Einreisebeschränkung	restriction on entering Switzerland
Risikofaktor	risk factor
Hygiene- und Abstandsregeln einhalten	rules on hygiene and social distancing
SARS coronavirus-2	SARS coronavirus-2
SARS-Coronavirus-2	SARS-Coronavirus-2
Seco	Seco
Isolation	self-isolation
SBFI	SERI
Konzept	set of precautionary measure
Institution im Pflegebereich	social care institution
Kundgebung	society demonstration
spontane Menschenansammlung	spontaneous gathering
Ausbreitung des Coronavirus	spread of the coronavirus
besonders gefährdete Person	staff at especially high risk
Staatssekretariats für Wirtschaft	State Secretariat for Economic Affairs
Staatssekretariat für Bildung, Forschung und Innovation	State Secretariat for Education, Research and Innovation
Lockerungsschritt	step to ease measures
Bund	Swiss Confederation
Erziehungsdirektorenkonferenz	Swiss Conference of Cantonal Directors of Education
Bund	Swiss federal government
Swiss National COVID-19 Science Task Force	Swiss National COVID-19 Science Task Force
wissenschaftliche Science Task Force	Swiss National COVID-19 Science Task Force
schweizer Wissenschaftsgemeinschaft	Swiss scientific community
SwissCovid App	SwissCovid app
positiv getestet werden	test positive
Testergebnis	test result
Testkriterium	testing criteria

Pandemie	threat
Rückverfolgung von Infektionsketten	tracing of infection chains
Übertragung	transmission
Übertragungskette	transmission chain
behandeln	treat
Sekundarstufe II	upper secondary level
Schule der Sekundarstufe II	upper secondary level school
Mittelschule	upper secondary school
Impfstoff	vaccine
Impfung gegen COVID-19	vaccine against COVID-19
Impfstoff-Unternehmen	vaccine company
Impfstoffprojekt	vaccine project
Berufsschule	vocational school
Hände waschen	wash your hands thoroughly
Maskentragpflicht	wear a mask is mandatory
Maske tragen	wear masks
Home-Office	work from home
von zu Hause aus arbeiten	work from home
besonders gefährdeter Arbeitnehmender	worker at especially high risk
Homeoffice-Empfehlungen	working from home recommended

Tableau 56 Liste des termes inclus dans le « glossaire » de DeepL Pro.

9.8 Corpus de test

Phrase n°	Src	Réf	Syst sans glos	Syst avec glos
1.	Wer sich auf Anordnung eines Arztes oder einer Behörde in Quarantäne begeben muss , hat Anrecht auf den Corona-Erwerbsersatz .	Anyone instructed by a doctor or the authorities to go into quarantine is entitled to compensation for loss of earnings related to the corona pandemic .	Anyone who has to go into quarantine by order of a doctor or an authority is entitled to the Corona compensation .	Anyone who has to go into quarantine by order of a doctor or an authority is entitled to the compensation for loss of earnings related to the corona pandemic .
2.	Schweizer Diplomattinnen und Diplomaten sind neu ebenfalls für notwendige dienstliche Tätigkeiten von der Quarantäne befreit .	Swiss diplomats will now also be exempt from the requirement to go into quarantine when returning from official business trips.	Swiss diplomats are now also exempt from quarantine for necessary official activities.	Swiss diplomats are now also exempt from the requirement to go into quarantine for necessary official activities.
3.	Weitere Impfstoffprojekte in Evaluation	Further vaccine projects under evaluation	Further vaccine projects under evaluation	Other vaccine projects under evaluation
4.	Der Zivilschutz übernimmt Aufgaben in verschiedenen Bereichen: Insbesondere kann er das Gesundheitswesen und die Institutionen im Pflegebereich unterstützen, indem er beispielsweise bei der ambulanten Betreuung von Betagten und Pflegebedürftigen mithilft, Mahlzeiten verteilt oder Fahrdienste übernimmt.	Civil defence units will carry out tasks in various areas: in particular, they can support the health service and social care institutions , for example by assisting in the outpatient care of the elderly and those in need of care, distributing meals or providing transport services.	Civil protection takes on tasks in various areas: In particular, it can support the health care system and institutions in the care sector , for example by assisting in the outpatient care of the elderly and people in need of care, distributing meals or providing transport services.	The civile defence unit performs tasks in various areas: In particular, it can support the healthcare and social care institutions , for example by helping with outpatient care for the elderly and those in need of care, distributing meals or providing transport services.
5.	Die Massnahmen gelten ab Donnerstag, 29. Oktober, und sind nicht befristet.	The measures will apply from Thursday, 29 October, for an indefinite period of time.	The measures will apply from Thursday 29 October and are not limited in time.	The measures apply from Thursday 29 October and are not limited in time.
6.	«Diese Pandemie kann nur gemeinsam bekämpft werden», sagte Bundesrat Alain Berset.	"This threat can only be dealt with by working together", stressed Federal Councillor Alain Berset.	"This pandemic can only be combated together", said Federal Councillor Alain Berset.	"This threat can only be dealt with by working together", said Federal Councilor Alain Berset.
7.	Die Sujets appellieren mit dem Aufruf «Mach's einfach!» daran, bei einem positiven Test mit Covid-19 unbedingt die Isolation oder im Falle einer Rückkehr aus einem Risikogebiet und bei einem ausstehenden Test , die Quarantäne einzuhalten .	Along with the call to "go on and do it!" the themes emphasise the importance of self-isolation after a positive test result and of going into quarantine upon return from a high-risk country and while test results are pending .	The subjects appeal with the call "Just do it!" to strictly observe isolation in case of a positive test with Covid-19 or quarantine in case of return from a risk area and in case of a pending test .	The subjects are calling for self-isolation in the event of a positive test result with Covid-19 or, in the event of a return from a high-risk country and pending test results , to go into quarantine .
8.	Neu muss auch in den Aussenbereichen von Einrichtungen und Betrieben eine Maske getragen werden , wie zum Beispiel in Aussenbereichen von Läden und an Veranstaltungsorten, auf	Masks must now also be worn outside establishments and facilities, for example in outdoor areas of shops and at event venues, on terraces of restaurants and bars, and at farmers' and Christmas markets.	Now, masks must also be worn in the outdoor areas of institutions and businesses, such as in the outdoor areas of shops and at event locations, on terraces in restaurants and bars and at weekly and Christmas markets.	Newly, wear masks must also be worn in the outdoor areas of institutions and businesses, such as outside shops and event venues, on terraces in restaurants and bars and at weekly and Christmas markets.

Terrassen in Restaurants und Bars und auf Wochen- und Weihnachtsmärkten.

- | | | | | |
|-----|---|---|---|---|
| 9. | Aufgrund ihres Mandats und ihrer Erfahrung könne GAVI wesentlich dazu beitragen, dass eine künftige Impfung gegen COVID-19 auch die Menschen in den ärmsten Ländern erreiche. | Due to its mandate and experience, Gavi is well placed to ensure that any future vaccine against COVID-19 will also reach people in the poorest countries | With its mandate and experience, GAVI can make a significant contribution to ensuring that future vaccination against COVID-19 reaches people in the poorest countries. | With its mandate and experience, Gavi says it can play a key role in ensuring that future vaccines against COVID-19 reach people in the poorest countries. |
| 10. | Ab Montag, 2. November, müssen Hochschulen auf Präsenzunterricht verzichten. | From Monday, 2 November, higher education institutions will have to forgo face-to-face teaching . | As of Monday, 2 November, universities will have to do without attendance classes . | From Monday, 2 November, higher education institutions will have to forgo face-to-face teaching . |
| 11. | Der Bundesrat wird das System ausser Betrieb nehmen, sobald es für die Bekämpfung des Coronavirus nicht mehr erforderlich ist. | As soon as the system is no longer required to combat the coronavirus , the Federal Council will take it out of service. | The Federal Council will take the system out of service as soon as it is no longer necessary for the control of the coronavirus . | The Federal Council will take the system out of service as soon as it is no longer needed to deal with the Coronavirus . |
| 12. | Covid-19-Impfstoff: Bund unterzeichnet Vertrag mit Biotech-Unternehmen Moderna | Covid-19 vaccine: federal government signs agreement with biotechfirm Moderna | Covid-19 vaccine: German government signs contract with biotech company Moderna | Covid-19 vaccine: Swiss federal government signs contract with biotechfirm Moderna |
| 13. | Das kann bedeuten, dass gewisse Truppen vorübergehend mobilisiert und vorsorglich für den Einsatz ausgebildet werden. | This may mean that certain troops will be mobilised only temporarily, so that they can be trained for deployment as a precautionary measure . | This may mean that certain troops are temporarily mobilised and trained for deployment as a precautionary measure . | This may mean that certain troops are temporarily mobilised and trained for deployment as a precautionary measure . |
| 14. | Identifizierung von Spezialmassnahmen im Bereich der Innovationsförderung mit dem Ziel, aus Schweizer Wissenschafts-Know-how rasch Produkte oder Dienstleistungen hervorbringen und damit wesentlich zur Bekämpfung von Covid-19 beitragen zu können. | To identify special measures that can be introduced to encourage innovation, with the aim of exploiting Swiss scientific know-how to rapidly generate products and services in order to effectively combat COVID-19 . | Identification of special measures in the field of innovation promotion with the aim of rapidly generating products or services from Swiss scientific know-how and thus making a significant contribution to the fight against Covid-19 . | Identify special measures to promote innovation with the aim of rapidly producing products or services from Swiss scientific know-how and thus making a significant contribution to the deal with combat COVID-19 . |
| 15. | Am 8. Juni sollen in einem dritten Schritt die Mittel-, Berufs- und Hochschulen wieder Präsenzveranstaltungen abhalten dürfen. | In phase three, set to commence on 8 June, upper secondary schools, vocational schools and higher education institutions will be allowed to resume face-to-face teaching . | In a third step, secondary schools, vocational colleges and universities will be allowed to hold face-to-face events again on 8 June. | In a third phase three, on 8 June, secondary schools, vocational schools and higher education institutions will again be allowed to hold face-to-face teaching . |
| 16. | Die Armeeapotheke hat bis heute 21 Millionen Hygienemasken an die Kantone verteilt. | To date, the army pharmacy has distributed 21 million hygienic masks to the cantons. | To date, the army pharmacy has distributed 21 million hygiene masks to the cantons. | To date, the army pharmacy has distributed 21 million hygienic masks to the cantons. |
| 17. | Neu gelten dieselben Vorgaben für alle Konzepte; Musterschutzkonzepte gibt es keine mehr. | The same requirements will apply for all sets of precautionary measures ; there are to be no more model sets of precautionary measures . | Now the same requirements apply to all concepts; design protection concepts no longer exist. | The same requirements now apply to all sets of precautionary measures ; there are no longer any |

model sets of precautionary measures.

18.	Alle Läden, Märkte, Restaurants, Bars sowie Unterhaltungs- und Freizeitbetriebe wie Museen, Bibliotheken, Kinos, Konzert- und Theaterhäuser, Sportzentren, Schwimmbäder und Skigebiete werden geschlossen .	All shops, markets, restaurants, bars and entertainment and leisure facilities, such as museums, libraries, cinemas, concert halls and theatres, sports centres, swimming pools and ski areas are to close .	All shops, markets, restaurants, bars and entertainment and leisure facilities such as museums, libraries, cinemas, concert and theatre halls, sports centres, swimming pools and ski resorts will be closed .	All shops, markets, restaurants, bars and entertainment and leisure facilities such as museums, libraries, cinemas, concert halls and theatres, sports centres, swimming pools and ski resorts will be closed .
19.	Über die weiteren Schritte entscheidet der Bundesrat am 27. Mai.	The Federal Council will decide on further steps to be taken on 27 May.	The Federal Council will decide on the next steps on 27 May.	The Federal Council will decide on the next steps on 27 May.
20.	Der Arbeitgeber ist in der Pflicht, besonders gefährdete Personen von zu Hause aus arbeiten zu lassen, wenn nötig durch eine angemessene Ersatzarbeit.	Employers have a duty to allow staff at especially high risk to work from home , if necessary by allowing them to perform appropriate alternative tasks.	The employer is obliged to let particularly vulnerable persons work from home , if necessary by providing appropriate replacement work.	Employers are obliged to allow staff at particularly high risk to work from home , if necessary by means of appropriate replacement work.
21.	Parallel dazu ist der Bund mit weiteren Impfstoff-Unternehmen im Gespräch.	In parallel to this, the federal government is in talks with other vaccine companies .	In parallel, the Federal Government is in talks with other vaccine companies .	At the same time, the Swiss federal government is in talks with other vaccine companies .
22.	Coronavirus: Bundesrat verbietet grosse Veranstaltungen	Coronavirus: Federal Council bans large-scale events	Coronavirus: Federal Council bans major events	Coronavirus: Federal Council not permit major gatherings
23.	Um die Schliessung von Baustellen zu verhindern und die Angestellten besser zu schützen, verpflichtet der Bundesrat die Arbeitgeber im Baugewerbe und in der Industrie, die Empfehlungen des Bundes betreffend Hygiene und Abstandhalten einzuhalten .	In order to prevent the closure of construction sites and to offer better protection to employees, the Federal Council is requiring employers in the construction sector and industry to comply with the federal recommendations on hygiene and social distancing .	In order to prevent the closure of construction sites and to better protect employees, the Federal Council obliges employers in the construction sector and industry to comply with the federal recommendations on hygiene and spacing .	In order to prevent the closure of construction sites and to better protect employees, the Federal Council obliges employers in the construction sector and industry to comply with the Swiss federal government's recommendations on hygiene and social distancing .
24.	Multilateralismus, Solidarität und internationale Zusammenarbeit seien die Basis des Schweizer Engagements und in der aktuellen Lage unverzichtbar, sagte Bundespräsidentin Sommaruga .	Multilateralism, solidarity and international cooperation form the basis of Switzerland's activities and are vital in the current situation, said President Sommaruga .	Multilateralism, solidarity and international cooperation are the basis of Switzerland's commitment and are indispensable in the current situation, said President Sommaruga .	Multilateralism, solidarity and international cooperation are the basis of Switzerland's commitment and are indispensable in the current situation, said President Sommaruga .
25.	Die Versorgung der gesamten Bevölkerung mit Lebensmitteln, Medikamenten und Waren des täglichen Gebrauchs ist sichergestellt, es sind genügend Vorräte angelegt.	Provision of the entire population with foodstuffs, medicinal products and essential everyday items is assured; there are sufficient stocks.	The supply of the entire population with food, medicines and goods for daily use is assured and sufficient stocks have been built up.	The provision of the entire population with food, medicines and goods for daily use is guaranteed, and sufficient stocks have been built up.

26.	Die Pflicht zum Tragen einer Gesichtsmaske gilt in obligatorischen Schulen, Schulen der Sekundarstufe II und der Tertiärstufe , in Institutionen der familienergänzenden Kinderbetreuung sowie in den Trainingsbereichen von Sport- und Fitnessseinrichtungen nur dann, wenn sie im betreffenden Schutzkonzept vorgesehen ist.	Face masks must also be worn in all schools and higher education establishments, in child-care facilities and in the training areas of sport and fitness centres if part of the centre's precautionary measures .	The obligation to wear a face mask applies in compulsory schools, upper secondary and tertiary schools, in institutions providing supplementary childcare and in the training areas of sports and fitness facilities only if it is provided for in the relevant protection concept .	Face masks must be worn in compulsory schools, upper secondary level schools and higher education establishments, in child-care facilities and in the training areas of sports and fitness facilities only if this is provided for in the precautionary measure concerned.
27.	Noch ist weltweit kein Impfstoff für den Markt zugelassen; die Entwicklung und Forschung laufen aber intensiv.	While no vaccine has yet been approved for the market anywhere in the world, intensive development and research work is underway.	No vaccine has yet been approved for the market worldwide, but development and research is ongoing.	No vaccine has yet been approved for the market worldwide; however, intensive research and development is underway.
28.	Voraussetzung ist, dass der Aufenthalt im Ausland nicht mehr als fünf Tage dauert und ein Schutzkonzept erstellt und umgesetzt wird.	The persons concerned must not remain abroad for more than five days, and precautionary measures must be planned and taken.	The prerequisite is that the stay abroad does not exceed five days and that a protection concept is drawn up and implemented.	The prerequisite is that the stay abroad does not exceed five days and that a precautionary measure is drawn up and implemented.
29.	Der Bundesrat will weiter die wirtschaftlichen Schäden möglichst gering halten und die Einschränkungen der Grundrechte wo möglich verringern.	The Federal Council wishes to minimise the economic impact and relax restrictions on basic rights .	The Federal Council intends to continue to keep economic damage to a minimum and to reduce restrictions on basic rights where possible.	The Federal Council continues to aim to minimise the economic impact and relax restrictions on basic rights where possible.
30.	Der Bundesrat hat zudem den Schutz besonders gefährdeter Arbeitnehmenden präzisiert.	The Federal Council also issued further details with regard to protecting workers at especially high risk .	The Federal Council has also specified the protection of particularly vulnerable workers .	Protection for employees at especially high risk has also been specified by the Federal Council .
31.	Die Covid-19-Pandemie zeigt deutlich, wie wichtig die Prävention und die Kontrolle von Infektionen sind.	The COVID-19 pandemic clearly shows the importance of infection prevention and control .	The Covid 19 pandemic clearly shows the importance of preventing and controlling infections .	The COVID-19 pandemic clearly shows how important the prevention and control of infections is.
32.	Schutz von besonders gefährdeten Arbeitnehmenden	Protection for employees at especially high risk	Protection of particularly exposed workers	Protection for employees at especially high risk
33.	Deshalb sicherte sich der Bund im Sommer 2020 ein Immunotherapeutikum von Molecular Partners.	In summer 2020, the Swiss federal government therefore secured access to an immunotherapeutic agent from Molecular Partners.	This is why the German government secured an immunotherapeutic drug from Molecular Partners in the summer of 2020.	This is why the Swiss federal government secured an immunotherapeutic agent from Molecular Partners in the summer of 2020.
34.	Der Arbeitgeber kann ein ärztliches Attest verlangen, das aufzeigt, weshalb eine angestellte Person zu einer besonders gefährdeten Personengruppe gehört.	The employer may request a medical certificate stating why the employee is considered to be at especially high risk .	The employer may require a medical certificate showing why an employed person belongs to a particularly vulnerable group .	The employer can demand a medical certificate that shows why an employee belongs to a staff at especially high risk employee .

35.	Ab Montag, 19. Oktober, sind im öffentlichen Raum spontane Menschenansammlungen von mehr als 15 Personen verboten .	From Monday, 19 October spontaneous gatherings of more than 15 persons are not permitted in public.	From Monday, 19 October, spontaneous gatherings of more than 15 people are prohibited in public places.	From Monday, 19 October, spontaneous gatherings of more than 15 people are not permitted in public places.
36.	Veranstaltungen mit bis zu 1000 Personen möglich	Events for up to 1000 people possible	Events with up to 1000 people possible	Gatherings of up to 1000 people are possible
37.	Coronavirus: Bundesrat erklärt die « ausserordentliche Lage » und verschärft die Massnahmen	Coronavirus: Federal Council declares 'extraordinary situation ' and introduces more stringent measures	Coronavirus: Federal Council declares "extraordinary situation " and tightens measures	Coronavirus: Federal Council declares "extraordinary situation " and introduces more stringent measures
38.	Der Bundesrat geht davon aus, dass der Bedarf der zivilen Behörden nach Unterstützung durch die Armee in den nächsten Tagen und Wochen markant steigen wird.	The Federal Council is expecting that the civilian authorities will require significant support from the armed forces in the coming days and weeks.	The Federal Council assumes that the need of the civilian authorities for support by the armed forces will increase significantly in the coming days and weeks.	The Federal Council assumes that the civilian authorities' need for support from the armed force will increase markedly in the coming days and weeks.
39.	Coronavirus: Bundesrat lockert weitere Massnahmen ab dem 11. Mai 2020	Coronavirus: Federal Council to ease further measures from 11 May	Coronavirus: Federal Council relaxes further measures from 11 May 2020	Coronavirus: Federal Council eases further measures from 11 May 2020
40.	Um eine gerechte Verteilung von COVID-19-Impfstoffen an alle Länder zu gewährleisten, haben sich Staaten weltweit zur COVAX-Initiative (COVID-19 Vaccine Global Access Facility) zusammengeschlossen und einen gemeinsamen Beschaffungsmechanismus installiert.	In order to secure a fair distribution of COVID-19 vaccines across the globe, nations across the world have formed the COVAX initiative (COVID-19 Vaccine Global Access Facility) and created a pooled procurement mechanism.	To ensure fair distribution of COVID-19 vaccines to all countries, countries around the world have joined forces to form the COVAX initiative (COVID-19 Vaccine Global Access Facility) and have established a common procurement mechanism.	In order to ensure a fair distribution of COVID-19 vaccines to all countries, countries around the world have joined together to form the COVAX initiative (COVID-19 Vaccine Global Access Facility) and set up a joint procurement mechanism.
41.	Damit folgt das Bundesamt für Gesundheit den Empfehlungen des Europäischen Zentrums für die Prävention und die Kontrolle von Krankheiten (ECDC) .	The Federal Office of Public Health is thus following the recommendations of the European Center for Disease Prevention and Control (ECDC) .	The Federal Office of Public Health is thus following the recommendations of the European Centre for Disease Prevention and Control (ECDC) .	The Federal Office of Public Health is thus following the recommendations of the European Center for Disease Prevention and Control (ECDC) .
42.	Auch in Schulen ab der Sekundarstufe II gilt neu eine Maskentragpflicht .	Wearing a mask is now also mandatory in schools from upper secondary level (level II) upwards.	The wearing of masks is now also compulsory in schools from upper secondary level .	A requirement to wear a mask also applies in schools from the upper secondary level onwards.
43.	Bern, 30.04.2020 - Das Bundesamt für Gesundheit passt seine Kampagne zum Schutz vor dem neuen Coronavirus an und schaltet in der Grundfarbe auf Pink.	Berne, 30.04.2020 - The Federal Office of Public Health is adapting its public awareness campaign to protect against the new coronavirus and is switching the base colour to pink.	Bern, 30.04.2020 - The Federal Office of Public Health is adapting its campaign to protect against the new coronavirus and switching to pink in the basic colour.	Bern, 30.04.2020 - The Federal Office of Public Health is adapting its campaign for protection against the new Coronavirus and switching to pink in the basic colour.
44.	Parallel dazu bereitet der Bundesrat die Überführung der relevanten Covid-Bestimmungen in ein dringliches und befristetes	In parallel to that, the Federal Council is preparing to transfer the relevant COVID provisions into an urgent and limited COVID-19 Act , which is	At the same time, the Federal Council is preparing the transfer of the relevant Covid provisions into an urgent and temporary Covid 19 Act ,	At the same time, the Federal Council is preparing the transfer of the relevant COVID provisions into an urgent and temporary

	Covid-19-Gesetz vor, das voraussichtlich am 19. Juni 2020 in die Vernehmlassung geschickt werden soll.	expected to be submitted for consultation on 19 June.	which is expected to be submitted for consultation on 19 June 2020.	COVID-19 Act , which is expected to be submitted for consultation on 19 June 2020.
45.	Ausserordentliche Lage endet am 19. Juni 2020	Extraordinary situation to end on 19 June	Extraordinary situation ends on 19 June 2020	Extraordinary situation ends on 19 June 2020
46.	Bis Ende Mai gilt noch der Fristenstillstand bei eidgenössischen Volksbegehren .	Deadlines for submitting federal popular requests remain suspended until the end of May.	Until the end of May, the deadline for federal referendums will be suspended.	Until the end of May, the deadline for federal popular requests still applies.
47.	Die Schweizer Wissenschaftsgemeinschaft hat ein hohes Potenzial, den Ausgang dieser Krise positiv zu beeinflussen, sei es durch Forschung, Bildung oder Wissenstransfer.	The Swiss scientific community has a high potential to positively influence on how this crisis develops in a number of ways, including through research, education and knowledge transfer.	The Swiss scientific community has a high potential to positively influence the outcome of this crisis, whether through research, education or knowledge transfer.	The Swiss scientific community has great potential to positively influence the outcome of this crisis, whether through research, education or knowledge transfer.
48.	Der Abstand kann weiterhin unterschritten werden, wenn eine Maske getragen wird oder Trennwände vorhanden sind.	If masks are worn or partitions are in place, the distance may be further reduced.	The distance can still be reduced if a mask is worn or if there are partitions .	The distance can still be reduced if wear masks or partitions are available.
49.	Aufgrund der epidemischen Entwicklung und gestützt auf Empfehlungen der Wissenschaft, lockert der Bundesrat ab dem 27. April die Massnahmen zum Schutz der Bevölkerung vor dem neuen Coronavirus .	In view of the epidemiological trend and based on expert recommendations, from 27 April the Federal Council will begin to ease the measures put in place to protect the public from the new coronavirus .	In view of the epidemic development and based on scientific recommendations, the Federal Council will relax measures to protect the population from the new coronavirus from 27 April.	In view of the epidemiological trend and based on scientific recommendations, the Federal Council is putting in place from 27 April onwards an easing of measures to protect the public from the new Coronavirus .
50.	Der starke Anstieg der Fallzahlen in den letzten Tagen ist besorgniserregend.	The rapid rise in coronavirus cases in the last few days is a cause for great concern.	The sharp increase in the number of cases in recent days is worrying.	The sharp increase in the number of cases of people infected with the virus in recent days is a cause for concern.
51.	Angesichts der anhaltend tiefen Fallzahlen hat der Bundesrat am 19. Juni 2020 die ausserordentliche Lage aufgehoben und weitere Lockerungen beschlossen.	In view of the continuing low number of coronavirus cases , the Federal Council downgraded the situation from 'extraordinary' on 19 June and eased further measures.	In view of the persistently low number of cases , the Federal Council lifted the extraordinary situation on 19 June 2020 and decided on further easing .	In view of the continuing low number of cases of people infected with the virus , the Federal Council no longer applied the extraordinary situation on 19 June 2020 and decided on further ease measures .
52.	Der Bundesrat verabschiedete zudem die Verordnung über das Proximity-Tracing-System	The Federal Council also adopted the Ordinance on the Proximity Tracing System ,	The Federal Council also approved the ordinance on the proximity tracing system	The Federal Council also adopted the Ordinance on the Proximity Tracing System
53.	Damit ermöglicht er den Start der SwissCovid App .	thereby making it possible to start using the SwissCovid app	thus enabling the launch of the SwissCovid App .	This enables the launch of the SwissCovid app .
54.	Das NCSC nimmt auch nach der öffentlichen Lancierung der SwissCovid App Meldungen von	The NCSC will continue to accept reports of test results following the public launch of the SwissCovid app so that data	Even after the public launch of the SwissCovid app , the NCSC will continue to receive reports of test results in order to guarantee the privacy and	Even after the public launch of the SwissCovid app , the NCSC will continue to receive reports of test results in order to ensure the privacy

	Testergebnissen entgegen, um den Datenschutz und die Sicherheit der App laufend zu gewährleisten.	protection and the security of the app can be assured.	security of the app on an ongoing basis.	and security of the app on an ongoing basis.
55.	Coronavirus: Keine Quarantäne bei der Einreise aus den Grenzregionen	Coronavirus: No quarantine for persons entering from border regions	Coronavirus: No quarantine on entry from border regions	Coronavirus: No quarantine for persons entering from border regions
56.	Coronavirus: weitere Massnahmen zur Eindämmung der Epidemie, Einführung von Schnelltests, Reisequarantäne neu geregelt.	Coronavirus: further measures to contain the epidemic, introduction of rapid testing, new rules on travel quarantine.	Coronavirus: further measures to contain the epidemic, introduction of rapid tests, new rules on travel quarantine.	Coronavirus: further measures to contain the epidemic, introduction of rapid tests, travel quarantine reorganised
57.	In den kommenden Wochen wird die App in einer Pilotphase getestet	In the coming weeks the app will be tested in a pilot phase .	In the coming weeks the app will be tested in a pilot phase	In the coming weeks the app will be tested in a pilot phase
58.	Kriterien sind die Anzahl Neuinfektionen , Spitaleinweisungen und Todesfälle sowie die Spitalbelegungszahlen.	The criteria are the number of new infections , hospital admissions and deaths, and hospital occupancy rates.	The criteria are the number of new infections , hospital admissions and deaths, and hospital occupancy rates.	The criteria are the number of new infections , hospital admissions and deaths, and hospital occupancy rates.
59.	Die Regel, dass solche Reisen höchstens fünf Tage dauern dürfen, wird aufgehoben .	The rule whereby such trips must not last longer than five days will no longer apply .	The rule that such journeys may not last more than five days is abolished .	The rule that such trips may last a maximum of five days will no longer apply .
60.	Der Bundesrat setzt nach den erfolgten Lockerungsschritten noch verstärkt auf eigenverantwortliches Handeln	Following the steps to ease measures , the Federal Council is relying even more heavily on the public to behave responsibly	Following the relaxation steps that have been taken, the Federal Council is placing even greater emphasis on independent action	Following the steps taken to relax the rules , the Federal Council is placing even greater emphasis on public to behave responsibly
61.	Die Menschen sollen weiterhin die Hygiene- und Abstandsregeln einhalten .	people should continue to follow the rules on hygiene and social distancing .	people should continue to comply with hygiene and distance regulations	People should continue to respect rules on hygiene and social distancing .
62.	Eine starke, effiziente und globale Organisation für den Gesundheitsschutz sei lebensnotwendig.	A strong, efficient and global organisation is therefore essential to protect our health .	A strong, efficient and global organisation for health protection was vital.	A strong, efficient and global organisation to protect our health is vital.
63.	Dabei nahm das Nationale Zentrum für Cybersicherheit (NCSC) seit dem 28. Mai Meldungen zur Sicherheit von Fachleuten und interessierten Personen entgegen.	The National Cybersecurity Centre (NCSC) has been receiving reports on security issues from specialists and interested parties since 28 May.	In this context, the National Centre for Cyber Security (NCSC) has been receiving reports on security from professionals and interested parties since 28 May.	To this end, the National Cybersecurity Centre (NCSC) has been receiving reports on the safety of professionals and interested persons since 28 May.
64.	Maske schützt die anderen	Mask protects others	Mask protects the others	Mask protects others
65.	Die Fallzahlen sowie die Anzahl Hospitalisationen nehmen drastisch zu.	The number of cases of people infected with the virus and the number of people requiring hospital treatment are rising sharply.	The number of cases as well as the number of hospitalisations is increasing dramatically.	The number of cases of people infected with the virus as well as the number of people requiring hospital treatment is increasing dramatically.
66.	Damit erhält die Schweizer Bevölkerung Zugang zu einem weiteren Covid-19-	The Swiss public will therefore gain access to a further COVID-19 vaccine - provided it	This will give the Swiss population access to a further Covid 19 vaccine - provided	This will give the Swiss population access to another Covid-19 vaccine - provided

	Impfstoff – sofern die klinische Testphase und die Zulassung durch Swissmedic erfolgreich verlaufen.	successfully passes the clinical trial phase and is approved by Swissmedic.	that the clinical trial phase and Swissmedic's approval are successful.	that the clinical trial phase and Swissmedic's approval are successful.
67.	Diese können den Unterricht flexibel gestalten und die Möglichkeiten von Fernunterricht weiter nutzen.	Teaching can be organised in a flexible manner and they and the may continue to make use of distance-learning channels.	They can organise their lessons flexibly and continue to use the possibilities of distance learning .	They will be able to organise their teaching flexibly and continue to make use of the possibilities offered by distance-learning .
68.	Derzeit ist es jedoch entscheidend, sich auf das Bewältigen dieser Krise zu konzentrieren	At the moment however, the priority is on managing this crisis .	However, at this stage it is crucial to focus on tackling this crisis	However, it is currently crucial to focus on managing the crisis
69.	Er hat das EJPD damit beauftragt, gemeinsam mit dem EDI und dem EDA die Ausdehnung der Grenzkontrollen und Einreiseverbote auf Reisende aus weiteren, stark von der Ausbreitung des Coronavirus betroffenen Länder zu prüfen.	It has tasked the FDJP to work with the FDHA and the F DFA to consider extending border checks and entry bans to travellers from other countries severely affected by the spread of the coronavirus .	It instructed the FDJP to examine, together with the FDHA and the F DFA , the extension of border controls and entry bans to travellers from other countries severely affected by the spread of the coronavirus .	It instructed the FDJP , together with the FDHA and the F DFA , to examine the extension of border controls and entry bans to travellers from other countries severely affected by the spread of the Coronavirus .
70.	Dabei gelten die Testkriterien Bundesamts für Gesundheit .	The testing criteria of the Federal Office of Public Health apply.	The test criteria of the Federal Office of Public Health apply.	The test criteria of the Federal Office of Public Health will apply.
71.	Mit ihrem freiwilligen und unentgeltlichen Engagement hat sie die Entscheidungsfindung des bundesrätlichen Krisenstabs im Kontext von COVID-19 massgeblich unterstützt.	This voluntary and unremunerated work has considerably aided federal crisis management of the COVID-19 pandemic .	With her voluntary and unpaid commitment, she has significantly supported the decision-making of the Federal Council's crisis management team in the context of COVID-19 .	Through its voluntary and unpaid commitment, it has provided substantial support for the Federal Council's crisis management of the COVID-19 pandemic by making decisions about action to be taken in the COVID-19 crisis.
72.	Diese Veranstaltungen sollen wenn möglich vermieden werden	Gatherings should be avoided where possible.	These events should be avoided if possible	These gatherings should be avoided if possible
73.	Der Bundesrat empfiehlt weiterhin, wo möglich im Home-Office zu arbeiten, auch um Spitzenauslastungen im öffentlichen	The Federal Council continues to recommend that people continue to work from home , not least to avoid overcrowding on public transport.	The Federal Council continues to recommend working in the home office wherever possible, also in order to avoid peak workloads in the public	The Federal Council continues to recommend working from home wherever possible, also in order to avoid peak workloads in the public
74.	Beratung von Politik und Behörden, um die politischen Behörden und Entscheidungsträger bei der Entscheidungsfindung im Kontext von COVID-19 wissenschaftlich zu unterstützen.	To provide appropriate scientific advice to politicians and administrators making decisions about action to be taken in the COVID-19 crisis .	Advice to policy and public authorities to provide scientific support to political authorities and decision-makers in making decisions in the context of COVID-19	Providing advice to policy and public authorities to provide scientific support to political authorities and decision-makers in making decisions about action to be taken in the COVID-19 crisis .
75.	Der Trainingsbetrieb ist für alle Sportarten ab dem 6. Juni ohne Einschränkung	Training activities for all sports may resume without any	Training is permitted again for all sports from 6 June onwards	Training will be allowed again for all sports from 6

	der Gruppengrösse wieder erlaubt.	restrictions on group sizes from 6 June.	without any restrictions on group size.	June onwards without any restriction on group size.
76.	Die Änderung tritt am 29.10.2020 in Kraft .	The changes come into force on 29 October.	The amendment will enter into force on 29.10.2020.	The change will come into force on 29.10.2020.
77.	Organisierte Veranstaltungen im öffentlichen Raum sind mit den entsprechenden Schutzmassnahmen weiterhin erlaubt, etwa politische und zivilgesellschaftliche Kundgebungen.	Organised outdoor events such as political and civil society demonstrations are permitted provided precautionary measures are observed.	Organised events in public spaces are still permitted with the appropriate protective measures , such as political and civil society rallies.	Organised gatherings in public spaces will continue to be permitted with the appropriate precautionary measures , such as political and civil society demonstrations.
78.	Parallel zu diesen Öffnungsschritten werden die Einreisebeschränkungen gelockert .	In parallel to this easing of measures, restrictions on entering Switzerland will also be relaxed .	. In parallel with these opening steps, entry restrictions will be relaxed .	In parallel with these easing of measures, restrictions on entering Switzerland will be eased .
79.	Wenn es die Entwicklung der Lage zulässt, sollen am 11. Mai die obligatorischen Schulen und die Läden wieder öffnen.	If the situation allows, schools for children of compulsory school age and shops will be able to reopen on 11 May.	If developments allow, compulsory schools and shops should reopen on 11 May.	If the development of the situation permits, compulsory schools and shops are to reopen on 11 May.
80.	Das sind künstliche Proteine , die Antigene erkennen und binden können.	These are designed proteins that can recognise and bind antigens .	These are artificial proteins that can recognise and bind antigens .	These are designed proteins that can recognise and bind antigens .
81.	Der Bundesrat hat insgesamt 300 Millionen Franken für die Beschaffung eines Impfstoffes gesprochen.	The Federal Council has committed a total of 300 million Swiss francs for the procurement of a vaccine .	The Federal Council has approved a total of CHF 300 million for the procurement of a vaccine .	The Federal Council has allocated a total of CHF 300 million for the procurement of a vaccine .
82.	Alle Betriebe und Organisatoren von Veranstaltungen müssen ein Schutzkonzept haben, das sich entweder auf ein Branchenkonzept oder auf die Vorgaben des BAG und des Staatssekretariats für Wirtschaft (Seco) stützt.	All companies and event organizers must have a protection concept that is based either on an industry concept or on the requirements of the FOPH and the State Secretariat for Economic Affairs (Seco) .	All companies and organisers of events must have a protection concept based either on a sector concept or on the specifications of the FOPH and the State Secretariat for Economic Affairs (Seco) .	All businesses and organisers of gatherings must have a precautionary measure based either on a sectoral concept or on the specifications of the FOPH and the State Secretariat for Economic Affairs (Seco) .
83.	Der Bundesrat hat zudem die Sperrstunde für Restaurationsbetriebe, Discos und Nachtclubs aufgehoben.	The Federal Council is also lifting the fixed closing time for restaurants, discos and nightclubs.	The Federal Council has also lifted the curfew for restaurants, discos and nightclubs.	The Federal Council has also no longer applied the curfew for restaurants, discos and nightclubs.
84.	Der Bundesrat hat deshalb beschlossen , den Kantonen ein Kontingent von maximal 850'000 Dienstagen zur Verfügung zu stellen.	The Federal Council has therefore decided to make up to 850,000 days of civil defence service available to the cantons.	The Federal Council has therefore decided to make a quota of a maximum of 850,000 working days available to the cantons.	The Federal Council has therefore decided to make a quota of a maximum of 850,000 working days available to the cantons.
85.	Ausgenommen von der Maskentragpflicht sind Kinder bis zu 12 Jahren, Personen, die aus medizinischen Gründen keine Maske tragen können und Gäste in	Children under the age of 12, persons who are unable to wear a mask for medical reasons, and guests in restaurants and bars when seated at a table are not required to wear a mask .	Children up to 12 years of age, persons who cannot wear a mask for medical reasons and guests in restaurants and bars are exempt from the obligation to wear a mask when sitting at a table.	Excluded from the requirement to wear a mask are children up to the age of 12, persons who cannot wear a mask for medical reasons and guests in

Restaurants und Bars, wenn sie am Tisch sitzen.

restaurants and bars when seated at a table.

86.	Vor diesem Hintergrund haben der Krisenstab des Bundesrates zur Bewältigung der Corona Krise (KSBC) , das Staatssekretariat für Bildung, Forschung und Innovation (SBFI) und das Bundesamt für Gesundheit (BAG) eine wissenschaftliche Science Task Force einberufen.	The Swiss National COVID-19 Science Task Force has therefore been set up by the Federal Council Coronavirus Crisis Unit , the State Secretariat for Education, Research and Innovation (SERI) and the Federal Office of Public Health (FOPH) .	Against this background, the Federal Council's Crisis Management Team for the Corona Crisis (KSBC) , the State Secretariat for Education, Research and Innovation (SERI) and the Federal Office of Public Health (FOPH) have convened a scientific science task force .	Against this background, the Federal Council Coronavirus Crisis Unit (KSBC) , the State Secretariat for Education, Research and Innovation (SERI) and the Federal Office of Public Health (FOPH) have convened a Swiss National COVID-19 Science Task Force to manage the crisis.
87.	Die kantonalen Gymnasien können dieses Jahr auch auf die schriftlichen Maturitätsprüfungen verzichten, nachdem die Erziehungsdirektorenkonferenz EDK bereits beschlossen hatte, die mündlichen Prüfungen nicht durchzuführen.	Cantonal baccalaureate schools will be permitted to dispense with written baccalaureate examinations this year, following a decision of the Swiss Conference of Cantonal Directors of Education (EDK) not to hold oral examinations (see separate press release).	This year, the cantonal Gymnasiums will also be able to dispense with the written Matura examinations, after the EDK Education Directors' Conference had already decided not to hold the oral examinations.	The cantonal baccalaureate schools will also be able to dispense with the written Matura examinations this year, after the Swiss Conference of Cantonal Directors of Education (EDK) had already decided not to hold the oral examinations.
88.	Bund beteiligt sich an COVAX- Initiative zur Beschaffung von COVID-19-Impfstoffen	The Swiss Confederation joins the COVAX initiative for the procurement of COVID-19 vaccines	Federal government participates in COVAX initiative for the procurement of COVID 19 vaccines	Swiss federal government participates in COVAX initiative for the procurement of COVID-19 vaccines
89.	Homeoffice-Empfehlungen	Working from home recommended	Home office recommendations	Working from home recommended
90.	Deshalb kommt die Maske nur ergänzend zu den Distanz- und Hygieneregeln zum Einsatz.	That is why the mask is only useful in combination with physical distancing and hygienic rules .	Therefore, the mask is only used in addition to the distance and hygiene rules .	Therefore, the mask is only used in addition to the distance and hygiene rules .
91.	Das Nachverfolgen von Kontakten muss aber stets möglich sein.	However, contact tracing must be possible at all times.	However, it must always be possible to track contacts .	However, contact tracing must always be possible.
92.	Der Bundesrat wird die Auswirkungen der Lockerungen auf die Entwicklung der Epidemie mit einem Monitoring genau beobachten.	The Federal Council will closely monitor the impact of easing measures on the development of the epidemic .	The Federal Council will closely monitor the impact of the relaxation of the rules on the development of the epidemic .	The Federal Council will closely monitor the impact of the ease measures on the development of the epidemic .
93.	Alle Personen, die mittels eines Schnelltests positiv getestet wurden, sollten sich umgehend in Isolation begeben.	However, anyone who tests positive following a rapid test should go into isolation immediately.	All persons who have tested positive using a rapid test should be placed in isolation immediately.	All persons who have tested positive using a rapid test should go into isolation immediately.
94.	Wettkämpfe in Sportarten mit ständigem, engem Körperkontakt, wie Schwingen, Judo, Boxen oder Paartanz sind voraussichtlich bis am 6. Juli 2020 untersagt .	Competitions in sports involving continuous, close physical contact, such as Swiss-style wrestling, judo, boxing and dancing in pairs are likely to remain prohibited until 6 July.	Competitions in sports involving constant, close physical contact, such as swinging, judo, boxing or couple dance, are expected to be prohibited until 6 July 2020.	Competitions in sports involving constant, close physical contact, such as swinging, judo, boxing or couple dance, are expected to remain prohibited until 6 July 2020.

95.	Bis zu 8000 Armeeangehörige für Gesundheitswesen , Logistik und Sicherheit	Up to 8000 members of the armed forces to assist with healthcare , logistics and security	Up to 8000 military personnel for health , logistics and security	Up to 8000 military personnel for healthcare , logistics and security
96.	Gleichzeitig soll ein Beitrag für einen fairen Zugang aller Länder zu einem solchen Impfstoff geleistet werden.	At the same time, Switzerland wants to contribute to fair access to such a vaccine for all countries.	At the same time, the aim is to contribute to fair access for all countries to such a vaccine .	At the same time, the aim is to contribute to fair access for all countries to such a vaccine .
97.	Er ruft insbesondere auch die ältere Bevölkerung dazu auf, zu Hause zu bleiben .	It is calling on the elderly in particular to remain at home .	It also calls in particular on the elderly to stay at home .	In particular, it also calls on the elderly population to remain at home .
98.	An privaten Veranstaltungen mit über 15 Personen darf künftig nur sitzend konsumiert werden.	At private events attended by over 15 people , food and drink may only be consumed while seated.	At private events with more than 15 people , consumption will in future only be allowed seated.	In future, consumption at private gatherings attended by over 15 people will only be allowed in a seated position.
99.	Einreisebeschränkungen bis am 20. Juli aufrechterhalten	Entry restrictions to remain in place until 20 July	Entry restrictions maintained until 20 July	Restrictions on entering Switzerland maintained until 20 July
100.	Die Massnahmen zur Bekämpfung des Virus werden gut umgesetzt und zeigen Wirkung.	The measures put in place to combat the virus are being followed well by the public and are having the desired effect.	The measures to combat the virus are being well implemented and are proving effective.	The measures put in place to combat the virus are being well implemented and are having an impact.
101.	Betriebe mit personenbezogenen Dienstleistungen mit Körperkontakt können ebenfalls wieder öffnen, zum Beispiel Coiffeurgeschäfte, Massagepraxen, Tattoo-Studios und Kosmetiksalons.	Businesses offering personal services involving physical contact may also reopen, for example hairdressing salons, massage practices, tattoo and cosmetic studios.	Companies offering personalised services with physical contact can also reopen, for example hairdressers, massage parlours, tattoo studios and beauty salons.	Businesses providing personal services involving physical contact can also reopen, for example hairdressers' shops, massage parlours, tattoo studios and beauty salons.
102.	Bern, 16.10.2020 - Der Bund will der Schweizer Bevölkerung einen raschen Zugang zu COVID-19-Impfstoffen gewährleisten.	Berne, 16.10.2020 - The Swiss federal government wants to assure members of the Swiss public rapid access to COVID-19 vaccines .	Bern, 16.10.2020 - The Swiss Confederation wants to guarantee the Swiss population rapid access to COVID-19 vaccines .	Bern, 16.10.2020 - The Swiss federal government wants to guarantee the Swiss population rapid access to Covid-19 vaccines .
103.	Für die Reihenfolge der Lockerungen hat der Bundesrat mehrere Risikofaktoren berücksichtigt.	The Federal Council has taken into account a range of risk factors in determining the phases for easing the measures .	The Federal Council has taken several risk factors into account for the order of loosening .	The Federal Council has taken several risk factors into account for the order of ease measures .
104.	Dieses Veranstaltungsverbot gilt ab sofort bis mindestens bis zum 15. März. Für den Vollzug und die Kontrolle sind die Kantone zuständig	This ban on events comes into immediate effect and will apply at least until 15 March.	This ban on events is valid from now until at least 15 March. The cantons are responsible for enforcement and control	This ban on events applies with immediate effect until at least 15 March. The cantons are responsible for enforcement and control
105.	Ziel ist auch, den Anstieg der Fallzahlen so stark zu bremsen, dass die Kantone das Contact Tracing weiterhin konsequent und	The intention is also to considerably curb the number of new cases so that the cantons can continue to contract trace effectively.	The aim is also to slow down the increase in the number of cases to such an extent that the cantons can continue to ensure contact tracing consistently and comprehensively.	The aim is also to slow down the increase in the number of cases of people infected with the virus to such an extent that the cantons can continue to ensure contact

	umfassend sicherstellen können.			tracing consistently and comprehensively.
106.	Den Test auf das SARS-Coronavirus-2 vergütet der Bund mit einer Pauschale von 169, jenen für den Antikörpernachweis mit 113 Franken.	It will reimburse the cost of tests for SARS-Coronavirus-2 at a flat rate of 169 francs, and of antibody tests at a rate of 113 francs.	The Swiss Confederation pays a lump sum of CHF 169 for the test for SARS coronavirus-2 and CHF 113 for the detection of antibodies .	The Swiss federal government will pay a flat-rate fee of CHF 169 for the test for SARS-Coronavirus-2 and CHF 113 for the detection of antibodies .
107.	Dieses Verbot ist eine der zentralen Massnahmen, mit der die Zahl der Infektionen mit dem Coronavirus reduziert werden konnte.	The ban is one of the key measures that has helped to reduce the number of coronavirus infections in Switzerland.	This ban is one of the key measures to reduce the number of infections with coronavirus .	This ban is one of the key measures which has made it possible to reduce the number of coronavirus infections .
108.	Die Zahl der Neuinfektionen ist seit ein paar Wochen auf tiefem Niveau stabil, ebenso die Zahl der Hospitalisationen und der Todesfälle.	The number of new cases of infection has remained stable at a low level for a number of weeks now, as has the number of hospital admissions and deaths.	The number of new infections has remained stable at a low level for a few weeks, as have the number of hospitalisations and deaths.	The number of new cases of infection has remained stable at a low level for a few weeks, as has the number of hospitalisations and deaths.
109.	Diese Änderung für politische und zivilgesellschaftliche Kundgebungen gelten bereits ab Samstag, 20. Juni 2020.	This change concerning political and civil society demonstrations will apply from Saturday, 20 June.	This change for political and civil society rallies will apply from Saturday 20 June 2020.	This change for political and society demonstrations will apply from Saturday 20 June 2020.
110.	Sie sind zudem ebenfalls verpflichtet, Menschenansammlungen von mehr als fünf Personen in Pausenräumen und Kantinen zu verhindern.	Employers must limit the number of people present on construction sites or in other workplaces, and prevent groups of more than five people from gathering in break rooms and canteens.	They are also obliged to prevent the gathering of more than five people in break rooms and canteens.	They are also obliged to prevent groups of more than five people from gathering in break rooms and canteens.
111.	Sobald die Fallzahlen in der Schweiz ausreichend gesunken sind, werden die Kantone die konsequente Rückverfolgung von Infektionsketten wieder aufnehmen	As soon as the number of cases in Switzerland has fallen sufficiently, the cantons will resume the consistent tracing of infection chains	As soon as the number of cases in Switzerland has fallen sufficiently, the cantons will resume the consistent tracing of infection chains	As soon as the number of cases of people infected with the virus in Switzerland has fallen sufficiently, the cantons will resume the consistent tracing of infection chains
112.	infizierte Personen sollen frühzeitig entdeckt, behandelt und isoliert , die Übertragungsketten eruiert und weitere Übertragungen verhindert werden.	infected persons should be detected, treated and placed in isolation at an early stage, transmission chains should be identified and further transmissions prevented.	infected persons are to be detected, treated and isolated at an early stage, the transmission chains are to be determined and further transmissions prevented.	The aim is to detect, treat and place in isolation infected persons at an early stage, identify transmission chains and prevent further transmission .
113.	Ab wann Grossveranstaltungen wieder möglich sein werden, entscheidet er in einer seiner nächsten Sitzungen.	A decision as to when large-scale events will again be allowed to go ahead will be taken at one of its next meetings.	It will decide at one of its next meetings when major events will be possible again.	At one of its next meetings, the Council will decide when large-scale events will be possible again.
114.	Wird das Produkt in der Schweiz zugelassen, dient es in erster Linie der Behandlung von mit dem neuen Coronavirus infizierten Personen .	If the product is approved in Switzerland, it will primarily be used to treat people infected with the new coronavirus .	If the product is approved in Switzerland, it will primarily be used to treat people infected with the new coronavirus .	If the product is approved in Switzerland, it will primarily be used to treat people infected with the new Coronavirus .

115.	Im Sinne einer Anschubversorgung liefert die Armeeapotheke ab nächster Woche, während zwei Wochen, täglich eine Million Hygienemasken an führende Detailhändler.	As an initial impulse, the army pharmacy will be supplying one million hygienic masks to leading retailers every day for two weeks starting next week.	In order to provide an initial supply, the army pharmacy will supply one million hygiene masks daily to leading retailers from next week for a fortnight.	In order to provide an initial supply, the army pharmacy will supply one million hygienic masks daily to leading retailers from next week for a fortnight.
116.	COVID-19-Impfstoff: Bund unterzeichnet Vertrag mit AstraZeneca	COVID-19 vaccine: Swiss federal government signs agreement with AstraZeneca	COVID 19 vaccine: Federal government signs contract with AstraZeneca	COVID-19 vaccine: Swiss federal government signs contract with AstraZeneca
117.	Dazu zählt unter anderem die interdepartementale Abstimmung sowie die Abstützung auf wissenschaftliche Expertise. Entsprechend statten EDI und BAG die « Swiss National COVID-19 Science Task Force » mit einem angepassten Rahmenmandat aus.	These include interdepartmental coordination and drawing on scientific expertise. Accordingly, the FDHA and the FOPH has adapted the mandate of the Swiss National COVID-19 Science Task Force , the general objective being to obtain independent scientific advice from the team of experts activated by the SN-STF.	This includes interdepartmental coordination and relying on scientific expertise. Accordingly, EDI and FOPH are providing the " Swiss National COVID-19 Science Task Force " with an adapted framework mandate.	This includes interdepartmental coordination and relying on scientific expertise. Accordingly, FDHA and FOPH are providing the " Swiss National COVID-19 Science Task Force " with an adapted framework mandate.
118.	Mit den ersten Lockerungen zum Schutz der Bevölkerung vor dem neuen Coronavirus ist manches wieder erlaubt aber es gelten weiterhin die gleichen Regeln:	While various activities are now permitted with the first phase of measures to protect the public being eased , the same rules still apply :	With the first relaxations to protect the population from the new coronavirus , some things are allowed again - but the same rules still apply :	With the first ease measures to protect the public , some things are allowed again, but the same rules still apply :
119.	Abstand halten, gründlich Hände waschen, in die Armbeuge niesen	Keep your distance, wash your hands thoroughly, and sneeze into the crook of your arm.	Keep your distance, wash your hands thoroughly, sneeze into the crook of your arm.	Keep your distance, wash your hands thoroughly, sneeze into the crook of your arm
120.	Bern, 08.04.2020 - Die Massnahmen zur Bekämpfung des Coronavirus werden gut umgesetzt und zeigen Wirkung.	Bern, 08.04.2020 - The measures introduced to combat the coronavirus epidemic are being followed well and are having the desired effect.	Bern, 08.04.2020 - The measures to combat coronavirus are being well implemented and are showing results.	Bern, 08.04.2020 - The measures introduced to combat the Coronavirus are being well implemented and are proving effective.
121.	Ist die Präsenz vor Ort unabdingbar, muss er die betreffende Person schützen , indem er die Abläufe oder den Arbeitsplatz entsprechend anpasst.	If it is essential for someone to work on site, the employer must take steps to protect that person by adapting working processes or the workplace.	If the presence on site is indispensable, he must protect the person concerned by adapting the procedures or the workplace accordingly.	If a presence on the ground is essential, it must protect the person concerned by adapting procedures or the workplace accordingly.
122.	Ausgenommen von der Quarantänepflicht sind neu auch Kulturschaffende nach einem kulturellen Anlass, Sportlerinnen und Sportler nach einem Wettkampf sowie Teilnehmende von Fachkongressen.	Under the new rules, creative artists returning from a cultural event, athletes returning from competitions and persons who have been attending professional conferences will be exempted from the quarantine requirement .	Excluded from the quarantine obligation are now also cultural workers after a cultural event, sportsmen and women after a competition, and participants of specialist congresses.	Exempted from the quarantine requirement are now also cultural workers after a cultural event, sportsmen and women after a competition and participants in specialist congresses.
123.	Ihr Gesundheitszustand ist gut.	Their health condition is good.	Your state of health is good.	Their health condition is good.
124.	Modernas mRNA-Impfstoff setzt auf eine neuartige Technologie:	Moderna's mRNA vaccine is based on an innovative technology:	Modern mRNA vaccine is based on a novel technology:	Modern's mRNA vaccine is based on a novel technology:

125.	Die mRNA ist eine Art Botenmolekül , die die Bauanleitung zur Herstellung von Proteinen mit sich trägt.	the mRNA is a kind of messenger molecule that carries the building instructions for the production of proteins.	mRNA is a kind of messenger molecule that carries the building instructions for the production of proteins.	The mRNA is a kind of messenger molecule that carries the building instructions for the production of proteins.
126.	Die Spitäler müssen zudem einen ausreichenden Bestand an wichtigen Arzneimitteln sowohl für COVID-19-Patientinnen und Patienten als auch für weitere medizinisch dringende Behandlungen halten.	The hospitals must also keep a sufficient supply of important medical products for COVID-19 patients and for other urgent medical treatments.	Hospitals must also maintain a sufficient stock of essential medical drugs for COVID-19 patients and other medically urgent treatments.	Hospitals must also maintain a sufficient stock of essential medical products both for COVID-19 patients and for other medically urgent treatments.
127.	Bundesrat erlaubt Sport-Trainings ab 11. Mai	Federal Council permits sports training sessions from 11 May	Federal Council permits sports training from 11 May	Federal Council permits sports training from 11 May
128.	Der Bundesrat hebt in einem vierten Schritt die verbliebenen Einschränkungen per 22. Juni weitgehend auf.	In a fourth phase, the Federal Council is now lifting most of the remaining restrictions as of 22 June.	In a fourth step, the Federal Council largely lifts the remaining restrictions as of 22 June.	In a fourth step, the Federal Council will largely no longer apply the remaining restrictions as of 22 June.