

# Les séquences préfabriquées dans les productions écrites des étudiants du français L2 et du suédois L2

Susanna Salminen

Mémoire de master

Département de français, parcours enseignement et apprentissage des langues

Institut de langues et de traduction

Université de Turku

Mai 2020

*Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.*

UNIVERSITÉ DE TURKU

Institut de langues et de traduction / Faculté des Lettres

SALMINEN, SUSANNA : Les séquences préfabriquées dans les productions écrites des étudiants du français L2 et du suédois L2

Mémoire de master, 113 p. (24 p. d'annexes)

Département de français, parcours enseignement et apprentissage des langues

Mai 2020

-----

Dans ce mémoire de master, nous étudions les séquences préfabriquées (SP) dans les processus d'écriture et dans les produits écrits finaux des étudiants finnophones du français L2 et du suédois L2. Notre objectif a été d'examiner la quantité de séquences préfabriquées dans les productions écrites produites par quatre étudiants multilingues et de savoir s'il existe des différences dans l'utilisation des SP entre les deux langues. Nous nous sommes également intéressée à la qualité des séquences préfabriquées et nous avons voulu savoir quelles catégories des SP les étudiants utilisent dans leurs productions en français et en suédois. De plus, notre intention a été d'étudier la variation individuelle dans l'usage des SP et d'examiner les profils d'utilisateurs des étudiants.

Tous les quatre étudiants ont écrit un texte en français et un texte en suédois, et les productions écrites ont été recueillies par l'intermédiaire du logiciel *ScriptLog* qui est un logiciel d'enregistrement des frappes. A l'aide de ce logiciel, nous avons pu étudier non seulement le produit écrit final mais aussi le processus d'écriture. Les étudiants ont verbalisé leurs processus d'écriture tout de suite après les séances d'écriture, ce qui veut dire que notre corpus a compris les verbalisations rétrospectives avec un rappel stimulé. De plus, les étudiants ont rempli un questionnaire qui nous a offert les informations de base des étudiants.

L'analyse du corpus a révélé que la quantité de SP a été élevée, même de 30 % dans le produit final en français et de 24 % en suédois, et la différence entre les langues a été notable. Ces pourcentages ont toutefois été semblables aux pourcentages présentés dans les études précédentes. Quant aux catégories des SP utilisées, les catégories des SP lexicales et des SP discursives ont été les plus grandes dans les deux langues, mais les occurrences des SP ont varié dans les catégories. Nous avons pu distinguer les profils d'utilisateurs individuels, et l'analyse individuelle a montré que les SP automatisées facilitent le processus d'écriture. Etant consciente du fait que le nombre de participants a été petit dans cette étude, ces résultats ne peuvent pas être généralisés. Effectivement, les séquences préfabriquées dans les productions écrites en langues étrangères représentent un sujet qui mériterait plus d'études ultérieures.

Mots-clés : séquence préfabriquée, production écrite, processus rédactionnel, langue étrangère, enregistrement de frappe

## TABLE DES MATIÈRES

|   |            |
|---|------------|
| <b>1. Introduction .....</b>  | <b>1</b>   |
| 1.1. Objet d'étude .....  | 2          |
| 1.2. Plan des chapitres .....   | 4          |
| <b>2. Les processus d'écriture en L2 et le rôle des séquences préfabriquées .....</b> | <b>4</b>   |
| 2.1. Les modèles de description des processus d'écriture en L2 .....                  | 5          |
| 2.2. La charge cognitive pendant le processus d'écriture .....                        | 7          |
| 2.3. Séquences préfabriquées et automatisation des processus cognitifs .....          | 8          |
| <b>3. Les séquences préfabriquées .....</b>   | <b>10</b>  |
| 3.1. Le phénomène des séquences préfabriquées .....                                   | 10         |
| 3.2. La terminologie liée aux séquences préfabriquées .....                           | 11         |
| 3.2.1. Définitions .....  | 14         |
| 3.2.2. L'approche linguistique et l'approche psycholinguistique .....                 | 16         |
| 3.3. La catégorisation des séquences préfabriquées .....                              | 19         |
| <b>4. Corpus et méthode d'analyse .....</b>   | <b>24</b>  |
| 4.1. Le profil des participants basé sur le questionnaire .....                       | 25         |
| 4.2. Les productions écrites à l'aide du logiciel <i>ScriptLog</i> .....              | 27         |
| 4.3. Les verbalisations .....   | 30         |
| 4.4. Méthode d'analyse .....  | 31         |
| <b>5. Analyse .....</b>   | <b>33</b>  |
| 5.1. La quantité de séquences préfabriquées dans les productions écrites .....        | 33         |
| 5.1.1. Les productions en français .....  | 34         |
| 5.1.2. Les productions en suédois .....   | 36         |
| 5.1.3. En somme .....   | 38         |
| 5.2. Les catégories des séquences préfabriquées utilisées .....                       | 43         |
| 5.2.1. Les catégories des SP utilisées en français .....                              | 44         |
| 5.2.2. Les catégories des SP utilisées en suédois .....                               | 48         |
| 5.2.3. En somme .....   | 52         |
| 5.3. Les profils d'utilisateurs dans l'usage des séquences préfabriquées .....        | 54         |
| 5.3.1. L'étudiant A .....   | 55         |
| 5.3.2. L'étudiant B .....   | 68         |
| 5.3.3. L'étudiant C .....   | 81         |
| 5.3.4. L'étudiant D .....   | 92         |
| 5.3.5. En somme .....   | 104        |
| <b>6. Conclusion .....</b>  | <b>106</b> |
| <b>BIBLIOGRAPHIE .....</b>  | <b>110</b> |
| <b>ANNEXES .....</b>  | <b>i</b>   |
| Annexe 1. Questionnaire de base .....   | i          |
| Annexe 2. Listes des SP utilisées par chaque étudiant en français et en suédois ..... | ix         |
| Annexe 3. Explications des marques utilisées dans des exemples .....                  | xvi        |
| Annexe 4. Résumé en finnois – Suomenkielinen tiivistelmä .....                        | xvii       |

## Liste des tableaux

|   |    |
|---|----|
| Tableau 1. Les termes utilisés dans le domaine du langage préfabriqué .....                                   | 13 |
| Tableau 2. Les catégories et les exemples de Forsberg (2006) et de Rasmussen (2015) .....                     | 20 |
| Tableau 3. Les catégories et les exemples de Cordier (2013).....  | 20 |
| Tableau 4. Les catégories et les exemples de Jaworska, Krummes et Ensslin (2015).....                         | 21 |
| Tableau 5. Les catégories et les exemples de Sundman (2011).....  | 22 |
| Tableau 6. Catégorisation des SP dans notre étude .....   | 23 |
| Tableau 7. Les informations de base des participants.....   | 26 |
| Tableau 8. Les nombres de mots produits et de mots faisant partie des SP au total en français                 | 35 |
| Tableau 9. Les mots produits et les mots faisant partie des SP au niveau de chaque étudiant en français ..... | 35 |
| Tableau 10. Les nombres de mots produits et de mots faisant partie des SP au total en suédois                 | 37 |
| Tableau 11. Les mots produits et les mots faisant partie des SP au niveau de chaque étudiant en suédois.....  | 38 |
| Tableau 12. Le nombre de SP différentes et leurs occurrences dans les productions en français                 | 44 |
| Tableau 13. Les SP ayant au moins deux occurrences dans les productions en français .....                     | 45 |
| Tableau 14. La distribution des SP en français en catégories.....   | 46 |
| Tableau 15. Le nombre de SP différentes et leurs occurrences dans les productions en suédois                  | 48 |
| Tableau 16. Les SP ayant au moins deux occurrences dans les productions en suédois .....                      | 49 |
| Tableau 17. La distribution des SP en suédois en catégories.....  | 50 |
| Tableau 18. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant A .....                 | 55 |
| Tableau 19. La distribution des SP en français en catégories ; l'étudiant A.....                              | 56 |
| Tableau 20. Les SP les plus utilisées en français par l'étudiant A.....                                       | 56 |
| Tableau 21. Les types des SP discursives en français de l'étudiant A.....                                     | 57 |
| Tableau 22. Les types de SP lexicales en français de l'étudiant A.....  | 59 |
| Tableau 23. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant A .....                  | 62 |
| Tableau 24. La distribution des SP en suédois en catégories ; l'étudiant A .....                              | 62 |
| Tableau 25. Les SP les plus utilisées en suédois par l'étudiant A .....                                       | 63 |
| Tableau 26. Les types de SP lexicales en suédois de l'étudiant A.....   | 64 |
| Tableau 27. Les types de SP discursives en suédois de l'étudiant A.....                                       | 65 |
| Tableau 28. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant B.....                  | 69 |
| Tableau 29. La distribution des SP en français en catégories ; l'étudiant B.....                              | 69 |
| Tableau 30. Les SP les plus utilisées en français par l'étudiant B.....                                       | 70 |
| Tableau 31. Les types de SP lexicales en français de l'étudiant B .....                                       | 70 |
| Tableau 32. Les types de SP discursives en français de l'étudiant B .....                                     | 72 |
| Tableau 33. Les types de SPIDI en français de l'étudiant B.....   | 73 |
| Tableau 34. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant B .....                  | 75 |
| Tableau 35. La distribution des SP en suédois en catégories ; l'étudiant B .....                              | 76 |
| Tableau 36. Les SP les plus utilisées en suédois par l'étudiant B .....                                       | 77 |
| Tableau 37. Les types de SP lexicales en suédois de l'étudiant B.....   | 78 |
| Tableau 38. Les types de SP discursives en suédois de l'étudiant B .....                                      | 79 |
| Tableau 39. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant C.....                  | 82 |
| Tableau 40. La distribution des SP en français en catégories ; l'étudiant C.....                              | 82 |
| Tableau 41. Les SP les plus utilisées en français par l'étudiant C.....                                       | 83 |
| Tableau 42. Les types de SP lexicales en français de l'étudiant C .....                                       | 83 |

|   |     |
|---|-----|
| Tableau 43. Les types de SP discursives en français de l'étudiant C .....                     | 85  |
| Tableau 44. Les types de SP modifiées en français de l'étudiant C .....                       | 86  |
| Tableau 45. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant C .....  | 87  |
| Tableau 46. La distribution des SP en suédois en catégories ; l'étudiant C .....              | 88  |
| Tableau 47. Les SP les plus utilisées en suédois par l'étudiant C .....                       | 88  |
| Tableau 48. Les types de SP lexicales en suédois de l'étudiant C .....                        | 89  |
| Tableau 49. Les types de SP grammaticales en suédois de l'étudiant C .....                    | 90  |
| Tableau 50. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant D ..... | 93  |
| Tableau 51. La distribution des SP en français en catégories ; l'étudiant D.....              | 93  |
| Tableau 52. Les SP les plus utilisées en français par l'étudiant D.....                       | 94  |
| Tableau 53. Les types de SP discursives en français de l'étudiant D .....                     | 95  |
| Tableau 54. Les types de SP lexicales en français de l'étudiant D .....                       | 96  |
| Tableau 55. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant D .....  | 99  |
| Tableau 56. La distribution des SP en suédois en catégories ; l'étudiant D .....              | 99  |
| Tableau 57. Les SP les plus utilisées en suédois par l'étudiant D .....                       | 100 |
| Tableau 58. Les types de SP lexicales en suédois de l'étudiant D.....                         | 101 |
| Tableau 59. Les types de SP discursives en suédois de l'étudiant D.....                       | 102 |

### **Liste des figures**

|  |    |
|--|----|
| Figure 1. Le problème terminologique dans le domaine des séquences préfabriquées .....                 | 12 |
| Figure 2. Les nombres de mots produits et de mots faisant partie des SP .....                          | 39 |
| Figure 3. Les nombres de mots produits par chaque étudiant.....  | 40 |
| Figure 4. La quantité de SP selon les étudiants individuels.....                                       | 41 |
| Figure 5. La distribution des SP en français en catégories dans le processus.....                      | 47 |
| Figure 6. La distribution des SP en français en catégories dans le produit final .....                 | 47 |
| Figure 7. La distribution des SP en catégories dans le processus en suédois .....                      | 51 |
| Figure 8. La distribution des SP en catégories dans le produit final en suédois .....                  | 51 |
| Figure 9. La quantité de SP dans chaque catégorie dans les productions en français et en suédois ..... | 52 |

### **Liste des illustrations**

|   |    |
|---|----|
| Illustration 1. Exemple d'un extrait du produit textuel final de l'étudiant C ..... | 29 |
| Illustration 2. Exemple d'un extrait du texte linéaire de l'étudiant C .....        | 29 |
| Illustration 3. Exemple des statistiques du texte en français de l'étudiant C ..... | 30 |

### **Liste des captures d'écran**

|  |    |
|--|----|
| Capture d'écran 1. Le processus d'écriture en français de l'étudiant C ..... | 31 |
|--|----|

## 1. Introduction

Le cadre théorique de ce mémoire de maîtrise s'inscrit dans le domaine des études des langues étrangères. Plus précisément, cette étude se situe aussi bien dans le domaine des activités rédactionnelles que dans le domaine des séquences préfabriquées, car nous étudierons le rôle que jouent les séquences préfabriquées (par exemple, des suites de mots comme *il y a, faire du sport, c'est*) dans les processus d'écriture et dans les produits écrits des apprenants multilingues.

Comme les séquences préfabriquées (SP) se situent au centre de notre intérêt, il est essentiel de s'attarder d'abord sur leur définition. La terminologie dans le domaine des SP est très riche et nous donnerons des définitions plus précises plus tard (*cf.* chapitre 3.2.), mais pour le moment nous nous contentons de dire qu'il s'agit des suites de mots qui semblent être soudées, c'est-à-dire qu'elles sont apprises et gardées en mémoire comme un tout, globalement (Wray, 2002, p. 9). Par exemple, les suites de mots comme *parler de, avoir X ans, je (ne) sais pas* et *beaucoup de* sont des séquences préfabriquées que l'on trouve souvent chez des apprenants universitaires du français langue étrangère (L2) (Forsberg, 2006, p. 84, 87, 91).

Nous étudierons donc les séquences préfabriquées aussi bien dans les processus d'écriture que dans les produits écrits finaux. Pour cette raison, nous présenterons des modèles de description des processus d'écriture, dont celui, bien connu, de Hayes et Flower (1980). Dans leur modèle, les deux chercheurs proposent que les processus d'écriture sont composés de trois processus principaux, à savoir de la planification (*planning*), de la mise en texte (*translating*) et de la révision (*reviewing*) (*id.*, p. 12). Dans notre travail, il est important de traiter les modèles de description des processus d'écriture parce que nous voulons indiquer de quelle manière les séquences préfabriquées peuvent faciliter la production écrite (*cf.* aussi Perez-Bettan, 2015 ; Cordier, 2013).

Avant de présenter les objets d'étude de ce mémoire, nous tenons à préciser que nous utiliserons les termes de *séquence préfabriquée (SP)* et de *langage préfabriqué* comme synonymes (*cf.* chapitre 3.2.).

## 1.1. Objet d'étude

Les séquences préfabriquées sont étudiées aussi bien dans le langage des locuteurs natifs (L1) que dans le langage des locuteurs non-natifs (L2), et de même, elles sont étudiées à l'oral et à l'écrit (Myles & Cordier, 2017 ; Perez-Bettan, 2015). Cependant, les études portant sur les SP dans la langue parlée semblent être plus nombreuses (*cf.* chapitre 3.1.) et, par conséquent, cela nous a menée à choisir d'étudier la langue écrite, et notamment celle en L2 pour compléter la connaissance de leur utilisation par les apprenants de langue.

Dans ce mémoire, notre intention est ainsi de trouver comment les étudiants multilingues utilisent les séquences préfabriquées dans les productions écrites en français L2 et en suédois L2. Nous nous intéressons aussi bien à la quantité totale de SP dans le corpus qu'à la quantité de SP dans les productions de chaque étudiant. Egalement, nous voulons savoir si la quantité est différente entre les deux langues. En ce qui concerne la qualité des SP, à savoir la catégorisation des SP (*cf.* chapitre 3.3.), notre but est de trouver quels types les étudiants utilisent et quelles sont des différences dans l'utilisation des deux langues. Comme il est possible pour nous d'étudier non seulement les produits écrits finaux mais aussi les processus d'écriture à l'aide du logiciel *ScriptLog* qui enregistre l'activité du scripteur (*cf.* chapitre 4.2.), nous nous intéressons à la variation individuelle que les étudiants ont dans l'usage des séquences préfabriquées dans leurs processus de rédaction.

En suivant les objectifs de recherche, nos questions de recherche sont ainsi les suivantes :

1. Quelle est la quantité de séquences préfabriquées dans les productions écrites des étudiants du français L2 et du suédois L2 ? Existe-il des différences entre les deux langues ?
2. Quelles catégories des séquences préfabriquées les étudiants utilisent-ils ? Utilisent-ils les mêmes catégories dans les deux langues ?
3. Quels sont les profils d'utilisateurs dans l'usage des séquences préfabriquées des étudiants individuels ?

En ce qui concerne notre première hypothèse, nous supposons que la quantité de SP dans les productions écrites est d'environ 25 %. Cette hypothèse se base sur les études

de Forsberg (2006), de Cordier (2013), de Rasmussen (2015), qui ont trouvé que la quantité de SP dans la production orale en L2 a été de 25 % à 27 %, et sur l'étude d'Erman et Warren (2000), qui ont trouvé que la quantité de SP dans les produits écrits en L1 a été même d'environ 52 %. Comme les textes que les étudiants ont écrits sont assez courts dans notre étude et comme nous nous concentrons sur la L2, nous estimons que la quantité de SP n'atteint pas celle présentée par Erman et Warren (ibid.). Il est aussi noté qu'en général, la proportion de SP en L2 est plus petite que celle en L1 (Perez-Bettan, 2015). De plus, nous supposons que la quantité totale de SP chez les étudiants est similaire dans les deux langues, c'est-à-dire en français et en suédois, car les étudiants ont des compétences langagières assez semblables dans les deux langues (*cf.* 4.1.).

Notre deuxième hypothèse comprend que les étudiants utilisent au moins les catégories des SP qui sont présentées dans les études antérieures sur la L2, c'est-à-dire les SP lexicales, discursives et grammaticales (Forsberg, 2006 ; Rasmussen, 2015). Nous estimons également que nous pouvons trouver dans notre corpus des SP qui forment une nouvelle catégorie à part dans le cadre de ce mémoire (*cf.* Rasmussen, 2015). Nous supposons en outre que les catégories des séquences préfabriquées que nous trouverons peuvent varier selon la langue et selon l'individu.

Finalement, l'hypothèse pour la dernière question est que nous pouvons distinguer des profils d'utilisateurs individuels et que la manière d'un étudiant d'utiliser des SP dans les deux langues peut varier. Nous basons cette hypothèse sur Hayes et Flower (1980, p. 19-20) qui proposent que des scripteurs différents utilisent de manières diverses lors de la rédaction ainsi que sur la supposition que les processus d'écriture semblent contenir beaucoup de variations individuelles (entre autres, Mutta, 2007).

Pour vérifier ces hypothèses et répondre à nos questions de recherche, nous nous baserons sur des textes écrits par quatre étudiants finnophones de l'Université de Turku qui étudient le français et le suédois, soit comme matière principale soit comme matière(s) secondaire(s). En conséquence, les étudiants ont tous écrit deux textes, un en français et un en suédois, que nous analyserons en nous appuyant sur des théories et sur des concepts fondamentaux dans les domaines.

Même si la plupart des études que nous utiliserons dans ce mémoire concernent la langue parlée, nous pouvons supposer que les SP, qui sont présentes dans le langage

d'un locuteur, soient présentes aussi bien à l'oral qu'à l'écrit, et en raison de cela, il nous est convenable d'utiliser les études existantes. Une autre justification pour l'utilisation de ces études est le fait que nous pouvons examiner si des catégories des SP qui sont présentées dans les recherches existantes se réalisent aussi dans notre étude. Il est ainsi possible que notre analyse du corpus révèle une catégorie des SP entièrement nouvelle. En ce qui concerne le choix des langues dans cette étude, c'est-à-dire le français et le suédois, nous y reviendrons quand nous présenterons le corpus dans le chapitre 4.

## **1.2. Plan des chapitres**

Dans notre mémoire, nous commencerons le cadre théorique en nous focalisant sur les modèles de description des processus d'écriture. Nous présenterons les modèles notamment du point de vue de la L2. Après cela, nous traiterons des aspects que nous offre l'étude du processus d'écriture. Nous continuerons en faisant le lien entre les séquences préfabriquées et l'automatisation des processus cognitifs.

Ensuite, à partir du chapitre 3, nous introduirons les séquences préfabriquées de manière plus exhaustive. Nous donnerons un aperçu sur les SP en tant que phénomène et après cela, nous nous intéresserons à la terminologie riche liée aux SP. Nous donnerons également les définitions pour les termes et nous présenterons la distinction entre l'approche linguistique et l'approche psycholinguistique. A la fin de la partie théorique, nous présenterons les catégorisations différentes des séquences préfabriquées et la catégorisation que nous utiliserons.

Après le cadre théorique, nous passerons à la présentation du corpus de cette étude. Nous couvrirons aussi bien la méthode de collecte des données que la méthode d'analyse. Ensuite, à partir du chapitre 5, nous commencerons à présenter l'analyse des productions écrites et nous avancerons dans l'ordre des questions de recherche. A la fin de ce mémoire, nous résumerons les résultats dans la conclusion.

## **2. Les processus d'écriture en L2 et le rôle des séquences préfabriquées**

Nous commencerons ce chapitre en décrivant des modèles de description des processus cognitifs d'écriture. Ensuite, nous traiterons l'étude du processus d'écriture du point de vue de la charge cognitive. A la fin, nous aborderons l'automatisation des processus, et nous montrerons quel rôle les séquences préfabriquées jouent dans cette automatisation.

## 2.1. Les modèles de description des processus d'écriture en L2

Dans le domaine des activités rédactionnelles, un modèle fondamental qui présente le processus de rédaction est celui de Hayes et Flower (1980), modèle créé pour décrire l'utilisation de la langue maternelle (voir aussi Mutta, 2007, p. 39-40, 46). Dans ce modèle, Hayes et Flower (1980, p. 12) proposent trois processus principaux qui forment le processus de rédaction : **la planification** (*planning*), **la mise en texte** (*translating*) et **la révision** (*reviewing*). Mutta (2007, p. 39) résume l'essentiel de ces processus de la manière suivante :

« [L]a planification permet d'établir des plans globaux de contenus et de traitements pour effectuer la réalisation du texte, la mise en texte développe chaque partie des plans et traduit linguistiquement le contenu sémantique de ces plans, et la révision tend à améliorer la qualité du texte écrit. »

Hayes et Flower (1980, p. 12) développent leur modèle à l'aide de sous-catégories, c'est-à-dire qu'ils suggèrent que le processus de la planification comprend également la génération (*generating*), l'organisation (*organizing*) et l'établissement de buts (*goal setting*), et que le processus de la révision comporte la lecture (*reading*) et la correction (*editing*). Quant au processus de la mise en texte, les auteurs n'ont pas prévu des sous-catégories pour lui. Pour décrire les sous-catégories nommées, nous nous appuyons sur la synthèse présentée par Mutta (2007, p. 39) :

### La planification

- La génération : « permet de récupérer depuis la mémoire à long terme des informations nécessaires pour effectuer la tâche rédactionnelle »
- L'organisation : « permet de sélectionner le matériel le plus approprié, récupéré par les processus de génération, et de l'organiser en plan »
- L'établissement de buts : « permet de créer les critères avec lesquels le texte est jugé en fonction des objectifs de production »

### La révision

- La lecture : « permet de lire le texte déjà écrit »
- La correction : « permet de repérer et corriger les violations de conventions d'écriture et les erreurs sémantiques, et d'évaluer l'adéquation entre le texte écrit et les buts déterminés à l'avance »

Le modèle de Hayes et Flower a reçu des critiques (Mutta, 2007, p. 41), et Hayes et Flower eux-mêmes (1980, p. 29) ont remarqué qu'il est possible que le processus

d'écriture ne se passe pas toujours comme leur théorie le représente, parce qu'il existe aussi d'autres démarches de rédaction. En effet, Roca de Larios, Manchón, Murphy et Marín (2008, p. 31) trouvent que le processus d'écriture est caractérisé par le dynamisme, ce qui veut dire que les scripteurs passent constamment d'un processus à l'autre. Par exemple, ils passent de la planification à l'évaluation (*i.e.* la révision), puis à la transcription (*i.e.* la mise en texte), et inversement (*ibid.*). Néanmoins, ce modèle, selon lequel la rédaction se compose de trois processus cognitifs principaux, est resté populaire, car nous pouvons trouver cette tripartition dans d'autres recherches dans le domaine (Mutta, 2007, p. 41 ; *cf.* aussi Cislaru & Olive, 2018a, 2018b). Selon Mutta (2007, p. 59), parmi les modèles d'écriture existants, celui de Hayes et Flower est, en effet, le modèle le plus adapté et le plus élaboré.

Nous continuons par les modèles d'écriture concernant la langue étrangère. Selon Ellis et Yuan (2004, p. 61), il a été reconnu dans des recherches antérieures qu'il n'existe ni de théorie de l'écriture en L2 universellement acceptée, ni de théorie qui concerne spécifiquement l'écriture en L2. Cependant, Ellis et Yuan (*ibid.*) trouvent qu'il existe des théories (par exemple celles citées par Ellis & Yuan [*ibid.*], à savoir Bereiter & Scardamalia, 1987 ; Flower & Hayes, 1980 ; Grabe & Kaplan, 1996 ; Kellogg, 1996 ; Zimmermann, 2000) qui représentent des modèles des processus d'écriture qui sont similaires entre eux. Ce fait mène à l'idée généralement acceptée que le processus d'écriture en L1 est semblable à celui en L2 (*ibid.*). (*cf.* aussi Mutta, 2007, p. 47 ; Mutta, 2020, p. 64.) Également, Roca de Larios *et al.* (2008, p. 31) indiquent que les modèles des processus d'écriture en L1 sont largement utilisés en L2. En disant cela et en nous basant sur Ellis et Yuan (2004) et sur Mutta (2007), nous pouvons nous servir du modèle d'écriture de Hayes et Flower (1980) dans notre étude qui concerne l'écriture en langue étrangère. Remarquons qu'il est évident que de nouveaux modèles plus élaborés sont apparus depuis les années 80, mais pour l'étendue de notre étude, le modèle de Hayes et Flower est suffisant comme un modèle de base en ce qui concerne l'écriture.

Dans ce chapitre nous avons présenté le modèle de description des processus d'écriture de Hayes et Flower (1980) qui est un modèle fondamental dans le domaine. Nous avons aussi discuté le cas des modèles des processus d'écriture en L2, et nous pouvons conclure que les modèles sont semblables en L1 et en L2. Dans le chapitre suivant, nous traiterons le processus d'écriture du point de vue de la charge cognitive.

## 2.2. La charge cognitive pendant le processus d'écriture

Traditionnellement, la recherche portant sur la production écrite a concerné des produits textuels finaux, complètement terminés, et cette méthode a été presque la seule à être utilisée (Mutta, 2007, p. 59). Quant à l'étude du produit final, l'évaluation de la production est faite par le lecteur, par exemple en examinant la cohérence et les fautes dans le texte (Mutta, 2020, p. 64 ; *cf.* aussi Lintunen, Mutta et Peltonen, 2020, p. 8). Pourtant, depuis des années 80, on a pris une nouvelle direction, car des chercheurs ont commencé à s'intéresser aux processus cognitifs dans l'activité d'écriture (Mutta, 2007, p. 38, 59). Le modèle de Hayes et Flower (1980), que nous avons présenté dans le chapitre précédent, est en effet considéré comme étant pionnier dans ce domaine s'intéressant au processus d'écriture (*ibid.*). Comme dans notre étude le processus d'écriture joue un rôle aussi important que le produit écrit, il nous est donc essentiel de traiter la question des nouveaux aspects que nous offre l'étude du processus.

Lorsque l'on étudie le processus d'écriture, on est souvent confronté aux pauses. Selon Mutta (2007, p. 52-53), qui se base sur plusieurs recherches, il est généralement présumé que les pauses indiquent la charge cognitive qui est liée aux processus que l'apprenant traite au moment de l'écriture (*cf.* aussi Cislaru & Olive, 2018b, p. 227). Les pauses sont liées à la fluence verbale, et les pauses sont donc étudiées en analysant leur fréquence, leur durée et leur localisation, car ces points « indiquent pas à pas les opérations de gestion des processus cognitifs » (Mutta, 2007, p. 61, 63). D'après Mutta (*id.*, p. 61), les études portant sur les pauses dans les productions écrites ont été moins nombreuses dans le passé que celles des productions orales (par exemple, dû au manque de moyens technologiques), mais aujourd'hui nous pouvons étudier les pauses dans la rédaction par exemple à l'aide des logiciels qui permettent l'enregistrement des actions du scripteur, comme le logiciel *ScriptLog* que nous utiliserons dans notre étude.

Mutta (2007, p. 52-53) présente, en se basant sur les recherches antérieures, la localisation des pauses selon la durée de la pause, et souligne le fait que, selon les recherches, nous trouvons les pauses les plus courtes à l'intérieur d'un mot. Entre deux mots, la durée des pauses est plus longue, et ensuite entre deux phrases, la durée est encore plus longue (*ibid.*). La durée des pauses est donc la plus longue entre deux paragraphes, au moins dans cette présentation de l'ordre, mais il faut noter que nous trouvons des pauses assez longues également au début et à la fin de la rédaction (*ibid.*). Au début, l'apprenant prend le temps pour la planification initiale et, à la fin, il relit et

révise le texte pour la dernière fois (ibid.). (cf. aussi Cislaru & Olive, 2018a, p. 38-40.) Spelman Miller (2000, p. 137) remarque également que le scripteur en L2 fait des pauses à l'intérieur d'un mot plus souvent que le scripteur en L1, et cela indique que l'apprenant se concentre par exemple sur l'orthographe ou sur la formation de mots, ou qu'il a des doutes sur ces sujets. En général, nous pouvons dire que les pauses faites par le scripteur en L2 sont plus longues que celles faites par le scripteur en L1, dans n'importe quel endroit (Spelman Miller, 2000, p. 137 ; cf. aussi Mutta, 2020, p. 64). Comme Mutta (2007, p. 61) le dit, les longues pauses dans la production indiquent des traitements cognitifs « longs et coûteux ». En outre, selon Roca de Larios, Manchón et Murphy (2006, p. 108), les scripteurs utilisent deux fois plus de temps dans la rédaction en L2 qu'en L1, et cela présente le fait que les scripteurs en L2 font de grands efforts pour obtenir le niveau qu'ils veulent avoir dans le texte.

Dans ce chapitre, nous avons traité quelques aspects que nous offre l'étude du processus d'écriture et nous les avons examinés du point de vue de la charge cognitive. Nous continuerons en faisant le lien entre l'automatisation des processus cognitifs et les séquences préfabriquées.

### **2.3. Séquences préfabriquées et automatisation des processus cognitifs**

Selon Coirier, Gaonac'h et Passerault (1996, p. 163), les activités de compréhension et de production des textes exigent la mise en œuvre d'opérations diverses, qui nécessitent encore l'utilisation de ressources cognitives importantes. Ils continuent en disant qu'il semble que ce système ne puisse pas fonctionner, si « chacune des opérations qui doivent être mises en œuvre nécessite qu'un contrôle explicite et coûteux soit exercé sur leur réalisation » (id., p. 164). Cela veut dire que certaines opérations devraient être réalisées de manière automatisée, c'est-à-dire sans l'engagement de ressources cognitives coûteuses (ibid.). Selon les auteurs (ibid.), « une telle automatisation [...] permet au système de fonctionner de manière efficace et relativement économique. »

Comme le fonctionnement de la mémoire de travail est fortement lié à l'automatisation des processus (Mutta, 2007, p. 51), notons que la capacité de la mémoire de travail est limitée et, par conséquent, elle ne peut pas traiter une trop large quantité d'information à la fois (Cordier, 2013, p. 57 ; Baddeley, 2000, p. 418). D'après Mutta (2007, p. 51), pour que la mémoire de travail fonctionne efficacement lors de la rédaction, les processus de l'encodage et de la récupération linguistique devraient être automatisés (au

moins dans une certaine mesure). Si certains processus cognitifs ne sont pas automatisés, les activités d'écriture ne peuvent pas se passer de manière fluide, car « le coût cognitif augmente, surcharge la mémoire de travail et perturbe la gestion de divers processus cognitifs » (ibid.). Autrement dit, les processus automatisés sont essentiels pour que la production verbale soit aisée (Olkkonen & Mutta, 2020, p. 36). Coirier *et al.* (1996, p. 209) remarquent également que les processus qui sont en général très automatisés en langue maternelle « peuvent redevenir des contraintes importantes » dans l'utilisation d'une langue étrangère parce qu'elle suscite une situation considérablement coûteuse au plan cognitif.

En ce qui concerne les séquences préfabriquées dans la production verbale (*cf.* chapitre 3.2.), on affirme en général que l'utilisation des séquences préfabriquées est « un moyen aisé de s'exprimer avec un minimum d'effort » (Perez-Bettan, 2015, p. 21). L'aisance qui découle de l'utilisation des SP concerne aussi bien les locuteurs natifs que les locuteurs non-natifs, et nous la trouvons à l'oral comme à l'écrit (id., p. 29, 55). Selon Perez-Bettan (id., p. 21), qui fait référence à Fillmore (1979)<sup>1</sup>, il n'est pas possible pour l'apprenant de progresser dans une langue étrangère s'il ne maîtrise pas les séquences préfabriquées. Mais pourquoi le langage préfabriqué facilite-t-il la production verbale de cette manière ? Nous trouvons la réponse par exemple dans la définition de la séquence préfabriquée proposée par Cordier (2013, p. 21), car selon cette auteure, dans le traitement langagier, les SP fonctionnent à l'avantage d'un locuteur parce qu'elles sont soit gardées en mémoire comme un tout, soit extrêmement **automatisées**. Comme nous l'avons évoqué, l'automatisation des processus limite la charge de la mémoire de travail, et par conséquent, l'utilisation des SP est un grand avantage pour le locuteur (Perez-Bettan, 2015, p. 55).

Les séquences préfabriquées jouent donc un rôle essentiel dans la production langagière en L2, car elles aident l'apprenant à s'exprimer de manière plus aisée et elles diminuent la charge cognitive (Perez-Bettan, 2015, p. 25, 22). Selon Perez-Bettan (id., p. 55), qui se base sur les études antérieures, les séquences préfabriquées rendent la compréhension et la production langagière plus faciles aussi bien en L1 qu'en L2. Underwood, Schmitt et Galpin (2004, p. 153) remarquent que la production verbale est facilitée par l'utilisation des SP parce qu'elles peuvent être récupérées de la mémoire avec un minimum d'effort, car elles sont gardées en mémoire comme un tout. Comme Perez-

---

<sup>1</sup> « A large portion of a person's ability to get along in a language consists in the mastery of formulaic utterances. » (Fillmore, 1979, p. 92) [cité par Perez-Bettan, 2015, p. 21].

Bettan (2015, p. 21) l'a également évoqué, l'utilisation des séquences préfabriquées exige moins d'effort de l'apprenant dans la production langagière, et cette utilisation libère donc du temps que l'apprenant peut utiliser pour la planification (id., p. 25). L'auteure (id., p. 22) note que nous pouvons voir l'importance de ces avantages notamment en langue étrangère.

Dans ce chapitre, nous avons souligné l'importance de l'automatisation des processus cognitifs dont les processus rédactionnels, et nous avons jeté un coup d'œil sur les séquences préfabriquées que nous traiterons de manière plus exhaustive dans les chapitres suivants.

### **3. Les séquences préfabriquées**

Nous commencerons ce chapitre en présentant le phénomène des séquences préfabriquées en général. Nous continuerons par la présentation de la terminologie riche liée aux SP. Par la suite, nous donnerons des définitions concernant les SP et nous ferons ainsi la distinction entre l'approche linguistique et l'approche psycholinguistique. A la fin de ce chapitre, nous présenterons des catégorisations des SP et nous montrerons la catégorisation que nous utiliserons dans cette étude.

#### **3.1. Le phénomène des séquences préfabriquées**

Les théories linguistiques sur lesquelles se fonde l'étude des séquences préfabriquées sont essentiellement la grammaire générative de Chomsky (1965), la grammaire de construction de Fillmore et al. (1988) et les deux principes de Sinclair (1991), à savoir le principe du libre choix (*open-choice principle*) et le principe idiomatique (*idiom principle*) (cités par Perez-Bettan, 2015, p. 31-32 ; Rasmussen, 2015, p. 4-7). En ce qui concerne le phénomène des séquences préfabriquées, il est devenu un sujet d'intérêt bien avant les années 80 et, en fait, l'on trouve des termes qui ont été utilisés pour référer au langage préfabriqué déjà dans les années 60 (Perez-Bettan, 2015, p. 29, 36).

L'intérêt pour ce phénomène a varié au fil des années, mais au cours des dernières décennies, la recherche autour des séquences préfabriquées a été étendue et elle a concerné aussi bien les locuteurs natifs que les locuteurs L2 (Myles & Cordier, 2017, p. 4 ; Perez-Bettan, 2015, p. 29 ; cf. aussi Nyqvist, 2013, p. 54). Les approches que l'on trouve dans la recherche sur les séquences préfabriquées sont diverses, s'inscrivant par exemple dans le cadre de la psycholinguistique, de la linguistique formelle, de la

linguistique de corpus et de la pragmatique (Myles & Cordier, 2017, p. 4). Par exemple, pour des chercheurs dans le cadre de la pragmatique, le contexte interactionnel dans lequel nous pouvons trouver des séquences préfabriquées est un sujet d'étude principal (Cordier, 2013, p. 6).

Selon Perez-Bettan (2015, p. 33, 35), il est reconnu que la préfabrication joue un rôle large et important dans la langue, non seulement à l'oral mais aussi à l'écrit, et il concerne aussi bien le langage d'un locuteur L1 que celui d'un locuteur L2 (la proportion en L2 est néanmoins plus petite). Par exemple, selon les résultats d'Erman et Warren (2000, p. 37), les séquences préfabriquées comprennent environ 58 % à l'oral et 52 % à l'écrit en L1. Perez-Bettan (2015, p. 35) note aussi que le phénomène des séquences préfabriquées est trouvé dans plusieurs langues qui sont souvent assez éloignées entre elles, comme l'anglais et l'hébreu. Elle propose donc qu'en raison de cela, nous puissions parler d'un phénomène universel qui n'est pas lié à une seule langue (ibid.).

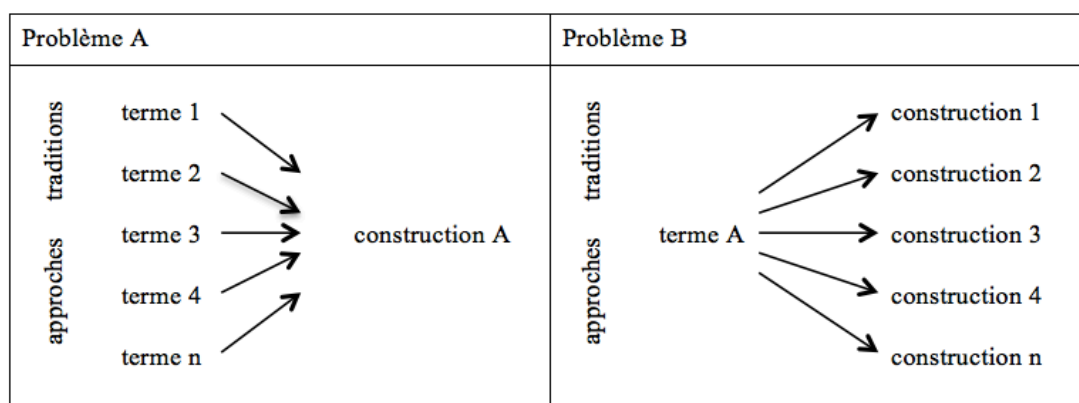
Les séquences préfabriquées sont ainsi présentes à l'oral et à l'écrit, mais la plupart des études que nous connaissons et des études que nous utiliserons dans ce mémoire traitent seulement les séquences préfabriquées dans la langue parlée (par exemple celles de Wray, 2002 ; Forsberg, 2006 ; Cordier, 2013 ; Perez-Bettan, 2015 ; Rasmussen, 2015). Néanmoins, on constate que la maîtrise des séquences préfabriquées améliore aussi la production écrite d'un scripteur L1 ou L2 (Perez-Bettan, 2015, p. 29), comme nous l'avons déjà évoqué dans les chapitres précédents. Par exemple, les chercheurs suivants ont étudié les SP dans la langue écrite, à savoir Erman et Warren (2000), Conklin et Schmitt (2008), Sundman (2011) et Jaworska, Krummes et Ensslin (2015). Nous reviendrons à ces dernières études quand nous présenterons la catégorisation des SP dans le chapitre 3.3.

Comme les séquences préfabriquées sont étudiées selon des approches différentes, la diversité dans les recherches cause le problème que la terminologie utilisée dans le domaine est très riche (Myles & Cordier, 2017, p. 4), ce que nous aborderons plus dans le chapitre suivant.

### **3.2. La terminologie liée aux séquences préfabriquées**

Le problème lié à la terminologie utilisée dans le domaine des séquences préfabriquées est souvent discuté dans les recherches et par exemple Wray (2002), Forsberg (2006),

Cordier (2013), Perez-Bettan (2015) et Myles et Cordier (2017) attirent l'attention sur cela (*cf.* aussi Siyanova-Chanturia & Pellicer-Sánchez, 2019). Selon Wray (2002, p. 8), la terminologie dans le domaine est riche, mais les phénomènes derrière les termes ne sont pas toujours les mêmes. Myles et Cordier (2017, p. 4) expliquent ce fait par ce qu'il existe plusieurs termes différents qui se réfèrent à la même construction (problème A dans la figure 1), mais qu'il est aussi possible que le même terme se réfère à des constructions différentes (problème B). Nous avons créé la figure 1 pour illustrer ce problème terminologique qui est persistant dans le domaine.



**Figure 1. Le problème terminologique dans le domaine des séquences préfabriquées**

A l'instar des autres chercheurs, Perez-Bettan (2015, p. 35-36) affirme que le problème est que les chercheurs du domaine étudient différents types de séquences et de plus, ils n'utilisent pas les mêmes termes pour ces types différents. Perez-Bettan remarque aussi que les fonctions et les critères d'identification peuvent varier selon les types de séquences (*ibid.*), ce qui cause plus de confusion terminologique. Notons alors ici que la terminologie riche reflète au fait que les sujets d'études principaux varient d'une approche à l'autre (Cordier, 2013, p. 2).

A ces affirmations, nous pouvons ajouter un problème terminologique lié aux traditions qui est présenté par Forsberg (2006, p. 3). Selon cette chercheuse (*ibid.*), il existe au moins deux traditions différentes dans le domaine des séquences préfabriquées, à savoir la tradition française et la tradition anglo-saxonne, et à cause de cela, la tâche de trouver une terminologie appropriée est difficile. Quant à ces traditions, d'après Forsberg (*ibid.*), les études dans la tradition française se sont beaucoup concentrées sur le figement en linguistique générale et, les études dans la tradition anglo-saxonne se sont majoritairement concentrées sur les fréquences et sur les collocations.

Pour concrétiser la terminologie riche en anglais et en français, Wray (2002, p. 9) et Perez-Bettan (2015, p. 36), par exemple, proposent des listes exhaustives des termes utilisés dans le domaine du langage préfabriqué. Dans le tableau récapitulatif ci-dessous, nous présenterons ces listes, et de plus, les termes qui sont mentionnés dans les études et dans les œuvres sur lesquelles nous nous appuyons dans notre mémoire. Notons que les termes présentés ne sont pas forcément des synonymes exacts entre eux.

**Tableau 1. Les termes utilisés dans le domaine du langage préfabriqué**

| Termes français   |  |
|---|--|
| Forsberg (2006)   | séquence préfabriquée, expression idiomatique, expression figée, stéréotype, locution, collocation, figement, préfabrication, unités non-analysées, langage préfabriqué  |
| Rasmussen (2015)  | langage préfabriqué, séquence préfabriquée, expression figée, locution, stéréotype, expression idiomatique   |
| Perez-Bettan (2015)                                     | langage stéréotypé, langage préfabriqué, langage préconstruit, expression figée, expression idiomatique, construction idiomatique, suite de mots   |
| Liste de termes présenté par Perez-Bettan (2015, p. 36) | phrasillons (Tesnière, 1965), holophrases (Corder, 1973), gambits (Keller, 1979), composites (Cowie, 1988), phrasèmes (Mel'čuk, 1995), séquences figées (Mejri, 2004), unités lexicales préconstruites (Grossmann et Tutin, 2003), séquences préfabriquées (Forsberg, 2006), unité de construction préformée (Schmäle, 2013)   |
| Sitri et Tutin (2016)                                   | séquence préfabriquée, « prêtes à parler », « prêtes à écrire », suites de formes et de lemmes, langage formulaire, préfabrication, collocation, segment répété, blocs lexicaux, figement  |
| Termes anglais  |  |
| Wray (2002)   | formulaic sequence, formulaic language, formulaic, formulaic word string, formulaic word, formulaic morpheme   |
| Liste de termes présenté par Wray (2002, p. 9)          | amalgams, automatic, chunks, clichés, co-ordinate constructions, collocations, complex lexemes, composites, conventionalized forms, fixed expressions including idioms, fixed expressions, formulaic language, formulaic speech, formulas/formulae, fossilized forms, frozen metaphors, frozen phrases, gambits, gestalt, holistic, holophrases, idiomatic, idioms, irregular, lexical simplex, lexical(ized) phrases, lexicalized sentence stems, listemes, multiword items/units, multiword lexical phenomena, noncompositional, noncomputational, nonproductive, nonpropositional, petrifications, phrasemes, praxons, preassembled speech, precoded conventionalized routines, prefabricated routines and patterns, ready-made expressions, ready-made utterances, recurring utterances, rote, routine formulae, schemata, semipreconstructed phrases that constitute single choices, sentence builders, set phrases, stable and familiar expressions with specialized subsenses, stereotyped phrases, stereotypes, stock utterances, synthetic, unanalyzed chunks of speech, unanalyzed multiword chunks, units |
| Forsberg (2006)   | chunks, prefabs, fixed expressions, formulae, gambits, prefabricated routines, prefabricated patterns, collocation   |
| Cordier (2013)  | formulaic sequence, formulaic strings, prefab, prefabricated chunks, formulaic multiword strings   |

Aux traditions française et anglo-saxonne, Forsberg (2006, p. 3) ajoute la tradition acquisitionnelle, qui est aussi considérée comme étant d'origine anglo-saxonne. Pour

souligner quelques termes présentés dans le tableau 1 ci-dessus, notons que selon Forsberg (ibid.), en linguistique française, les termes les plus souvent utilisés sont, entre autres : *expression idiomatique*, *expression figée*, *stéréotype* et *locution*. En revanche, en linguistique anglo-saxonne et acquisitionnelle, les termes fréquemment employés sont par exemple *chunk*, *prefab*, *fixed expression*, *formulae* et *collocation* (ibid.). En ce qui concerne le terme de *séquence préfabriquée* que nous avons utilisé dans ce mémoire, c'est une traduction du terme de Wray (2002), c'est-à-dire *formulaic sequence*, que Forsberg (2006, p. 3-4 ; 2008, p. 13-14) a présenté. Forsberg (2006, p. 3-4) a opté pour ce terme parce qu'il a une portée très vaste (cf. chapitre 3.2.1.), et en effet, le *formulaic sequence* de Wray est devenu une sorte de terme parapluie dans le domaine au fil des années (Myles & Cordier, 2017 ; voir aussi le chapitre 3.2.2.).

Nous avons outre Forsberg (2006) choisi d'utiliser le terme de *séquence préfabriquée* dans notre étude, comme l'ont fait entre autres Perez-Bettan (2015) et Rasmussen (2015), et nous donnerons des justifications plus précises pour notre choix dans les deux chapitres suivants où nous présenterons les définitions des termes.

### 3.2.1. Définitions

En quoi consiste en effet une séquence préfabriquée ? Wray (2002, p. 4) emploie le terme *formulaic* pour parler des mots et des séquences de mots qui sont traités sans une analyse profonde sur la composition des mots. Selon Wray (id., p. 9), le mot *formulaic* peut être associé à l'unité et à l'habitude, tandis que le mot *sequence* indique qu'il y existe plus d'une seule unité. Par conséquent, la définition que Wray (id., p. 9) donne pour le terme de *formulaic sequence* est la suivante :

« A sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. »

En étudiant des séquences préfabriquées dans la production orale des apprenants du français langue étrangère, Forsberg (2006, p. 3-4) a opté pour ce terme de Wray et elle explique que les autres termes, comme *expression figée* ou *stéréotype*, ont une portée trop restreinte pour décrire le phénomène qu'elle étudie. Par contre, la portée de la *séquence préfabriquée* est très vaste (ibid.), ce qui est souhaitable également dans notre étude. La *séquence préfabriquée* n'est donc pas définie par des critères stricts, et Forsberg évoque que « [l]es phénomènes et les termes visés par la tradition française

sont inclus sous le terme de SP, mais tout ce qui rentre dans le terme de SP ne rentre pas dans les termes français » (ibid.). Une raison qui explique cela est le fait que la *séquence préfabriquée* comprend aussi des séquences qui semblent être préfabriquées parce qu'elles sont trop complexes à l'égard du niveau de langue de l'apprenant (par exemple l'expression *en ce qui concerne* dans une production d'un apprenant débutant) ou parce que les effets pragmatiques sont trop forts (par exemple des marqueurs de discours comme *tu sais, en effet*) (ibid.). Autrement dit, selon Forsberg (id., p. 43), la définition de Wray (2002, p. 9) est tellement extensive en raison du fait qu'il est important que tous les phénomènes « qui semblent acquis en tant qu'unités et qui ne sont pas générés par la mise en œuvre de règles créatives » soient inclus sous un même terme.

Forsberg (2006, p. 44) a également formulé sa propre définition pour la séquence préfabriquée, qu'elle a rédigée d'après celles de Wray (2002, p. 9) et d'Erman et Warren (2000, p. 31) qui ont donné une définition concernant seulement les séquences du répertoire natif. Pour cela, avant de passer à la définition de Forsberg, il nous paraît important de présenter aussi la définition proposée par Erman et Warren (ibid.), même si nous nous concentrons seulement sur les locuteurs non-natifs. Nous trouvons donc leur définition dans ce qui suit :

« A prefab is a combination of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization. » (Erman & Warren, 2000, p. 31)

En conséquence, Forsberg définit donc la séquence préfabriquée de la manière suivante :

« Une séquence préfabriquée est composée d'au moins deux mots graphiques. Elle est préférée, c'est-à-dire plus fréquente dans un contexte donné, chez les locuteurs natifs, qu'une autre combinaison qui aurait pu être équivalente si la première n'avait pas été conventionnalisée. Or le répertoire des séquences préfabriquées chez les locuteurs L2 diverge de celui des locuteurs natifs. Par conséquent, une séquence préfabriquée peut aussi être une séquence préférée par un locuteur L2, identifiée à cause de son emploi idiosyncrasique (morphosyntaxique, sémantique, pragmatique) et récurrent dans l'interlangue d'un locuteur L2. » (2006, p. 44)

Comme nous pouvons le voir dans cette citation, Forsberg (ibid.) inclut dans sa définition non seulement les séquences du répertoire des locuteurs L1, mais aussi celles du répertoire des locuteurs L2. Selon Forsberg (ibid.), cette définition indique qu'il faut prendre en compte le contexte où les séquences préfabriquées apparaissent, car on peut

les identifier par rapport à un contexte donné. Nous trouvons ce fait par exemple dans l'étude de Rasmussen (2015, p. 26), car cette auteure a trouvé dans son corpus des séquences récurrentes qui sont fortement liées à la thématique de l'étude (par exemple *matière secondaire et à l'école* en parlant des études), et par conséquent, Rasmussen les regroupe sous une catégorie des séquences préfabriquées thématiques. Les séquences préfabriquées sont en effet souvent liées à des contextes spécifiques et les mots qui construisent ces séquences peuvent avoir des combinaisons différentes dans d'autres contextes (Forsberg, 2006, p. 44). Forsberg (ibid.) démontre cet aspect par l'exemple d'*un dossier*, qui se combine avec le verbe *déposer* dans un contexte administratif, tandis que dans un contexte universitaire ou scolaire, *un dossier* est employé avec le verbe *rendre*.

Dans notre étude, il est également important de noter le contexte, car nous pouvons nous attendre à ce que, dans les productions écrites des étudiants, il existe des séquences préfabriquées qui sont liées aux sujets d'écriture. En raison de cela et en raison de la présence de l'emploi idiosyncrasique des SP chez un locuteur L2, le terme de séquence préfabriquée et la définition proposée par Forsberg (2006, p. 44) nous paraissent convenables pour notre étude, même si Forsberg a étudié les SP dans la langue parlée. Notamment, un point important pour nous dans l'étude de Forsberg est le fait que nous trouvons des productions verbales des apprenants de FLE dans son corpus. Cependant, nous avons aussi besoin d'une définition qui concerne uniquement des locuteurs L2, et par conséquent, nous continuerons à présenter d'autres définitions dans le chapitre suivant.

### **3.2.2. L'approche linguistique et l'approche psycholinguistique**

Comme nous l'avons déjà indiqué dans le chapitre 3.2., Cordier (2013, p. 2) ainsi que Myles et Cordier (2017, p. 4) reconnaissent le problème terminologique dans le domaine des séquences préfabriquées et selon elles, le fait que parfois le même terme se réfère à des constructions différentes est un problème plus grand. D'après Cordier (2013, p. 2) et Myles et Cordier (2017, p. 5), ce problème concerne notamment le terme de Wray (2002), c'est-à-dire *formulaic sequence*, car ce terme est adopté par plusieurs chercheurs et la manière d'utiliser le terme est devenue très générique – en effet, *formulaic sequence* (ou séquence préfabriquée) est devenu un terme parapluie.

Depuis la parution de son œuvre bien connue de 2002, Wray a élaboré ses idées concernant les séquences préfabriquées. Cela veut dire que Wray (2008, p. 11) fait la distinction entre (1) ce qui est préfabriqué en général dans la langue et (2) ce qui est préfabriqué dans le langage d'un individu ou d'un groupe particulier. Ainsi, selon Wray (ibid.), il faut remarquer que ce qui est préfabriqué pour un locuteur ne l'est pas forcément pour l'autre. Selon Cordier (2013, p. 1, 18 ; cf. aussi Myles & Cordier, 2017), la distinction faite par Wray représente deux approches différentes, à savoir (1) l'approche linguistique centrée sur la langue (*speaker-external*) et (2) l'approche psycholinguistique centrée sur le locuteur (*speaker-internal*). Cordier (ibid.) se concentre dans son étude sur les séquences préfabriquées psycholinguistiques et elle donne, en s'appuyant sur la distinction proposée par Wray, une explication exhaustive qui montre en quoi cette distinction entre les deux approches est essentielle :

« [...] it is not because a sequence is a semantic unit or occurs in corpora at a high level of frequency that it will necessarily have a psycholinguistic reality for a particular individual. Conversely, a multiword sequence which presents a processing advantage for a given language user might not be a frequent bundle or in the case of an L2 learner, it could even be a sequence which is grammatically incorrect. » (Cordier, 2013, p. 18)

Nous pouvons voir dans cette citation que Cordier (ibid.) est en accord avec Wray (2008), c'est-à-dire avec l'idée que ce qui est considéré préfabriqué varie selon le locuteur. Cordier (2013, p. 18) rappelle aussi qu'il est possible que les deux approches se chevauchent.

La raison pour laquelle Myles et Cordier (2017, p. 4, 5, 12) voient l'utilisation générique de *formulaic sequence* comme un problème est le fait qu'elles reconnaissent également la distinction entre les deux approches et elles la considèrent essentielle. Pour donner des descriptions plus étendues de ces approches, nous nous appuyons sur les caractéristiques présentées par Myles et Cordier (2017). Selon les auteures, l'approche linguistique centrée sur la langue (*speaker-external*) concerne ce qui est préfabriqué en général dans la langue cible, par exemple des collocations et des expressions idiomatiques (id., p. 4, 5). Il s'agit donc du phénomène linguistique en soi, indépendamment du locuteur (id., p. 12). Quant à l'approche psycholinguistique centrée sur le locuteur (*speaker-internal*), elle comprend ce qui est préfabriqué dans le langage du locuteur/de l'apprenant (c'est-à-dire quels mots il traite comme une unité dans son langage) et qui, par conséquent, est à l'avantage de l'apprenant dans le processus

d'apprentissage (id., p. 4). Dans cette approche, il s'agit donc d'un processus cognitif propre au locuteur (id., p. 12).

Comme Myles et Cordier (2017, p. 12) font la distinction entre ces deux types différents de séquences préfabriquées, elles proposent qu'au lieu d'utiliser le terme de *formulaic sequence* comme un terme générique, nous devrions utiliser deux termes distincts (notamment quand il s'agit de la L2) : les séquences préfabriquées en soi, externes à l'apprenant (*learner-external formulaic sequences*), c'est-à-dire les regroupements linguistiques (*linguistic clusters*), et les séquences préfabriquées propres à l'apprenant (*learner-internal formulaic sequences*), c'est-à-dire les unités de traitement (*processing units*). Selon Myles et Cordier (id., p. 5, 12), ces deux types des séquences préfabriquées représentent des phénomènes différents au niveau conceptuel et, par conséquent, il est nécessaire que nous étudions ces phénomènes séparément, particulièrement dans un contexte L2. Pour ces nouveaux termes, Myles et Cordier (id., p. 12) proposent leur propres définitions, et notons que la définition qu'elles donnent aux unités de traitement est beaucoup inspirée par la définition que Cordier (2013, p. 21)<sup>2</sup> a donné pour la séquence préfabriquée psycholinguistique<sup>3</sup> dans son étude. Les définitions présentées par Myles et Cordier (2017, p. 12) sont donc les suivantes :

Les regroupements linguistiques (les SP externes à l'apprenant) :

« multimorphemic clusters which are either semantically or syntactically irregular, or whose frequent co-occurrence gives them a privileged status in a given language as a conventional way of expressing something ».

Les unités de traitement (les SP propres à l'apprenant) :

« a multiword semantic/functional unit that presents a processing advantage for a given speaker, either because it is stored whole in their lexicon or because it is highly automatized ».

Etant donné que Wray (2008) a fait la distinction entre l'approche linguistique et l'approche psycholinguistique, Cordier (2013, p. 19) et Myles et Cordier (2017, p. 19), entre autres, pensent que la définition de *formulaic sequence* donnée par Wray (2002)

---

<sup>2</sup> « A formulaic sequence (FS) is a multiword semantic / functional unit that presents a processing advantage for a given speaker, either because it might be stored whole in their lexicon or because it is highly automatized. » (Cordier, 2013, p. 21)

<sup>3</sup> Dans l'étude de Cordier (2013, p. 1), elle utilise les SP psycholinguistiques (*psycholinguistic FS*) et les SP propres au locuteur (*speaker-internal FS*) comme synonymes. De même, Cordier (ibid.) utilise les SP externes au locuteur (*speaker-external FS*), les SP linguistiques (*linguistic FS*) et les SP idiomatiques (*idiomatic FS*) comme synonymes.

représente la définition psycholinguistique (*speaker-internal*) la plus largement utilisée dans le domaine.

Pour finir la partie terminologique, nous présenterons la définition, ou plutôt, les définitions de SP que nous utiliserons dans cette étude. Comme nous l'avons évoqué dans le chapitre précédent, nous utilisons en général le terme de *séquence préfabriquée* et la définition proposée par Forsberg (2006) est très convenable pour nous. Néanmoins, pour que nous ayons une définition qui concerne spécifiquement un apprenant individuel d'une langue, nous choisissons d'utiliser aussi la définition que Myles et Cordier (2017) proposent pour les unités de traitement (les SP propres à l'apprenant). Cette définition nous semble convenable parce qu'elle comprend toutes les séquences qui sont préfabriquées notamment dans le langage de l'apprenant et elle souligne aussi le fait que les SP sont à l'avantage dans les processus cognitifs. Dans ce mémoire, nous utiliserons donc les deux définitions et nous nous appuierons sur elles lors du travail d'analyse. Dans le chapitre suivant, nous continuerons en présentant les catégories différentes des séquences préfabriquées.

### **3.3. La catégorisation des séquences préfabriquées**

Comme nous l'avons vu, le domaine des séquences préfabriquées comprend des approches et des définitions diverses, et nous pouvons dire que la variété dans le nombre des catégorisations possibles des SP ne fait pas d'exception. Entre autres, Cordier (2013, p. 145) propose que les SP peuvent être catégorisées soit selon la forme grammaticale d'une séquence préfabriquée, c'est-à-dire que l'on se concentre à la forme, soit selon la fonction communicative d'une séquence préfabriquée. Dans l'étude de Forsberg (2006, p. 46), la catégorisation concerne principalement la forme des SP, car les catégories utilisées sont les SP lexicales, les SP grammaticales, les SP discursives, les SP situationnelles et les SP idiosyncrasiques intérimaires, c'est-à-dire fautives/erronées ; notons ici que Forsberg base les trois premières catégories sur la catégorisation d'Erman et Warren (2000). Pour nous, cette catégorisation est principalement formelle parce que les SP discursives peuvent aussi représenter une catégorie fonctionnelle, comme nous verrons dans la suite de ce chapitre. Rasmussen (2015, p. 22) utilise les mêmes catégories que Forsberg (2006), mais elle y ajoute les SP thématiques liées à la discussion sur les études de français. Toutes les catégories mentionnées sont encore divisées en plusieurs sous-catégories, mais pour le moment nous nous contentons de traiter seulement les catégories principales. Dans le tableau 2

ci-dessous, nous présentons les catégories et des exemples de Forsberg (2006) et de Rasmussen (2015).

**Tableau 2. Les catégories et les exemples de Forsberg (2006) et de Rasmussen (2015)**

|                     | <b>Catégorie</b>                                | <b>Exemples</b>   |
|---------------------|---|---|
| Forsberg<br>(2006)  | Les SP Lexicales                                | <i>autre chose, par semaine, avoir X ans</i>  |
|                     | Les SP Grammaticales                            | <i>beaucoup de, quelqu'un, pas du tout</i>  |
|                     | Les SP Discursives                              | <i>c'est, je crois, parce que</i>   |
|                     | Les SP Situationnelles                          | <i>je m'appelle, je ne comprends pas, répétez s'il vous plaît</i>                             |
|                     | Les SP Idiosyncrasiques<br>intérimaires (SPIDI) | <i>c'est très chaud (il fait très chaud), <u>au premier</u> je dois dire que (en premier)</i> |
| Rasmussen<br>(2015) | Les SP Thématiques                              | <i>langue maternelle, à l'écrit, à l'université</i>   |

Forsberg (2006, p. 48) ainsi que Cordier (2013, p. 145) remarquent qu'il n'est pas toujours facile de classer une séquence préfabriquée dans une seule catégorie, car elle peut avoir des traits qui pourraient faire partie de plusieurs catégories. Cordier (ibid.) a choisi d'utiliser une catégorisation où les séquences préfabriquées sont classées non seulement selon la forme grammaticale mais aussi selon la fonction communicative. Dans cette catégorisation présentée par Cordier (ibid.), nous trouvons trois grandes catégories, à savoir les SP référentielles (*referential FS*), les SP métadiscursives (*meta-discursive FS*) et les SP qui construisent des phrases (*sentence builders FS*). Ces trois catégories comprennent également plusieurs sous-catégories, et dans le tableau 3 ci-dessous nous montrons des exemples des séquences que les trois catégories peuvent inclure.

**Tableau 3. Les catégories et les exemples de Cordier (2013)**

|                | <b>Catégorie</b>                    | <b>Exemples</b>  |
|----------------|-------------------------------------|--|
| Cordier (2013) | Les SP Référentielles               | <i>je suis d'accord, c'est intéressant, l'année dernière, à l'université, transports en commun, parler anglais</i> |
|                | Les SP Métadiscursives              | <i>je sais pas, je crois, à mon avis, par exemple, en fait, et puis</i>  |
|                | Les SP Qui construisent des phrases | <i>je pense que, il y a, je m'appelle, en train de + verbe, beaucoup de, avoir besoin de</i>                       |

Selon Cordier (2013, p. 151), sa catégorisation peut être comparée à celle de Forsberg (2006), car les SP référentielles correspondent aux SP lexicales, une partie des sous-catégories des SP qui construisent des phrases correspondent aux SP grammaticales, tandis qu'une partie de ces SP en plus des SP métadiscursives correspondent aux SP discursives.

Comme les catégorisations présentées en haut ont été utilisées dans les études qui concernent les séquences préfabriquées dans la langue parlée, il nous paraît important de donner des exemples des catégorisations employées dans les études de la langue écrite. Jaworska, Krummes et Ensslin (2015) ont étudié les séquences préfabriquées dans des essais écrits en allemand par des étudiants avancés d'allemand L2 et par des étudiants natifs. Les chercheurs classent les séquences préfabriquées selon la fonction des SP<sup>4</sup> et les catégories utilisées sont donc les moyens discursifs (*discourse-structuring devices*), les séquences référentielles (*reference-marking sequences*), les expressions de positionnement (*stance expressions*) et les expressions liées au sujet d'essai (*topic-related expressions*) (Jaworska *et al.*, 2015, p. 514). Ces catégories ne comprennent pas de sous-catégories, et dans le tableau 4 ci-dessous nous présentons collectivement les catégories et leurs exemples (en anglais, car les chercheurs ont étudié la langue allemande et ils ont traduit quelques exemples en anglais) (Jaworska *et al.*, 2015, p. 515-518).

**Tableau 4. Les catégories et les exemples de Jaworska, Krummes et Ensslin (2015)**

|                                     | Catégorie                              | Exemples   |
|-------------------------------------|--|--|
| Jaworska, Krummes et Ensslin (2015) | Les moyens discursifs                  | <i>there is, in this essay, for example, all in all, on the other hand</i> |
|                                     | Les séquences référentielles           | <i>the whole day, in the past, for the future, at uni</i>                  |
|                                     | Les expressions de positionnement      | <i>in my opinion, it is clear, one could say</i>                           |
|                                     | Les expressions liées au sujet d'essai | Les séquences copiées du sujet d'essai : <i>for society</i>                |

Cette catégorisation est un bon exemple pour montrer que l'on peut approcher des séquences préfabriquées d'un point de vue différent. Pourtant, notons que par exemple

<sup>4</sup> Jaworska *et al.* (2015, p. 514) utilisent le terme de *formulaic sequence*, mais ils utilisent aussi les termes de *3-gram types* et de *3-gram tokens* pour parler des séquences préfabriquées. Nous référons à tous les termes par la séquence préfabriquée.

les séquences référentielles sont très semblables aux SP référentielles de Cordier (2013). Une autre catégorisation qui traite des séquences préfabriquées dans les productions écrites est celle de Sundman (2011). Sundman (2011, p. 329) a étudié les SP dans les textes écrits en suédois par des élèves finnophones au collège. La catégorisation que Sundman (2011, p. 332) utilise est semblable à celle de Forsberg (2006), car elle comprend les phrases nominales (*substantivfraser*), les phrases prépositionnelles (*prepositionsfraser*), les expressions de temps (*tidsuttryck*) et les phrases verbales (*verbfraser*). Dans le tableau 5 qui suit, les catégories et leurs exemples sont groupés.

**Tableau 5. Les catégories et les exemples de Sundman (2011)**

|                | Catégorie                     | Exemples  |
|----------------|-------------------------------|---|
| Sundman (2011) | Les phrases nominales         | <i>mina kompisar, min familj, det bästa jag vet</i><br>‘mes amis’, ‘ma famille’, ‘le meilleur que je sais’                  |
|                | Les phrases prépositionnelles | <i>med mina kompisar, i skolan, på sommaren</i><br>‘avec mes amis’, ‘à l’école’, ‘en été’                                   |
|                | Les expressions de temps      | <i>varje dag, x gånger i veckan, i framtiden</i><br>‘tous les jours’, ‘x fois par semaine’, ‘dans l’avenir’                 |
|                | Les phrases verbales          | <i>spela musik, gilla att resa, åka skidor, tycka om</i><br>‘jouer de la musique’, ‘aimer voyager’, ‘faire du ski’, ‘aimer’ |

Même si les apprenants du suédois L2 dans l’étude de Sundman (2011) ont un niveau de langue plus bas que les étudiants dans notre étude et, par conséquent, les séquences préfabriquées sont plus simples, nous trouvons que cette catégorisation et les exemples peuvent nous aider dans notre travail de catégorisation. En effet, cela nous emmène à présenter la catégorisation que nous utiliserons dans cette étude.

Comme notre objet d’étude est de trouver la quantité et les types différents des séquences préfabriquées, il nous semble convenable que nous prenions un point de vue principalement grammatical en ce qui concerne la catégorisation, à savoir que nous nous concentrons plutôt à la forme de SP. Nous nous basons donc notre catégorisation largement à celles de Forsberg (2006) et de Rasmussen (2015), et aussi de Sundman (2011). La catégorisation de Forsberg est utile pour nous particulièrement parce que les séquences préfabriquées idiosyncrasiques intérimaires (c’est-à-dire les SPIDI) sont incluses dans les catégories, et nous estimons que nous les trouverons aussi dans notre corpus. Cependant, nous pensons que la catégorie des SP situationnelles n’est pas

pertinente pour nous parce que cette catégorie concerne la situation de l'interview et la langue parlée dans l'étude de Forsberg. Dans le tableau 6 ci-dessous, nous montrons alors la catégorisation des SP que nous avons choisie.

**Tableau 6. Catégorisation des SP dans notre étude**

| Catégorie                        | Sous-catégorie   | Abréviation | Exemples (fr, sué)   |
|----------------------------------|--|-------------|--|
| SP Lexicales                     |  | L           |  |
|                                  | Syntagmes nominaux   | SN          | Etudes en échange, tid för att 'temps pour'                              |
|                                  | Syntagmes adjectivaux                                      | SAdj        | Väldigt bra 'très bon'   |
|                                  | Syntagmes adverbiaux                                       | SAdv        | A l'université, per vecka 'par semaine'                                  |
|                                  | Syntagmes verbaux  | SV          | Faire des études, ha tillräckligt med 'avoir assez de'                   |
|                                  | Syntagmes prépositionnels                                  | SPré        | A cause de   |
| SP Discursives                   |  | D           |  |
|                                  | Marqueurs textuels   | MT          | Il y a, till exempel 'par exemple'                                       |
|                                  | Marqueurs autorégulateurs et expressions de positionnement | MAEP        | Je pense que, enligt min åsikt 'à mon avis'                              |
|                                  | Marqueurs interactifs                                      | MI          | Bien sûr   |
| SP Grammaticales                 |  | G           |  |
|                                  | Expressions de quantité et de degré                        | QD          | Beaucoup de, för mycket 'trop beaucoup'                                  |
|                                  | Pronoms indéfinis  | PI          | Pour quelqu'un, någon annan 'quelqu'un autre'                            |
| SP Idiosyncrasiques intérimaires |  | SPIDI       |  |
|                                  | Lexicales  | L           | <u>Tous</u> les <u>cho</u> ces, på <u>sommarna</u> 'en été'              |
|                                  | Grammaticales  | G           | Tellement <u>beaucoup</u>  |
|                                  | Discursives  | D           | <u>Il est aussi le</u> possibilité que, <u>a min asivkt</u> 'à mon avis' |
| SP Thématiques                   |  | T           |  |
|                                  | Syntagmes adverbiaux                                       | SAdv        | En France, i Finland 'en Finlande'                                       |
|                                  | Syntagmes nominaux   | SN          | En français  |
| SP Individuelles                 |  | I           |  |
|                                  | Toutes les catégories sont possibles                       |             | Ce pays, extra jobbet 'extra travail'                                    |
| SP Modifiées                     |  | M           |  |
|                                  | Toutes les catégories sont possibles                       |             | Dans le cas → dans le meilleur cas, avsikter om → åsikter om 'avis'      |

Comme nous pouvons le voir dans ce tableau, les catégories des SP lexicales, discursives, grammaticales et SPIDI viennent de la catégorisation de Forsberg (2006).

Nous avons choisi des sous-catégories de Forsberg qui sont pertinentes pour nous, mais nous avons modifié une sous-catégorie, à savoir « marqueurs autorégulateurs et expressions de positionnement ». La sous-catégorie de Forsberg (ibid.) s'intitule « marqueurs autorégulateurs », mais nous avons voulu que cette catégorie comprenne toutes les expressions d'opinion et de positionnement en général parce que telles expressions sont fréquentes dans notre corpus. Par conséquent, une catégorie de Jaworska *et al.* (2015) nous a inspirée et nous avons ajouté « expressions de positionnement » à cette sous-catégorie. La catégorie et les sous-catégories des SP thématiques sont de Rasmussen (2015), et dans notre étude, les SP T veulent dire les SP qui sont liées aux sujets d'écriture, mais qui ne seraient pas autrement classifiées comme les SP. Comme nous l'avons estimé (*cf.* chapitre 1.1.), nous formons deux nouvelles catégories dans le cadre de notre corpus. Les SP individuelles comprennent des séquences qui paraissent comme les SP pour un apprenant individuel (*cf.* Myles & Cordier, 2017), mais qui n'ont pas autrement un statut de SP dans le corpus, par exemple *ce pays* et *man måste 'on doit'*. En ce qui concerne les SP modifiées, elles incluent les séquences qui apparaissent comme les SP dans le processus, mais qui sont modifiées dans le produit final.

Après la présentation de notre catégorisation des SP, nous continuerons à traiter le corpus et la méthode d'analyse de cette étude.

#### **4. Corpus et méthode d'analyse**

Notre corpus est composé de quatre parties : les processus d'écriture, les productions écrites, les verbalisations et le questionnaire. Les productions écrites et les processus d'écriture seront l'objet d'analyse principal, tandis que les verbalisations et le questionnaire nous donnent des informations qui serviront pour le travail d'analyse. Nous avons effectué le recueil du corpus en novembre et en décembre 2018. Les productions écrites (*i.e.* le processus et le produit fini) ont été recueillies par l'intermédiaire du logiciel *ScriptLog*, qui est un logiciel d'enregistrement des frappes, avec lequel il est possible d'enregistrer des touches du clavier et des mouvements de la souris. Pour créer le questionnaire de base, nous avons utilisé *Webropol* qui nous a permis de recueillir des réponses en ligne. Les verbalisations ont été réalisées à l'aide du logiciel *QuickTime Player*. Nous avons donc recueilli notre corpus auprès de quatre étudiants qui ont tous rempli le questionnaire, écrit un texte en français et un texte en suédois, et verbalisé leurs processus d'écriture.

Dans les chapitres suivants, nous présenterons plus en détail les participants de notre étude, le logiciel *ScriptLog* et les matériaux de l'étude. Nous traiterons également la méthode d'analyse de cette étude.

#### **4.1. Le profil des participants basé sur le questionnaire**

Nous avons décidé de réaliser notre étude auprès d'apprenants au niveau universitaire, car nous pouvons nous attendre à ce qu'ils aient des compétences linguistiques assez avancées pour écrire les textes comportant des séquences préfabriquées. Nous basons notre hypothèse sur le fait que les étudiants ont dû passer un examen d'entrée pour pouvoir étudier ces langues au niveau universitaire et le niveau des compétences linguistiques qu'on attend dans cet examen est au moins le niveau B1<sup>5</sup>.

Dans cette étude de cas, nous avons donc quatre participants qui sont des étudiants âgés de 22 à 26 ans. Ils étudient à l'Université de Turku et ils ont tous le finnois comme langue maternelle. Deux étudiants ont comme matière principale le français et les deux autres, l'anglais. Tous les quatre ont le suédois comme matière secondaire et les étudiants ayant l'anglais comme matière principale étudient aussi le français comme matière secondaire. De plus, certains ont aussi d'autres matières secondaires, c'est-à-dire, la psychologie et les sciences politiques. Comme les étudiants ont dû écrire les textes en français et en suédois, nous avons choisi les participants selon le critère qu'ils doivent avoir le français et le suédois soit comme matière principale soit comme matière secondaire (ils ont donc passé l'examen d'entrée pour les deux langues). En disant cela, nous avons exclu de notre étude les étudiants qui ont commencé à étudier le français ou le suédois à l'université, par exemple dans un centre de langues où il est possible de commencer à étudier des langues sans connaissances antérieures. Dans le tableau 7, nous présentons un résumé des informations de base des participants de notre étude.

---

<sup>5</sup> [https://www.utu.fi/sites/default/files/public%3A//media/file/valintaopas-19-turun-yliopisto\\_2.pdf](https://www.utu.fi/sites/default/files/public%3A//media/file/valintaopas-19-turun-yliopisto_2.pdf)  
consulté le 13 avril 2019

**Tableau 7. Les informations de base des participants**

| <b>Etudiant</b> | <b>Année de naissance</b> | <b>Langue maternelle</b> | <b>Matière principale</b> | <b>Matière(s) secondaire(s)</b>  |
|-----------------|---------------------------|--------------------------|---------------------------|--|
| <b>A</b>        | 1996                      | le finnois               | le français               | les langues scandinaves (le suédois) <sup>6</sup>                          |
| <b>B</b>        | 1992                      | le finnois               | l'anglais                 | le français, les langues scandinaves (le suédois), la psychologie          |
| <b>C</b>        | 1996                      | le finnois               | le français               | les langues scandinaves (le suédois)                                       |
| <b>D</b>        | 1996                      | le finnois               | l'anglais                 | le français, les langues scandinaves (le suédois), les sciences politiques |

Quant aux principes de codage du corpus, nous avons décidé de désigner les participants par « étudiant A, B, C et D ». Nous ne distinguons pas les étudiants selon leur sexe parce que cela n'est pas pertinent pour notre étude, donc nous nous référons aux étudiants toujours par le pronom *il*.

En ce qui concerne le questionnaire qui nous a offert les informations de base présentées, nous l'avons déjà indiqué au début du chapitre 4 que nous avons effectué le questionnaire en ligne. Nous avons créé ce questionnaire pour avoir plus d'information sur les participants, sur leurs études et sur leurs compétences en français et en suédois. Le questionnaire (en annexe 1) comprend 24 questions en finnois, dont les sept premières traitent des informations de base des participants : les coordonnées, l'année de naissance, la langue maternelle, la matière principale et la/les matière/s secondaire/s à l'université, ainsi que l'autorisation d'utiliser les informations dans la recherche. Dans les questions 8 et 9, nous avons demandé aux étudiants d'indiquer quelle langue étrangère ils maîtrisent le mieux et quelles autres langues ils maîtrisent. Nous avons aussi demandé aux étudiants de motiver leurs réponses et évaluer leurs compétences linguistiques.

---

<sup>6</sup> Le nom de la matière est les langues scandinaves parce que les études comprennent aussi des cours de norvégien et de danois, mais le suédois est la langue principalement étudiée. <https://www.utu.fi/fi/opiskelijaksi/pohjoismaiset-kielet-humanististen-tieteiden-kandidaatti-ja-filosofian-maisteri-3-v-2> consulté le 1er octobre 2019

Ensuite, les questions de 10 à 16 concernent les études du français. Nous avons voulu savoir à quel âge les étudiants ont commencé à étudier le français et combien de temps ils l'ont étudié. Nous avons continué par les questions sur des séjours à l'étranger : les étudiants ont-ils passé du temps dans un pays ou dans une région francophone ? Si oui, nous avons voulu savoir dans quel pays ou dans quelle région ils ont fait le séjour, quelle a été la raison de leur séjour et combien de temps ils ont passé dans le pays ou dans la région. Nous avons aussi posé des questions sur l'usage de la langue française, notamment sur la fréquence avec laquelle l'étudiant parle, écrit, écoute ou lit en français. Nous avons précisé cet aspect par la question suivante qui traite les situations où l'étudiant utilise la langue française, par exemple avec des amis ou à l'université. Les questions de 17 à 23 sont les mêmes que les questions de 10 à 16, mais elles concernent les études du suédois, les séjours dans un pays ou dans une région suédophone et l'utilisation du suédois. Quant à la question 24, il s'agit de l'écriture. Nous avons posé la question « Quel type de scripteur êtes-vous ? » aux étudiants et cela veut dire que nous nous sommes attendue à ce que les étudiants décrivent comment ils écrivent, s'ils pensent qu'ils écrivent lentement ou vite, s'ils planifient l'écriture à l'avance ou s'ils utilisent la technique de l'écriture spontanée, etc.

Nous avons déjà mentionné que le questionnaire nous aidera dans notre travail d'analyse et qu'il ne sera pas analysé de manière approfondie. Nous utiliserons donc les réponses des étudiants pour soutenir l'analyse de leurs productions écrites et leurs processus d'écriture.

#### **4.2. Les productions écrites à l'aide du logiciel *ScriptLog***

Les productions écrites et les processus d'écriture forment la partie la plus importante de notre corpus. Comme nous l'avons déjà dit, les quatre étudiants ont écrit deux textes en utilisant le logiciel *ScriptLog* : l'un en français et l'autre en suédois. Ils ont eu 30 minutes pour l'écriture d'un texte et cela veut dire que nous ne leur avons pas demandé d'écrire un certain nombre de mots – ils ont écrit tant qu'ils ont pu écrire en 30 minutes. La séance d'écriture s'est déroulée dans une salle à la bibliothèque de l'université. L'étudiant a été seul dans la salle pendant le temps de rédaction et il n'a pas pu utiliser de dictionnaire ou d'autres outils. Le nombre de jours entre la première et la seconde séance d'écriture a varié d'un jour à quatre jours, selon la disponibilité des étudiants. Nous avons décidé que deux étudiants (aléatoirement choisis) ont commencé avec le texte en français et les deux autres ont commencé avec le texte en suédois pour que

toutes les productions soient comparables (*cf.* chapitre 5.1.). Dans ce qui suit, nous présentons les sujets sur lesquels les étudiants ont écrit :

1) Qu'est-ce que vous pensez des études en échange ? Quels sont les avantages ? Quels sont les inconvénients ?

2) Vad tycker du om att jobba vid sidan av studierna? Vilka fördelar eller nackdelar finns det? (Trad. Qu'est-ce que vous pensez du fait de travailler à côté des études ? Quels sont les avantages ou les inconvénients ?)

Nous avons choisi ces sujets parce qu'ils sont proches de la vie estudiantine, nous avons donc pensé que tous les étudiants auraient une opinion sur ces sujets et ils pourraient écrire quelque chose sur les sujets. Nous basons notre choix également sur le fait que dans le Cadre européen commun de référence pour les langues (CECR, 2001), il est dit qu'au niveau B1 l'apprenant peut écrire au moins des textes qui sont proches de sa vie personnelle (CECR, 2001, p. 26). Quant au thème de notre étude, nous avons pensé que dans les textes d'opinion nous pouvons trouver un bon nombre de séquences préfabriquées à analyser.

Comme le choix des langues (le français et le suédois) dans notre mémoire de maîtrise en FLE peut sembler particulier, il nous paraît important d'expliquer pourquoi nous avons choisi d'inclure les textes écrits en français et en suédois dans notre étude. Comme nous l'avons dit, les étudiants sont multilingues, à savoir qu'ils parlent tous au moins trois langues étrangères. Le suédois est la deuxième langue officielle en Finlande, mais tous les étudiants ont néanmoins étudié l'anglais plus longtemps que le suédois, et trois sur quatre pensent que l'anglais est la langue qu'ils maîtrisent le mieux. Comme nous pouvons nous attendre à ce que les étudiants ont un niveau de langue assez similaire en français et en suédois, nous avons voulu choisir le français et le suédois, au lieu de l'anglais, pour pouvoir comparer les productions écrites en deux langues. De plus, les chercheurs dans le département des langues scandinaves à l'Université de Turku sont en train d'étudier les séquences préfabriquées dans les productions écrites (en suédois), et ils utilisent aussi le logiciel *ScriptLog* dans leur recherche<sup>7</sup>. Cela nous a également inspirée, et dans ce qui suit, nous présentons ce logiciel plus en détail.

Le logiciel *ScriptLog* est donc un logiciel d'enregistrement des frappes (*keystroke logging*), c'est-à-dire qu'il enregistre les touches du clavier et les mouvements de la souris. Le logiciel nous offre des manières différentes pour analyser les productions

---

<sup>7</sup> <https://blogit.utu.fi/nordiskaforskning/helsekvenser-i-tre-morfologiskt-olika-inlararsprak/> consulté le 14 avril 2019

écrites : la visualisation du processus d'écriture, le produit textuel final, le texte linéaire (qui montre le processus du texte) et les statistiques. (Entre autres, Cislaru & Olive, 2018a ; Johansson, 2017 ; Mutta, 2007, p. 62, 72-75 ; cf. aussi Lindgren & Sullivan, 2019.) Ci-dessous, nous donnons des exemples d'un extrait du produit textuel final, d'un extrait du texte linéaire et des statistiques.

Les études en échange sont très utiles. La chose la plus important est bien sûr l'apprentissage d'une langue. En échange, l'apprentissage se déroule en contexte authentique avec les contacts avec les locuteurs natifs. On peut améliorer l'écriture, la compréhension, la parole etc. On entend la langue étrangère tout le temps. Dans le meilleur cas, on peut apprendre une langue étrangère totalement nouvelle et en même temps améliorer les compétences de l'anglais, comme on rencontre des personnes qui viennent de tous les pays dans le monde.

### **Illustration 1. Exemple d'un extrait du produit textuel final de l'étudiant C**

Dans cette illustration 1, nous voyons un extrait du texte final. Nous utilisons ce même extrait aussi dans les deux exemples suivants. Lorsque nous voulons étudier le processus de la rédaction, nous prenons le texte linéaire (illustration 2) qui nous montre par exemple les durées des pauses (en secondes) et les mots/les lettres que le scripteur a effacé.

<META totalRecTime=1810.364 /META><START><29.000>Les études en échange sont <60.618>très utiles. <2.926>Le <BACKSPACE2><4.979>a chose la plus important est <6.757>bien sûr l'apprentid<BACKSPACE1>ssage d'une langue. <10.403>En échange <BACKSPACE1>, on doit <4.968><BACKSPACE5><8.620><BACKSPACE3>l'apprentissage se <BACKSPACE3>déroule <3.631>dans <BACKSPACE5>en contexte authentique<5.744><MOUSECLICK>se <MOUSECLICK><14.295> avec les contact<3.246>s avec les locuteurs natifs. <41.538>On <19.442>peut améliorer <BACKSPACE5>iorer <14.772>l'écriture <BACKSPACE1>, <3.209>la compréhension, <2.783>la parole etc. <16.259>Dans le cas <2.651><BACKSPACE5> meilleur cas, <2.220>on peut apprendre <6.061>une langue<3.742> <3.948>étrangèrre<BACKSPACE2>e totalement nouvelle<3.635>, <BACKSPACE2> et <12.005>en même temps <23.342>améliorer l'<BACKSPACE1>es compétences de l'anglais, comme <4.717>on <4.547>rencotre <BACKSPACE3>ntrre <BACKSPACE6>ntrre des personnes <27.595>qui viennent <6.811>de tous les pays dans le monde.

### **Illustration 2. Exemple d'un extrait du texte linéaire de l'étudiant C**

Par exemple, nous pouvons voir de l'illustration 2 que le scripteur a eu une pause de 29 secondes avant qu'il ait commencé à écrire la phrase « Les études en échanges sont... ». Quant aux statistiques que le logiciel nous offre, elles montrent par exemple le nombre des caractères dans le produit textuel final et le nombre des caractères utilisés lors de la rédaction.

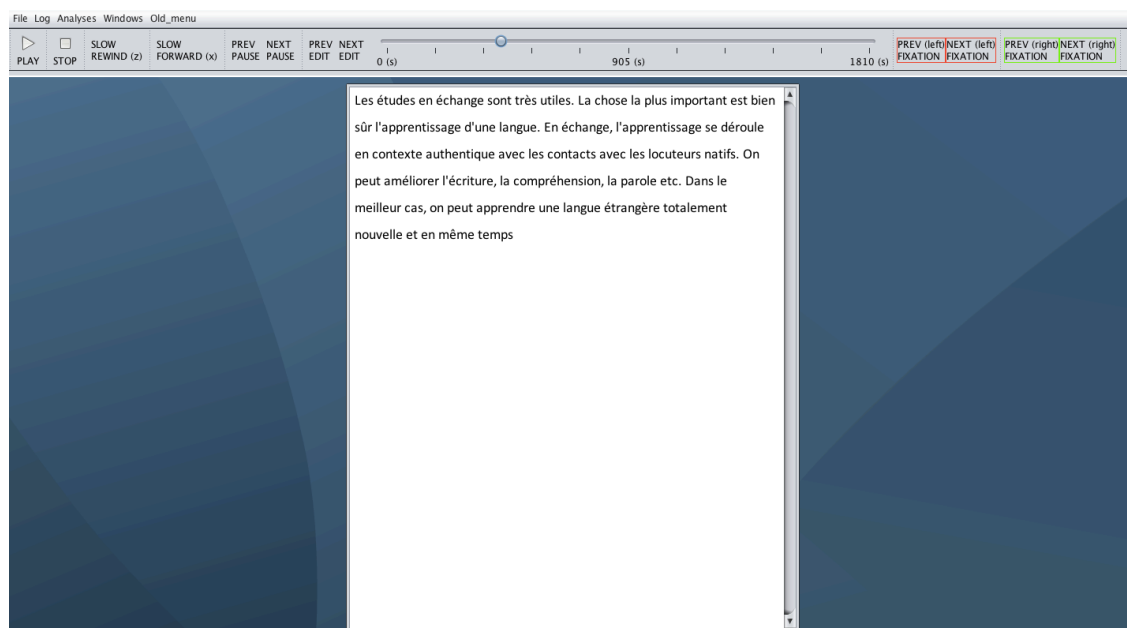
--SUMMARY--  
Tokens in Final Text: 1423  
Tokens in Linear Text: 1588  
KeyStrokes in Linear Text: 1751  
Total Number of Events: 1778  
Total recording time (s): 1810,364  
Time until first key press (s): 28,797  
Total typing time (s): 1663,380  
Time after last key release (s): 118,186  
Pause criteria (s): 2,000  
Number of pauses, Key: 61  
Total pause time, Key (s): 701,794  
Mean pause length, Key (s): 11,505  
Sd pause length, Key (s): 12,813  
Median pause length, Key (s): 6,422  
Number of pauses, Cursor: 113  
Total pause time, Cursor (s): 1105,304  
Mean pause length, Cursor (s): 9,781  
Sd pause length, Cursor (s): 11,672  
Median pause length, Cursor (s): 4,763  
Median Transition Time (s): 0,184  
5%-trimmed Mean Transition Time (s): 0,180

### **Illustration 3. Exemple des statistiques du texte en français de l'étudiant C**

Par exemple, nous voyons que dans le texte final, il existe 1423 caractères (*tokens in final text*) et que ce scripteur a écrit pendant 1663,38 secondes, c'est-à-dire pendant 27 minutes (*total typing time*).

### **4.3. Les verbalisations**

Dans cette section, nous continuons par l'explication de la situation de verbalisation. Nous sommes revenue dans la salle après les 30 minutes pendant lesquelles l'étudiant a écrit son texte, afin d'effectuer la verbalisation. Comme le logiciel *ScriptLog* produit une sorte d'enregistrement vidéo du processus d'écriture, nous avons regardé cette vidéo ensemble et nous avons demandé à l'étudiant de commenter (en finnois) ce qu'il a pensé pendant les pauses et quand il a fait des corrections ou des révisions. Nous avons aussi posé des questions à l'étudiant pendant la verbalisation, par exemple pourquoi il a fait tel ou tel choix dans l'écriture. La durée de la verbalisation a aussi été de 30 minutes, parce que la « vidéo » que nous avons vue comprend tout ce que l'étudiant avait fait pendant le temps d'écriture. Les étudiants ont donc verbalisé leurs processus d'écriture tout de suite après les séances d'écriture, c'est-à-dire qu'il s'agit de la verbalisation rétrospective avec un rappel stimulé (*stimulated recall*). Pour illustrer ce que les étudiants ont vu pendant la verbalisation, nous montrons une capture d'écran de la « vidéo ». Cette capture d'écran 1 ci-dessous présente le processus d'écriture en français de l'étudiant C.



**Capture d'écran 2. Le processus d'écriture en français de l'étudiant C**

La fonction des verbalisations dans notre étude est de nous servir dans le travail d'analyse. Par exemple, les verbalisations peuvent nous donner des vérifications ou des explications dans le cas où nous analysons les productions écrites et nous ne pouvons pas savoir ce que l'étudiant a pensé pendant le processus d'écriture. Une transcription complète de la verbalisation n'est donc pas nécessaire pour l'étude, mais nous avons transcrit des points intéressants dans les verbalisations que nous avons utilisés dans la partie d'analyse.

#### **4.4. Méthode d'analyse**

Afin d'analyser des séquences préfabriquées, notre premier travail a été de les identifier dans les productions écrites. Dans les études concernant des séquences préfabriquées, l'identification des SP est souvent considérée comme problématique, car le travail d'identification est toujours subjectif ; le chercheur doit décider si une séquence de mots est en effet une séquence préfabriquée ou pas dans la production (entre autres, Wray, 2009, p. 28 ; Sundman, 2011, p. 329 ; Gheitasi, 2017, p. 52). Par exemple, comme Sundman (2011, p. 329) le dit, une séquence peut représenter une séquence préfabriquée pour un étudiant mais pas pour l'autre (*cf.* aussi Nyqvist, 2013). De plus, quand il s'agit d'étudiants avancés des langues étrangères, le travail d'identification vient plus difficile parce qu'ils sont capables de produire des phrases aisées et grammaticalement correctes (Myles & Cordier, 2017, p. 14).

Etant consciente de ces problèmes, nous avons effectué l'identification des SP à l'aide des définitions de SP que nous utilisons (*cf.* chapitres 3.2.1. et 3.2.2.), et aussi à l'aide du logiciel *ScriptLog*. L'utilisation des définitions comme un outil semble convenable pour nous, car selon Wray (2009, p. 28), les définitions choisies influencent largement l'identification des SP. En ce qui concerne le logiciel *ScriptLog*, nous avons profité de la possibilité de localiser des pauses dans les processus d'écriture. Nous avons déterminé que le logiciel montre des pauses ayant une durée supérieure à 2 secondes. Nous nous basons cette décision sur Cislaru et Olive (2018a, p. 43-44) qui ont utilisé la pause minimale de 2 secondes pour segmenter des séquences textuelles, c'est-à-dire des *jets textuels*, comme ils les appellent. En plus, selon Cislaru et Olive (*ibid.*), l'utilisation de la pause minimale de 2 secondes est commune dans le domaine pour identifier des séquences de texte séparées par des pauses (*cf.* aussi Cislaru & Olive, 2018b). Comment avons-nous donc profité des pauses ? Nous avons analysé comment les pauses se localisent autour les séquences de mots et s'il existe des séquences préfabriquées entre ces pauses. Notons qu'il ne devrait pas exister de pauses au milieu d'une SP, car la définition de la SP suggère que les SP devraient être automatisées dans le langage d'un locuteur (*cf.* chapitre 2.3.). Les SP peuvent donc être entourées directement par les pauses, par exemple *<pause> en même temps <pause>*, ou elles peuvent se placer au milieu d'une phrase plus longue entourée par les pauses, comme *<pause> des études en échange sont une bonne manière de <pause>*. Comme nous l'avons dit, les définitions des SP nous ont servie dans la sélection des SP, et c'est notamment le cas avec les phrases longues. En outre, nous avons beaucoup bénéficié des exemples des SP possibles présentées par exemple par Forsberg (2006), Sundman (2011), Cordier (2013) et Rasmussen (2015). Egalement, nous avons utilisé les verbalisations pour identifier les SP propres aux étudiants individuels et nous y reviendrons plus en détail dans le chapitre 5.3.

En ce qui concerne la quantité de séquences préfabriquées dans les productions et le comptage des mots, nous nous basons largement sur Forsberg (2006) et Rasmussen (2015). A l'instar de ces chercheuses, pour trouver la quantité de SP, nous avons compté séparément le nombre total de mots produits dans le processus et dans le produit final, et aussi le nombre total de mots faisant partie des SP dans le processus et dans le produit final. Certes, nous avons effectué ces calculs pour chaque étudiant et pour les deux langues. Quant aux mots séparés par des apostrophes, nous les avons comptés comme deux mots, d'après Forsberg (2006, p. 67). Par exemple, *c'est* est composé de deux mots

et à *l'université* de trois mots dans notre corpus. Les mots composés par un trait d'union représentent un seul mot, mais notons que *peut-être* et *là-bas* sont néanmoins considérés comme les SP dans notre étude, comme l'ont fait Forsberg (2006) et Rasmussen (2015). Remarquons aussi ici que les apostrophes et les traits d'union ne concernent pas la langue suédoise dans notre corpus, nous avons donc compté chaque mot comme un seul mot en suédois, ce qui peut influencer les résultats (*cf.* chapitre 5.1.3.).

Dans le travail de la catégorisation des séquences préfabriquées, nous avons bénéficié des catégorisations antérieures que nous avons présentées dans le chapitre 3.3. et aussi des exemples antérieures des SP. Quand nous avons eu un problème avec la catégorisation d'une SP, nous avons vérifié comment les autres chercheurs l'ont catégorisée. Nous sommes consciente du fait qu'il existe des SP qui pourraient faire partie de plusieurs catégories, mais nous avons suivi Forsberg (2006, p. 48) en ce cas et classifié telles séquences selon les traits les plus forts, par exemple *il est possible* dans la catégorie des SP discursives au lieu des SP lexicales. Comme nous l'avons montré (*cf.* chapitre 3.3.), nous avons formé deux nouvelles catégories parce que nous avons trouvé les SP qui sont propres à notre corpus et qui méritent d'être analysées.

Les méthodes d'analyse présentées ici nous mènent à commencer l'analyse de notre corpus dans le chapitre suivant.

## **5. Analyse**

Dans les chapitres suivants, nous présenterons notre analyse dans l'ordre des questions de recherche (*cf.* chapitre 1.1.). Nous commencerons avec la quantité de séquences préfabriquées dans les productions écrites et nous traiterons la quantité selon la langue utilisée. Après cela, nous aborderons les types de SP que les étudiants ont utilisés et nous examinerons les catégories au niveau général. Ensuite, nous continuerons à traiter individuellement les étudiants et leur usage des SP dans les deux langues. Remarquons ici que toutes les SP utilisées sont listées en annexe 2, et nous y présentons séparément les listes des SP de chaque étudiant en français et en suédois.

### **5.1. La quantité de séquences préfabriquées dans les productions écrites**

Pour étudier la quantité de SP dans le corpus, nous avons formulé la première question de recherche comme suit :

Quelle est la quantité de séquences préfabriquées dans les productions écrites des étudiants du français L2 et du suédois L2 ? Existe-il des différences entre les deux langues ?

Dans les sous-chapitres suivants, nous analyserons la quantité de SP dans les processus d'écriture et dans les produits finaux, et nous diviserons les chapitres selon la langue (5.1.1. et 5.1.2.). Nous traiterons aussi bien la quantité totale de SP que la quantité produite par chaque étudiant séparément. Nous finirons l'analyse de la quantité dans le chapitre où nous comparerons nos résultats en français et en suédois entre eux et aussi aux autres études (5.1.3.). Pour pouvoir comparer les productions en français et en suédois entre elles, nous avons réalisé des analyses similaires dans les deux langues. Rappelons que deux étudiants (B et D, aléatoirement choisis) ont commencé avec le texte en suédois et, par conséquent, les deux autres ont commencé avec le français (*cf.* chapitre 4.2.). Nous avons fait un tel choix pour que le texte écrit en second lieu ne souffre pas de « manque d'intérêt ». Par exemple, si tous les étudiants avaient commencé avec le texte en français, cela aurait pu influencer négativement le nombre de mots produits dans les textes en suédois. Nous sommes consciente que le nombre de participants dans notre étude n'est pas grand, mais nous pensons qu'il est important à noter ce fait quand nous étudions la quantité de SP dans les productions.

### **5.1.1. Les productions en français**

Nous commençons cette partie au niveau général, c'est-à-dire que nous traitons les productions de tous les étudiants ensemble. Comme les étudiants ont eu 30 minutes pour la séance d'écriture (*cf.* chapitre 4.2.), ils ont réussi à produire un bon nombre de mots. En effet, les quatre étudiants ont produit 1588 mots au total en français dans le processus d'écriture, et dans ce processus, ils ont produit 462 mots qui font partie des séquences préfabriquées. Comme les processus comprennent des révisions et des effacements, il est normal que le nombre de mots produits dans le texte final soit plus bas que celui dans le processus. Les quatre étudiants ont donc produit 1315 mots au total dans le produit final. Egalement, le produit final comporte moins des mots faisant partie des SP, c'est-à-dire que les étudiants en ont produit 393 au total. Pour trouver ce qui est la quantité totale de SP dans les productions, nous avons effectué quelques calculs et dans le tableau 8 ci-dessous nous avons groupé le nombre total de mots produits dans le processus et dans le produit final ainsi que les nombres de mots faisant partie des SP respectivement et leurs pourcentages.

**Tableau 8. Les nombres de mots produits et de mots faisant partie des SP au total en français**

|                      | Mots produits dans le processus | Mots faisant partie des SP dans le processus | Pourcentage des SP dans le processus | Mots produits dans le produit final | Mots faisant partie des SP dans le produit final | Pourcentage des SP dans le produit final |
|----------------------|---------------------------------|--|--------------------------------------|-------------------------------------|--|--|
| Au total en français | 1588                            | 462  | 29 %                                 | 1315                                | 393  | 30 %                                     |

En divisant le nombre de mots faisant partie des SP par le nombre de mots produits, nous avons eu la quantité de SP dans le processus d'écriture et dans le produit final. Comme ce tableau nous indique, la quantité totale de séquences préfabriquées en français est de 29 % dans le processus d'écriture et de 30 % dans le produit final. Le pourcentage des SP plus élevé dans le produit final (quoique seulement une unité de pourcentage) résulte du fait que le nombre de mots produits dans le produit final a diminué de 17 %, tandis que celui de mots faisant partie des SP a diminué de 15 %. Toutefois, nous pouvons dire que les étudiants ont produit un bon nombre de séquences préfabriquées dans leurs productions écrites. Pour analyser plus en détail ces nombres, nous présentons le tableau 9 suivant qui montre les nombres au niveau de chaque étudiant. Notons que les chiffres dans la ligne « moyenne » sont arrondis.

**Tableau 9. Les mots produits et les mots faisant partie des SP au niveau de chaque étudiant en français**

| français | Etudiant | Mots produits dans le processus | Mots faisant partie des SP dans le processus | Pourcentage des SP dans le processus | Mots produits dans le produit final | Mots faisant partie des SP dans le produit final | Pourcentage des SP dans le produit final |
|----------|----------|---------------------------------|--|--------------------------------------|-------------------------------------|--|--|
|          | A        | 548                             | 147  | 27 %                                 | 452                                 | 128  | 28 %                                     |
|          | B        | 528                             | 170  | 32 %                                 | 421                                 | 150  | 36 %                                     |
|          | C        | 271                             | 71   | 26 %                                 | 248                                 | 60   | 24 %                                     |
|          | D        | 241                             | 74   | 31 %                                 | 194                                 | 55   | 28 %                                     |
|          | Moyenne  | 397                             | 116  | 29 %                                 | 329                                 | 98   | 30 %                                     |

Nous pouvons voir à partir de ce tableau qu'au niveau des étudiants individuels, la quantité de SP varie de 26 % à 32 % dans le processus d'écriture. L'étudiant A a produit le plus de mots dans le processus, mais son pourcentage des SP est assez bas (27 %) si

nous le comparons au pourcentage des SP de l'étudiant B (32 %) qui n'a produit que 20 mots moins que lui au total. Les étudiants C et D ont produit notablement moins de mots dans le processus que les deux autres étudiants, mais leurs pourcentages des SP (26 % et 31 %) suivent ceux des étudiants A et B. En ce qui concerne le produit final, la quantité de SP varie de 24 % à 36 %. Quand le pourcentage des SP est plus élevé dans le produit final que dans le processus (le cas des étudiants A et B), cela veut dire que le nombre de mots produits au total a diminué plus par rapport à la diminution du nombre de mots faisant partie des SP. Par exemple, le nombre de mots produits par l'étudiant A a diminué de 18 % et le nombre de mots faisant partie des SP de 13 %. Pareillement, quand le pourcentage des SP est plus bas dans le produit final que dans le processus (le cas des étudiants C et D), cela indique des changements contraires dans des rapports des nombres. Par exemple, le nombre de mots produits par l'étudiant D a diminué de 20 % dans le produit final, tandis que les mots faisant partie des SP ont diminué de 26 %. Comme nous l'avons dit, tous les changements que l'étudiant a faits dans le processus ont influencé ces chiffres et pour cette raison les pourcentages sont différents au niveau de chaque étudiant.

Le tableau 9 au-dessus indique aussi la moyenne du pourcentage des SP dans le processus et dans le produit final, et ces taux sont évidemment les mêmes que dans le tableau 8. Pour répondre à notre question « quelle est la quantité de séquences préfabriquées » en français, nous pouvons donc dire que la quantité totale de SP que les étudiants ont produites en moyenne est de 29 % dans le processus et de 30 % dans le produit final. Dans le chapitre suivant, nous continuerons à étudier la quantité de SP dans les productions en suédois, pour pouvoir comparer nos résultats du français et du suédois entre eux, et aussi les comparer aux résultats des études précédentes.

### **5.1.2. Les productions en suédois**

Pendant les 30 minutes de la séance d'écriture, les quatre étudiants ont produit 1519 mots au total en suédois dans le processus. Dans ce processus, nous trouvons 346 mots qui font partie des séquences préfabriquées et c'est un nombre respectable de mots des SP. Comme nous l'avons dit, le nombre de mots produits dans le produit final est souvent plus bas que dans le processus à cause des changements dans le texte. Ce fait concerne aussi les productions en suédois, comme le produit final comporte 1350 mots écrits par les quatre étudiants. Le nombre de ces mots a donc diminué de 11 % du processus au produit final. Ensuite, dans le produit final nous trouvons 321 mots faisant

partie des SP, ce qui équivaut à la diminution de 7 % dans le nombre de mots. Ces pourcentages ne sont pas très élevés, ce qui veut dire que les étudiants n'ont pas fait tellement de changements dans leurs textes. Comme nous l'avons fait dans la partie du français, nous avons voulu trouver la quantité totale de SP dans les productions en suédois. Dans le tableau 10 ci-dessous, nous avons groupé les chiffres présentés ainsi que les pourcentages des SP.

**Tableau 10. Les nombres de mots produits et de mots faisant partie des SP au total en suédois**

|                     | Mots produits dans le processus | Mots faisant partie des SP dans le processus | Pourcentage des SP dans le processus | Mots produits dans le produit final | Mots faisant partie des SP dans le produit final | Pourcentage des SP dans le produit final |
|---------------------|---------------------------------|--|--------------------------------------|-------------------------------------|--|--|
| Au total en suédois | 1519                            | 346  | 23 %                                 | 1350                                | 321  | 24 %                                     |

Comme marqué dans le tableau 10, la quantité totale de SP en suédois est de 23 % dans le processus d'écriture et de 24 % dans le produit final. Nous avons déjà expliqué dans la partie du français (5.1.1.) pourquoi le pourcentage des SP peut être plus élevé dans le produit final que dans le processus, et dans ce cas des productions en suédois, nous trouvons le même phénomène. La différence entre les pourcentages n'est qu'une unité de pourcentage, mais elle montre que les étudiants ont révisé les textes. En ce qui concerne ces pourcentages des SP notablement plus bas que ceux dans les productions en français, nous y reviendrons dans le chapitre suivant (5.1.3.) où nous ferons la comparaison entre les langues. Comme nous voulons savoir comment chaque étudiant a utilisé des SP, nous présentons également les nombres de mots produits au niveau des étudiants individuels. Ces chiffres sont groupés dans le tableau 11, et notons aussi ici que les chiffres dans la ligne « moyenne » sont arrondis.

**Tableau 11. Les mots produits et les mots faisant partie des SP au niveau de chaque étudiant en suédois**

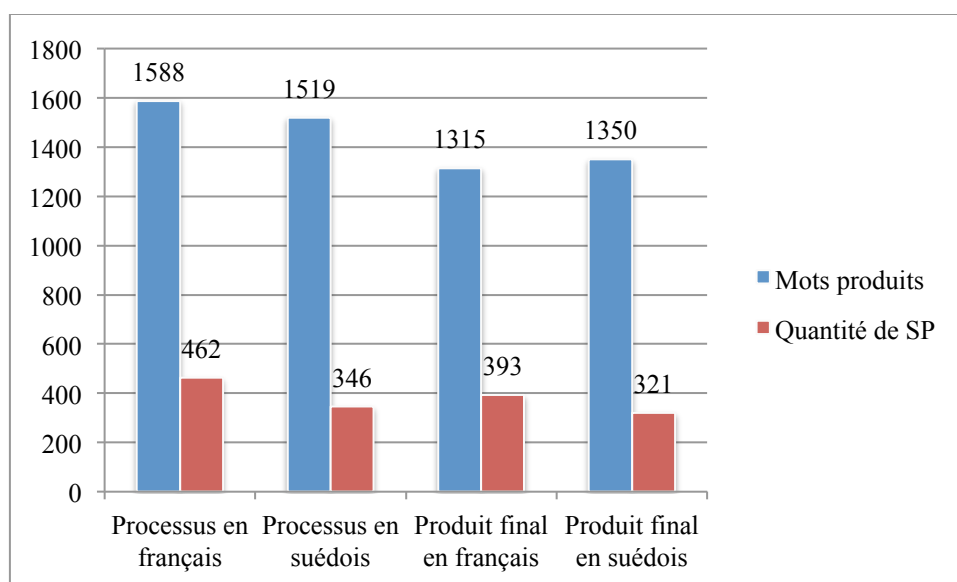
| suédois | Etudiant | Mots produits dans le processus | Mots faisant partie des SP dans le processus | Pourcentage des SP dans le processus | Mots produits dans le produit final | Mots faisant partie des SP dans le produit final | Pourcentage des SP dans le produit final |
|---------|----------|---------------------------------|--|--------------------------------------|-------------------------------------|--|--|
|         | A        | 552                             | 122  | 22 %                                 | 484                                 | 112  | 23 %                                     |
|         | B        | 483                             | 106  | 22 %                                 | 434                                 | 97   | 22 %                                     |
|         | C        | 214                             | 47   | 22 %                                 | 206                                 | 47   | 23 %                                     |
|         | D        | 270                             | 71   | 26 %                                 | 226                                 | 65   | 29 %                                     |
|         | Moyenne  | 380                             | 87   | 23 %                                 | 338                                 | 80   | 24 %                                     |

Comme le tableau 11 montre, le nombre de mots produits dans le processus varie de 214 mots à 552 mots. Cela veut dire que l'étudiant C ayant le nombre le plus bas a produit 61 % moins de mots que l'étudiant A ayant le nombre le plus élevé. Evidemment, le nombre de mots produits au total influence le nombre de mots qui font partie des SP, et nous voyons que le nombre de ces mots varie de 47 à 122 mots. Pourtant, la quantité de SP dans le processus est de 22 % pour les étudiants A, B et C, et de 26 % pour l'étudiant D ayant le pourcentage le plus grand. Quant au produit final, nous voyons plus de variation dans les pourcentages des SP, mais les taux d'A, B et D sont toujours plus bas que celui de C. Les chiffres de l'étudiant C montrent un processus intéressant, car le nombre de mots faisant partie des SP reste le même et le nombre de mots produits change très peu. Cela indique qu'il a beaucoup planifié ce qu'il veut écrire (le fait qu'il a aussi révélé dans le questionnaire) et pour cette raison, il a fait peu de changements. Nous voyons également de ce tableau que les étudiants ont au niveau individuel les pourcentages plus bas en suédois qu'en français, mais nous traiterons ce fait plus dans le chapitre suivant. En somme, pour répondre à notre question « quelle est la quantité de séquences préfabriquées » en suédois, nous pouvons dire que la quantité totale de SP est en moyenne de 23 % dans le processus d'écriture et de 24 % dans le produit final. Comme nous l'avons évoqué, nous continuerons avec la comparaison de nos résultats dans le chapitre suivant.

### **5.1.3. En somme**

Dans cette partie, notre objectif est de répondre à notre sous-question qui s'intitule « Existe-il des différences entre les deux langues ? ». Dans les deux chapitres en haut,

nous avons présenté les nombres de mots produits dans les processus et dans les produits finaux ainsi que les nombres de mots faisant partie des SP et leurs pourcentages, c'est-à-dire les quantités de SP. Au niveau général, nous avons montré que les étudiants ont produit 1588 mots au total dans le processus d'écriture en français et 1519 mots dans le processus en suédois. La différence n'est que 69 mots et en raison de cela, il semble que les étudiants aient des compétences de niveau semblable dans les deux langues, au moins en ce qui concerne la production du texte. Les nombres de mots produits ont diminué du processus au produit final dans les deux langues, et dans le produit final, ces nombres sont donc assez près l'un de l'autre, c'est-à-dire 1315 mots en français et 1350 en suédois. Cependant, nous avons trouvé qu'il existe des différences dans les nombres de mots faisant partie des SP et dans leurs pourcentages entre les deux langues. Nous montrons ces différences dans la figure 2 ci-dessous en nombres de mots.

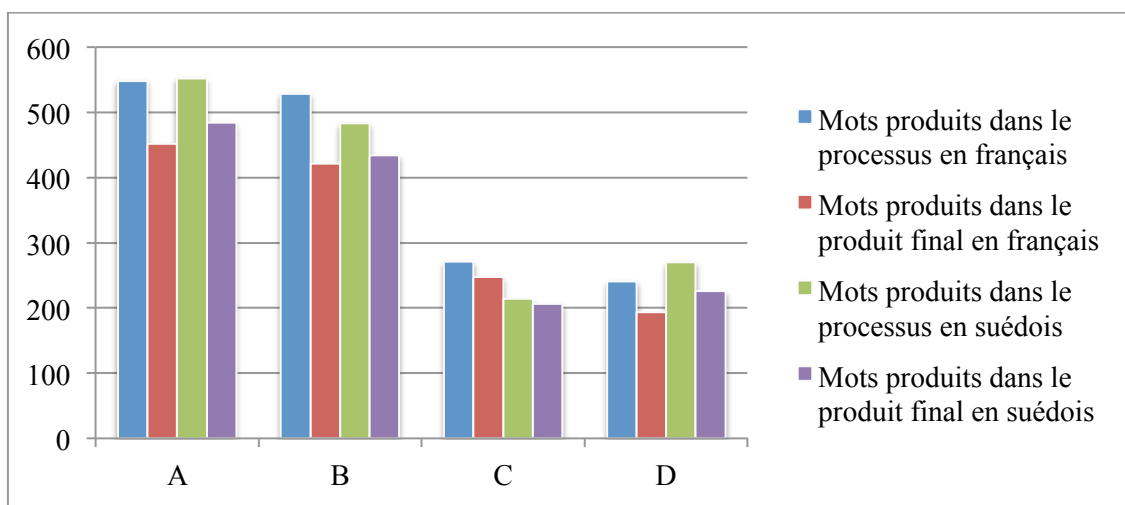


**Figure 2. Les nombres de mots produits et de mots faisant partie des SP**

La quantité de SP en français a donc été de 29 % dans le processus et de 30 % dans le produit final, tandis qu'en suédois les chiffres ont été de 23 % dans le processus et de 24 % dans le produit final. La différence est six unités de pourcentage dans les deux cas entre les langues, et les pourcentages du français sont alors notablement plus élevés que ceux du suédois. Qu'est-ce qui pourrait expliquer ce fait ? Nous trouvons que les pourcentages du français, c'est-à-dire les nombres de mots faisant partie des SP, sont plus grands parce que l'utilisation d'articles et la déclinaison influencent le nombre de mots qui font partie des SP. En suédois, l'on utilise des articles, mais des mots ont aussi une forme définie sans l'article, par exemple *en bil* 'une voiture', *bilen* 'la voiture'.

Egalement, la forme du pluriel est souvent exprimée sans d'article, comme *bilar* 'des voitures', *bilarna* 'les voitures'. Même si le mot *bil* n'est pas très pertinent dans notre corpus, nous pouvons voir que des formes différentes des noms peuvent agir sur le nombre de mots que nous trouvons dans les productions. Egalement, comme nous l'avons évoqué dans le chapitre 4.4., l'utilisation des apostrophes en français a pu influencé les résultats, car elles ne sont pas utilisées en suédois de telle manière. Une autre raison peut être que les étudiants n'ont pas de compétences langagières de même niveau dans les deux langues en ce qui concerne les séquences préfabriquées. Cela indiquerait que les étudiants n'aient pas tellement de SP automatisées à leur disposition (c'est-à-dire gardées en mémoire comme un tout) en suédois qu'en français. Rappelons que nous n'avons pas compté de séquences des mots où il existe des pauses au milieu comme les SP (cf. ch. 4.4.).

Au niveau individuel, si nous étudions le nombre de mots qu'un étudiant a écrit en français et en suédois, les nombres sont relativement semblables dans les productions de l'étudiant. Nous illustrons ce fait dans la figure 3 ci-dessous, où les nombres de mots produits sont présentés selon chaque étudiant.

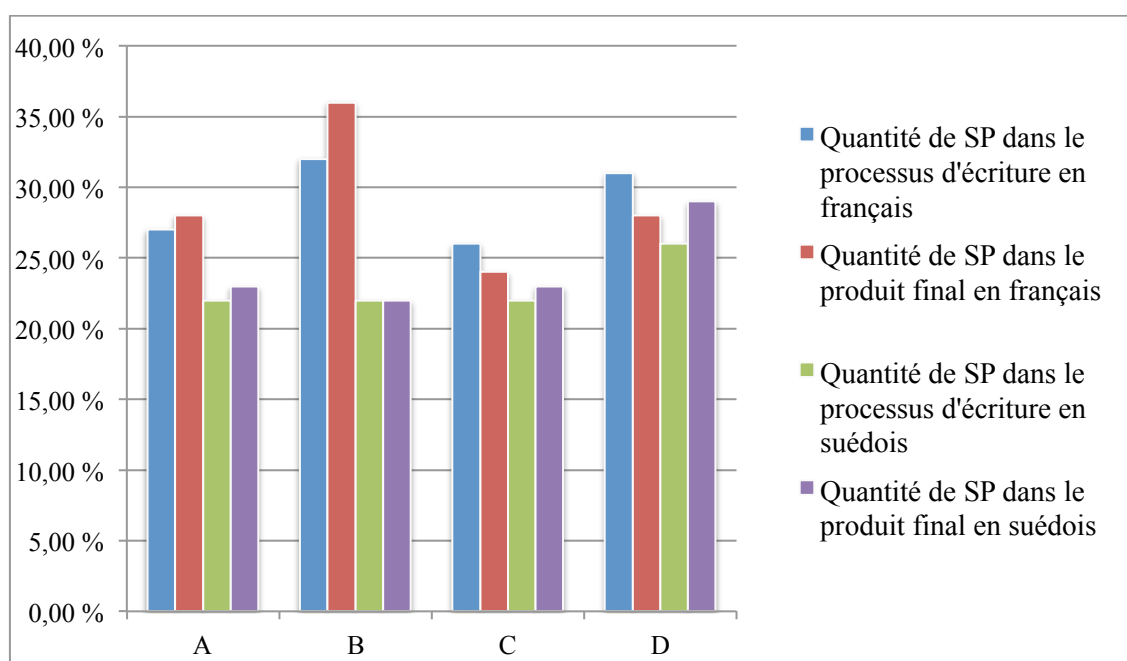


**Figure 3. Les nombres de mots produits par chaque étudiant**

Comme la figure 3 l'indique, il n'existe pas de différences remarquables entre les deux langues quand il s'agit des nombres de mots produits par un étudiant. Cependant, nous voyons que les étudiants A et B ont écrit considérablement plus de mots que les étudiants C et D. Cela peut être expliqué par le fait que les étudiants A et B ont répondu dans le questionnaire (cf. ch. 4.1.) et dans les verbalisations qu'ils ont l'habitude d'écrire du courant de conscience, c'est-à-dire qu'ils produisent du texte de manière

fluide. En revanche, les étudiants C et D ont dit qu'ils écrivent lentement en langues étrangères, car ils ont l'habitude de planifier beaucoup ce qu'ils écrivent et aussi d'utiliser des dictionnaires pendant le processus d'écriture, ce qui n'était pas possible dans notre étude. Comme les étudiants ont eu un temps limité pour la séance d'écriture, il semble naturel que C et D n'aient pas pu produire autant de mots. Nous pensons qu'il est important de connaître les nombres de mots produits pour comprendre les quantités de SP, ce que nous traiterons par la suite.

Nous avons montré dans ce chapitre les pourcentages des SP dans les productions et nous avons expliqué pourquoi les pourcentages en français ont été plus élevés que ceux en suédois. Toutefois, il est aussi intéressant d'étudier des différences dans les quantités entre les langues au niveau des étudiants individuels. Dans la figure 4 ci-dessous, nous avons groupé les quantités de SP de toutes les productions des étudiants.



**Figure 4. La quantité de SP selon les étudiants individuels**

Comme marqué dans cette figure, pour l'étudiant A, les pourcentages des SP varient de 22 % à 28 % dans toutes ses productions. Il a raconté dans le questionnaire qu'il pense qu'il a de meilleures compétences en français qu'en suédois, et pareillement, il utilise du français plus souvent que du suédois, par exemple en temps libre. Selon ses résultats, il a produit plus de mots en suédois, mais la quantité de SP est plus grande en français, ce qui correspond à ses propres estimations de ses compétences. Dans le cas de l'étudiant B, la quantité de SP varie de 22 % à 36 % dans ses productions. Selon cet

étudiant, ses compétences langagières sont meilleures en suédois qu'en français, et il utilise du suédois chaque semaine, par exemple avec des amis, et du français quelquefois par mois. Il est surprenant qu'il a pourtant produit plus de mots en français et la quantité de SP est très élevée dans les productions en français comparé à celles en suédois. Les différences entre les deux langues sont en effet les plus grandes parmi les étudiants. Quant à l'étudiant C, les pourcentages des SP varient de 22 % à 26 % dans ses productions, la différence étant donc seulement quatre unités de pourcentage. Selon ses réponses dans le questionnaire, ses compétences langagières sont meilleures en français qu'en suédois, mais il utilise les deux langues chaque semaine, en cours à l'université pour la plupart des cas. Cet étudiant a produit plus de mots en français et la quantité de SP est également plus élevée dans les productions en français, ce qui correspond à ses descriptions de ses compétences. En tout cas, il n'existe pas beaucoup de différence dans ses résultats entre les langues, autrement dit qu'il a les nombres de mots et les quantités de SP assez similaires dans les deux langues. Dans les productions de l'étudiant D, la quantité de SP varie de 26 % à 31 %, ce qui montre que ses pourcentages sont, somme toute, élevés. Selon cet étudiant, il utilise les deux langues chaque semaine, mais il a de meilleures compétences en suédois qu'en français ; il utilise du français seulement en cours, mais du suédois aussi en temps libre. Les résultats montrent que l'étudiant D a produit notablement plus de mots dans les productions en suédois qu'en français, mais la quantité de SP est cependant élevée aussi dans les productions en français. Il est intéressant, car cela pourrait indiquer qu'il a pourtant plus de SP automatisées dans la langue française que suédoise.

Il est intéressant à noter que les pourcentages sont plus élevés dans notre étude, à savoir dans les productions écrites en français, que dans les études de Forsberg (2006) et de Rasmussen (2015) qui ont étudié le français parlé. Dans leurs études, le pourcentage moyen des SP dans le langage d'apprenants universitaires a été de 25 %, ce qui a aussi été notre hypothèse pour le taux. Comme nous l'avons évoqué au début de ce chapitre, la quantité totale de SP a été en suédois de 23 % dans le processus et de 24 % dans le produit final, et en français de 29 % dans le processus et de 30 % dans le produit final. La moyenne totale dans toutes les productions ensemble est donc de 27 % dans notre étude, ce qui est toujours plus élevé que dans les études de Forsberg et de Rasmussen. Cela pourrait indiquer que la quantité de séquences préfabriquées est plus grande dans l'écriture que dans la langue parlée, mais évidemment nous aurions dû analyser des productions écrites et parlées par les mêmes étudiants pour pouvoir le déclarer. En effet,

nous trouvons que c'est un sujet qui mériterait des études ultérieures. Dans l'étude de Cordier (2013), le pourcentage des SP a été de 27 % dans le langage parlé des étudiants avancés en FLE. Cependant, Cordier (id., p. 141) remarque que la quantité de SP est toujours influencée par la définition de la SP et par la méthode d'indentification des SP. Pour cette raison, il est difficile de comparer les résultats entre eux. Toutefois, la quantité présentée par Cordier suggère que les étudiants d'une langue étrangère peuvent atteindre la quantité de SP plus élevée que celle de 25 %, ce qui soutient nos résultats. En ce qui concerne des études des productions écrites, nous avons remarqué qu'elles sont moins nombreuses que celles de la langue parlée. Dans le chapitre 1.1., nous avons noté que dans l'étude d'Erman et Warren (2000), ils ont trouvé que la quantité de SP a été d'environ 52 % dans les productions écrites, mais ils ont étudié la langue maternelle. Par conséquent, nous ne pouvons pas comparer nos résultats à cette quantité, mais le pourcentage très élevé indique que la quantité de SP peut en effet être plus grande dans l'écriture que dans la langue parlée.

Dans ce chapitre, nous avons présenté la quantité de séquences préfabriquées dans les productions écrites en français et en suédois et nous avons montré qu'il existe des différences entre les deux langues. Nous continuerons le travail d'analyse dans le chapitre suivant en étudiant des types différents de SP que les étudiants ont produits.

## **5.2. Les catégories des séquences préfabriquées utilisées**

Dans ce chapitre, notre objectif est d'étudier l'usage des séquences préfabriquées et de répondre à nos questions de recherche qui concernent les catégories de SP :

Quelles catégories des séquences préfabriquées les étudiants utilisent-ils ? Utilisent-ils les mêmes catégories dans les deux langues ?

Dans le chapitre précédent, nous avons analysé la quantité de séquences préfabriquées dans les productions en calculant les nombres de mots produits et de mots faisant partie des SP. Par contre, lorsque nous étudions les catégories et les types des SP, nous calculons les occurrences et les nombres de SP différentes ; par exemple, *études en échange*, *pour moi* et *c'est* représentent des SP différentes. Cela veut dire que nous traitons et calculons les SP comme les unités, malgré le nombre de mots qui font partie des SP ; *beaucoup de* est une SP et *à mon avis* est l'autre. Remarquons ici que nous considérons les SP qui contiennent un verbe et sont présentées avec la négation ou avec l'inversion comme les SP séparées par rapport aux SP à la forme de base, c'est-à-dire que *il y a* et *il n'y a pas* sont des SP différentes pour nous. Egalement, nous séparons

telles SP selon le temps ; *il est* et *il était* représentent des SP différentes. Ces faits peuvent donc influencer le calcul des SP différentes dans les chapitres suivants. Il faut aussi noter que le nombre de SP différentes et le nombre d'occurrences des SP ne sont pas la même chose, c'est-à-dire qu'une SP peut avoir, par exemple, six occurrences au total et l'autre seulement une occurrence.

Nous commencerons cette partie en examinant le rapport entre le nombre de SP différentes et leurs occurrences et en présentant la distribution des SP en catégories au niveau général. Certes, nous présenterons séparément la distribution dans les deux langues. Après, nous comparerons nos résultats en français et en suédois entre eux et aussi à ceux des études précédentes.

### 5.2.1. Les catégories des SP utilisées en français

Nous commençons avec les productions en français au niveau général, et premièrement nous traitons le rapport entre le nombre de SP différentes et leurs occurrences dans toutes les productions. Ces chiffres sont présentés dans le tableau 12 ci-dessous.

**Tableau 12. Le nombre de SP différentes et leurs occurrences dans les productions en français**

| français      |           | Nombre de SP différentes au total | Nombre d'occurrences des SP au total |
|---------------|-----------|-----------------------------------|--------------------------------------|
|               | Processus | 108                               | 167                                  |
| Produit final | 97        | 149                               |                                      |

Nous avons donc trouvé 108 séquences préfabriquées différentes dans le processus d'écriture et le nombre de leurs occurrences est de 167. Dans le produit final, les chiffres sont plus bas, à savoir 97 pour le nombre de SP différentes et 149 pour les occurrences, ce qui est en ligne avec le fait que les produits finaux comprennent moins de mots que les processus (*cf.* ch. 5.1.1.). Nous pouvons alors conclure que les étudiants ont effacé des séquences de mots entières qu'ils avaient produites comme les SP. Comme les nombres d'occurrences sont plus grands que ceux de SP différentes dans les deux cas, cela indique qu'il est probable que certaines SP ont plus d'occurrences que d'autres. En fait, les productions en français comprennent 21 séquences préfabriquées différentes qui se trouvent dans le corpus au moins deux fois. Autrement dit, ces SP sont les SP les plus utilisées dans les productions des quatre étudiants ensemble. Nous avons listé les SP dans le tableau 13 ci-dessous, et notons que nous les avons groupées selon

les occurrences dans le processus, car le processus inclut toutes les SP qui ont été produites. Le tableau 13 montre également que le nombre d'occurrences dans le cas de la SP la plus utilisée, à savoir *en échange*, n'est pas très haut, et pour cette raison nous avons décidé que nous prenons toutes les SP ayant au moins deux occurrences sur cette liste. Notons ici que toutes les SP utilisées par chaque étudiant sont listées en annexe 2.

**Tableau 13. Les SP ayant au moins deux occurrences dans les productions en français**

| SP                   | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type   |
|----------------------|------------------------|----------------------------|-----------|--------|
| En échange           | 8                      | 8                          | L         | SAdv   |
| Par exemple          | 7                      | 6                          | D         | MT     |
| A l'université       | 7                      | 7                          | L         | SAdv   |
| Il y a               | 6                      | 5                          | D         | MT     |
| En Finlande          | 6                      | 6                          | T         | SAdv   |
| Etudes en échange    | 6                      | 6                          | L         | SN     |
| Faire des/les études | 6                      | 6                          | L         | SV     |
| A mon avis           | 3                      | 3                          | D         | MAEP   |
| Bien sûr             | 3                      | 2                          | D         | MI     |
| Pour moi             | 3                      | 3                          | D         | MAEP   |
| Parce que            | 3                      | 3                          | D         | MT     |
| Beaucoup des         | 3                      | 2                          | SPIDI     | G (QD) |
| Ce pays              | 3                      | 2                          | I         | L (SN) |
| En France            | 2                      | 2                          | T         | SAdv   |
| Toutes les fois      | 2                      | 2                          | L         | SAdv   |
| J'ai appris          | 2                      | 1                          | L         | SV     |
| Ce n'est (pas)       | 2                      | 2                          | D         | MT     |
| Il n'y a pas         | 2                      | 2                          | D         | MT     |
| Quelques autres      | 2                      | 2                          | G         | PI     |
| C'est                | 2                      | 0                          | D         | MT     |
| C'est difficile      | 2                      | 2                          | D         | MAEP   |

Catégories : L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, SPIDI= SP idiosyncrasiques intérimaires  
Types : SAdv= syntagmes adverbiaux, SN= syntagmes nominaux, SV= syntagmes verbaux, MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement, MI= marqueurs interactifs, QD= expressions de quantité et de degré, PI= pronoms indéfinis

Comme les catégories des SP sont visibles dans le tableau 13, cela nous mène à présenter la distribution des SP en catégories. Dans le chapitre 4.4., nous avons expliqué comment nous avons choisi une catégorie (et un type) pour une séquence préfabriquée. En ce qui concerne la distribution des SP en catégories, elle est faite selon les occurrences des SP et pas selon les types d'occurrences, et cela veut dire qu'une catégorie peut contenir plusieurs fois la même SP. Dans les productions écrites en

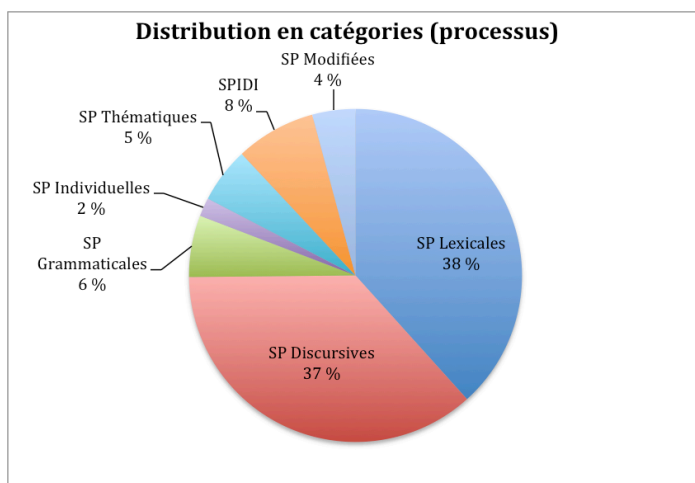
français, les quatre étudiants ont donc produit 167 SP dans le processus et 149 SP dans le produit final. Pour les distribuer en catégories, nous avons utilisé des sept catégories que nous avons présentées dans le chapitre 3.3., à savoir les SP lexicales (L), discursives (D), grammaticales (G), thématiques (T), individuelles (I), modifiées (M) et idiosyncrasiques intérimaires (les SPIDI). Dans le tableau 14 ci-dessous, nous montrons comment les SP en français sont distribuées dans ces catégories.

**Tableau 14. La distribution des SP en français en catégories**

| français | SP            | L  | D  | G  | T | I | M | SPIDI | Au total |
|----------|---------------|----|----|----|---|---|---|-------|----------|
|          | Processus     | 64 | 61 | 10 | 9 | 3 | 7 | 13    | 167      |
|          | Produit final | 58 | 53 | 10 | 9 | 2 | 7 | 10    | 149      |

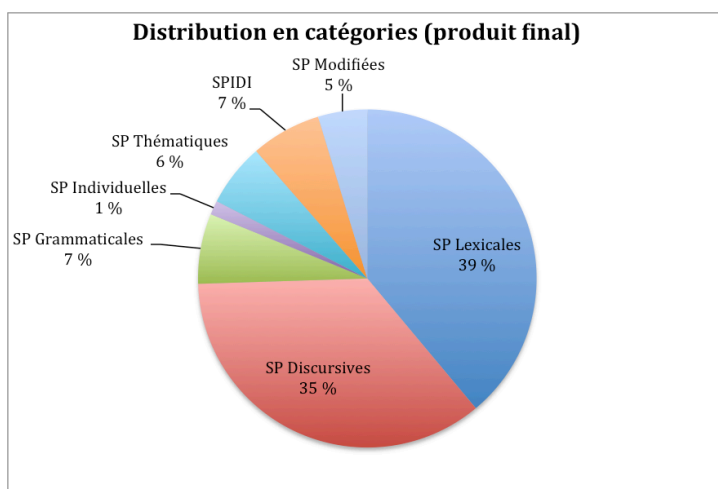
SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires

A partir de ce tableau, nous pouvons voir que les nombres dans la ligne du processus sont assez près des nombres dans la ligne du produit final, et de plus, trois catégories (G, T et M) comprennent les mêmes nombres dans les deux lignes. Cependant, comme le nombre de SP au total est plus bas dans le produit final que dans le processus, cela indique qu'il existe des changements dans quelques catégories. Clairement, deux catégories sont dominantes dans cette distribution, c'est-à-dire les SP lexicales ayant 64 SP dans le processus et 58 SP dans le produit final, et les SP discursives ayant 61 SP dans le processus et 53 SP dans le produit final. Dans les catégories des SP grammaticales et des SP thématiques, les SP sont restées les mêmes, mais dans le cas des SP modifiées, le nombre de SP est le même dans les deux lignes parce que nous avons décidé de marquer les SP aussi dans le produit final, même si elles y sont modifiées. Nous traiterons ce fait plus en détail au niveau des étudiants individuellement (*cf.* ch. 5.3.). En tout, les catégories des SP modifiées et des SP individuelles comportent le moins d'occurrences, ce qui n'est pas surprenant, car elles sont les nouvelles catégories que nous avons formées. Pour illustrer la quantité de SP dans chaque catégorie, nous présentons les figures 5 et 6 qui montrent la distribution des SP dans le processus et dans le produit final en pourcentages.



**Figure 5. La distribution des SP en français en catégories dans le processus**

Comme la figure 5 l'indique, dans le processus les SP lexicales couvrent donc 38 % de toutes les SP utilisées. Avec une différence d'une unité de pourcentage, le taux de SP discursives est de 37 %. Ces catégories sont visiblement les plus grandes parmi toutes les catégories, car les pourcentages des autres catégories varient de 2 % à 8 %.



**Figure 6. La distribution des SP en français en catégories dans le produit final**

La figure 6 montre que les pourcentages sont semblables dans le produit final par rapport au processus. La différence entre les catégories les plus grandes est pourtant quatre unités de pourcentage dans le produit final, au lieu de seulement une unité. Les pourcentages des autres catégories ont changé un peu, mais ils sont toutefois bas, en variant de 1 % à 7 %. Nous pouvons donc voir à partir de ces figures que l'ordre des catégories les plus utilisées est presque le même dans le processus et dans le produit

final. Pour pouvoir comparer ces résultats à ceux du suédois, nous commencerons à traiter des données des productions écrites en suédois dans le chapitre suivant.

### 5.2.2. Les catégories des SP utilisées en suédois

Dans cette partie, nous continuons en examinant des productions écrites en suédois. Comme nous l'avons premièrement fait avec les productions en français, nous présentons le rapport entre le nombre de SP différentes et le nombre de leurs occurrences dans toutes les productions des quatre étudiants ensemble. Le tableau 15 ci-dessous montre ces chiffres du suédois.

**Tableau 15. Le nombre de SP différentes et leurs occurrences dans les productions en suédois**

| suédois       |           | Nombre de SP différentes au total | Nombre d'occurrences des SP au total |
|---------------|-----------|-----------------------------------|--------------------------------------|
|               | Processus | 83                                | 137                                  |
| Produit final | 82        | 128                               |                                      |

Comme le tableau 15 l'indique, le processus d'écriture en suédois comprend 83 séquences préfabriquées différentes et le nombre de leurs occurrences est de 137. Le nombre de SP différentes dans le produit final est de 82, la différence n'étant qu'une SP par rapport au processus, c'est-à-dire qu'une SP (ayant une ou plusieurs occurrences) a été effacée dans le produit final. Le nombre d'occurrences des SP est de 128 dans le produit final, ce qui prouve cette disparition d'une SP et aussi prouve que quelques SP ont probablement plus d'occurrences que d'autres.

Comme nous l'avons fait avec les productions en français, nous avons groupé les SP qui ont eu au moins deux occurrences dans toutes les productions en suédois. C'est-à-dire, ces SP représentent les SP les plus utilisées en suédois par les quatre étudiants ensemble, et dans le corpus, nous en trouvons 24. Dans le tableau 16 ci-dessous, nous présentons la liste de ces SP et aussi leurs traductions.

**Tableau 16. Les SP ayant au moins deux occurrences dans les productions en suédois**

| SP  | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type     |
|---|------------------------|----------------------------|-----------|----------|
| Vid sidan av studier/na, 'à côté d'/des études'   | 13                     | 13                         | L         | SAdv     |
| Det är, 'il est/c'est'  | 8                      | 5                          | D         | MT       |
| Till exempel, 'par exemple'   | 6                      | 5                          | D         | MT       |
| För mig, 'pour moi'   | 6                      | 5                          | D         | MAEP     |
| Så mycket, 'tellement'  | 5                      | 5                          | G         | QD       |
| I Finland, 'en Finlande'  | 3                      | 2                          | T         | SAdv     |
| Hur mycket, 'combien'   | 3                      | 3                          | G         | QD       |
| Jag tänker att, 'je pense que'  | 2                      | 2                          | D         | MAEP     |
| Ha tillräckligt med, 'avoir assez de'   | 2                      | 2                          | L         | SV       |
| Jag tror (att), 'je trouve (que)'   | 2                      | 1                          | D         | MAEP     |
| Var det, 'était-il'   | 2                      | 2                          | D         | MT       |
| Det var, 'il était'   | 2                      | 2                          | D         | MT       |
| Hitta tid för, 'trouver du temps pour'  | 2                      | 2                          | L         | SV       |
| Dag/ar per vecka, 'un jour/des jours par semaine'   | 2                      | 2                          | L         | SAdv     |
| Kundservice industrin, 'le service client'  | 2                      | 2                          | I         | L (SN)   |
| För mycket, 'trop'  | 2                      | 2                          | G         | QD       |
| Även om, 'même si'  | 2                      | 2                          | D         | MT       |
| Under studietiden, 'pendant le temps d'études'  | 2                      | 2                          | L         | SAdv     |
| Är det, 'est-ce'  | 2                      | 2                          | D         | MT       |
| Extra pengar, 'extra d'argent'  | 2                      | 2                          | L         | SN       |
| Man måste, 'on doit/il faut'  | 2                      | 2                          | I         | D (MAEP) |
| Väldigt bra, 'très bon'   | 2                      | 2                          | L         | SAdj     |
| På sommarna, 'en été'   | 2                      | 2                          | SPIDI     | L (SAdv) |
| Extra jobbet, 'extra travail'   | 2                      | 1                          | I         | L (SN)   |
| Catégories : L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, SPIDI= SP idiosyncrasiques intérimaires<br>Types : SAdv= syntagmes adverbiaux, SN= syntagmes nominaux, SV= syntagmes verbaux, SAdj= syntagmes adjectivaux, MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement, MI= marqueurs interactifs, QD= expressions de quantité et de degré, PI= pronoms indéfinis |                        |                            |           |          |

Comme nous voyons à partir de ce tableau, la SP *vid sidan av studier/na* 'à côté d'/des études' a l'occurrence la plus élevée, à savoir 13, et pour la même raison que dans la liste des SP en français, nous avons inclut toutes les SP qui ont au moins deux

occurrences dans le processus. Notons également ici que toutes les SP produites en suédois par les étudiants sont listées en annexe 2.

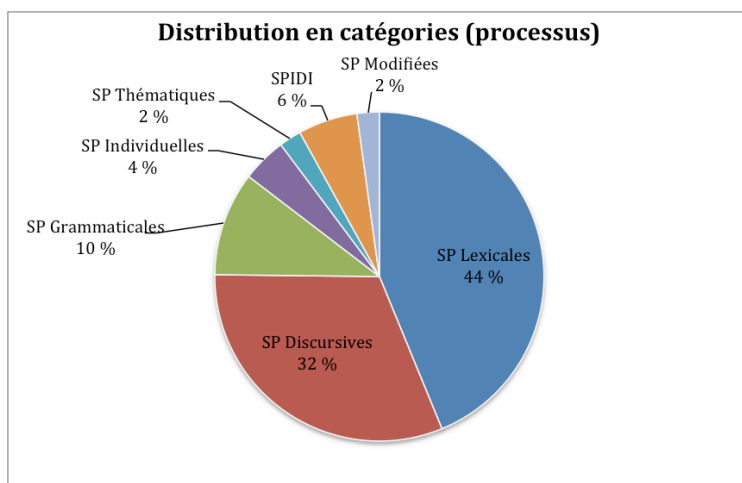
Comme le tableau 15 a montré, les étudiants ont produit au total 137 SP dans le processus et 128 SP dans le produit final en suédois ; remarquons donc qu'il s'agit des nombres d'occurrences des SP utilisées. Pour la catégorisation des SP, nous avons naturellement utilisé les mêmes catégories en français et en suédois. Dans ce qui suit, nous présentons la distribution des SP en suédois en ces catégories dans le tableau 17.

**Tableau 17. La distribution des SP en suédois en catégories**

| suédois | SP            | L  | D  | G  | T | I | M | SPIDI | Au total |
|---------|---------------|----|----|----|---|---|---|-------|----------|
|         | Processus     | 60 | 43 | 14 | 3 | 6 | 3 | 8     | 137      |
|         | Produit final | 60 | 36 | 14 | 2 | 5 | 3 | 8     | 128      |

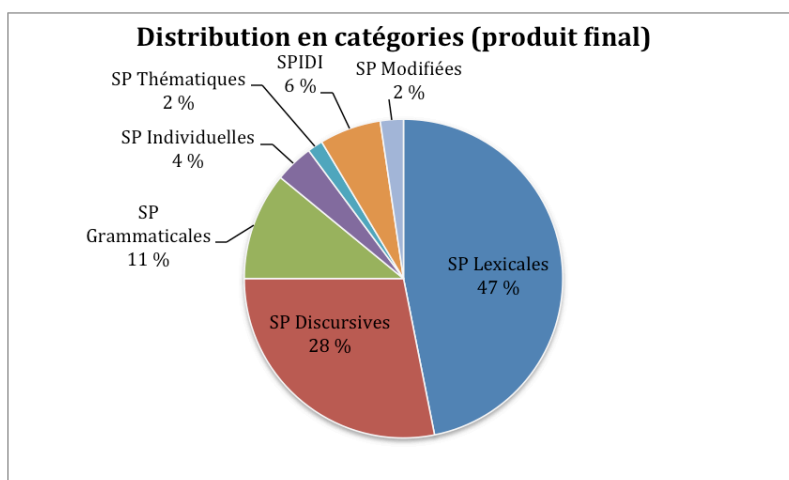
SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires

Le tableau 17 montre qu'il n'existe pas beaucoup de variation à l'intérieur d'une catégorie en ce qui concerne les nombres de SP dans le processus et dans le produit final. Dans quatre catégories (à savoir L, G, M et SPIDI), les nombres sont restés les mêmes dans les deux lignes. Cela indique que les étudiants n'ont pas eu besoin d'effacer des séquences de mots qu'ils ont produites, probablement parce que les SP ont été tellement automatisées. Les catégories les plus utilisées en suédois sont les SP lexicales ayant 60 SP dans le processus et dans le produit final, et les SP discursives comportant 43 SP dans le processus et 36 SP dans le produit final. En effet, la catégorie des SP D est la catégorie dans laquelle nous trouvons le plus de variation entre les chiffres. Ensuite, les catégories des SP grammaticales et des SP individuelles ont un bon nombre de SP, mais dans toute la distribution, ces catégories sont pourtant assez petites. En ce qui concerne les SP modifiées, nous avons déjà expliqué dans la partie du français qu'elles ont toujours le même nombre d'occurrences dans les deux lignes. Quant à la catégorie des SPIDI, il est intéressant à noter que le nombre de SP soit de 8 aussi bien dans le processus que dans le produit final. Cela veut dire que les étudiants n'ont pas remarqué aucune des formes fautes, et nous étudierons ce fait plus en détail quand nous analyserons les étudiants individuels (*cf.* ch. 5.3.). Pour pouvoir comparer les productions en français et en suédois, nous présentons la distribution des SP en suédois en catégories également en pourcentages. La figure 7 montre les pourcentages dans le processus d'écriture et la figure 8 dans le produit final.



**Figure 7. La distribution des SP en catégories dans le processus en suédois**

Comme nous l'avons noté, les catégories dominantes sont les SP lexicales avec un pourcentage de 44 % et les SP discursives avec un taux de 32 %. La catégorie des SP grammaticales atteint le taux de 10 %, alors que les autres catégories ont des pourcentages plus bas.



**Figure 8. La distribution des SP en catégories dans le produit final en suédois**

Même si les nombres de SP ont été les mêmes dans le processus et dans le produit final dans quelques catégories, le nombre total de SP a changé dans le produit final, et pour cette raison les pourcentages sont différents dans la figure 8. Les SP lexicales couvrent donc 47 % de toutes les SP dans le produit final et le pourcentage des SP discursives a diminué à 28 %. Autrement dit, la différence entre la quantité de SP lexicales et de SP discursives est notablement grande dans le produit final, car les SP L couvrent presque une moitié de toutes les SP utilisées. La partie des SP grammaticales a grandi d'une unité de pourcentage, mais les taux des autres catégories sont restés les mêmes.

Nous pouvons tirer à partir de ces figures que l'ordre des catégories utilisées est identique aussi bien dans le processus que dans le produit final en suédois.

### 5.2.3. En somme

En comparant les résultats des productions en français et en suédois, nous avons trouvé que les nombres de SP différentes et les nombres d'occurrences ont été notablement plus élevés dans les productions en français qu'en suédois. Cela est évidemment en ligne avec les autres résultats entre le français et le suédois que nous avons trouvés dans la partie 5.1 concernant la quantité de SP. Cependant, il est intéressant que le nombre de SP différentes a diminué par 11 SP du processus au produit final en français, tandis qu'en suédois, la différence n'est qu'une SP entre le processus et le produit final (*cf.* le tableau 12, ch. 5.2.1. et le tableau 15, ch. 5.2.2.). Cela montre que les étudiants ont fait plus de changements dans leurs productions en français.

Pour comparer les catégories utilisées en suédois et en français aux autres études, nous avons groupé ensemble les pourcentages présentés dans les parties en haut. La figure 9 ci-dessous montre donc la quantité de SP dans chaque catégorie, et toutes les productions de notre corpus sont incluses.

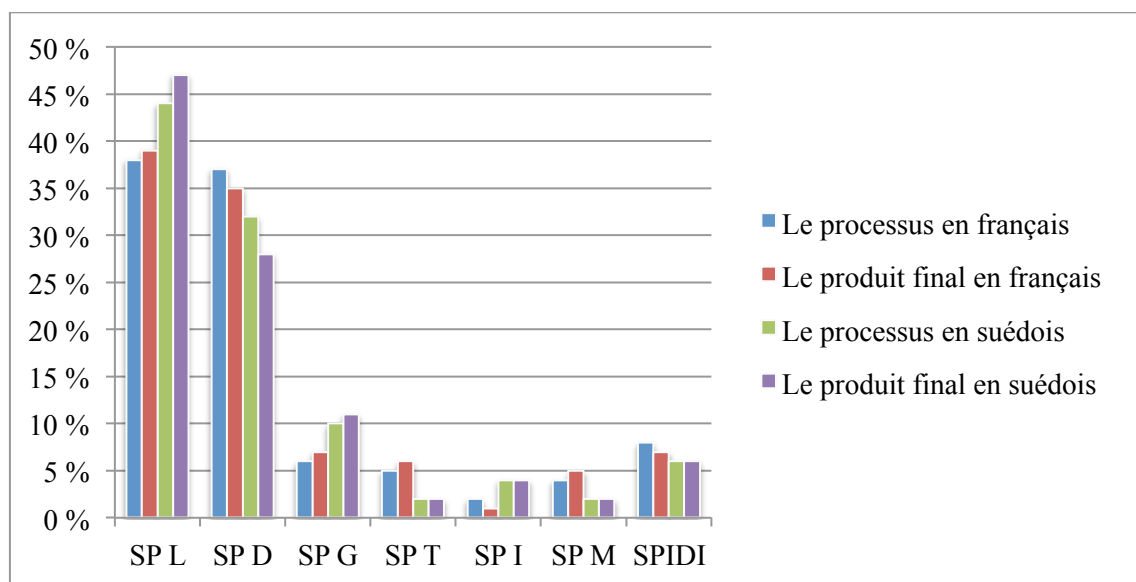


Figure 9. La quantité de SP dans chaque catégorie dans les productions en français et en suédois

Lorsque nous analysons des pourcentages des catégories utilisées, les SP lexicales et les SP discursives sont clairement les catégories les plus utilisées dans les deux langues. Il est pourtant intéressant que les taux de SP lexicales soient notablement plus élevés dans

les productions en suédois qu'en français. En revanche, les taux des SP discursives sont plus élevés en français qu'en suédois, mais la différence n'est pas aussi grande que dans le premier cas. Pour expliquer ces différences entre les langues, il faut que nous examinions les résultats au niveau de chaque étudiant (*cf.* ch. 5.3.). Dans les études de Forsberg (2006) et de Rasmussen (2015), les SP lexicales et les SP discursives sont également les catégories les plus grandes, mais pas dans le même ordre que dans notre étude. Il faut rappeler que ces chercheuses ont étudié le français parlé, mais il est intéressant d'examiner comment nos résultats se représentent par rapport à leurs. Forsberg (2006, p. 78 ; *cf.* aussi Forsberg, 2008) a trouvé que dans des productions orales d'apprenants universitaires, les SP discursives couvrent 69 % de la production, les SP lexicales 18 %, les SP grammaticales 9 % et les SPIDI 1 %. Quant à l'étude de Rasmussen (2015, p. 40), les SP discursives comportent 58 % de toutes les SP utilisées, les SP lexicales 18 %, les SP grammaticales 13 %, les SPIDI 5 % et les SP thématiques 4 %. Notons que Forsberg et Rasmussen ont aussi utilisé une catégorie des SP situationnelles, ce que nous n'avons pas fait (*cf.* ch. 3.3.). Pareillement, il faut rappeler que nous avons formé deux nouvelles catégories, ce qui a évidemment influencé et diminué les nombres de SP dans nos autres catégories. Cependant, nous pouvons voir que les pourcentages que nous avons présentés dans la figure 9, c'est-à-dire aussi les pourcentages du suédois, suivent à peu près les taux des études précédentes. Selon Forsberg (2006, p. 56-57), les SP discursives sont en général plus utilisées à l'oral qu'à l'écrit, ce qui peut expliquer ce fait que nous avons les taux plus élevés dans les SP lexicales que dans les SP discursives. Dans nos résultats, la quantité de SPIDI est plus grande que dans les deux autres études, ce qui pourrait signifier que les étudiants dans notre étude aient un niveau langagier plus bas. Nous examinerons cela plus en détail quand nous analyserons les étudiants individuellement (*cf.* ch. 5.3.). Comme nous l'avons évoqué dans le chapitre 3.3., Cordier (2013), qui a étudié le français parlé, a proposé que sa catégorisation puisse être comparée à celle de Forsberg (2006). Toutefois, Cordier (2013) n'a pas étudié la distribution des SP en catégories *per se*, et en raison de cela, nous ne pouvons pas comparer nos résultats à ceux de son étude. Cordier (*id.*, p. 151-152) a pourtant indiqué que dans son étude, les SP métadiscursives et les SP qui construisent des phrases (à savoir les SP discursives) ont été utilisées notablement plus que les SP référentielles (à savoir les SP lexicales). Ces résultats sont donc en ligne avec ceux de Forsberg (2006) et de Rasmussen (2015).

Dans ce chapitre, nous avons examiné le rapport entre le nombre de SP différentes et leurs occurrences, d'où nous avons continué à présenter la distribution des SP en catégories au niveau général. Nous avons traité séparément les productions en français et les productions en suédois, et vers la fin du chapitre, nous avons comparé les résultats entre eux et aussi aux autres études. Dans les chapitres suivants, nous continuerons à analyser la distribution des SP en catégories au niveau de chaque étudiant et nous présenterons l'analyse selon les langues.

### **5.3. Les profils d'utilisateurs dans l'usage des séquences préfabriquées**

Nous avons déjà présenté au niveau général les catégories des SP utilisées, mais dans ce chapitre, nous examinerons l'usage des SP par les étudiants individuellement. Notre objectif est de répondre à notre troisième question de recherche :

Quels sont les profils d'utilisateurs dans l'usage des séquences préfabriquées des étudiants individuels ?

Chez chaque étudiant, nous présenterons la distribution des SP en catégories et nous traiterons plus en détail quelques catégories et types de SP qui décrivent le mieux l'utilisation des SP par l'étudiant. Nous examinerons séparément les productions en français et en suédois, et après, nous comparerons les résultats de deux langues entre eux. Notre objectif dans ce chapitre n'est donc pas de comparer les étudiants entre eux, mais plutôt de nous concentrer à la variation individuelle chez les étudiants.

Rappelons ici que les sujets sur lesquels les étudiants ont écrit ont concerné les avantages et les inconvénients des études en échange (le texte en français), et les avantages et les inconvénients du fait de travailler à côté des études (le texte en suédois). En ce qui concerne des exemples que nous traiterons dans l'analyse, des explications de marques utilisées dans les exemples sont présentées en annexe 3. Dans les cas des exemples en suédois, nous donnerons des traductions en français, mais les traductions seront simplifiées ; elles ne comprendront que les marques essentielles, comme <BACKSPACE> qui indique l'effacement du texte.

Nous continuerons dans l'ordre des étudiants d'A à D. Le chapitre de l'étudiant en question sera alors divisé selon les deux langues. A la fin de ce chapitre, nous présenterons en résumé dans lequel nous discuterons nos résultats par rapport aux autres études.

### 5.3.1. L'étudiant A

Dans ce chapitre, nous analyserons l'usage des SP par l'étudiant A et nous diviserons ce chapitre selon les langues. L'étudiant A étudie le français comme matière principale et en revanche le suédois est une matière secondaire pour lui. Selon cet étudiant, il a de meilleures compétences en français qu'en suédois et il utilise du français plus en temps libre (*cf.* ch. 5.1.3.). L'étudiant A a passé 10 mois en France en échange universitaire, mais il n'a pas eu de séjours dans une région suédophone. Nous commencerons alors avec la production en français et après nous la comparerons à celle en suédois.

#### En français

Nous avons présenté dans le chapitre 5.1.1. que l'étudiant A a produit 548 mots en français dans le processus d'écriture, la quantité de SP étant de 27 %, et dans le produit final, il a produit 452 mots dont les mots faisant partie des SP ont couvert 28 %. Nous avons donc pu nous attendre à ce que les productions de l'étudiant A comporte un bon nombre de SP différentes et que des occurrences des SP sont nombreuses. Dans le tableau 18 suivant, nous présentons ces nombres de cet étudiant.

**Tableau 18. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant A**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 42                       | 53                                   |
| Produit final | 38                       | 47                                   |

Le tableau 18 montre que le nombre de SP différentes a diminué par quatre dans le produit final, ce qui indique que l'étudiant a fait des changements notables, et que les nombres d'occurrences sont plus grands que ceux de SP différentes, ce qui veut dire que l'étudiant a utilisé certaines SP plus que d'autres. Nous continuons avec les nombres d'occurrences des SP, car nous présentons comment les SP que l'étudiant A a utilisées sont distribuées dans les catégories. Le tableau 19 montre donc la quantité de SP dans chaque catégorie principale en nombres d'occurrences et en pourcentages. Notons que les pourcentages indiquent alors le nombre d'occurrences des SP dans une catégorie par rapport au nombre total d'occurrences, soit dans le processus soit dans le produit final.

**Tableau 19. La distribution des SP en français en catégories ; l'étudiant A**

| SP   | L    | D    | G    | T    | I   | M   | SPIDI | Au total |
|--|------|------|------|------|-----|-----|-------|----------|
| Processus  | 17   | 23   | 5    | 6    | 0   | 1   | 1     | 53       |
| Pourcentage  | 32 % | 43 % | 10 % | 11 % | 0 % | 2 % | 2 %   | 100 %    |
| Produit final  | 13   | 21   | 5    | 6    | 0   | 1   | 1     | 47       |
| Pourcentage  | 28 % | 45 % | 10 % | 13 % | 0 % | 2 % | 2 %   | 100 %    |
| SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires |      |      |      |      |     |     |       |          |

Dans le cas de l'étudiant A, la catégorie des SP discursives est utilisée le plus, aussi bien dans le processus que dans le produit final. Deuxièmement, nous trouvons la catégorie des SP lexicales. Ces catégories sont notablement plus grandes que les autres, et elles sont aussi les seules catégories où il existe des changements dans les nombres d'occurrences entre le processus et le produit final. Ces changements expliquent donc la diminution que nous avons vue parmi les nombres dans le tableau 18. A partir du tableau 19, nous pouvons aussi remarquer que cet étudiant n'a pas du tout produit de SP individuelles, mais les autres petites catégories ont pourtant quelques occurrences. En ce qui concerne les SP particulières que l'étudiant a employées le plus dans toute la production en français, nous les illustrons dans le tableau 20 suivant. Comme l'évoqué dans les chapitres précédents, nous avons employé le nombre minimal de 2 occurrences dans la sélection de ces SP (*cf.* ch. 5.2.1.).

**Tableau 20. Les SP les plus utilisées en français par l'étudiant A**

| SP  | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type |
|---|------------------------|----------------------------|-----------|------|
| A mon avis  | 3                      | 3                          | D         | MAEP |
| Il y a  | 3                      | 3                          | D         | MT   |
| En Finlande   | 3                      | 3                          | T         | SAdv |
| En France   | 2                      | 2                          | T         | SAdv |
| Toutes les fois   | 2                      | 2                          | L         | SAdv |
| J'ai appris   | 2                      | 1                          | L         | SV   |
| Bien sûr  | 2                      | 1                          | D         | MI   |
| Pour moi  | 2                      | 2                          | D         | MAEP |
| Catégories : L= lexicales, D= discursives, T= thématiques<br>Types : SAdv= syntagmes adverbiaux, SV= syntagmes verbaux, MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement, MI= marqueurs interactifs |                        |                            |           |      |

Le tableau 20 indique que l'étudiant A a utilisé ces 8 SP plus que les autres, et naturellement, les SP discursives ont eu le plus d'occurrences. La répétition de ces 8 SP et le nombre d'occurrences des SP au total montrent toutefois que cet étudiant a des compétences langagières avancées et il a clairement des SP automatisées en français.

Pour analyser plus en détail l'usage des SP par l'étudiant A, nous examinons les catégories les plus grandes, à savoir les SP discursives et les SP lexicales, et aussi les autres catégories et types des SP qui décrivent cet usage individuel, c'est-à-dire les SP thématiques, modifiées et les SPIDI dans ce cas du français. Nous illustrons les catégories avec des exemples tirés de la production de l'étudiant.

### Les SP discursives

Nous commençons alors avec la catégorie la plus grande, c'est-à-dire les SP discursives. Dans le tableau 21, nous montrons quels types de SP D l'étudiant A a employé (*cf.* aussi ch. 3.3. pour toutes les sous-catégories).

**Tableau 21. Les types des SP discursives en français de l'étudiant A**

| SP  | Nombre d'occurrences |      |    |          |
|---|----------------------|------|----|----------|
|   | MT                   | MAEP | MI | Au total |
| Discursives   |                      |      |    |          |
| Processus   | 9                    | 11   | 3  | 23       |
| Produit final   | 8                    | 11   | 2  | 21       |
| MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement, MI= marqueurs interactifs |                      |      |    |          |

Le tableau 21 indique que l'étudiant A a utilisé le plus des SP D qui font partie des marqueurs autorégulateurs et des expressions de positionnement (MAEP), comme la SP *à mon avis* dans le tableau 20. Pour illustrer d'autres SP du type de MAEP, nous présentons exemples 1 et 2 de la production de cet étudiant. Toutes les SP présentes dans les exemples sont soulignées et les SP du type que nous traitons (à présent, MAEP) sont mises en gras. Des chiffres entre les marques <n> montrent la durée de la pause en secondes.

(1)

... <11.101> **Il est non plus pas possible** <2.087> ...

L'exemple 1 montre une SP qui est complètement entourée par les pauses. Dans la verbalisation, l'étudiant A commente pendant cette pause de 11 secondes qu'il a pensé à ce quoi il dirait par la suite. Ce commentaire et la position des pauses indiquent que l'étudiant a effectivement produit la suite de mots comme une unité ; il n'existe pas de pauses au milieu de l'unité, donc il s'agit d'une SP. Cette SP particulière pourrait aussi être classifiée comme une SP individuelle, mais nous trouvons qu'elle est plutôt une expression de positionnement, comme *il est possible*. L'étudiant A a donc utilisé la SP une fois et elle est restée la même dans le processus et dans le produit final.

(2)

... <6.164> En plus, les examens <2.155> était super <BACKSPACE6> facile **pour moi** <3.754>. ...

L'exemple 2 présente en revanche une SP du type de MAEP, à savoir *pour moi*, qui fait partie d'une phrase plus longue. L'étudiant a employé cette SP deux fois dans le processus et dans le produit final, et nous pouvons voir qu'elle est une SP automatisée pour l'étudiant.

Des SP discursives, nous présentons encore un exemple du type de marqueurs textuels (MT). L'exemple 3 illustre une SP que nous trouvons une fois dans le processus d'écriture de l'étudiant A, mais pas dans le produit final ; la marque <BACKSPACE13> indique que l'étudiant a effacé ce qu'il venait d'écrire.

(3)

... <2.125> mardi et jeudi **il y avait** une <9.829> <BACKSPACE13> ...

Cette SP *il y avait* représente donc le type de MT, comme la SP *il y a* que l'étudiant a utilisée même trois fois dans la production (voir le tableau 20). Comme nous l'avons évoqué (*cf.* ch. 5.2.), nous avons calculé telles formes différentes séparément, et cela influence le nombre d'occurrences des SP. En effet, l'étudiant A a produit trois fois *il y a*, une fois *il y avait* et une fois *il y aurait* dans le processus, ce qui prouve qu'il s'agit d'une forme automatisée chez lui.

### **Les SP lexicales**

Nous continuons avec la catégorie des SP lexicales qui est deuxièmement la catégorie la plus grande dans la production de l'étudiant A. Le tableau 22 ci-dessous montre les types de SP L que l'étudiant a utilisés.

**Tableau 22. Les types de SP lexicales en français de l'étudiant A**

| SP   | Nombre d'occurrences |    |      |      |          |
|--|----------------------|----|------|------|----------|
|  | SN                   | SV | SAdv | SPré | Au total |
| Lexicales  |                      |    |      |      |          |
| Processus  | 2                    | 6  | 8    | 1    | 17       |
| Produit final  | 2                    | 4  | 6    | 1    | 13       |
| SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux, SPré= syntagmes prépositionnels |                      |    |      |      |          |

Comme le tableau l'indique, l'étudiant A a produit notablement plus de syntagmes adverbiaux (SAdv) et verbaux (SV) que de syntagmes nominaux (SN) et prépositionnels (SPré). Nous avons vu dans le tableau 20 que l'étudiant a utilisé deux fois la SP *toutes les fois* qui représente le type de SAdv, et aussi deux fois la SP *j'ai appris* qui fait partie du type de SV. Autrement, les SP lexicales que l'étudiant a produites ont eu seulement une occurrence dans la production. Comme le type de SAdv est le plus populaire parmi ces quatre types, nous donnerons un exemple d'une SP, en plus de *toutes les fois*. Cet exemple 4 est présent aussi bien dans le processus que dans le produit final de l'étudiant A.

(4)

... <BACKSPACE29> il ya a<BACKSPACE3> a des <3.765> gesn  
<BACKSPACE3>ns qui ont commencé une grève **à la fac**. <4.381> ...

La SP du type de SAdv *à la fac* est produite à la fin d'une phrase et elle est clairement une SP pour cet étudiant, car il s'agit d'une expression qui est souvent utilisée au contexte universitaire. L'étudiant a également employé l'expression *à l'université* comme une SP, ce qui est semblable à cette SP.

Des SP lexicales, nous soulevons aussi le type de syntagmes nominaux (SN) qui a compris deux occurrences. L'exemple 5 ci-dessous illustre une de ces SP.

(5)

... <9.931> une année en échange <11.791> ...

L'exemple 5 indique clairement que cette unité est produite entre les pauses longues qui marquent le temps de réflexion. Dans le cas de cette SP, nous avons exclu l'article indéfini parce qu'il ne joue pas un rôle important dans la SP ; l'article pourrait aussi être défini, mais *année en échange* est néanmoins la base de la SP.

## Les SP thématiques

L'étudiant A a produit le plus de SP thématiques en français parmi tous les quatre étudiants, à savoir 6 SP dans le processus et dans le produit final. Ce nombre total est notablement plus bas que les nombres dans les deux catégories présentées en haut, mais il est pourtant intéressant à examiner les types de SP T de l'étudiant A parce que les autres étudiants n'en ont pas beaucoup produites. Nous nous sommes servie de mêmes sous-catégories que Rasmussen (2015) a utilisées, et par conséquent, il n'en existe que deux ; les syntagmes adverbiaux (SAdv) et les syntagmes nominaux (SN). Les SP T de l'étudiant A sont donc distribuées en ces types comme le suit : 5 occurrences pour les SAdv et une occurrence pour les SN. Les occurrences sont les mêmes dans le processus et dans le produit final. Comme marqué dans le tableau 20, l'étudiant A a employé plusieurs fois les SP T *en France* et *en Finlande* qui sont du type de SAdv. Nous les illustrons dans les exemples 6 et 7.

(6)

... <6.354> Par exemple, <4.271> quand <BACKSPACE5> quand j'ér<BACKSPACE2>étais en France <5.579> ...

(7)

... <BACKSPACE2> 'aurais pas pu apprendre en restant <3.947> en Finlande. <3.486> ...

Ces exemples 6 et 7 montrent donc la manière dont l'étudiant a utilisé les SP contenant un nom du pays. En effet, nous n'aurions pas autrement considéré ces expressions comme les SP, mais comme elles sont répétées plusieurs fois dans tout le corpus, nous avons décidé de les classifiées comme les SP thématiques.

## Les SP modifiées

Comme nous l'avons évoqué (*cf.* ch. 3.3.), la catégorie des SP modifiées est une des deux catégories que nous avons formées d'après notre corpus. Les SP modifiées sont donc les SP que l'étudiant a produites dans le processus, mais qu'il a modifié dans le produit final. Dans la production de l'étudiant A, nous trouvons une SP modifiée. La catégorie originale de cette SP, à savoir *pour une étudiante*, a été discursive et le type a été de MAEP. Pourtant, comme l'étudiant a fait des modifications sur cette SP, nous n'avons pas pu calculer la SP comme la SP D, mais nous avons plutôt classifié l'occurrence dans le processus et l'occurrence dans le produit final comme les SP M.

Nous illustrons ce fait par l'exemple 8 dans lequel nous présentons une partie plus longue du processus de l'étudiant A pour montrer comment la modification est faite.

(8)

... <2.572> C <BACKSPACE1> **Pour une étudiant** <BACKSPACE1>e <2.280>  
d'échange cette période de <2.186> « chômage » on<BACKSPACE2> <4.174> était  
<4.212> <MOUSECLICK> <2.908> <HIGHLIGHT 1248 :1261> <BACKSPACE1>  
<2.646> <MOUSECLICK> <MOUSECLICK> <HIGHLIGHT 1258 :1249>  
<MOUSECLICK> <BACKSPACE2> **moi**, <MOUSECLICK> stressant. ...

Ce que nous voyons dans l'exemple 8, cela équivaut donc à cette phrase dans le produit final « Pour moi, cette période de « chômage » était stressant. » Cette SP représente donc le même type de SP que *pour moi*, ce que nous avons traité dans la partie des SP discursives de l'étudiant A. Cependant, puisque la SP *pour moi* n'est pas construite ici comme une unité, elle représente une SP modifiée.

### Les SPIDI

Dans la production de l'étudiant A, nous ne trouvons qu'une SPIDI, ce qui soutient le fait que cet étudiant a des compétences langagières avancées en français. La catégorie de cette SPIDI, à savoir *à l'étragère\**, est lexicale et le type est de syntagmes adverbiaux (SAdv). La SPIDI est restée la même aussi bien dans le processus que dans le produit final, mais lorsque nous étudions l'exemple 9 ci-dessous, nous pouvons voir qu'il s'agit probablement d'une faute d'orthographe.

(9)

... <4.808> pendant les études **à l'étragÉre**<BACKSPACE3>**ère** <3.823> ...

Il est intéressant que l'étudiant A a aussi produit la forme correcte de cette SP, à savoir *à l'étranger*. Cela indiquerait que cette SP n'est pas complètement automatisée chez l'étudiant A parce qu'il existe de la variation.

Dans cette partie, nous avons étudié les catégories et les exemples qui décrivent l'usage des SP en français de l'étudiant A. Nous continuerons alors avec les catégories que nous trouvons dans la production en suédois de l'étudiant A.

### En suédois

Comme nous l'avons présenté dans le chapitre 5.1.2., l'étudiant A a produit 552 mots en suédois dans le processus d'écriture et 484 mots dans le produit final, les nombres étant

donc plus élevés que les nombres dans ses productions en français. Cependant, la quantité de SP en suédois a été de 22 % dans le processus et de 23 % dans le produit final, c'est-à-dire que les pourcentages sont plus bas en suédois qu'en français. D'après ces résultats, nous avons prévu que les nombres de SP différentes et leurs occurrences sont moins nombreux en suédois. Dans le tableau 23 suivant, nous présentons le rapport entre ces nombres dans la production en suédois de l'étudiant A.

**Tableau 23. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant A**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 33                       | 48                                   |
| Produit final | 32                       | 44                                   |

Le tableau 23 nous indique que le nombre de SP différentes dans le produit final de l'étudiant A n'a changé qu'une SP par rapport au processus, ce qui est en ligne avec les résultats au niveau général que nous avons présentés (*cf.* ch. 5.2.2.). Toutefois, ces nombres sont plus bas que les nombres en français (le tableau 18), comme nous l'avons prévu. L'étudiant A a pourtant produit tellement de SP en suédois qu'il est bien possible de les distribuer dans les catégories. Ensuite, nous présentons dans le tableau 24 la quantité de SP dans chaque catégorie principale en nombres d'occurrences et en pourcentages.

**Tableau 24. La distribution des SP en suédois en catégories ; l'étudiant A**

| SP   | L    | D    | G   | T   | I   | M   | SPIDI | Au total |
|--|------|------|-----|-----|-----|-----|-------|----------|
| Processus  | 22   | 19   | 2   | 1   | 0   | 1   | 3     | 48       |
| Pourcentage  | 46 % | 40 % | 4 % | 2 % | 0 % | 2 % | 6 %   | 100 %    |
| Produit final  | 22   | 15   | 2   | 1   | 0   | 1   | 3     | 44       |
| Pourcentage  | 50 % | 34 % | 5 % | 2 % | 0 % | 2 % | 7 %   | 100 %    |
| SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires |      |      |     |     |     |     |       |          |

Nous pouvons tirer à partir de ce tableau 24 que la catégorie la plus utilisée en suédois par l'étudiant A est clairement la catégorie des SP lexicales couvrant même une moitié de toutes les SP utilisées dans le produit final. La catégorie des SP discursives est en

deuxième lieu et il est intéressant à noter qu'elle est la seule catégorie dans laquelle nous trouvons de la variation entre les nombres de SP dans le processus et dans le produit final. L'ordre des deux catégories les plus utilisées est donc inverse en suédois par rapport à ce qu'il a été en français, mais dans les deux langues, ces catégories ont les pourcentages notablement plus élevés que les autres catégories. En ce qui concerne les autres catégories dans la distribution, leurs occurrences sont minimales. Cet étudiant n'a produit aucune SP individuelle, ce qui a aussi été le cas en français. Il est néanmoins remarquable que la catégorie des SP thématiques n'a qu'une occurrence en suédois, tandis qu'en français nous avons trouvé 6 occurrences. Cela indique que le sujet d'écriture en français, c'est-à-dire des études en échange, a fait naître plus de SP T chez l'étudiant A. La quantité de SPIDI est aussi une particularité qui apparaît dans la distribution, car le nombre est plus élevé en suédois qu'en français. Cela soutient le fait que selon l'étudiant A, ses compétences langagières sont meilleures en français qu'en suédois. En raison de cela, il est probable que les SPIDI soient plus nombreuses en suédois. Quant aux SP différentes que l'étudiant A a employées le plus en suédois, nous présentons une liste de ces SP et leurs traductions dans le tableau 25 suivant.

**Tableau 25. Les SP les plus utilisées en suédois par l'étudiant A**

| SP   | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type |
|--|------------------------|----------------------------|-----------|------|
| Vid sidan av studier, 'à côté d'études'  | 5                      | 5                          | L         | SAdv |
| För mig, 'pour moi'  | 5                      | 4                          | D         | MAEP |
| Det är, 'il est'   | 3                      | 1                          | D         | MT   |
| Jag tänker att, 'je pense que'   | 2                      | 2                          | D         | MAEP |
| Till exempel, 'par exemple'  | 2                      | 2                          | D         | MT   |
| Ha tillräckligt med, 'avoir assez de'  | 2                      | 2                          | L         | SV   |
| Var det, 'était-il'  | 2                      | 2                          | D         | MT   |
| Det var, 'il était'  | 2                      | 2                          | D         | MT   |
| Catégories : L= lexicales, D= discursives<br>Types : SAdv= syntagmes adverbiaux, SV= syntagmes verbaux, MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement |                        |                            |           |      |

L'étudiant A a donc utilisé 8 SP différentes qui ont eu au moins deux occurrences dans la production. Par hasard, le nombre est le même que dans la production en français, mais cela indique aussi que l'étudiant a eu l'habitude de répéter quelques SP dans les deux langues. Il s'agit donc des SP automatisées chez l'étudiant. Nous continuerons à étudier plus en détail l'usage des SP en suédois par l'étudiant A et nous commencerons

avec les catégories les plus grandes, à savoir les SP lexicales et les SP discursives. Des autres catégories, nous traiterons les SPIDI.

### Les SP lexicales

L'étudiant A a produit le plus de SP lexicales en suédois et ces SP sont distribuées dans quatre types. Nous les présentons dans le tableau 26 suivant.

**Tableau 26. Les types de SP lexicales en suédois de l'étudiant A**

| SP   | Nombre d'occurrences |    |      |      |          |
|--|----------------------|----|------|------|----------|
|  | SN                   | SV | SAdv | SAdj | Au total |
| Lexicales  |                      |    |      |      |          |
| Processus  | 2                    | 7  | 12   | 1    | 22       |
| Produit final  | 2                    | 7  | 12   | 1    | 22       |
| SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux, SAdj= syntagmes adjectivaux |                      |    |      |      |          |

A partir de ce tableau, nous voyons que les types les plus utilisés parmi les SP L sont les syntagmes adverbiaux (SAdv) et les syntagmes verbaux (SV). Dans les SP L en français, ces deux types ont aussi été les plus populaires et dans le même ordre. Nous avons marqué dans le tableau 25 que l'étudiant A a utilisé le plus dans toute la production en suédois la SP du type de SAdv *vid sidan av studier* 'à côté d'études', à savoir 5 fois. En effet, cette SP est la seule parmi les SP du type de SAdv qui a eu plus qu'une occurrence. Dans l'exemple 10 ci-dessous, nous illustrons comment l'étudiant a utilisé cette SP.

(10)

... <4.290> JA g <BACKSPACE3> ag t<BACKSPACE1> skulle inte vilja jobba **vid sidan av studier** just eftersom <4.276> ...

'Je ne voudrais pas travailler **à côté d'études** parce que'

L'étudiant a écrit cette SP du type de SAdv chaque fois en liaison avec le verbe *jobba* 'travailler', mais le verbe n'existe pas toujours près de la SP comme dans cet exemple ; il existe même des pauses entre le verbe et la SP dans quelques phrases, et pour cette raison, nous pouvons dire que *vid sidan av studier* est la SP pour l'étudiant A. Il faut pourtant noter que la suite de mots *jobba vid sidan av studier* 'travailler à côté des études' a été présente dans la question qui a été le sujet d'écriture, il est donc possible qu'en raison de cela, l'étudiant a employé plusieurs fois cette SP. Dans la verbalisation, l'étudiant a également commenté au moment de cette SP que les phrases (les questions)

sont utiles parce qu'il peut les employer pour l'inspiration. Cependant, il est intéressant à noter que l'étudiant a toujours utilisé la forme indéfinie *studier* 'des études' même si la forme définie *studierna* 'les études' était aussi possible. Cela pourrait indiquer que l'étudiant a déjà acquis la SP auparavant.

Pour donner un autre exemple des SP L, nous traitons les SP du type de syntagmes verbaux (SV) dont l'étudiant a utilisé nombreuses. Nous prenons la SP *ha erfarenhet* 'avoir d'expérience' que l'étudiant a employé une fois dans le processus et dans le produit final, et nous l'illustrons dans l'exemple 11.

(11)

... <2.555> <MOUSECLICK>redan <MOUSECLICK> du har erfarenhet .  
 <BACKSPACE2>. <3.317> ...  
 '<MOUSECLICK>déjà <MOUSECLICK> tu as d'expérience'

Comme les marques de <MOUSECLICK> indiquent, l'étudiant a déplacé la souris dans le texte et après cela, il a produit la suite de mots *har erfarenhet* 'as d'expérience' comme une unité, c'est-à-dire comme une SP, car nous ne trouvons pas de pause au milieu de cette expression.

### Les SP discursives

La deuxième catégorie que l'étudiant A a utilisée le plus a été les SP discursives. L'étudiant s'est servi de deux types de SP D, et nous les présentons dans le tableau 27 ci-dessous.

**Tableau 27. Les types de SP discursives en suédois de l'étudiant A**

| SP   | Nombre d'occurrences |      |          |
|--|----------------------|------|----------|
|  | MT                   | MAEP | Au total |
| Discursives  |                      |      |          |
| Processus  | 11                   | 8    | 19       |
| Produit final  | 8                    | 7    | 15       |
| MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement |                      |      |          |

Les chiffres dans ce tableau 27 sont en ligne avec ce que nous avons vu dans la liste des SP les plus utilisées (le tableau 25), c'est-à-dire que l'étudiant a utilisé plus de SP du type de marqueurs textuels (MT), mais le nombre de SP du type de marqueurs autorégulateurs et expressions de positionnement (MAEP) est aussi élevé. Des

marqueurs textuels, nous traitons la SP *det är* ‘il est’ qui a eu trois occurrences dans le processus, mais seulement une occurrence dans le produit final. Comme nous l’avons dit, la catégorie des SP D est la seule dans laquelle le nombre d’occurrences a diminué dans le produit final, et il est donc intéressant à trouver la raison derrière ces changements. L’exemple 12 ci-dessous illustre alors un cas où la SP a été effacée.

(12)

... <4.359> det är lätt <BACKSPACE12>jobbet är lätt för mig och <BACKSPACE5>.  
...  
'il est facile <BACKSPACE12>le travail est facile pour moi et <BACKSPACE5>.'

Dans le produit final, cette phrase est donc « ... jobbet är lätt för mig. » ‘le travail est facile pour moi’. Nous pouvons voir à partir de l’exemple 12 que l’étudiant a produit *det är* ‘il est’ comme une SP, mais il a décidé de remplacer cette expression impersonnelle par le mot défini *jobbet* ‘le travail’, et pour cette raison, la SP n’existe pas dans le produit final. Toutefois, il est intéressant à noter que l’étudiant A a employé plusieurs fois « la forme de base » de cette SP, car il a utilisé deux fois les SP *var det* ‘était-il’ et *det var* ‘il était’. Nous avons évoqué que nous avons calculé telles SP séparément (cf. ch. 5.2.), mais il est clair que cette SP est très automatisée chez l’étudiant A. Cet exemple 12 est aussi marquant parce qu’il montre l’avantage de l’étude du processus ; si nous avons étudié seulement le produit final, nous n’aurions pas pu savoir que l’étudiant a utilisée cette SP particulière plusieurs fois.

Nous présentons un autre exemple des SP D qui est du type de MAEP. Nous traitons la SP *för mig* ‘pour moi’ qui a aussi été présente dans la liste des SP les plus utilisées, mais il est intéressant à l’analyser parce que l’étudiant A a employé la même SP *pour moi* aussi en français plusieurs fois. Dans l’exemple 13, nous illustrons un cas de cette SP qui est présent aussi bien dans le processus que dans le produit final.

(13)

... <3.065> <BACKSPACE5> Jag tänker <BACKSPACE11> För mig är helgen  
<5.484> ...  
'Je [pense] <BACKSPACE11> Pour moi le week-end est'

Nous avons inclus dans cet exemple « *jag tänker\** », qui veut probablement dire *jag tänker* ‘je pense’, pour montrer comment l’étudiant a utilisé la SP *för mig* ‘pour moi’ pour remplacer une autre expression. En effet, cela indique que l’expression *för mig* est effectivement une SP pour cet étudiant et qu’il s’agit d’une SP du type de MAEP. En

somme, il semble que la même SP soit très automatisée chez l'étudiant aussi bien en suédois qu'en français.

### Les SPIDI

Dans la production en suédois, la catégorie des SPIDI est marquante chez l'étudiant A parce qu'il en a produit plus en suédois qu'en français. Toutes les trois SPIDI en suédois sont restées les mêmes dans le processus et dans le produit final, dont deux représentent les SP L du type de SV et une SP représente la SP D du type de MAEP. Nous présentons dans l'exemple 14 l'une des SP L, à savoir la SPIDI *lägga märk\* till* 'remarquer'.

(14)

... J<BACKSPACE1>Jag **har lagt märk till** att det passar mig <2.970> ...  
'J'**ai remarqué** que ça va pour moi'

Comme la forme correcte de cette expression serait *har lagt märke till*, il semble qu'il s'agit d'une faute d'orthographe. Pourtant, il est intéressant que l'étudiant a produit la même expression correctement dans le processus, mais dans ce cas, il existe une pause au milieu d'un mot, à savoir « ... jag har l<2.622>agt märke till att... ». Cela et l'exemple 14 montrent donc que la SP *lägga märke till* 'remarquer' n'est pas automatisée chez l'étudiant A, et il est même possible que nous ne puissions pas parler de SP dans ce cas.

La SPIDI qui a été une SP D du type de MAEP a bien montré que la langue française est supérieure à la langue suédoise chez l'étudiant A, comme il l'a aussi constaté lui-même. Nous illustrons ce fait dans l'exemple 15 ci-dessous.

(15)

... <6.399> Ja<BACKSPACE2> **A min asivkt** är <2.917> ...  
'[Je]<BACKSPACE2> **A mon avis** est'

A partir de cet exemple, nous pouvons clairement voir que les suites de mots soulignées sont semblables, et cela montre que l'expression en français a influé sur celle en suédois. En suédois, l'expression *å min åsikt* 'à mon avis' est possible mais rare, tandis que l'expression *enligt min åsikt* 'selon mon avis= à mon avis' est commune<sup>8</sup>. Dans le processus de l'étudiant A, nous trouvons aussi un autre point où l'étudiant commence à

---

<sup>8</sup> Nous avons exécuté de recherches dans un dictionnaire en ligne <https://svenska.se/>, consulté le 26 mars 2020.

écrire « A min », mais il l'efface tout de suite. Dans la verbalisation, il commente qu'il a en effet pensé comment dire « à mon avis » en suédois et il constate aussi qu'il s'est souvenu plus de mots en français qu'en suédois. Tout cela est en ligne avec le fait que selon l'étudiant A, ses compétences langagières sont meilleures en français qu'en suédois.

Dans cette partie, nous avons examiné les deux catégories que l'étudiant a utilisées le plus en suédois et également la catégorie des SPIDI qui a montré un point intéressant de l'usage des SP par l'étudiant A. Avant tout, nous avons comparé les productions en français et en suédois entre eux pour mieux décrire la manière de cet étudiant d'utiliser des SP. Dans le chapitre suivant, nous continuerons avec l'étudiant B.

### **5.3.2. L'étudiant B**

L'étudiant B étudie l'anglais comme matière principale et, par conséquent, le français et le suédois sont tous les deux ses matières secondaires. Selon l'étudiant B, ses compétences langagières sont meilleures en suédois qu'en français et il utilise du suédois plus souvent ; nous avons cependant trouvé des résultats à peu près inverses dans le chapitre 5.1.3. Cet étudiant a passé un mois en France sans interruption et il a travaillé trois mois dans une région suédophone, mais il n'a pourtant pas été en échange universitaire. Dans ce qui suit, nous commencerons avec la partie du français.

#### **En français**

Nous avons présenté dans le chapitre 5.1.1. que l'étudiant B a produit beaucoup de mots en français, à savoir 528 mots dans le processus et 421 mots dans le produit final, et la quantité de SP a été la plus élevée parmi toutes les productions en français, c'est-à-dire 32 % dans le processus et même 36 % dans le produit final. Cela nous a indiquée que la production de l'étudiant B comportera nombreuses SP différentes et notamment le nombre d'occurrences des SP sera haut. Dans le tableau 28 suivant, nous montrons ce rapport entre le nombre de SP différentes et leurs occurrences dans la production de cet étudiant.

**Tableau 28. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant B**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 42                       | 60                                   |
| Produit final | 38                       | 55                                   |

Le tableau 28 indique que l'étudiant B a effacé dans le produit final même quatre suites de mots qu'il avait produit comme les SP. Cela est ainsi en ligne avec le fait que l'étudiant a effacé beaucoup de texte en général dans le produit final par rapport au processus. Comme nous l'avons prévu, les nombres d'occurrences des SP sont pourtant très élevés, en effet les plus élevés dans notre étude. Ces occurrences pourraient signaler que la distribution des SP en catégories serait riche, mais en fait, la distribution de l'étudiant B comprend deux catégories sans aucune occurrence. La distribution des SP est illustrée dans le tableau 29 ci-dessous.

**Tableau 29. La distribution des SP en français en catégories ; l'étudiant B**

| SP   | L    | D    | G   | T   | I   | M   | SPIDI | Au total |
|--|------|------|-----|-----|-----|-----|-------|----------|
| Processus  | 24   | 22   | 4   | 0   | 0   | 2   | 8     | 60       |
| Pourcentage  | 40 % | 37 % | 7 % | 0 % | 0 % | 3 % | 13 %  | 100 %    |
| Produit final  | 23   | 20   | 4   | 0   | 0   | 2   | 6     | 55       |
| Pourcentage  | 42 % | 36 % | 7 % | 0 % | 0 % | 4 % | 11 %  | 100 %    |
| SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires |      |      |     |     |     |     |       |          |

En ce qui concerne les catégories les plus utilisées par l'étudiant B, cette distribution suit le même modèle que nous avons vu dans les chapitres précédents. Dans le cas de l'étudiant B, la catégorie la plus grande est les SP lexicales et la deuxième catégorie, à savoir les SP discursives, contient presque les mêmes pourcentages. Quant aux autres catégories, la proportion de SPIDI est remarquable, ce qui pourrait refléter le fait que les compétences langagières de l'étudiant ne sont pas très avancées en français. Comme nous l'avons évoqué, cet étudiant n'a produit aucune SP thématique ou individuelle en français, ce qui peut dépendre du sujet d'écriture, à savoir des études en échange, qui n'a pas réveillé telles SP chez l'étudiant B. En ce qui concerne les SP différentes que l'étudiant a employées au moins deux fois dans toute sa production en français, la liste

comprend des SP de trois catégories les plus grandes, comme illustré dans le tableau 30 suivant.

**Tableau 30. Les SP les plus utilisées en français par l'étudiant B**

| SP   | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type   |
|--|------------------------|----------------------------|-----------|--------|
| A l'université   | 6                      | 6                          | L         | SAdv   |
| En échange   | 4                      | 4                          | L         | SAdv   |
| Par exemple  | 4                      | 4                          | D         | MT     |
| Faire des/les études   | 3                      | 3                          | L         | SV     |
| Parce que  | 3                      | 3                          | D         | MT     |
| Etudes en échange  | 2                      | 2                          | L         | SN     |
| Il n'y a pas   | 2                      | 2                          | D         | MT     |
| Beaucoup des   | 2                      | 1                          | SPIDI     | G (QD) |
| Catégories : L= lexicales, D= discursives, SPIDI= SP idiosyncrasiques intérimaires, G= grammaticales   |                        |                            |           |        |
| Types : SAdv= syntagmes adverbiaux, SV= syntagmes verbaux, SN= syntagmes nominaux, MT= marqueurs textuels, QD= expressions de quantité et de degré |                        |                            |           |        |

A partir de ce tableau, nous voyons que cet étudiant a en effet répété quelques SP beaucoup par rapport aux autres SP qu'il a produites. Cela montre que ces SP sont effectivement automatisées chez l'étudiant B. Pour examiner plus en détail l'usage des SP en français par cet étudiant, nous continuons à analyser les catégories les plus grandes, c'est-à-dire les SP lexicales, les SP discursives et aussi les SPIDI, et en plus, nous traitons les SP modifiées.

### Les SP lexicales

L'étudiant B a produit en français le plus de SP qui font partie des SP lexicales. Ces SP sont distribuées dans les types différents, et nous les présentons dans le tableau 31 ci-dessous.

**Tableau 31. Les types de SP lexicales en français de l'étudiant B**

| SP   | Nombre d'occurrences |    |      |      |          |
|--|----------------------|----|------|------|----------|
|  | SN                   | SV | SAdv | SPré | Au total |
| Lexicales  |                      |    |      |      |          |
| Processus  | 3                    | 6  | 14   | 1    | 24       |
| Produit final  | 3                    | 5  | 14   | 1    | 23       |
| SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux, SPré= syntagmes prépositionnels |                      |    |      |      |          |

Le tableau 31 indique clairement que le type de syntagmes adverbiaux (SAdv) est le plus populaire parmi tous les types que l'étudiant a utilisés. Même si notre objectif est d'étudier l'usage individuel dans ce chapitre, il est pourtant intéressant à noter que les types les plus utilisés par l'étudiant B dans la catégorie des SP L, à savoir les SAdv et les SV, ont été les mêmes dans les distributions de l'étudiant A. Dans la liste des SP les plus utilisées par l'étudiant B (le tableau 30), les SP du type de SAdv *à l'université et en échange* sont présentes à la tête de la liste. Nous illustrons l'usage de la première SP par l'exemple 16 ci-dessous.

(16)

... <13.834> <MOUSECLICK> <LEFT1> j'<2.022>éta<BACKSPACE10> Quand j'ai commencé d'etu<BACKSPACE3>étur<BACKSPACE1>dier à l'université, ...

Cette SP est clairement une SP automatisée chez l'étudiant B, car il l'utilise dans des phrases diverses dans sa production. Par exemple, dans le produit final, nous trouvons d'autres phrases comme « pendant la deuxième année à l'université », « je fais plusieurs matières à l'université » et « passer par exemple huit années à l'université », et dans toutes ces phrases l'étudiant produit *à l'université* sans pauses, c'est-à-dire comme une SP.

L'étudiant B a produit plusieurs SP qui font partie du type de syntagmes verbaux (SV). Dans le tableau 30, nous avons aussi vu que l'étudiant a utilisé trois fois dans le processus et dans le produit final la SP du type de SV *faire des/les études*, et nous l'illustrons dans l'exemple 17.

(17)

... <BACKSPACE4>j'ai pensé que j<BACKSPACE15>e pensais que je voudrais faire des études en échange. Mais <2.180> ...

Cet exemple 17 comprend en effet quatre SP différentes, mais la SP L *faire des études* représente le type de SV. Dans deux cas sur trois, l'étudiant a utilisé cette SP en liaison avec une autre expression, à savoir *en échange* qui est une SP L du type de SAdv. Néanmoins, comme l'étudiant a aussi produit ces SP séparément, il est clair que nous les traitons comme les SP séparées. Nous avons compris aussi bien l'article indéfini que l'article défini dans cette SP parce que la base de la SP est pourtant la même, et nous avons évidemment procédé de la même manière chez les autres étudiants.

Des SP lexicales, nous prenons encore le type de syntagmes prépositionnels (SP<sub>Pré</sub>), car les SP de ce type sont assez rares dans tout notre corpus. L'étudiant B a employé une fois la SP *à cause de* et l'exemple 18 suivant la montre.

(18)

... <4.464> passé<BACKSPACE1>er p<BACKSPACE1> <10.038> par exemple huit années à l'unv<BACKSPACE1>iversité<BACKSPACE1>é **à cause de** mes études en échn<BACKSPACE1>ange. <14.010> ...

La phrase qui équivaut à cet exemple dans le produit final est donc « passer par exemple huit années à l'université à cause de mes études en échange ». Nous voyons que cette suite de mots se trouve au milieu de la phrase, mais elle est cependant produite comme une unité.

### Les SP discursives

Avec une petite différence à la première catégorie, les SP discursives représentent la deuxième catégorie que l'étudiant B a utilisée le plus. Le tableau 32 suivant montre comment des occurrences sont distribuées parmi les types différents.

**Tableau 32. Les types de SP discursives en français de l'étudiant B**

| SP   | Nombre d'occurrences |      |          |
|--|----------------------|------|----------|
|  | MT                   | MAEP | Au total |
| Processus  | 18                   | 4    | 22       |
| Produit final  | 17                   | 3    | 20       |
| MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement |                      |      |          |

Comme marqué dans ce tableau, le type de marqueurs textuels (MT) a eu notablement plus d'occurrences que le type de marqueurs autorégulateurs et expressions de positionnement (MAEP). Dans la liste des SP les plus utilisées (le tableau 30), nous trouvons trois SP du type de MT que l'étudiant a employées plusieurs fois, à savoir *par exemple*, *parce que* et *il n'y a pas*. Nous présentons l'exemple 19 qui comprend la SP *il n'y a pas* et aussi l'autre SP du type de MT *il y a*.

(19)

... <17.915> <BACKSPACE3> il peut être que dans <BACKSPACE7>'**il y a** beaucoup d'autres choses de faire en échange et **il n'y a pas** <3.017> ...

L'étudiant B a donc produit deux fois la SP *il n'y a pas*, une fois la SP *il y a* et une fois la SP *il y a aussi* que nous avons calculée séparément. Ces occurrences indiquent que la forme de base *il y a* est clairement une SP automatisée chez l'étudiant et qu'il a l'habitude de l'utiliser dans sa production.

En ce qui concerne le type de MAEP, nous avons antérieurement montré dans l'exemple 17 que l'étudiant a produit l'expression *j'ai pensé que* comme une SP et puis il l'a tout de suite remplacée par *je pensais que* qui représente aussi une SP pour l'étudiant. Ces SP sont alors du type de MAEP et dans l'exemple 20 ci-dessous, nous présentons une autre telle SP que l'étudiant a produite.

(20)

... **Pour moi**, ca serait très difficile de déménager tout qu<BACKSPACE1> mes choses  
...

Dans cet exemple, la SP *pour moi* se distingue clairement parce que l'étudiant commence la nouvelle phrase avec celle-ci et il la produit comme une unité. Toutefois, l'étudiant l'a utilisée seulement une fois dans le processus et dans le produit final.

### Les SPIDI

La catégorie des SPIDI est remarquable chez l'étudiant B, car la proportion des SPIDI a été de 13 % dans le processus et de 11 % dans le produit final. La distribution en catégories et en types montre également que nous trouvons des SPIDI dans plusieurs catégories dans la production de cet étudiant. Le tableau 33 ci-dessous le montre.

**Tableau 33. Les types de SPIDI en français de l'étudiant B**

| SP            | Nombre d'occurrences |        |        |          |          |
|---------------|----------------------|--------|--------|----------|----------|
|               | G (QD)               | G (PI) | L (SN) | D (MAEP) | Au total |
| Processus     | 4                    | 2      | 1      | 1        | 8        |
| Produit final | 3                    | 2      | 0      | 1        | 6        |

G= grammaticales, QD= expressions de quantité et de degré, PI= pronoms indéfinis, L= lexicales, SN= syntagmes nominaux, D= discursives, MAEP= marqueurs autorégulateurs et expressions de positionnement

Le nombre élevé des SPIDI grammaticales du type de QD est causé par le fait que l'étudiant a eu de difficultés avec l'expression *beaucoup de*. Nous l'illustrons dans les exemples 21 et 22.

(21)

... parce qu'il n'y a pas beau<BACKSPACE4>**beaucoup des** études ...

(22)

... j'ai **tellement beaucoup** des choses. <3.186> ...

L'étudiant a donc produit la SPIDI *beaucoup des\** deux fois dans le processus et une fois dans le produit final, et la SPIDI *tellement beaucoup\** une fois dans le processus et dans le produit final. Dans l'exemple 22, nous trouvons également *beaucoup des\**, mais dans ce cas, nous nous sommes concentrée à l'expression *tellement beaucoup\**. Néanmoins, nous avons vu dans l'exemple 19 que l'étudiant a aussi utilisé de l'expression correcte *beaucoup de* comme une SP et en effet, dans la verbalisation il commente qu'il note qu'il a écrit *beaucoup des* même s'il sait que la forme correcte est *beaucoup de*. Ses commentaires et sa production indiquent toutefois que la SP *beaucoup de* n'est pas automatisée chez lui. Quant à la SPIDI *tellement beaucoup*, l'étudiant le commente également dans la verbalisation et il se demande si c'est une expression que l'on peut utiliser ou pas. Cela montre qu'il effectivement s'agit d'une SPIDI pour cet étudiant.

### Les SP modifiées

Dans le cadre de la production en français, nous traitons dernièrement les SP modifiées dont l'étudiant B a produit deux différentes. Il est intéressant que les deux SP que l'étudiant a produites dans le processus ont été correctes, mais à cause de modifications, ces SP sont devenues les SPIDI. La catégorie de ces SP est donc les SP modifiées, mais les deux SP représentent une SPIDI lexicale du type de SV et une SPIDI lexicale du type de SAdv. Nous les illustrons dans les exemples 23 et 24.

(23)

... <5.077> **commencer des etude**<BACKSPACE5>**études** ou on <14.676> ...

... <5.461> <MOUSECLICK> <BACKSPACE1>é <10.061> ...

L'exemple 23 montre que l'étudiant B a produit une SP L du type de SV *commencer des études*, mais à la fin de sa production quand il a révisé le texte, il l'a changée. Dans le produit final, nous trouvons alors la SPIDI *comméncer des études\**.

(24)

... <12.699> utiliser une langue èt<BACKSPACE2>+e<BACKSPACE2>étrangère  
de<BACKSPACE2> **tout le jour** <5.852> ...  
... <5.465> <MOUSECLICK> s <MOUSECLICK> s <MOUSECLICK> s <11.858>  
...

L'exemple 24 illustre la SP L du type de SAdv *tout le jour* dont l'étudiant a modifiée également à la fin de sa production, comme les marques <MOUSECLICK> et les lettres *s* indiquent. Dans le produit final, cette expression représente donc la SPIDI *touts les jours\**, mais il semble clair que l'étudiant a essayé d'écrire *tous les jours*. Ces deux exemples indiquent toutefois qu'il s'agit des SP qui ne sont pas automatisées chez l'étudiant B ; elles représentent plutôt la généralisation des règles de la langue cible. Il est aussi intéressant à noter la manière dont l'étude du processus est avantageuse pour nous ; sans l'étude du processus, nous n'aurions pas su que l'étudiant B avait aussi produit les formes correctes de ces SP.

Dans cette partie, nous avons analysé l'usage des SP en français par l'étudiant B et nous avons examiné les catégories qui lui décrivent le mieux. Nous continuerons dans la partie suivante avec la production en suédois de cet étudiant.

### **En suédois**

Comme nous l'avons montré dans le chapitre 5.1.2., l'étudiant B a produit 483 mots dans le processus en suédois, ayant la quantité de SP de 22 %, et 434 mots dans le produit final, la quantité de SP étant aussi de 22 %. Ces chiffres sont notablement plus bas dans la production de l'étudiant B en suédois qu'en français, ce qui est en ligne avec le fait que nous avons montré dans le chapitre 5.1.3., c'est-à-dire que les pourcentages des SP en suédois ont été plus bas aussi au niveau général. Cet étudiant a pourtant évalué que ses compétences langagières sont meilleures en suédois, et en raison de cela, il est intéressant à étudier les nombres de SP différentes et leurs occurrences dans sa production. Ces chiffres sont présentés dans le tableau 34 ci-dessous.

**Tableau 34. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant B**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 29                       | 43                                   |
| Produit final | 28                       | 40                                   |

A partir du tableau 34, nous pouvons voir que les nombres de SP différentes sont considérablement plus petits que ceux en français (à savoir 42 et 38 respectivement), mais les nombres d'occurrences sont cependant assez élevés en suédois. En somme, tous les nombres dans le tableau montrent que l'étudiant n'a pas fait beaucoup de changements dans le produit final, ce qui indique que nous pouvons trouver des SP automatisées dans sa production en suédois. Nous continuons en présentant la distribution des SP utilisées en catégories, et nous montrons la quantité de SP dans chaque catégorie en nombres d'occurrences et en pourcentages dans le tableau 35 suivant.

**Tableau 35. La distribution des SP en suédois en catégories ; l'étudiant B**

| SP   | L    | D    | G    | T   | I   | M   | SPIDI | Au total |
|--|------|------|------|-----|-----|-----|-------|----------|
| Processus  | 17   | 16   | 6    | 0   | 2   | 1   | 1     | 43       |
| Pourcentage  | 40 % | 37 % | 14 % | 0 % | 5 % | 2 % | 2 %   | 100 %    |
| Produit final  | 17   | 13   | 6    | 0   | 2   | 1   | 1     | 40       |
| Pourcentage  | 42 % | 32 % | 15 % | 0 % | 5 % | 3 % | 3 %   | 100 %    |
| SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires |      |      |      |     |     |     |       |          |

Comme le tableau 35 l'indique, les catégories que l'étudiant B a utilisées le plus en suédois sont les SP lexicales et les SP discursives. Les pourcentages dans ces catégories sont très similaires en suédois et en français, et la différence entre les deux catégories est petite aussi en suédois. A partir de ce tableau, nous voyons que la catégorie des SP discursives est en fait la seule catégorie dans laquelle le nombre d'occurrences a diminué dans le produit final. Il est notable que la catégorie des SPIDI n'a eu qu'une occurrence, tandis que dans la production en français, la quantité de SPIDI a été même de 8 occurrences. Cela peut indiquer que les compétences langagières de l'étudiant B sont en effet meilleures en suédois, comme il l'a constaté. Quant aux autres catégories, l'étudiant B a produit des SP individuelles en suédois, alors qu'en français il ne l'a pas fait, mais la catégorie des SP thématiques est restée sans aucune occurrence. La proportion de SP grammaticales a en revanche augmenté par rapport à la production en français. En ce qui concerne les SP différentes que l'étudiant a employé au moins deux fois, nous en avons trouvé 9 dans sa production. Ces SP et leurs traductions sont présentées dans le tableau 36 suivant.

**Tableau 36. Les SP les plus utilisées en suédois par l'étudiant B**

| SP  | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type   |
|---|------------------------|----------------------------|-----------|--------|
| Till exempel, 'par exemple'   | 4                      | 3                          | D         | MT     |
| Det är, 'il est'  | 4                      | 3                          | D         | MT     |
| Så mycket, 'tellement'  | 3                      | 3                          | G         | QD     |
| Vid sidan av studierna, 'à côté des études'   | 2                      | 2                          | L         | SAdv   |
| Hitta tid för, 'trouver du temps pour'  | 2                      | 2                          | L         | SV     |
| Dag/ar per vecka, 'un jour/des jours par semaine'   | 2                      | 2                          | L         | SAdv   |
| Kundservice industrin, 'le service client'  | 2                      | 2                          | I         | L (SN) |
| Även om, 'même si'  | 2                      | 2                          | D         | MT     |
| Under studietiden, 'pendant le temps d'études'  | 2                      | 2                          | L         | SAdv   |
| Catégories : L= lexicales, D= discursives, G= grammaticales, I= individuelles<br>Types : SAdv= syntagmes adverbiaux, SN= syntagmes nominaux, SV= syntagmes verbaux, MT= marqueurs textuels, QD= expressions de quantité et de degré |                        |                            |           |        |

Les SP sur cette liste suivent la distribution que nous avons présentée, mais il est intéressant à noter que la SP G *så mycket* 'tellement' a acquis même trois occurrences, tandis que dans la liste des SP en français, il n'existe pas de SP G. Dans la partie qui suit, nous analyserons plus précisément comment l'étudiant B a utilisé les SP en suédois et nous examinerons non seulement les catégories les plus grandes, à savoir des SP lexicales et des SP discursives, mais aussi les catégories des SP grammaticales et des SP individuelles pour décrire son usage des SP.

### **Les SP lexicales**

L'étudiant B a produit en suédois le plus de SP qui font partie de la catégorie des SP lexicales, comme il l'a fait en français. Dans la production en français, les SP L sont distribuées dans quatre types différents, mais en suédois elles sont distribuées seulement dans deux types. Nous les présentons dans le tableau 37 ci-dessous.

**Tableau 37. Les types de SP lexicales en suédois de l'étudiant B**

| SP  | Nombre d'occurrences |      |          |
|---|----------------------|------|----------|
|   | SV                   | SAdv | Au total |
| Processus   | 7                    | 10   | 17       |
| Produit final                                     | 7                    | 10   | 17       |
| SV= syntagmes verbaux, SAdv= syntagmes adverbiaux |                      |      |          |

Comme les deux types utilisés sont les syntagmes adverbiaux (SAdv) et les syntagmes verbaux (SV), nous voyons que les SP L du type de SAdv sont plus nombreuses, ce qui a aussi été le cas dans la production en français. Dans la liste des SP les plus utilisées (le tableau 36), nous trouvons la SP L du type de SAdv *vid sidan av studiern* 'à côté des études' et nous avons noté (*cf.* ch. 5.3.1.) qu'il s'agit d'une expression qui a été présente dans le sujet d'écriture. Néanmoins, nous illustrons dans l'exemple 25 comment l'étudiant B l'a utilisée.

(25)

<START> <7.838> Många folk jobba <BACKSPACE1>r <2.568>  
vis<BACKSPACE1>d sida<BACKSPACE1>an av studiern. <MOUSECLICK>  
 <3.141> ...

'Plusieurs gens travaillent à côté des études.'

Dans l'exemple 25, la marque <START> signifie que cette phrase est le début de la production de l'étudiant. Il est donc intéressant à remarquer que l'étudiant a produit cette SP *vid sidan av studiern* 'à côté des études' déjà dans la première phrase, et l'expression se trouve clairement entre deux pauses comme une SP. Pourtant, comme l'étudiant a utilisé cette SP pour commencer sa production, il semble qu'il a été inspiré par le sujet d'écriture dans lequel cette expression a été présente. Quant aux marques <BACKSPACE> à l'intérieur de la SP, elles sont causées par le fait que l'étudiant a écrit avec un clavier auquel il n'est pas habitué, c'est-à-dire qu'il a toutefois produit la SP de manière fluide.

Des SP L du type de SV nous présentons la SP que l'étudiant a utilisée deux fois, à savoir *hitta tid för* 'trouver du temps pour'. Nous illustrons cette SP dans l'exemple 26 suivant.

(26)

... Därför kan det kännas svårt att hitta tid för <2.069> en<BACKSPACE1>tt jobb <4.143> <BACKSPACE8> att jobba <3.148> ...  
‘[En raison de cela] il peut [être] difficile de trouver du temps pour un travail <BACKSPACE8> travailler’

L'exemple 26 montre qu'il n'existe pas de pause au milieu de cette suite de mots, mais après l'étudiant a une pause de réflexion, ce qui renforce d'une certaine manière le fait que *hitta tid för* 'trouver du temps pour' est produit comme une unité.

### Les SP discursives

Dans la catégorie des SP discursives, l'étudiant B a utilisé des SP D de deux types différents, à savoir les types de marqueurs textuels (MT) et de marqueurs autorégulateurs et expressions de positionnement (MAEP). Leurs proportions sont indiquées dans le tableau 38.

Tableau 38. Les types de SP discursives en suédois de l'étudiant B

| SP   | Nombre d'occurrences |      |          |
|--|----------------------|------|----------|
|  | MT                   | MAEP | Au total |
| Processus  | 13                   | 3    | 16       |
| Produit final  | 11                   | 2    | 13       |
| MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement |                      |      |          |

Dans la production en français, l'étudiant B a employé notablement plus de SP du type de MT que du type de MAEP, et à partir de ce tableau, nous pouvons dire que la distribution est similaire également en suédois. Dans le tableau 36, nous avons vu que cet étudiant a utilisé le plus souvent les SP D du type de MT *till exempel* 'par exemple' et *det är* 'il est'. Il est intéressant que l'étudiant B a aussi employé la SP *par exemple* en français plusieurs fois, et en raison de cela, nous illustrons cette SP en suédois dans l'exemple 27 ci-dessous.

(27)

... <6.165> Att su<BACKSPACE1>tudera tar mycket tid, speciellt om man har t<BACKSPACE5> gör till exempel sex kurser samtidigt. ...  
‘A étudier prend beaucoup de temps, spécialement si on a <BACKSPACE5> fait par exemple six cours en même temps.’

A partir de l'exemple 27, nous pouvons remarquer qu'il n'existe pas de pause au milieu de cette expression. Nous avons également trouvé que dans le processus d'écriture, l'étudiant produit *till exempel* 'par exemple' toujours comme une SP.

Toutes les trois occurrences dans le type de MAEP représentent des SP D différentes. Nous montrons l'exemple 28 dans lequel deux de ces SP sont présentes.

(28)

... **jag to**<BACKSPACE1>**ror**<BACKSPACE8>**enligt min åsikt** d<BACKSPACE1>  
är det bra att...  
'**je trouve**<BACKSPACE8>**à mon avis** il est bon à'

L'exemple 28 illustre donc les SP D du type de MAEP *jag tror* 'je trouve' et *enligt min åsikt* 'à mon avis'. Comme l'étudiant B a premièrement produit une expression d'opinion et puis il l'a tout de suite remplacée par une autre, cela renforce le fait qu'il s'agit effectivement des SP du type de MAEP pour cet étudiant.

### Les SP grammaticales

Comme la quantité de SP grammaticales a été notablement plus élevée en suédois qu'en français, il nous paraît important d'examiner cette catégorie. L'étudiant B a donc produit 6 SP G dans le processus et elles sont restées les mêmes dans le produit final. Toutes ces SP représentent le type d'expressions de quantité et de degré (QD), et nous avons vu dans le tableau 36 que l'étudiant a employé trois fois la SP *så mycket* 'très beaucoup= tellement'. Les autres SP qui ont eu une occurrence sont *lite mer* 'un peu plus', *för mycket* 'trop beaucoup= trop' et *mycket mer* 'beaucoup plus'. Il est donc clair que cet étudiant a employé plusieurs SP qui comprennent le mot *mycket* 'beaucoup', et ces SP sont automatisées chez lui. Dans l'exemple 29, nous illustrons la SP G du type de QD *så mycket* 'tellement'.

(29)

... <3.821> '<BACKSPACE1> är inte ett bra idé att jobba **så mycket** att man måste återlämna <BACKSPACE2>a <BACKSPACE3>a pengar <2.285>. ...  
'n'est pas une bonne idée à travailler **tellement** que l'on doit rendre d'argent.'

L'étudiant B a donc utilisé cette SP trois fois, et chaque fois il l'a produite comme une unité et aussi chaque fois en liaison avec le verbe *jobba* 'travailler'. Cela peut indiquer que le sujet d'écriture en général a influencé l'usage de cette SP par l'étudiant B.

## Les SP individuelles

L'étudiant B n'a pas produit de SP individuelles dans la production en français, mais en suédois il l'a fait. L'étudiant a employé deux fois dans le processus et dans le produit final la SP L du type de SN *kundservice industrin* 'le service client'. Nous l'illustrons dans l'exemple 30 ci-dessous.

(30)

... Jag har jobbat som <2.172> <BACKSPACE4>vid kundservice industrin  
so<BACKSPACE1>å <2.338> ...

'J'ai travaillé comme <BACKSPACE4> au service client donc'

Comme la phrase dans l'exemple 30 le dit, il est clair qu'il s'agit d'une expression que l'étudiant a acquise dans un milieu du travail. En effet, l'étudiant B a répondu dans le questionnaire qu'il utilise le suédois de temps en temps à son travail, et cela indique que nous pouvons traiter cette SP comme une SP individuelle chez l'étudiant B.

Dans cette partie, nous avons examiné l'usage des SP en suédois par l'étudiant B, et notre objectif a été de décrire son usage des SP en général en comparant ses productions en français et en suédois. Dans le chapitre suivant, nous continuerons à traiter les données de l'étudiant C.

### 5.3.3. L'étudiant C

L'étudiant C étudie le français comme matière principale et le suédois comme matière secondaire, et selon lui, il a de meilleures compétences langagières en français qu'en suédois (*cf.* ch. 5.1.3.). Cet étudiant utilise les deux langues chaque semaine, mais dans la plupart de cas, il les emploie seulement en cours à l'université. Il n'a passé de séjours longs ni dans une région francophone ni dans une région suédophone et il n'a pas été en échange universitaire. Dans la partie qui suit, nous commencerons à traiter la production en français de l'étudiant C.

## En français

L'étudiant C a produit en français 271 mots dans le processus, dont la quantité de SP a été de 26 %, et 248 mots dans le produit final, ayant la quantité de SP de 24 % (*cf.* ch. 5.1.1.). Ces chiffres ont été notablement plus bas que les chiffres des étudiants A et B, ce qui nous a indiqué qu'il est probable que nous ne trouvons pas tellement de SP différentes dans cette production. Dans le tableau 39 ci-dessous, nous présentons donc les nombres de SP différentes que l'étudiant C a produites et aussi leurs occurrences.

**Tableau 39. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant C**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 22                       | 25                                   |
| Produit final | 21                       | 24                                   |

A partir de ce tableau, nous pouvons remarquer que les différences entre le processus et le produit final sont minimales dans les deux colonnes. Cela est en ligne avec le fait que l'étudiant a beaucoup planifié sa production, comme il l'a dit lui-même, et par conséquent, il a fait peu de changements. Il est aussi intéressant à noter que les nombres d'occurrences ne sont pas très grands par rapport aux nombres de SP différentes, ce qui veut dire que l'étudiant C a souvent employé une SP seulement une fois. Le nombre total d'occurrences est néanmoins suffisant pour catégoriser les SP que l'étudiant a utilisées, et dans ce qui suit, nous présentons la distribution des SP en catégories dans le tableau 40.

**Tableau 40. La distribution des SP en français en catégories ; l'étudiant C**

| SP   | L    | D    | G   | T   | I   | M    | SPIDI | Au total |
|--|------|------|-----|-----|-----|------|-------|----------|
| Processus  | 15   | 5    | 0   | 1   | 0   | 3    | 1     | 25       |
| Pourcentage  | 60 % | 20 % | 0 % | 4 % | 0 % | 12 % | 4 %   | 100 %    |
| Produit final  | 14   | 5    | 0   | 1   | 0   | 3    | 1     | 24       |
| Pourcentage  | 58 % | 21 % | 0 % | 4 % | 0 % | 13 % | 4 %   | 100 %    |
| SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires |      |      |     |     |     |      |       |          |

La distribution de l'étudiant C prouve le même phénomène que nous avons vu dans les chapitres précédents, c'est-à-dire que les catégories les plus utilisées sont les SP lexicales et les SP discursives. Pourtant, il est remarquable que le pourcentage des SP L soit beaucoup plus élevé que celui des SP D. Egalement, il est surprenant que la proportion de SP modifiées est grande dans la production de cet étudiant, car il s'agit de la proportion la plus élevée dans toute notre étude. A partir de ce tableau, nous pouvons aussi tirer la conclusion que l'étudiant C n'a produit aucune SP grammaticale ou individuelle. Le fait qu'il n'existe pas de SP G est assez surprenant parce que nous en trouvons dans toutes les autres productions dans cette étude. Comme nous l'avons

évoqué, l'étudiant C a employé la plupart de SP seulement une fois, mais nous trouvons quelques exceptions dans sa production. Les SP que cet étudiant a utilisées au moins deux fois sont donc présentées dans le tableau 41.

**Tableau 41. Les SP les plus utilisées en français par l'étudiant C**

| SP                | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type |
|-------------------|------------------------|----------------------------|-----------|------|
| Etudes en échange | 2                      | 2                          | L         | SN   |
| En échange        | 2                      | 2                          | L         | SAdv |
| Faire des études  | 2                      | 2                          | L         | SV   |

L= lexicales, SN= syntagmes nominaux, SAdv= syntagmes adverbiaux, SV= syntagmes verbaux

De ce tableau nous voyons qu'il n'existe que trois SP que l'étudiant C a utilisées deux fois. Toutes les SP représentent la catégorie des SP L, ce qui est vraisemblable d'après la distribution des SP en catégories. Nous continuerons en analysant plus en détail les catégories qui décrivent l'usage des SP de cet étudiant et nous traiterons alors les catégories des SP lexicales, discursives et modifiées.

### Les SP lexicales

Nous commençons avec la catégorie des SP lexicales que l'étudiant C a employée le plus dans sa production en français. Les SP L utilisées sont distribuées dans trois types différents et nous les montrons dans le tableau 42.

**Tableau 42. Les types de SP lexicales en français de l'étudiant C**

| SP            | Nombre d'occurrences |    |      |          |
|---------------|----------------------|----|------|----------|
|               | SN                   | SV | SAdv | Au total |
| Processus     | 6                    | 3  | 6    | 15       |
| Produit final | 5                    | 3  | 6    | 14       |

SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux

Comme ce tableau l'indique, les SP L du type de syntagmes nominaux (SN) et de syntagmes adverbiaux (SAdv) sont populaires. Nous avons vu dans le tableau 41 que les SP que l'étudiant C a utilisées deux fois représentent chacune un de ces trois types. Dans l'exemple 31, nous illustrons alors la SP L du type de SN *études en échange*.

(31)

<START> <29.000> Les études en échange sont <60.618> très utiles. <2.926> ...

La marque <START> dans l'exemple 31 indique que cette phrase est le début de la production de l'étudiant C. Comme nous l'avons évoqué (*cf.* ch. 5.3.1.), l'expression *études en échange* a été présente dans la question qui a été le sujet d'écriture, et il est donc possible que l'étudiant a produit cette expression comme une SP pour cette raison. Dans l'exemple 31, il est pourtant intéressant à noter que l'étudiant a produit cette unité clairement entre les pauses longues. Particulièrement, la pause de 60 secondes renforce le fait que l'étudiant a produit cette SP de manière aisée, car il a une pause de réflexion longue après la SP. Comme l'étudiant a employé cette SP aussi une autre fois dans sa production, il semble qu'elle puisse être une SP automatisée chez l'étudiant C.

Nous présentons une autre SP L du type de SN, à savoir *locuteurs natifs*, que l'étudiant C a utilisée dans le processus et dans le produit final. L'exemple 32 suivant l'illustre.

(32)

... <14.295> avec les contact<3.246>s avec les locuteurs natifs. <41.538> ...

Nous trouvons cette expression encore entre les pauses assez longues et nous voyons que l'étudiant l'a produite comme une unité, comme une SP. L'étudiant C a constaté qu'il utilise la langue française pour la plupart du temps en cours à l'université, et cette SP représente bien telle SP que l'étudiant a probablement acquis dans un milieu universitaire. Cet exemple décrit donc bien son usage des SP.

Des SP lexicales, nous traitons aussi le type de SAdv. Comme marqué dans le tableau 41, l'étudiant C a employé deux fois la SP du type de SAdv *en échange*, ce que nous illustrons dans l'exemple 33.

(33)

... <10.403> En échange <BACKSPACE1>, on doit <4.968> <BACKSPACE5>  
<8.620> <BACKSPACE3>l'apprentissage se <BACKSPACE3>déroule <3.631> ...

Cet exemple montre que l'étudiant C a aussi produit cette suite de mots comme une SP particulière, même si ces mots font aussi partie de la SP *études en échange* que nous avons traitée. De plus, l'étudiant a eu une pause de réflexion de 10 secondes avant qu'il a produit cette SP de manière aisée, comme une unité, surtout par rapport à toutes les pauses utilisées partout ailleurs dans la construction de la phrase.

## Les SP discursives

Le nombre des SP discursives que l'étudiant C a utilisées est resté le même, à savoir 5, dans le processus et dans le produit final, et la distribution des SP en types montre que l'étudiant s'est servi de trois types différents. Le tableau 43 suivant les présente.

**Tableau 43. Les types de SP discursives en français de l'étudiant C**

| SP            | Nombre d'occurrences |      |    |          |
|---------------|----------------------|------|----|----------|
|               | MT                   | MAEP | MI | Au total |
| Discursives   |                      |      |    |          |
| Processus     | 1                    | 3    | 1  | 5        |
| Produit final | 1                    | 3    | 1  | 5        |

MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement, MI= marqueurs interactifs

Le tableau 43 indique que le type de MAEP est utilisé le plus par cet étudiant. Les trois SP du type de MAEP utilisées ont été *c'est difficile, il peut être plus facile de* et *il peut être difficile de* ; il est intéressant que toutes ces SP marquent que la facilité ou la difficulté. L'exemple 34 ci-dessous illustre la dernière.

(34)

... <20.405> Pour quelque<2.483>s<BACKSPACE2> <5.607>'uns <3.999>  
<BACKSPACE1>, **il peut être difficile de** <5.183> retourner <3.921> ...

A partir de cet exemple, il est clair que cette SP est complètement entourée par les pauses. Cela indique qu'il s'agit d'une unité qui représente une SP automatisée pour l'étudiant C.

Des SP discursives, nous traitons aussi le type de marqueurs interactifs (MI), car ce type est assez rare dans toutes les productions. L'étudiant C a utilisé telle SP *bien sûr* et nous le présentons dans l'exemple 35 suivant.

(35)

... <2.926> Le<BACKSPACE2> <4.976>a chose la plus important est <6.757> **bien sûr** l'apprentid<BACKSPACE1>ssage d'une langue. <10.403> ...

Egalement dans ce cas, l'étudiant a produit la SP *bien sûr* après une pause assez longue. Il semble qu'une pause de réflexion ou de planification rende possible le fait que l'étudiant produit des suites de mots aisées.

## Les SP modifiées

Comme nous l'avons évoqué, la catégorie des SP modifiées est relativement grande chez l'étudiant C. Nous trouvons la réponse pour cela dans le tableau 44 qui présente la distribution des SP M en types.

**Tableau 44. Les types de SP modifiées en français de l'étudiant C**

| SP   | Nombre d'occurrences |              |          |
|--|----------------------|--------------|----------|
|  | SPIDI L (SAdv)       | SPIDI L (SV) | Au total |
| Processus  | 2                    | 1            | 3        |
| Produit final  | 2                    | 1            | 3        |
| SPIDI= SP idiosyncrasiques intérimaires, L= lexicales, SAdv= syntagmes adverbiaux, SV= syntagmes verbaux |                      |              |          |

Le tableau 44 nous montre que les trois SP M ont originellement représenté les SPIDI. Cela a donc influé sur le fait que dans la vraie catégorie des SPIDI, nous ne trouvons qu'une occurrence. Dans l'exemple 36, nous illustrons l'une des SP M qui sont des SPIDI du type de SAdv.

(36)

... <2.949> <BACKSPACE1> Si on a fait des études en étrangères<BACKSPACE1>  
 ...  
 ... <11.152> <MOUSECLICK> <BACKSPACE8>change<MOUSECLICK> <2.767>  
 ...

La phrase dans le produit final est donc « Si on a fait des études en échange ». L'exemple 36 montre que l'étudiant C est revenu à la SPIDI *en étrangère\** et il l'a modifiée de cette manière que dans le produit final il existe la SP M *en échange*. Notons ici que la SPIDI que nous trouvons dans la catégorie des SPIDI (le tableau 40) a aussi été *en étrangère\**. Comme cet étudiant a produit deux fois correctement la SP *en échange*, il semble que dans ce cas, il a essayé de produire l'expression à *l'étranger*. Cette SP M (la SPIDI) prouve de nouveau que l'étude du processus d'écriture est importante, car sans l'étude du processus nous aurions compté *en échange* dans cette phrase comme une SP que l'étudiant aurait produite.

Dans cette partie, nous avons examiné les catégories des SP lexicales, discursives et modifiées qui ont décrit l'usage des SP en français par l'étudiant C. Nous continuerons dans la partie suivante avec la production en suédois de cet étudiant.

### **En suédois**

Nous avons constaté dans le chapitre 5.1.2. que l'étudiant C a produit 214 mots dans le processus et 206 mots dans le produit final. Le nombre de mots faisant partie des SP a été de 47 dans les deux cas, mais le pourcentage des SP a été de 22 % dans le processus et de 23 % dans le produit final. Comme ces nombres ont été plus bas que ceux dans la production en français de cet étudiant, il a été clair que le nombre de SP différentes et leurs occurrences seront plus petits. Dans le tableau 45 suivant, nous montrons ces chiffres.

**Tableau 45. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant C**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 16                       | 18                                   |
| Produit final | 16                       | 18                                   |

Comme les nombres de mots faisant partie des SP sont restés les mêmes dans le processus et dans le produit final, les nombres de SP différentes et les occurrences sont également les mêmes dans les deux cas, comme indiqué dans le tableau 45. L'étudiant C a produit du texte de la même manière en français et en suédois, car il a utilisé beaucoup de temps pour planifier son texte et, en conséquence, il n'existe pas de changements remarquables entre le processus et le produit final. Pour la plupart, l'étudiant C a employé une SP seulement une fois, comme il l'a aussi fait en français. Malgré le petit nombre d'occurrences des SP, nous illustrons leur distribution en types dans le tableau 46 ci-dessous.

**Tableau 46. La distribution des SP en suédois en catégories ; l'étudiant C**

| SP   | L    | D    | G    | T   | I    | M   | SPIDI | Au total |
|--|------|------|------|-----|------|-----|-------|----------|
| Processus  | 9    | 2    | 3    | 0   | 2    | 0   | 2     | 18       |
| Pourcentage  | 50 % | 11 % | 17 % | 0 % | 11 % | 0 % | 11 %  | 100 %    |
| Produit final  | 9    | 2    | 3    | 0   | 2    | 0   | 2     | 18       |
| Pourcentage  | 50 % | 11 % | 17 % | 0 % | 11 % | 0 % | 11 %  | 100 %    |
| SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires |      |      |      |     |      |     |       |          |

Cette distribution révèle que la catégorie des SP lexicales est notablement la plus utilisée parmi toutes les catégories. Un changement intéressant par rapport à la distribution des SP en français (le tableau 40) est le fait que la catégorie des SP grammaticales a la quantité deuxièmement la plus grande. En revanche, l'étudiant C a produit autant de SP discursives, SP individuelles et SPIDI dans la production en suédois. La catégorie des SP modifiées a eu un nombre assez élevé en français, mais en suédois l'étudiant n'en a pas produit. En tout, la distribution des SP en suédois est considérablement différente par rapport à celle en français, ce qui fait l'étude de l'usage des SP par l'étudiant C intéressante. Comme nous l'avons évoqué, l'étudiant a souvent utilisé une SP une fois. Dans le tableau 47, nous présentons les SP qui ont eu au moins deux occurrences.

**Tableau 47. Les SP les plus utilisées en suédois par l'étudiant C**

| SP   | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type     |
|--|------------------------|----------------------------|-----------|----------|
| Vid sidan av studier/na, 'à côté d'/des études'  | 2                      | 2                          | L         | SAdv     |
| Man måste, 'on doit'   | 2                      | 2                          | I         | D (MAEP) |
| L= lexicales, I= individuelles, D= discursives, SAdv= syntagmes adverbiaux, MAEP= marqueurs autorégulateurs et expressions de positionnement |                        |                            |           |          |

L'une de ces SP dans le tableau 47 représente évidemment la catégorie des SP L, mais il est surprenant que l'autre SP fait partie de la catégorie des SP I. Cette liste est donc différente de celle en français. Ensuite, nous continuerons à examiner plus précisément les catégories qui décrivent la manière de l'étudiant C d'utiliser les SP en suédois. Nous traiterons les catégories les plus grandes, c'est-à-dire les SP lexicales et les SP grammaticales, et aussi les catégories des SP discursives et des SP individuelles.

## Les SP lexicales

L'étudiant C a employé en suédois le plus de SP qui font partie de la catégorie des SP lexicales. La distribution des SP en types comprend les mêmes trois types que l'étudiant a utilisés aussi en français, mais les occurrences sont distribuées différemment parmi ces types. Nous les présentons dans le tableau 48 suivant.

**Tableau 48. Les types de SP lexicales en suédois de l'étudiant C**

| SP  | Nombre d'occurrences |    |      |          |
|---|----------------------|----|------|----------|
|   | SN                   | SV | SAdv | Au total |
| Processus   | 2                    | 4  | 3    | 9        |
| Produit final   | 2                    | 4  | 3    | 9        |
| SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux |                      |    |      |          |

Dans cette distribution des SP L en suédois, le type de syntagmes verbaux (SV) est le plus employé. En français, l'étudiant a utilisé plusieurs SP du type de syntagmes nominaux (SN), mais en suédois leur proportion est la plus petite. Dans l'exemple 37 ci-dessous, nous illustrons une phrase qui comprend deux SP L du type de SV et aussi une SP L du type de SN que l'étudiant a produites une fois.

(37)

... <8.991> kna<BACKSPACE2>an man <5.419> få jobb som har något att göra med <8.432> si<14.437>n branch eller <6.176> blivande yrke. <8.344> ...  
 'on peut avoir un travail qui a quelque chose à faire avec son domaine ou futur métier.'

Les deux SP du type de SV *få jobb* 'avoir un travail' et *ha något att göra med* 'avoir quelque chose à faire avec' sont alors produites comme les unités, sans pauses au milieu de ces expressions. Ces SP sont automatisées chez l'étudiant C, et nous le voyons notamment dans le cas de la SP *ha något att göra med* parce qu'elle est construite de plusieurs mots. Quant à la SP L du type de SN *blivande yrke* 'futur métier', elle est intéressante parce que nous trouvons des pauses assez longues autour de cette expression, ce qui indique que c'est une SP pour l'étudiant C, et aussi parce qu'il s'agit d'une expression qui est construite de manière particulière en suédois. Le mot *blivande* 'futur' est un adjectif qui ne s'accorde pas et il est en effet formé comme un participe présent en suédois. Nous avons évoqué dans le chapitre 3.2.1. que les SP sont en général

produites sans une analyse profonde sur la composition des mots (*cf.* Wray, 2002). Pour cette raison, la SP *blivande yrke* ‘futur métier’ est un bon exemple d’une expression qui représente une SP pour l’étudiant C probablement à cause de son « difficulté ».

Nous avons vu dans le tableau 47 que l’étudiant C a employé deux fois la SP L du type de SAdv *vid sidan av studier/na* ‘à côté d’/des études’. Nous avons noté dans les chapitres précédents que les autres étudiants ont aussi utilisé cette SP beaucoup, car elle a été présente dans le sujet d’écriture. Dans le cas de l’étudiant C, il est intéressant qu’il ait également employé en français une SP qui a fait partie du sujet d’écriture. Néanmoins, nous présentons dans l’exemple 38 l’autre SP du type de SAdv que l’étudiant C a produite.

(38)

... studier <3.502> <BACKSPACE1>na <26.671> fortsätter ibland längre <3.500>, **i synnerhet** om man <43.399> ...  
 ‘les études continuent parfois plus longtemps, **en particulier** si on’

Egalement dans cet exemple, l’étudiant C a produit l’expression *i synnerhet* ‘en particulier’ après une pause, comme une unité sans pauses. Il est clair qu’il s’agit de la SP automatisée chez cet étudiant.

### Les SP grammaticales

L’étudiant C n’a produit au total que trois occurrences des SP grammaticales, mais cette catégorie est toutefois deuxièmement la plus grande. La distribution de ces SP en types est présentée dans le tableau 49 suivant.

**Tableau 49. Les types de SP grammaticales en suédois de l’étudiant C**

| SP   | Nombre d'occurrences |    |          |
|--|----------------------|----|----------|
|  | QD                   | PI | Au total |
| Processus  | 2                    | 1  | 3        |
| Produit final  | 2                    | 1  | 3        |
| QD= expressions de quantité et de degré, PI= pronoms indéfinis |                      |    |          |

Comme l’étudiant C n’a produit aucune SP grammaticale en français, il est marquant à noter à partir du tableau 49 que cet étudiant a utilisé des SP G de deux types en suédois. Les SP G *hur mycket* ‘comment beaucoup= combien’ et *så mycket* ‘très beaucoup= tellement’ représentent le type d’expressions de quantité et de degré (QD), tandis que la

SP G *någon annan* ‘quelqu’un autre’ représente le type de pronoms indéfinis (PI). Dans l’exemple 39, nous illustrons la SP *så mycket*.

(39)

... <8.045> för att man inte <60.048> heller jobbar **så mycket** att man <8.069> skulle inte få studiestöd. <84.377> ...

‘pour que l’on ne travaille pas non plus **tellement** que l’on n’aurait pas d’allocation d’aide aux étudiants.’

L’étudiant a produit cette expression comme une unité, sans pauses. Elle représente une SP automatisée pour cet étudiant, et il est intéressant à remarquer que l’étudiant C utilise une SP contenant le mot *mycket* ‘beaucoup’ en liaison avec le verbe *jobba* ‘travailler’, comme l’a aussi fait l’étudiant B (ch. 5.3.2.). Dans l’exemple 40 suivant, nous présentons comment l’étudiant C a employé la SP G du type de PI.

(40)

... <8.901> arbetstider <BACKSPACE1>, som **någon annan** har <28.576> satt. <54.167> ...

‘horaires de travail, que **quelqu’un autre** a mis.’

Puisque l’étudiant a employé cette expression comme un sujet dans la phrase, il semble clair qu’il s’agit d’une SP chez l’étudiant C. Cela veut dire que *någon* ‘quelque’ et *annan* ‘autre’ peuvent aussi fonctionner comme des déterminants séparément, mais dans ce cas, l’étudiant a produit une unité, une SP.

### Les SP discursives

Dans la production en français de l’étudiant C, nous avons trouvé cinq occurrences des SP D et elles ont été distribuées en trois types différents. En revanche, en suédois l’étudiant a produit deux SP D et les deux représentent le type de marqueurs textuels (MT). Ces SP sont *det är också* ‘il est aussi’ et *det är* ‘il est’ ; d’une certaine manière il s’agit de la même SP, mais nous avons compté les SP comme *il est aussi* et *c’est aussi* séparées des SP *il est* et *c’est* dans tout le corpus. Nous illustrons la SP *det är också* dans l’exemple 41 ci-dessous.

(41)

... <64.295> <ENTER1> **Det är också** <7.644> bra att få mer arbetserfa<2.495>rhet <6.875> ...

‘**Il est aussi** bon à avoir plus d’expérience de travail’

Cet exemple nous montre que l'étudiant C a commencé la phrase par cette expression et il l'a clairement produite comme une unité ; les pauses longues entourent la suite de mots directement. Ces raisons prouvent qu'il s'agit d'une SP automatisée chez l'étudiant C.

### Les SP individuelles

L'étudiant C n'a produit aucune SP individuelle en français, mais en suédois il a utilisé une SP I même deux fois. La SP I que nous trouvons est *man måste* 'on doit' et elle représente la catégorie des SP discursives et le type de MAEP. Dans les autres productions, nous n'avons pas considéré des constructions *man* 'on' + *verbe* comme des SP, mais dans cette production de l'étudiant C, il semble que l'étudiant utilise l'expression *man måste* de manière particulière. Nous l'illustrons dans l'exemple 42.

(42)

... <22.712> **Man måste** <3.412> vara ganska noggrann <8.045> ...  
'**On doit** être assez soigneux'

L'exemple 42 indique que l'étudiant C a produit cette suite de mots comme une unité après une pause de 22 secondes. Egalement, après cette unité, il a eu une pause de réflexion de 3 secondes avant qu'il ait produit le reste de la phrase. Ces faits montrent que l'étudiant a écrit la SP de manière aisée et après cela, il a eu besoin de planifier ce qu'il écrira.

Dans cette partie, nous avons examiné l'usage des SP dans la production en suédois par l'étudiant C. Nous avons traité les catégories qui décrivent cet usage le mieux, c'est-à-dire les catégories des SP lexicales, grammaticales, discursives et individuelles. Dans le chapitre suivant, nous analyserons les productions de l'étudiant D.

#### 5.3.4. L'étudiant D

L'étudiant D étudie l'anglais comme matière principale et le français et le suédois comme matières secondaires. Selon cet étudiant, ses compétences langagières sont meilleures en suédois qu'en français et, par exemple, il utilise du suédois plus en temps libre (*cf.* ch. 5.1.3.). L'étudiant D n'a pas passé de séjours longs dans une région francophone ou suédophone, et il n'a pas été en échange universitaire. Dans ce qui suit, nous continuerons avec la partie des SP en français.

## En français

L'étudiant D a produit 241 mots en français dans le processus d'écriture, la quantité de SP étant de 31 %, et 194 mots dans le produit final, ayant la quantité de SP de 28 % (cf. ch. 5.1.1.). Cet étudiant a donc produit notablement moins de mots que les étudiants A et B, mais pourtant les pourcentages des SP dans sa production ont été aussi élevés que ceux dans les productions d'A et de B. Cela nous a indiquée que nous pouvons bien trouver des SP différentes dans la production de l'étudiant D. Dans le tableau 50 suivant, nous présentons les nombres de SP différentes et leurs occurrences dans la production en français.

**Tableau 50. Le nombre de SP différentes en français et leurs occurrences ; l'étudiant D**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 22                       | 29                                   |
| Produit final | 19                       | 23                                   |

A partir de ce tableau, nous pouvons tirer que les nombres de SP que l'étudiant D a produites ne sont pourtant pas très grands, ce qui peut indiquer que l'étudiant a employé des SP qui sont construites de plusieurs mots (comme la quantité de SP a été calculée par le nombre de mots faisant partie des SP). Nous voyons également que l'étudiant D a fait de changements notables dans la production, car les chiffres sont diminués du processus au produit final. En ce qui concerne la distribution des SP en catégories, l'étudiant D a utilisé des SP de toutes les catégories. La distribution est présentée dans le tableau 51 ci-dessous, et notons que les pourcentages sont arrondis pour avoir le pourcentage au total.

**Tableau 51. La distribution des SP en français en catégories ; l'étudiant D**

| SP            | L    | D    | G   | T   | I    | M   | SPIDI | Au total |
|---------------|------|------|-----|-----|------|-----|-------|----------|
| Processus     | 8    | 11   | 1   | 2   | 3    | 1   | 3     | 29       |
| Pourcentage   | 28 % | 38 % | 4 % | 7 % | 10 % | 3 % | 10 %  | 100 %    |
| Produit final | 8    | 7    | 1   | 2   | 2    | 1   | 2     | 23       |
| Pourcentage   | 35 % | 30 % | 4 % | 9 % | 9 %  | 4 % | 9 %   | 100 %    |

SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires

Comme ce tableau l'indique, l'étudiant D a employé des SP de manière très variée. En effet, la production de cet étudiant est la seule dans laquelle nous trouvons des SP de toutes les catégories dans toute l'étude. Dans cette production, les catégories les plus grandes sont les SP lexicales et les SP discursives, mais elles ne restent pas dans le même ordre dans le processus et dans le produit final. Quant aux autres catégories, il est intéressant que les proportions dans les catégories des SP individuelles et des SPIDI soient assez élevées. Notamment les SPIDI peuvent indiquer le fait que l'étudiant D a de meilleures compétences en suédois qu'en français, comme il l'a constaté lui-même. Dans la liste des SP que cet étudiant a utilisées au moins deux fois, nous voyons également la distribution diversifiée. Ces SP sont montrées dans le tableau 52.

**Tableau 52. Les SP les plus utilisées en français par l'étudiant D**

| SP  | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type     |
|---|------------------------|----------------------------|-----------|----------|
| Ce pays   | 3                      | 2                          | I         | L (SN)   |
| En échange  | 2                      | 2                          | L         | SAdv     |
| Il y a  | 2                      | 1                          | D         | MT       |
| Par exemple   | 2                      | 1                          | D         | MT       |
| Etudes en échange   | 2                      | 2                          | L         | SN       |
| En Finlande   | 2                      | 2                          | T         | L (SAdv) |
| Catégories : I= individuelles, L= lexicales, D=discursives, T=thématiques<br>Types : SN= syntagmes nominaux, SAdv= syntagmes adverbiaux, MT= marqueurs textuels |                        |                            |           |          |

Dans ce tableau 52, il est surprenant qu'une SP individuelle est la plus utilisée parmi toutes les SP de cet étudiant. Egalement, une SP thématique a eu deux occurrences, ce qui veut dire que toute la catégorie des SP T ne comprend que cette SP (voir le tableau 51). Dans la partie qui suit, nous commencerons à examiner plus précisément les catégories qui décrivent le mieux l'usage des SP en français par l'étudiant D. Nous traiterons les catégories les plus grandes, à savoir les SP discursives et les SP lexicales, et aussi les catégories des SP individuelles, des SPIDI et des SP modifiées.

### **Les SP discursives**

Nous examinons premièrement la catégorie des SP discursives, car l'étudiant D en a produit le plus dans le processus en français. Le nombre total de SP D a diminué dans le produit final, et dans le tableau 53, nous présentons comment les SP D sont distribuées dans des types différents.

**Tableau 53. Les types de SP discursives en français de l'étudiant D**

| SP   | Nombre d'occurrences |      |          |
|--|----------------------|------|----------|
|  | MT                   | MAEP | Au total |
| Processus  | 9                    | 2    | 11       |
| Produit final  | 5                    | 2    | 7        |
| MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement |                      |      |          |

Comme marqué dans le tableau 53, l'étudiant D a produit plus des SP D du type de marqueurs textuels (MT) que des SP D du type de marqueurs autorégulateurs et expressions de positionnement (MAEP). Le type de MT est aussi celui dans lequel le nombre d'occurrences est diminué. Dans l'exemple 43 ci-dessous, nous illustrons la SP du type de MT *il y a* que l'étudiant a employée deux fois dans le processus, mais une fois dans le produit final.

(43)

... <11.872> **Il y a** plusieurs des <2.267> avantages <26.904> <LEFT25>  
<BACKSPACE6> Les études <4.737> d'échange <3.215> <LEFT1> <2.455>  
<RIGHT3> <LEFT2> <RIGHT1> <3.201> ont <RIGHT24> <BACKSPACE1>.  
<17.117> ...

L'exemple 43 montre la SP *il y a* que l'étudiant D a effacée dans le produit final. Nous voyons que l'étudiant a produit cette suite de mots après une pause longue, et puis il n'existe pas de pause au milieu de cette expression. La SP *il y a* représente donc la SP automatisée pour cet étudiant, notamment parce qu'il l'a produite deux fois de manière aisée. Dans cet exemple, nous voyons également comment l'étudiant a remplacé la SP ; la phrase dans le produit final est alors « Les études d'échange ont plusieurs des avantages ». L'étudiant a commenté dans la verbalisation pendant la pause de 26 secondes (après *avantages*) qu'il a pensé s'il pouvait utiliser une autre expression dans cette phrase. Ce fait prouve également que l'étudiant a produit l'expression *il y a* automatiquement, et quand il a voulu la remplacer, cela a pris plus de temps (de pauses). Notons ici que l'étudiant D a beaucoup utilisé de touches flèches pour se déplacer dans le texte et c'est marqué par <LEFT> ou <RIGHT>.

Des deux SP du type de MAEP, nous présentons la SP *je pense que* que l'étudiant a donc utilisée une fois. L'exemple 44 l'illustre ci-dessous.

(44)

... <2.251> <BACKSPACE1>, je pense qu'on <2.255> doit le faire. <12.048> ...

Cette SP est produite entre les pauses comme une unité, et à partir de cet exemple, il semble que l'étudiant a produit la SP de manière automatisée.

### Les SP lexicales

Le nombre d'occurrences dans la catégorie des SP lexicales est resté le même dans le processus et dans le produit final et, par conséquent, cette catégorie est la plus grande dans le produit final. Dans le tableau 54, la distribution des SP L en types est présentée.

**Tableau 54. Les types de SP lexicales en français de l'étudiant D**

| SP            | Nombre d'occurrences |    |      |          |
|---------------|----------------------|----|------|----------|
|               | SN                   | SV | SAdv | Au total |
| Lexicales     |                      |    |      |          |
| Processus     | 2                    | 3  | 3    | 8        |
| Produit final | 2                    | 3  | 3    | 8        |

SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux

Cette distribution est en ligne avec le fait que l'étudiant D a utilisé des SP de manière variée ; même si le nombre total d'occurrences est de 8, l'étudiant a employé des SP L de trois types différents. Nous traitons dans l'exemple 45 suivant la SP du type de SAdv *en échange* que nous avons aussi vu dans la liste des SP les plus utilisées (le tableau 52).

(45)

... <14.961> fair e<BACKSPACE2>e les études en échange <3.679> ...

Dans cet exemple, ces suites de mots sont clairement produites comme les unités, sans pauses au milieu. Nous trouvons dans l'exemple 45 aussi la SP L du type de SV *faire les études*, mais nous savons que la SP *en échange* est une SP séparée parce que l'étudiant D l'a utilisée aussi telle quelle. Egalement, nous trouvons dans la production de cet étudiant la SP L du type de SN *études en échange* dont l'étudiant a utilisé de manière différente comparé à ces deux autres SP.

L'étudiant D a alors employé trois SP L du type de SV différentes. De ces SP, nous présentons dans l'exemple 46 la SP *se familiariser avec*.

(46)

... <RIGHT8> <17.238> se familiariser avec <2.203> <DOWN3> <ENTER1> ...

Dans l'exemple 46, les touches flèches montrent que l'étudiant est revenu à une phrase qu'il a déjà écrite et il y a ajouté cette suite de mots. Néanmoins, l'exemple indique que l'étudiant a produit cette expression comme une SP entre les pauses.

### Les SP individuelles

Comme marqué dans le tableau 52, la SP que l'étudiant D a employée le plus a été la SP individuelle *ce pays*. Cette SP représente la catégorie lexicale et le type de SN, mais il faut noter que nous n'avons pas considéré des suites de mots comme *ce + nom* en général comme SP. Dans le cadre de la production de cet étudiant, il est pourtant clair qu'il utilise cette expression de manière particulière. Nous l'illustrons par les exemples 47 et 48.

(47)

... <3.243> <BACKSPACE1> et aussi le culu<BACKSPACE1>ture <3.002> de <2.145> <18.289> ce pays. <4.418> ...

(48)

... seulement le culture de pays <BACKSPACE5>ce pays qu'on peut <6.780> recevoir.  
...

Dans l'exemple 47, nous remarquons que l'étudiant a produit cette SP après une pause longue de réflexion et la SP est entourée par les pauses. L'exemple 48 montre en revanche que le mot *culture* est clairement lié à la SP, mais cet exemple indique également que l'étudiant a produit la SP *ce pays* comme une unité séparée.

### Les SPIDI

La proportion de la catégorie des SPIDI a été assez élevée dans la production en français de l'étudiant D, même si le nombre d'occurrences n'a été que 3 dans le processus et 2 dans le produit final. Les SPIDI que l'étudiant a utilisées ont été *plusieurs des\** et *beaucoup des\**, qui représentent des SP G du type de QD, et *faire la connaissance avec\**, qui représente une SP L du type de SV. La dernière est la SP que l'étudiant a effacée dans le produit final. Il est intéressant à noter que deux de trois SPIDI ont exprimé de la quantité, ce qui pourrait indiquer qu'il s'agit d'un point grammatical que l'étudiant n'a pas acquis complètement. Nous avons déjà présenté dans

l'exemple 43 comment l'étudiant a utilisé la SPIDI *plusieurs des\**, et dans l'exemple 49 suivant, nous illustrons la SPIDI *beaucoup des\**.

(49)

... <9.762> peut donner **beaucoup des** <2.245> <BACKSPACE2> <3.425>s <21.963>  
expériences et <3.611> memoires. <LEFT7> <BACKSPACE1>é<4.081>  
<BACKSPACE1>e<BACKSPACE1>é<4.030> ...

La phrase qui équivaut à cet exemple dans le produit final est donc « ... peut donner beaucoup des expériences et mémoires ». Dans l'exemple 49, il est marqué que l'étudiant a effacé la lettre *s* à la fin de *des*, mais puis il l'a ajoutée de nouveau. Cette hésitation montre qu'il s'agit effectivement d'une SPIDI chez l'étudiant D.

### Les SP modifiées

Nous traitons dernièrement la catégorie des SP modifiées, car la SP M que l'étudiant a produite représente aussi une SPIDI, et plus précisément une SPIDI L du type de SAdv. Nous illustrons cette SP dans l'exemple 50 ci-dessous.

(50)

... <6.359> à <BACKSPACE2> À <2.535> vivre <4.080> da<BACKSPACE2>  
<4.124> **dans un pays étrangere**<BACKSPACE1> <4.860> <LEFT18> <2.553>  
<BACKSPACE4> **en** <4.961> <RIGHT18> <9.762> ...

A partir de cet exemple, nous pouvons tirer que l'étudiant D a produit la suite de mots *dans un pays étranger* (à savoir, sans la lettre *e* à la fin parce qu'il l'a effacée tout de suite) comme une SP, car elle est entourée par les pauses et il n'existe pas de pauses à l'intérieur de cette expression. Cette SP telle quelle représente la SP L du type de SAdv, mais comme l'exemple 50 le révèle, l'étudiant D a hésité avec la préposition et finalement il l'a remplacée par *en*. Pour ces raisons, nous classifions cette SP M comme SPIDI chez l'étudiant D. Avec cet exemple, nous avons encore illustré l'importance de l'étude du processus, car dans le produit final, nous trouvons seulement la phrase « À vivre en un pays étranger ».

Cette partie de ce chapitre a traité l'usage des SP en français par l'étudiant D. Comme l'étudiant s'est servi de toutes les catégories possibles, nous avons examiné plusieurs catégories qui décrivent son usage. Nous continuerons dans la partie suivante en traitant la production en suédois de cet étudiant.

## En suédois

Nous avons montré dans le chapitre 5.1.2. que l'étudiant D a produit 270 mots en suédois dans le processus, la quantité de SP étant de 26 %, et 226 mots dans le produit final, avec la quantité de SP de 29 %. Les pourcentages des SP ont donc été à peu près de même niveau dans les productions en français et en suédois de cet étudiant. Par conséquent, nous avons pu nous attendre à ce que les nombres de SP différentes et les nombres d'occurrences soient assez similaires dans les deux langues. Dans le tableau 55 suivant, ces chiffres de la production en suédois sont présentés.

**Tableau 55. Le nombre de SP différentes en suédois et leurs occurrences ; l'étudiant D**

|               | Nombre de SP différentes | Nombre d'occurrences des SP au total |
|---------------|--------------------------|--------------------------------------|
| Processus     | 20                       | 28                                   |
| Produit final | 20                       | 26                                   |

Comme l'étudiant D a évalué ses compétences langagières en suédois meilleures que celles en français, il est intéressant à noter du tableau 55 que les nombres de SP produites en suédois dans le processus sont pourtant plus bas que les nombres en français (le tableau 50). Néanmoins, nous pouvons tirer à partir du tableau 55 que l'étudiant D n'a pas fait beaucoup de changements dans la production en suédois, tandis qu'en français il les a faits, ce qui pourrait indiquer que sa manière d'écrire est plus cohérente en suédois. Toutefois, cet étudiant a employé des SP en suédois de manière variée, et nous trouvons des SP dans toutes les catégories possibles, comme en français. La distribution des SP en catégories est donc illustrée dans le tableau 56 ci-dessous.

**Tableau 56. La distribution des SP en suédois en catégories ; l'étudiant D**

| SP            | L    | D    | G    | T   | I   | M   | SPIDI | Au total |
|---------------|------|------|------|-----|-----|-----|-------|----------|
| Processus     | 12   | 6    | 3    | 2   | 2   | 1   | 2     | 28       |
| Pourcentage   | 43 % | 21 % | 11 % | 7 % | 7 % | 4 % | 7 %   | 100 %    |
| Produit final | 12   | 6    | 3    | 1   | 1   | 1   | 2     | 26       |
| Pourcentage   | 46 % | 23 % | 11 % | 4 % | 4 % | 4 % | 8 %   | 100 %    |

SP= séquences préfabriquées, L= lexicales, D= discursives, G= grammaticales, T= thématiques, I= individuelles, M= modifiées, SPIDI= SP idiosyncrasiques intérimaires

Dans cette distribution des SP en suédois de l'étudiant D, la catégorie des SP lexicales est considérablement plus grande que les autres catégories, et même la différence entre les SP lexicales et les SP discursives est marquante. Dans sa production en français, ces deux catégories ont été plus ou moins égales. En ce qui concerne les autres catégories, l'étudiant a produit plus de SP grammaticales en suédois qu'en français, et la catégorie des SPIDI est par contre plus petite en suédois. Il est en effet intéressant que cet étudiant a utilisé toutes sortes des SP dans les deux langues. La distribution diversifiée est aussi visible dans la liste des SP que l'étudiant a employées au moins deux fois, et nous la présentons dans le tableau 57.

**Tableau 57. Les SP les plus utilisées en suédois par l'étudiant D**

| SP   | Occurrence (processus) | Occurrence (produit final) | Catégorie | Type     |
|--|------------------------|----------------------------|-----------|----------|
| Vid sidan av studier/na, 'à côté d'/des études'  | 4                      | 4                          | L         | SAdv     |
| Väldigt bra, 'très bon'  | 2                      | 2                          | L         | SAdj     |
| I Finland, 'en Finlande'   | 2                      | 1                          | T         | L (SAdv) |
| Hur mycket, 'combien'  | 2                      | 2                          | G         | QD       |
| På sommarna, 'en été'  | 2                      | 2                          | SPIDI     | L (SAdv) |
| Extra jobbet, 'extra travail'  | 2                      | 1                          | I         | L (SN)   |
| Catégories : L= lexicales, T= thématiques, G= grammaticales, SPIDI= SP idiosyncrasiques intérimaires, I= individuelles<br>Types : SAdv= syntagmes adverbiaux, SAdj= syntagmes adjectivaux, SN= syntagmes nominaux, QD= expressions de quantité et de degré |                        |                            |           |          |

Comme marqué dans le tableau 57, une SP de la catégorie L est présentée au début de la liste, ce qui est probable d'après la distribution. Il est aussi surprenant que la même SP *T i Finland* 'en Finlande' se trouve dans la liste du suédois et du français. Nous continuons l'analyse de l'usage des SP en examinant les catégories que l'étudiant D a utilisées le plus, c'est-à-dire les SP lexicales et les SP discursives, et aussi les catégories des SP grammaticales et des SPIDI.

### **Les SP lexicales**

Nous commençons avec la catégorie des SP lexicales que l'étudiant D a utilisée le plus. Dans cette catégorie, l'étudiant a employé même quatre types différents des SP, et nous présentons la distribution en types dans le tableau 58 ci-dessous.

**Tableau 58. Les types de SP lexicales en suédois de l'étudiant D**

| SP   | Nombre d'occurrences |    |      |      |          |
|--|----------------------|----|------|------|----------|
|  | SN                   | SV | SAdv | SAdj | Au total |
| Lexicales  |                      |    |      |      |          |
| Processus  | 1                    | 4  | 5    | 2    | 12       |
| Produit final  | 1                    | 4  | 5    | 2    | 12       |
| SN= syntagmes nominaux, SV= syntagmes verbaux, SAdv= syntagmes adverbiaux, SAdj= syntagmes adjectivaux |                      |    |      |      |          |

Ce tableau indique que les types le plus populaires en suédois ont été les syntagmes adverbiaux (SAdv) et les syntagmes verbaux (SV), comme ils l'ont été aussi en français. Dans la production en suédois, l'étudiant a produit des SP L du type de syntagmes adjectivaux (SAdj), le type qui a été assez rare dans toute l'étude. Nous avons vu dans le tableau 57 que cet étudiant a employé le plus la SP L du type de SAdv *vid sidan av studier/na* 'à côté d'/des études', et nous avons expliqué dans les chapitres précédents qu'il s'agit d'une suite de mots qui a été présente dans le sujet d'écriture. Cependant, nous illustrons dans les exemples 51 et 52 comment l'étudiant D s'est servi de cette SP.

(51)

<START> <33.621> Det finns många <6.414> avsikter om att jobba **vid sidan av studierna.** <11.147> ...

'Il y a plusieurs idées [du fait] de travailler **à côté des études.**'

(52)

... <29.456> men jag <8.028> skulle ku<BACKSPACE2> <2.378> inte kunna gå<BACKSPACE2>öra <BACKSPACE1> <BACKSPACE5>jobba <18.306> hela tiden v<3.150> <BACKSPACE1> **vid sidann**<BACKSPACE1> **av studierna.** <BACKSPACE1> <6.966> <3.091> ...

'mais je ne pourrait pas faire <BACKSPACE5>travailler tout le temps **à côté des études.**'

L'exemple 51 montre le début de la toute production, car il y existe la marque <START>. Nous voyons que l'étudiant a eu des pauses longues pour la planification du texte, et par conséquent, il semble que l'étudiant a produit la SP *vid sidan av studierna* 'à côté des études' en suivant la phrase dans le sujet d'écriture. Néanmoins, l'étudiant a employé cette expression plusieurs fois, ce qui indiquerait qu'il s'agit d'une SP, et l'exemple 52 illustre comment l'étudiant a utilisé cette SP vers la fin du texte. Nous pouvons tirer de cet exemple que le verbe *jobba* 'travailler' ne fait pas partie de la SP, même s'il se trouve à proximité de la SP ; nous avons fait la même remarque dans les productions des autres étudiants. Dans la verbalisation, l'étudiant D aussi prouve que

cette expression représente la SP automatisée pour lui ; il commente pendant les pauses après le mot *studierna* qu'il a pensé s'il existait une autre façon de dire *vid sidan av studierna* 'à côté des études', car il l'avait répétée quatre fois.

Des autres types, nous présentons une SP L du type de SV que l'étudiant D a produite. Nous l'illustrons par l'exemple 53 suivant.

(53)

... En nackdel som <2.635> kan <3.715> **komma fram** med samtidigt <2.686> arbetet och studiera<BACKSPACE1>na <13.332> ...

'Un inconvénient qui peut **apparaître** avec le travail et les études simultanés'

Dans cet exemple, il est clair que la suite de mots est produite comme une unité. Nous trouvons une pause avant l'expression et il n'existe pas de pauses à l'intérieur de cet SP, il semble donc qu'il s'agit d'une SP automatisée chez l'étudiant D.

### Les SP discursives

La deuxième catégorie est les SP discursives, et l'étudiant D s'est servi de deux types différents, comme il l'a fait en français. Le tableau 59 montre des occurrences dans ces types.

**Tableau 59. Les types de SP discursives en suédois de l'étudiant D**

| SP   | Nombre d'occurrences |      |          |
|--|----------------------|------|----------|
|  | MT                   | MAEP | Au total |
| Processus  | 2                    | 4    | 6        |
| Produit final  | 2                    | 4    | 6        |
| MT= marqueurs textuels, MAEP= marqueurs autorégulateurs et expressions de positionnement |                      |      |          |

Dans la production français, cet étudiant a utilisé notablement plus de SP D du type de MT, mais comme le tableau 59 l'indique, le type de MAEP est plus utilisé en suédois. Cependant, nous avons trouvé que l'étudiant a employé des SP semblables en français et en suédois, et dans l'exemple 54 suivant, nous illustrons telle SP du type de MAEP.

(54)

... <5.569> Men <5.764> **jag tycker att** i Finland <6.114> ...

'Mais **je pense qu'**en Finlande'

Cet exemple montre que l'étudiant D a produit cette expression d'opinion après une pause de réflexion, sans pauses à l'intérieur, c'est-à-dire comme une SP. Dans sa production en français, il a également employé de la même manière la SP *je pense que*, donc il semble que ces SP représentent les SP automatisées pour lui dans les deux langues.

Les deux SP D du type de MT en suédois ont été *det finns* 'il y a' et *är det* 'est-ce' (c'est). Elles sont toutes les deux présentes aussi dans la production en français, et nous traitons la SP *det finns* 'il y a' dans l'exemple 55 suivant.

(55)

<START> <33.621> **Det finns** många <6.414> ...  
'**Il y a** plusieurs'

Nous avons déjà montré ce début de la production de l'étudiant D dans l'exemple 51, mais il nous semble important de souligner le fait que l'étudiant a commencé toute la production avec la SP *det finns* 'il y a'. Cela montre qu'une SP automatisée peut aider l'étudiant à commencer à écrire, car la production d'une SP automatisée exige moins d'efforts cognitifs (*cf.* entre autres, Perez-Bettan, 2015).

### Les SP grammaticales

Comme l'étudiant D a produit des SP grammaticales plus en suédois qu'en français, nous présentons cette catégorie. Nous avons vu dans le tableau 57 que l'étudiant a employé deux fois la SP *hur mycket* 'comment beaucoup= combien', et l'autre SP que l'étudiant a utilisée une fois a été *för mycket* 'trop beaucoup= trop'. Les deux SP G représentent donc le type d'expressions de quantité et de degré (QD) et elles sont restées les mêmes dans le processus et dans le produit final. L'exemple 56 ci-dessous illustre la SP *hur mycket*.

(56)

... man måste inte <6.040> tänka på <2.856> gränd<BACKSPACE1>serna.<LEFT10>  
<5.910> <RIGHT9> <BACKSPACE9> <3.735> **hur mycket** man <3.809> tjänar  
<DOWN7> <15.780> ...  
'on ne doit pas penser aux limites <BACKSPACE9> **combien** on gagne'

Nous voyons dans cet exemple que l'étudiant a produit cette expression comme une unité entière entre les pauses, comme il l'a fait dans les cas de plusieurs SP. Il est intéressant qu'il ait produit le reste de la phrase après une pause (3.809), ce qui renforce le fait que la SP *hur mycket* 'combien' a été produite de manière aisée avant la pause.

## Les SPIDI

L'étudiant D a produit un type de SPIDI, mais il l'a utilisé deux fois dans la production. Le fait que les occurrences de la SPIDI sont restées les mêmes dans le processus et dans le produit final nous a indiquée qu'il s'agit effectivement d'une forme fautive que l'étudiant a acquise telle quelle. La SPIDI est donc *på sommarna*\* et elle représente une SP L du type de SAdv. Nous l'illustrons dans l'exemple 57 ci-dessous.

(57)

... <BACKSPACE13> Jag har få<BACKSPACE2>inte fått stöder **på sommarna** så  
<4.613> ...

'Je n'ai pas eu d'aide chaque été donc'

La forme correcte serait *på somrarna* et en effet, c'est la forme au pluriel. Dans cet exemple, la SP n'est pas entourée directement par des pauses, mais comme elle est produite comme une unité, et l'étudiant l'a produite deux fois, il est clair qu'il s'agit de la SPIDI chez l'étudiant D.

Cette partie de ce chapitre a traité l'usage des SP en suédois par l'étudiant D. Les catégories marquantes avec lesquelles nous avons décrit cet usage ont été les SP lexicales, discursives, grammaticales et les SPIDI. Comme nous avons examiné toutes les productions de tous les quatre étudiants, nous continuerons dans le chapitre suivant avec un résumé et nous discuterons nos résultats par rapport aux autres études.

### 5.3.5. En somme

Notre objectif dans ce chapitre 5.3. a été d'analyser l'utilisation individuelle des SP par chaque étudiant, et les résultats ont montré que nous pouvons distinguer des profils d'utilisateurs individuels dans l'usage des SP. L'étudiant A a produit beaucoup de SP dans les deux langues, mais il en a également effacé notablement. La distribution des SP en catégories a été un peu différente dans ses deux productions, mais il existe aussi des similarités dans ces productions. Par exemple, l'étudiant A a employé plusieurs fois des SP D du type de MAEP et de MT ainsi que des SP L du type de SAdv et de SV dans les deux langues. Il a donc considérablement telles SP automatisées aussi bien en français qu'en suédois. Quant à l'étudiant B, il a aussi produit beaucoup de SP dans ses deux productions, et la distribution des SP en catégories a été similaire dans les deux langues. Une exception a été la catégorie des SPIDI, car l'étudiant en a produit notablement plus en français. Autrement, l'utilisation remarquable des SP L du type de SAdv et des SP D

du type de MT a décrit son usage des SP dans les deux langues. En ce qui concerne l'étudiant C, il a eu beaucoup de pauses longues dans les processus d'écriture dans les deux langues. Cela a rendu l'identification des SP facile, mais les pauses ont évidemment influé sur le nombre de mots produits et sur le nombre d'occurrences des SP. Dans les deux productions, l'étudiant C a employé considérablement plus de SP L que les autres catégories, mais autrement, nous trouvons beaucoup de variation dans les catégories et les types utilisés entre les langues. Quant à l'étudiant D, ses deux productions ont montré que sa manière d'utiliser des SP est très variée ; il a été le seul étudiant qui a employé les SP de toutes les catégories dans les deux langues. Son usage des SP a aussi été diversifié de telle manière que la distribution des SP en catégories a été notablement différente en français et en suédois. Toutefois, l'étudiant D n'a pas produit tellement de mots dans ses productions, mais les proportions des SP ont été élevées, ce qui soulève la question si l'utilisation abondante des SP rend le texte final répétitif.

Même si notre objectif n'a pas été de comparer les étudiants entre eux, les résultats individuels ont pourtant soulevé des points intéressants qui se méritent d'être traités. Il est devenu clair que les catégories des SP lexicales et des SP discursives ont été utilisées le plus, mais dans les productions en suédois, la catégorie des SP L a été notablement supérieure aux autres catégories (*cf.* aussi ch. 5.2.3.). En effet, tous les quatre étudiants ont employé le plus de SP L dans leur production en suédois. Cela est remarquable parce qu'il s'agit des occurrences des SP et pas par exemple de nombre de mots faisant partie des SP ; les productions en suédois comprennent donc plus d'occurrences des SP L que des SP D. Ce fait peut indiquer que nous n'avons pas identifié toutes les SP D qui ont été produites, ou que des séquences discursives en suédois sont plus difficiles pour les étudiants que des séquences lexicales. Il faut noter que dans les productions en français, il n'existe pas de grande différence entre les nombres d'occurrences dans ces deux catégories. En tout, les étudiants ont produit beaucoup de SP L dans les deux langues, ce qui a été surprenant, car Forsberg (2008) a proposé que le taux élevé de SP L indique que le locuteur/l'étudiant a des compétences linguistiques très avancées (*cf.* aussi Rasmussen, 2015). Notons que Forsberg a étudié de la langue française dans ce cas, donc nous ne pouvons pas dire que c'est équivalent aussi en suédois. Toutefois, cela peut indiquer que nous avons identifié les SP L de manière différente que Forsberg (2008) et Rasmussen (2015), et en raison de cela, nous avons eu des occurrences élevées dans cette catégorie.

Un autre point marquant qui concerne la catégorie des SP L est le fait que le type de SAdv a été le plus utilisé dans les deux langues ; seulement dans la production en suédois de l'étudiant C, le type de SV a été plus grand que le type de SAdv. Dans l'étude de Rasmussen (2015), la proportion du type de SAdv a été élevée, mais les types de SN et de SV ont pourtant eu des taux plus grands. Le nombre haut d'occurrences des SP L du type de SAdv dans notre étude peut être causé par le fait que tous les étudiants se sont servis des expressions qui ont été présentes dans les sujets d'écriture et ils les ont produites comme les SP. Il s'agit donc des SP *en échange* et *vid sidan av studier/na* 'à côté d'/des études' qui représentent le type de SAdv, et les étudiants les ont employées plusieurs fois. Nous avons évoqué dans le chapitre 3.3. que dans l'étude de Jaworska *et al.* (2015), les chercheurs ont formé une catégorie pour les expressions liées au sujet d'essai (*topic-related expressions*). Ils ont donc séparé des autres SP produites les SP qui ont été directement copiées du sujet d'écriture, et en effet, ils n'ont pas effectué une analyse profonde sur ces SP copiées (id., p. 510, 516). Toutefois, Jaworska *et al.* (2015, p. 519) ont trouvé que le niveau des compétences langagières du scripteur n'influence pas tellement d'utilisation des SP copiées, car aussi bien les scripteurs L1 que les scripteurs L2 les ont employées dans leurs productions. Nous pouvons donc constater qu'il est naturel que les étudiants dans notre étude aient beaucoup utilisé de séquences qui ont fait partie des sujets d'écriture. Il n'a pourtant pas été nécessaire d'exclure ces SP « copiées » dans cette étude parce que l'étude du processus nous a montrée que les étudiants ont employé des séquences « prêtes » de manières variées.

Dans le chapitre 5, nous avons présenté de manière profonde l'analyse des productions écrites en français et en suédois. Nous avons répondu à nos questions de recherche et nous avons traité des résultats aussi bien au niveau général qu'au niveau des étudiants individuels. Nous terminerons ce mémoire dans le chapitre suivant avec la conclusion.

## **6. Conclusion**

Dans ce mémoire de maîtrise, notre objectif a été d'étudier les séquences préfabriquées dans les productions écrites en français et en suédois à l'aide du logiciel *ScriptLog* qui est un logiciel d'enregistrement des frappes. Nous nous sommes concentrée non seulement sur la quantité de SP dans les productions mais aussi sur les types différents de SP et sur les manières individuelles d'utiliser des SP dans les productions dans les deux langues. Egalement, nous avons voulu montrer la manière dont les SP automatisées influencent le processus d'écriture. Comme les quatre étudiants

multilingues ont tous produit un texte en français et un autre en suédois, nous avons pu comparer les résultats au niveau de tous les étudiants ensemble et aussi au niveau de chaque étudiant individuellement.

Notre première question de recherche a concerné la quantité de SP dans les productions écrites et nous l'avons formulée comme suit : « Quelle est la quantité de séquences préfabriquées dans les productions écrites des étudiants du français L2 et du suédois L2 ? Existe-il des différences entre les deux langues ? » Notre hypothèse a été que la quantité de SP est d'environ 25 %, car c'est le taux que par exemple Forsberg (2006) et Rasmussen (2015) ont trouvé dans leurs études des SP en français, et nous avons estimé que la quantité est similaire dans les deux langues. Nous avons calculé la quantité selon le nombre de mots faisant partie des SP, comme il a été fait dans les études antérieures. Dans notre étude, la quantité de SP a alors été de 29 % dans le processus en français et de 30 % dans le produit final en français, tandis que dans les productions en suédois, la quantité a été de 23 % dans le processus et de 24 % dans le produit final. Nous pouvons donc constater que notre hypothèse a été partiellement correcte ; les pourcentages sont d'environ 25 %, mais il existe assez grande différence entre les deux langues. Une raison pour ce fait a été par exemple l'orthographe différente dans les langues (*cf.* ch. 5.1.3.).

La deuxième question de recherche a concerné l'utilisation des catégories différentes des SP au niveau général. Les deux questions que nous avons posées ont été : « Quelles catégories des séquences préfabriquées les étudiants utilisent-ils ? Utilisent-ils les mêmes catégories dans les deux langues ? » Notre hypothèse a compris que les étudiants utilisent au moins les catégories des SP lexicales, discursives et grammaticales, et que les catégories utilisées peuvent varier selon la langue. Egalement, nous avons proposé qu'il soit possible de trouver des SP qui forment une nouvelle catégorie dans le cadre de cette étude. Les résultats montrent que les catégories les plus grandes ont été les SP lexicales et les SP discursives dans les deux langues, mais les catégories des SP grammaticales et des SPIDI ont aussi eu des nombres d'occurrences élevés. Les catégories les plus grandes ont alors été les mêmes dans les deux langues, mais les nombres d'occurrences dans ces catégories ont été différents entre les langues (*cf.* ch. 5.2.3.). Comme nous l'avons prévu, nous avons formé même deux nouvelles catégories d'après les SP que nous avons trouvées dans les productions. Dans cette partie, toutes les hypothèses ont donc été confirmées.

Notre troisième objectif dans cette étude a été d'examiner l'usage individuel des SP par chaque étudiant et en suivant cet objectif, la troisième question de recherche a été « Quels sont les profils d'utilisateurs dans l'usage des séquences préfabriquées des étudiants individuels ? » Notre hypothèse a été que nous pouvons distinguer des profils d'utilisateurs individuels et que la manière d'un étudiant d'utiliser des SP dans les deux langues peut varier, c'est-à-dire que l'étudiant peut employer des types différents de SP dans ses productions en français et en suédois. Les résultats montrent que cette hypothèse a été confirmée ; nous avons pu distinguer les profils d'utilisateurs individuels et nous avons trouvé que les étudiants ont varié leur utilisation des types différents de SP entre les deux langues, même si les catégories les plus utilisées avaient été les mêmes dans leurs deux productions. L'analyse individuelle a aussi montré que les SP automatisées facilitent la production écrite. Quoique nous n'ayons pas comparé les résultats individuels entre eux, nous avons pourtant trouvé quelques similarités dans les résultats, par exemple le fait que tous les étudiants ont utilisé le plus de SP lexicales dans leur production en suédois (*cf.* ch. 5.3.5.).

Comme nous l'avons évoqué (*cf.* ch. 4.4.), nous avons été consciente du fait que les étudiants avancés des langues étrangères peuvent produire des séquences aisées et grammaticalement correctes sans qu'elles représentent des SP pour les étudiants (*cf.* Myles & Cordier, 2017, p. 14). L'analyse des productions nous a cependant révélée que nous pouvons savoir à l'aide de l'étude du processus d'écriture si des séquences produites sont en effet les séquences préfabriquées. Notamment, si une séquence a été produite une fois dans le produit final, mais le processus montre qu'elle a en effet été produite plusieurs fois, cela indique qu'il s'agit effectivement d'une SP. Par exemple, l'étudiant A avait produit trois fois la SP *det är* 'il est' dans le processus en suédois, mais dans le produit final nous la trouvons seulement une fois. L'étude du processus a donc prouvé que l'étudiant avait produit cette séquence comme la SP. Egalement, l'étude du processus d'écriture a révélé si une séquence qui semble être produite comme une SP dans le produit final n'est pas une SP en réalité. Par exemple, l'étudiant C avait employé la SPIDI *en étrangère\** dans le processus en français, mais il l'avait modifiée de telle manière que dans le produit final, il existe la séquence *en échange*. Evidemment, nous n'avons pas pu compter cette séquence comme une SP. La localisation des pauses nous a aussi aidée dans le travail d'identification des SP, car nous avons trouvé des SP possibles dans les produits finaux, mais les processus ont indiqué qu'il existe des pauses à l'intérieur de ces séquences. Finalement, les

verbalisations rétrospectives nous ont servie comme un moyen de vérifier si une séquence représente une SP pour un étudiant. Par exemple, cela a été le cas avec l'étudiant D et la SP *vid sidan av studiern* 'à côté des études', comme nous l'avons vu dans le chapitre 5.3.4.

Dans ce mémoire, nous avons alors montré que les étudiants multilingues peuvent atteindre la quantité élevée de SP dans les productions écrites en L2. La comparaison entre les productions en français et en suédois a révélé qu'il existe de la variation dans l'usage des SP dans les deux langues chez l'étudiant individuel. Nous avons également vu que les compétences langagières de l'étudiant ont influé sur son utilisation des types différents de SP. L'étude du processus d'écriture à l'aide du logiciel *ScriptLog* a prouvé qu'il est pertinent, voire nécessaire, d'examiner aussi le processus lorsque nous étudions les SP dans les produits écrits. Le processus a révélé d'information importante par exemple des SP automatisées chez les étudiants. Toutefois, nous sommes consciente du fait que le nombre de participants dans cette étude a été petit, en raison de cela, ces résultats ne peuvent pas être généralisés. Nous avons cependant trouvé d'indices qu'il est possible que la quantité de SP soit plus grande dans l'écriture que dans la langue parlée d'un étudiant, mais évidemment cela mériterait plus d'études ultérieures. Il serait en effet intéressant de continuer cette étude en analysant les SP dans la langue parlée des mêmes étudiants afin d'avoir une image complète de leurs manières d'utiliser les SP. En définitive, nous attendons avec enthousiasme que plus d'études soient menées sur les séquences préfabriquées dans les productions écrites en langues étrangères.

## BIBLIOGRAPHIE

- Baddeley, A. (2000). The episodic buffer : a new component of working memory ? *Trends in Cognitive Science*, 4(11), 417-423.
- Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR). (2001). Disponible sur <https://rm.coe.int/16802fc3a8>
- Cislaru, G., & Olive, T. (2018a). *Le processus de textualisation : Analyse des unités linguistiques de performance écrite*. Louvain-la-Neuve : De Boeck Supérieur.
- Cislaru, G., & Olive, T. (2018b). Bursts of written language as performance units for the description of genre routines. Dans D. Legallois, T. Charnois & M. Larjavaara (éds.), *The grammar of genres and styles : from discrete to non-discrete units* (p. 219-246). Berlin : De Gruyter Mouton.
- Coirier, P., Gaonac'h, D., & Passerault, J.-M. (1996). *Psycholinguistique textuelle. Approche cognitive de la compréhension et de la production des textes*. Paris : Armand Colin.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences : Are they processed more quickly than nonformulaic language by native and nonnative speakers ? *Applied Linguistics*, 29(1), 72-89.
- Cordier, C. (2013). *The presence, nature and role of formulaic sequences in English advanced learners of French : a longitudinal study*. (Thèse de doctorat). Newcastle : Newcastle University. Disponible sur <https://theses.ncl.ac.uk/dspace/bitstream/10443/2265/1/Cordier%2013.pdf>
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26(1), 59-84.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29-62.
- Forsberg, F. (2006). *Le langage préfabriqué en français parlé L2 : Étude acquisitionnelle et comparative*. (Thèse de doctorat). Stockholm : Université de Stockholm.

- Forsberg, F. (2008). *Le langage préfabriqué : Formes, fonctions et fréquences en français parlé L2 et L1*. Bern : Peter Lang.
- Gheitasi, P. (2017). « *Say it fast, fluent and flawless* » : *Formulaicity in the oral language production of young foreign language learners*. (Thèse de doctorat). Umeå : Umeå University.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. Dans L. W. Gregg & E. R. Steinberg (éds.), *Cognitive processes in writing* (p. 3-30). Hillsdale, N.J. : Lawrence Erlbaum Associates.
- Jaworska, S., Krummes, C., & Ensslin, A. (2015). Formulaic sequences in native and non-native argumentative writing in German. *International Journal of Corpus Linguistics*, 20(4), 500-525.
- Johansson, V. (2017). *ScriptLog (short user introduction, April 2017)*.
- Lindgren, E., & Sullivan, K. P. H. (2019). *Observing writing : Insights from keystroke logging and handwriting*. Leiden : Brill.
- Lintunen, P., Mutta, M., & Peltonen, P. (2020). Defining fluency in L2 learning and use. Dans P. Lintunen, M. Mutta & P. Peltonen (éds.), *Fluency in L2 learning and use* (p. 1-15). Bristol : Multilingual Matters.
- Mutta, M. (2007). *Un processus cognitif peut en cacher un autre : étude de cas sur l'aisance rédactionnelle des scripteurs finnophones et francophones*. (Thèse de doctorat). Turku : Université de Turku.
- Mutta, M. (2020). L2 fluency and writer profiles. Dans P. Lintunen, M. Mutta & P. Peltonen (éds.), *Fluency in L2 learning and use* (p. 63-80). Bristol : Multilingual Matters.
- Myles, F., & Cordier, C. (2017). Formulaic sequence(FS) cannot be an umbrella term in SLA : Focusing on psycholinguistic FSs and their identification. *Studies in Second Language Acquisition*, 39(1), 3-28.
- Nyqvist, E.-L. (2013). *Species och artikelbruk i finskspråkiga grundskoleelevers inläraarsvenska. En longitudinell undersökning i årskurserna 7-9*. (Thèse de doctorat). Turku : Université de Turku.

- Olkkonen, S., & Mutta, M. (2020). Cognitive fluency in L2 : What inaccuracies can reveal about processing and proficiency. Dans P. Lintunen, M. Mutta & P. Peltonen (éds.), *Fluency in L2 learning and use* (p. 34-48). Bristol : Multilingual Matters.
- Perez-Bettan, A. (2015). *Apprentissage et utilisation du langage préfabriqué chez des apprenants de français langue étrangère*. (Thèse de doctorat). Paris : Université Paris 8 – Vincennes – Saint-Denis.
- Rasmussen, L. (2015). *L'usage des séquences préfabriquées dans le langage parlé d'étudiants finnophones du français L2 au niveau universitaire : Une étude de cas*. (Mémoire de Master inédit). Université de Turku.
- Roca de Larios, J., Manchón, R.M., & Murphy, L. (2006). Generating text in native and foreign language writing : a temporal analysis of problem-solving formulation processes. *The Modern Language Journal*, 90(1), 100-114.
- Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The Foreign Language Writer's Strategic Behaviour in the Allocation of Time to Writing Processes. *Journal of Second Language Writing*, 17(1), 30-47.
- Sitri, F., & Tutin, A. (2016). Présentation. *Lidil*, 53, 5-18.
- Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). Formulaic language : Setting the scene. Dans A. Siyanova-Chanturia & A. Pellicer-Sánchez (éds.), *Understanding formulaic language : A second language acquisition perspective* (p. 1-15). New York : Routledge.
- Spelman Miller, K. (2000). Academic writers on-line : investigating pausing in the production of text. *Language Teaching Research*, 4(2), 123-148.
- Sundman, M. (2011). Helsekvenser i finska grundskoleelevers skrivande. Dans A.-C. Edlund & I. Mellenius (éds.), *Svenskans beskrivning 31* (p. 327-336). Umeå : Umeå universitet.
- Svenska Akademiens ordböcker. Disponible sur <https://svenska.se/>
- Turun yliopisto, valintaopas 2019*. Turun yliopisto. Disponible sur [https://www.utu.fi/sites/default/files/public%3A//media/file/valintaopas-19-turun-yliopisto\\_2.pdf](https://www.utu.fi/sites/default/files/public%3A//media/file/valintaopas-19-turun-yliopisto_2.pdf) consulté le 13 avril 2019

- Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it : An eye-movement study into the processing of formulaic sequences. Dans N. Schmitt (éds.), *Formulaic Sequences: Acquisition, Processing and Use* (p. 153-172). Amsterdam : John Benjamins Publishing Company.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge : Cambridge University Press.
- Wray, A. (2008). *Formulaic language : Pushing the boundaries*. Oxford : Oxford University Press.
- Wray, A. (2009). Identifying formulaic language : Persistent challenges and new opportunities. Dans R. Corrigan, E. A. Moravcsik, H. Ouali & K. M. Wheatley (éds.), *Formulaic language : Volume 1, Distribution and historical change* (p. 27-51). Amsterdam : John Benjamins Publishing Company.

# ANNEXES

## Annexe 1. Questionnaire de base

### Taustatietokysely

1. Nimi \*

2. Sähköposti \*

3. Tutkimukseen osallistuminen \*

- Suostun siihen, että tietojani käytetään anonyymisti pro gradu -tutkimuksessa.
- En suostu siihen, että tietojani käytetään anonyymisti pro gradu -tutkimuksessa.

#### 4. Syntymävuosi \*

- 2000
- 1999
- 1998
- 1997
- 1996
- 1995
- 1994
- 1993
- 1992
- 1991
- 1990
- 1989
- 1988
- 1987
- 1986
- 1985
- 1984
- 1983
- 1982
- 1981
- 1980
- ennen vuotta 1980

5. Pääaine \*

- ranska
- pohjoismaiset kielet
- englantti
- muu, mikä? \_\_\_\_\_

6. Sivuaine/et \*

|   |  |
|---|--|
| * |  |
|---|--|

7. Äidinkieli/kielet \*

- suomi
- ruotsi
- kaksikielinen, kielet? \_\_\_\_\_
- muu, mikä? \_\_\_\_\_

**Eurooppalaisen viitekehyksen taitotasot A1-C2 (aukaise linkki uudelle välilehdelle): <https://europass.cedefop.europa.eu/sites/default/files/cefr-fi.pdf>**

**8. Mikä on vahvin vieras kielesi? Arvioi osaamistasi eurooppalaisen viitekehyksen taitotasojen avulla (A1-C2) ja/tai kerro omin sanoin. \***

---

---

---

---

---

---

---

---



Jos vastasit kyllä, vastaa kysymyksiin 13 ja 14.

**13. Mikä oli oleskelusi syy? (Esim. vaihto-opiskelu, työ, lomamatka...)**

---

---

---

---

---

---

---

---

**14. Kuinka kauan vietit siellä aikaa?**

---

---

---

---

---

---

---

---

**15. Ranskan kielen käyttö \***

|                                     | Päivittäin               | Viikoittain              | Kuukausittain            | Harvemmin<br>kuin kerran<br>kuukaudessa | En<br>milloinkaan        |
|-------------------------------------|--------------------------|--------------------------|--------------------------|---|--------------------------|
| Kuinka usein puhut ranskaa? *       | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |
| Kuinka usein kirjoitat ranskaksi? * | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |
| Kuinka usein kuuntelet ranskaa? *   | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |
| Kuinka usein luet ranskaksi? *      | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |

16. Millaisissa tilanteissa käytät ranskaa? (Esim. kavereiden kanssa, lukeminen, kirjoittaminen, elokuvien/tv-ohjelmien katselu...)

---

---

---

---

---

---

---

---

---

---

### **Ruotsin opiskelu**

17. Missä iässä olet aloittanut ruotsin opinnot? \*

---

---

18. Kuinka kauan olet opiskellut ruotsia? \*

---

---

---

19. Oletko oleskellut ruotsinkielisessä maassa/alueella? Missä? \*

---

---

---

---

---

---

---

---

Jos vastasit kyllä, vastaa kysymyksiin 20 ja 21.

**20. Mikä oli oleskelusi syy? (Esim. vaihto-opiskelu, työ, lomamatka...)**

---

---

---

---

---

---

---

---

**21. Kuinka kauan vietit siellä aikaa?**

---

---

---

---

---

---

---

---

**22. Ruotsin kielen käyttö \***

|                                     | Päivittäin               | Vikoittain               | Kuukausittain            | Harvemmin<br>kuin kerran<br>kuukaudessa | En<br>milloinkaan        |
|-------------------------------------|--------------------------|--------------------------|--------------------------|---|--------------------------|
| Kuinka usein puhut ruotsia? *       | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |
| Kuinka usein kirjoitat ruotsiksi? * | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |
| Kuinka usein kuuntelet ruotsia? *   | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |
| Kuinka usein luet ruotsiksi? *      | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/>                | <input type="checkbox"/> |



## Annexe 2. Listes des SP utilisées par chaque étudiant en français et en suédois

### L'étudiant A ; les SP en français

| SP                           | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type     |
|------------------------------|------------------------|----------------------------|--------------------------------|-----------|----------|
| A mon avis                   | 3                      | 3                          |                                | D         | MAEP     |
| Année en échange             | 1                      | 1                          |                                | L         | SN       |
| Il y a                       | 3                      | 3                          |                                | D         | MT       |
| Prendre du temps             | 1                      | 0                          |                                | L         | SV       |
| Un peu                       | 1                      | 1                          |                                | G         | QD       |
| Avant que                    | 1                      | 1                          |                                | D         | MT       |
| Il peut sembler que          | 1                      | 1                          |                                | D         | MAEP     |
| Par exemple                  | 1                      | 1                          |                                | D         | MT       |
| En France                    | 2                      | 2                          |                                | T         | SAdv     |
| Pas du tout                  | 1                      | 1                          |                                | G         | QD       |
| Toutes les fois              | 2                      | 2                          |                                | L         | SAdv     |
| En Finlande                  | 3                      | 3                          |                                | T         | SAdv     |
| J'ai appris                  | 2                      | 1                          |                                | L         | SV       |
| En français                  | 1                      | 1                          |                                | T         | SN       |
| Je crois que                 | 1                      | 1                          |                                | D         | MAEP     |
| Il n'est pas possible        | 1                      | 1                          |                                | D         | MAEP     |
| Parler une langue            | 1                      | 1                          |                                | L         | SV       |
| Il est non plus pas possible | 1                      | 1                          |                                | D         | MAEP     |
| Passer du temps              | 1                      | 1                          |                                | L         | SV       |
| Il est possible              | 1                      | 1                          |                                | D         | MAEP     |
| Bien sûr                     | 2                      | 1                          |                                | D         | MI       |
| Ce n'est (pas)               | 1                      | 1                          |                                | D         | MT       |
| Pour moi                     | 2                      | 2                          |                                | D         | MAEP     |
| Je pense pas                 | 1                      | 1                          |                                | D         | MAEP     |
| Beaucoup (+V)                | 1                      | 1                          |                                | G         | QD       |
| En plus                      | 1                      | 1                          |                                | D         | MT       |
| Pas de problèmes             | 1                      | 1                          |                                | L         | SN       |
| En passant les cours         | 1                      | 1                          |                                | L         | SV       |
| A l'université               | 1                      | 1                          |                                | L         | SAdv     |
| Au début                     | 1                      | 1                          |                                | L         | SAdv     |
| Au printemps                 | 1                      | 0                          |                                | L         | SAdv     |
| A la fac                     | 1                      | 1                          |                                | L         | SAdv     |
| Il y avait                   | 1                      | 0                          |                                | D         | MT       |
| A nouveau                    | 1                      | 1                          |                                | L         | SAdv     |
| Pour une étudiante           | 1                      | 1                          | Pour moi                       | M         | D (MAEP) |
| Pour quelqu'un               | 1                      | 1                          |                                | G         | PI       |
| Là-bas                       | 1                      | 1                          |                                | L         | SPré     |
| Ce que je veux dire          | 1                      | 1                          |                                | D         | MI       |

|               |   |   |  |       |          |
|---------------|---|---|--|-------|----------|
| Il y aurait   | 1 | 1 |  | D     | MT       |
| Plus de       | 1 | 1 |  | G     | QD       |
| A l'étrangère | 1 | 1 |  | SPIDI | L (SAdv) |
| A l'étranger  | 1 | 0 |  | L     | SAdv     |

### L'étudiant A ; les SP en suédois

| SP                       | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type     |
|--------------------------|------------------------|----------------------------|--------------------------------|-----------|----------|
| Jag tänker att           | 2                      | 2                          |                                | D         | MAEP     |
| Vid sidan av studier     | 5                      | 5                          |                                | L         | SAdv     |
| I Finland                | 1                      | 1                          |                                | T         | SAdv     |
| Van vid                  | 1                      | 1                          |                                | L         | SAdj     |
| Till exempel             | 2                      | 2                          |                                | D         | MT       |
| Nästa sommar             | 1                      | 1                          |                                | L         | SN       |
| För att                  | 1                      | 0                          |                                | D         | MT       |
| I framtiden              | 1                      | 1                          |                                | L         | SAdv     |
| Det finns också          | 1                      | 1                          |                                | D         | MT       |
| Ha tillräckligt med      | 2                      | 2                          |                                | L         | SV       |
| Ha inte tillräckligt med | 1                      | 1                          |                                | L         | SV       |
| I sin vardag             | 1                      | 1                          |                                | L         | SAdv     |
| Jag tror att             | 1                      | 1                          |                                | D         | MAEP     |
| Så mycket                | 1                      | 1                          |                                | G         | QD       |
| På helgen                | 1                      | 1                          |                                | L         | SAdv     |
| För mig                  | 5                      | 4                          |                                | D         | MAEP     |
| Tid för att              | 1                      | 1                          |                                | L         | SN       |
| Koppla sig av            | 1                      | 1                          |                                | SPIDI     | L (SV)   |
| Lägga märke till         | 1                      | 1                          |                                | SPIDI     | L (SV)   |
| Under veckan             | 1                      | 1                          |                                | L         | SAdv     |
| Det är                   | 3                      | 1                          |                                | D         | MT       |
| Ha erfarenhet            | 1                      | 1                          |                                | L         | SV       |
| Det finns                | 1                      | 1                          | Det kan finnas                 | M         | D (MT)   |
| Gå på toalett            | 1                      | 1                          |                                | L         | SV       |
| Ta sina pauser           | 1                      | 1                          |                                | L         | SV       |
| A min asivkt             | 1                      | 1                          |                                | SPIDI     | D (MAEP) |
| Vara rädd att            | 1                      | 1                          |                                | L         | SV       |
| Var det                  | 2                      | 2                          |                                | D         | MT       |
| Inte alls                | 1                      | 1                          |                                | G         | QD       |
| På jobbet                | 1                      | 1                          |                                | L         | SAdv     |
| Det var                  | 2                      | 2                          |                                | D         | MT       |
| På viktigt               | 1                      | 1                          |                                | L         | SAdv     |
| I början                 | 1                      | 1                          |                                | L         | SAdv     |

## L'étudiant B ; les SP en français

| SP                              | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type           |
|---------------------------------|------------------------|----------------------------|--------------------------------|-----------|----------------|
| Etudes en échange               | 2                      | 2                          |                                | L         | SN             |
| Langues étrangères              | 1                      | 1                          |                                | L         | SN             |
| Utiliser une langue étrangère   | 1                      | 1                          |                                | L         | SV             |
| Tout le jour                    | 1                      | 1                          | Touts les jours                | M         | SPIDI (L SAdv) |
| C'est aussi                     | 1                      | 1                          |                                | D         | MT             |
| Dans l'avenir                   | 1                      | 1                          |                                | L         | SAdv           |
| Il y a aussi                    | 1                      | 1                          |                                | D         | MT             |
| Perdre de temps                 | 1                      | 0                          |                                | L         | SV             |
| Il peut être que                | 1                      | 1                          |                                | D         | MT             |
| Il y a                          | 1                      | 1                          |                                | D         | MT             |
| Beaucoup de                     | 1                      | 1                          |                                | G         | QD             |
| Autres choces                   | 1                      | 1                          |                                | SPIDI     | G (PI)         |
| En échange                      | 4                      | 4                          |                                | L         | SAdv           |
| Il n'y a pas                    | 2                      | 2                          |                                | D         | MT             |
| Plain de temps                  | 1                      | 1                          |                                | SPIDI     | G (QD)         |
| Faire des/les études            | 3                      | 3                          |                                | L         | SV             |
| Commencer des études            | 1                      | 1                          | Commencer des études           | M         | SPIDI (L SV)   |
| A l'université                  | 6                      | 6                          |                                | L         | SAdv           |
| Sauf que                        | 1                      | 1                          |                                | D         | MT             |
| En fait                         | 1                      | 1                          |                                | D         | MT             |
| Il est aussi le possibilité que | 1                      | 1                          |                                | SPIDI     | D (MAEP)       |
| Dans le pays                    | 1                      | 1                          |                                | L         | SAdv           |
| Par exemple                     | 4                      | 4                          |                                | D         | MT             |
| Quelques autres                 | 1                      | 1                          |                                | G         | PI             |
| Parce que                       | 3                      | 3                          |                                | D         | MT             |
| J'ai pensé que                  | 1                      | 0                          |                                | D         | MAEP           |
| Je pensais que                  | 1                      | 1                          |                                | D         | MAEP           |
| Plus que                        | 1                      | 1                          |                                | G         | QD             |
| A cause de                      | 1                      | 1                          |                                | L         | SPré           |
| C'est                           | 1                      | 0                          |                                | D         | MT             |
| Quelque d'autres choces         | 1                      | 1                          |                                | SPIDI     | G (PI)         |
| Beaucoup des                    | 2                      | 1                          |                                | SPIDI     | G (QD)         |
| Dans les universités            | 1                      | 1                          |                                | L         | SAdv           |
| Le fait que                     | 1                      | 1                          |                                | D         | MT             |
| Tous les choces                 | 1                      | 0                          |                                | SPIDI     | L (SN)         |
| Ce que                          | 1                      | 1                          |                                | D         | MT             |
| Pour moi                        | 1                      | 1                          |                                | D         | MAEP           |
| Quelqu'un autre                 | 1                      | 1                          |                                | G         | PI             |

|                        |   |   |  |       |        |
|------------------------|---|---|--|-------|--------|
| Tellement beaucoup     | 1 | 1 |  | SPIDI | G (QD) |
| Je suis assez sûre que | 1 | 1 |  | D     | MAEP   |
| Avoir l'argent         | 1 | 1 |  | L     | SV     |
| Après l'échange        | 1 | 1 |  | L     | SAdv   |

### L'étudiant B ; les SP en suédois

| SP                        | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type   |
|---------------------------|------------------------|----------------------------|--------------------------------|-----------|--------|
| Vid sidan av studierna    | 2                      | 2                          |                                | L         | SAdv   |
| Ta mycket tid             | 1                      | 1                          |                                | L         | SV     |
| Till exempel              | 4                      | 3                          |                                | D         | MT     |
| Hitta tid för             | 2                      | 2                          |                                | L         | SV     |
| Per vecka                 | 1                      | 1                          |                                | L         | SAdv   |
| På lördag                 | 1                      | 1                          |                                | L         | SAdv   |
| D.v.s. (det vill säga)    | 1                      | 1                          |                                | D         | MT     |
| Dag/ar per vecka          | 2                      | 2                          |                                | L         | SAdv   |
| Kundservice industrin     | 2                      | 2                          |                                | I         | L (SN) |
| På förhand                | 1                      | 1                          |                                | L         | SAdv   |
| Med sidan av studierna    | 1                      | 1                          |                                | SPIDI     | L (SN) |
| Lite mer                  | 1                      | 1                          |                                | G         | QD     |
| Vara beroende av          | 1                      | 1                          | Vara helt beroende av          | M         | L (SV) |
| Komma ihåg                | 1                      | 1                          |                                | L         | SV     |
| För mycket                | 1                      | 1                          |                                | G         | QD     |
| Det är                    | 4                      | 3                          |                                | D         | MT     |
| Så mycket                 | 3                      | 3                          |                                | G         | QD     |
| Även om                   | 2                      | 2                          |                                | D         | MT     |
| Jag har också tänkt att   | 1                      | 1                          |                                | D         | MAEP   |
| Under studietiden         | 2                      | 2                          |                                | L         | SAdv   |
| Mycket mer                | 1                      | 1                          |                                | G         | QD     |
| Därför att                | 1                      | 1                          |                                | D         | MT     |
| Jag tror                  | 1                      | 0                          |                                | D         | MAEP   |
| Enligt min åsikt          | 1                      | 1                          |                                | D         | MAEP   |
| Är det                    | 1                      | 1                          |                                | D         | MT     |
| Studera vid universitetet | 1                      | 1                          |                                | L         | SV     |
| Ha en möjlighet           | 1                      | 1                          |                                | L         | SV     |
| Vara bra på               | 1                      | 1                          |                                | L         | SV     |
| I fortsättningen          | 1                      | 1                          |                                | L         | SAdv   |

### L'étudiant C ; les SP en français

| SP                          | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type           |
|-----------------------------|------------------------|----------------------------|--------------------------------|-----------|----------------|
| Etudes en échange           | 2                      | 2                          |                                | L         | SN             |
| Bien sûr                    | 1                      | 1                          |                                | D         | MI             |
| Apprentissage d'une langue  | 1                      | 1                          |                                | L         | SN             |
| En échange                  | 2                      | 2                          |                                | L         | SAdv           |
| Locuteurs natifs            | 1                      | 1                          |                                | L         | SN             |
| Dans le cas                 | 1                      | 1                          | Dans le meilleur cas           | M         | SPIDI (L SAdv) |
| En même temps               | 1                      | 1                          |                                | L         | SAdv           |
| Rencontrer des personnes    | 1                      | 1                          |                                | L         | SV             |
| Tous les pays               | 1                      | 1                          |                                | L         | SN             |
| Dans le monde               | 1                      | 1                          |                                | L         | SAdv           |
| De plus                     | 1                      | 1                          |                                | D         | MT             |
| Manière de faire            | 1                      | 0                          |                                | L         | SN             |
| Faire les amis              | 1                      | 1                          | Faire des nouveaux amis        | M         | SPIDI (L SV)   |
| C'est difficile             | 1                      | 1                          |                                | D         | MAEP           |
| Faire des études            | 2                      | 2                          |                                | L         | SV             |
| En Finlande                 | 1                      | 1                          |                                | T         | SAdv           |
| En étrangère                | 1                      | 1                          | En échange                     | M         | SPIDI (L SAdv) |
| En étrangère                | 1                      | 1                          |                                | SPIDI     | L (SAdv)       |
| Il peut être plus facile de | 1                      | 1                          |                                | D         | MAEP           |
| Au moins                    | 1                      | 1                          |                                | L         | SAdv           |
| Pas souvent                 | 1                      | 1                          |                                | L         | SAdv           |
| Il peut être difficile de   | 1                      | 1                          |                                | D         | MAEP           |

### L'étudiant C ; les SP en suédois

| SP                      | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type     |
|-------------------------|------------------------|----------------------------|--------------------------------|-----------|----------|
| Vid sidan av studier/na | 2                      | 2                          |                                | L         | SAdv     |
| Extra pengar            | 1                      | 1                          |                                | L         | SN       |
| I synnerhet             | 1                      | 1                          |                                | L         | SAdv     |
| Det är också            | 1                      | 1                          |                                | D         | MT       |
| I det bästa fallet      | 1                      | 1                          |                                | SPIDI     | L (SAdv) |
| Få jobb                 | 1                      | 1                          |                                | L         | SV       |
| Ha något att göra med   | 1                      | 1                          |                                | L         | SV       |
| Blivande yrke           | 1                      | 1                          |                                | L         | SN       |
| Se till att             | 1                      | 1                          |                                | L         | SV       |
| Ha tillräckligt         | 1                      | 1                          |                                | SPIDI     | L (SV)   |

|             |   |   |  |   |          |
|-------------|---|---|--|---|----------|
| Det är      | 1 | 1 |  | D | MT       |
| Få pengar   | 1 | 1 |  | L | SV       |
| Man måste   | 2 | 2 |  | I | D (MAEP) |
| Hur mycket  | 1 | 1 |  | G | QD       |
| Så mycket   | 1 | 1 |  | G | QD       |
| Någon annan | 1 | 1 |  | G | PI       |

### L'étudiant D ; les SP en français

| SP                         | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type           |
|----------------------------|------------------------|----------------------------|--------------------------------|-----------|----------------|
| En échange                 | 2                      | 2                          |                                | L         | SAdv           |
| Il y a                     | 2                      | 1                          |                                | D         | MT             |
| Plusieurs des              | 1                      | 1                          |                                | SPIDI     | G (QD)         |
| Par exemple                | 2                      | 1                          |                                | D         | MT             |
| Dans un pays étranger      | 1                      | 1                          | En un pays étranger            | M         | SPIDI (L SAdv) |
| Beaucoup des               | 1                      | 1                          |                                | SPIDI     | G (QD)         |
| Faire la connaissance avec | 1                      | 0                          |                                | SPIDI     | L (SV)         |
| Et aussi                   | 1                      | 0                          |                                | D         | MT             |
| Ce pays                    | 3                      | 2                          |                                | I         | L (SN)         |
| Etudes en échange          | 2                      | 2                          |                                | L         | SN             |
| Faire connaissance avec    | 1                      | 1                          |                                | L         | SV             |
| C'est                      | 1                      | 0                          |                                | D         | MT             |
| Ce n'est pas               | 1                      | 1                          |                                | D         | MT             |
| Se familiariser avec       | 1                      | 1                          |                                | L         | SV             |
| C'est difficile            | 1                      | 1                          |                                | D         | MAEP           |
| Peut-être                  | 1                      | 1                          |                                | D         | MT             |
| Mais si                    | 1                      | 1                          |                                | D         | MT             |
| Faire les études           | 1                      | 1                          |                                | L         | SV             |
| Je pense que               | 1                      | 1                          |                                | D         | MAEP           |
| Dans un échange            | 1                      | 1                          |                                | L         | SAdv           |
| Quelques autres            | 1                      | 1                          |                                | G         | PI             |
| En Finlande                | 2                      | 2                          |                                | T         | L (SAdv)       |

### L'étudiant D ; les SP en suédois

| SP                      | Occurrence (processus) | Occurrence (produit final) | Modifiée dans le produit final | Catégorie | Type         |
|-------------------------|------------------------|----------------------------|--------------------------------|-----------|--------------|
| Det finns               | 1                      | 1                          |                                | D         | MT           |
| Avsikter om             | 1                      | 1                          | Åsikter om                     | M         | SPIDI (L SN) |
| Vid sidan av studier/na | 4                      | 4                          |                                | L         | SAdv         |
| Extra pengar            | 1                      | 1                          |                                | L         | SN           |

|                      |   |   |  |       |          |
|----------------------|---|---|--|-------|----------|
| Är det               | 1 | 1 |  | D     | MT       |
| Väldigt bra          | 2 | 2 |  | L     | SAdj     |
| Jag tycker att       | 1 | 1 |  | D     | MAEP     |
| I Finland            | 2 | 1 |  | T     | SAdv     |
| Bero mycket på       | 1 | 1 |  | L     | SV       |
| Hur mycket           | 2 | 2 |  | G     | QD       |
| För mycket           | 1 | 1 |  | G     | QD       |
| På sommarna          | 2 | 2 |  | SPIDI | L (SAdv) |
| För mig              | 1 | 1 |  | D     | MAEP     |
| Komma fram           | 1 | 1 |  | L     | SV       |
| Min åsikt är (att)   | 1 | 1 |  | D     | MAEP     |
| Det är viktigt (att) | 1 | 1 |  | D     | MAEP     |
| Extra jobbet         | 2 | 1 |  | I     | L (SN)   |
| Hela tiden           | 1 | 1 |  | L     | SAdv     |
| Tänka på             | 1 | 1 |  | L     | SV       |
| Njuta av             | 1 | 1 |  | L     | SV       |

### Annexe 3. Explications des marques utilisées dans des exemples

|                             |   |
|-----------------------------|---|
| <u>une SP</u>               | Toutes les SP présentes dans un exemple sont soulignées.                                  |
| <b><u>une SP</u></b>        | La SP qui est analysée dans un exemple est soulignée et mise en gras.                     |
| ... un exemple ...          | Trois points signifient qu'il existe du texte avant/après l'exemple.                      |
| <2.123>                     | La durée d'une pause en secondes.   |
| <BACKSPACE><br><BACKSPACE5> | Retour arrière. Le scripteur a effacé du texte. Le chiffre signifie le nombre de touches. |
| <MOUSECLICK>                | Clic de souris. Le scripteur s'est déplacé dans le texte.                                 |
| <HIGHLIGHT>                 | Le scripteur a sélectionné du texte.  |
| <START>                     | Le début de la production.  |
| <ENTER><br><ENTER1>         | Touche d'entrée. Le chiffre signifie le nombre de touches.                                |
| <LEFT><br><LEFT15>          | Touche flèche gauche. Le chiffre signifie le nombre de touches.                           |
| <RIGHT><br><RIGHT7>         | Touche flèche droite. Le chiffre signifie le nombre de touches.                           |
| <DOWN><br><DOWN3>           | Touche flèche bas. Le chiffre signifie le nombre de touches.                              |
| <UP><br><UP2>               | Touche flèche haut. Le chiffre signifie le nombre de touches.                             |
| 'traduction'                | La traduction en français d'un exemple en suédois.  |

## **Annexe 4. Résumé en finnois – Suomenkielinen tiivistelmä**

### **Sanaryypäät ranskan ja ruotsin kielen yliopisto-opiskelijoiden kirjallisissa tuotoksissa**

#### **1. Johdanto**

Tämän pro gradu -tutkielman aiheena on sanaryypäät vieraalla kielellä kirjoitetuissa tuotoksissa. Sanaryypäät ovat sanojen keskittymiä, jotka koostuvat vähintään kahdesta sanasta ja jotka kielenkäyttäjä oletettavasti omaksuu yhtenä kokonaisuutena – näin ollen ne myös tuotetaan yhtenä kokonaisuutena (Wray 2002; Forsberg 2006). Sanaryypäitä esiintyy sekä äidinkielessä että vieraassa kielessä, ja niitä tutkitaan niin puheessa kuin kirjoituksessa (Myles & Cordier 2017; Perez-Bettan 2015). Esimerkiksi *à mon avis* 'minun mielestäni' sekä *vid sidan av studierna* 'opintojen ohessa' edustavat sanaryypäitä tässä tutkielmassa.

Tutkimuksen kohteena on neljän yliopisto-opiskelijan kirjoittamat tekstit näppäilyntallennusohjelma ScriptLogin avulla. Kaikki neljä opiskelijaa opiskelevat ranskaa ja ruotsia (ts. pohjoismaisia kieliä) joko molempia sivuaineina tai toista kieltä pääaineena ja toista sivuaineena. Näin ollen kaikki opiskelijat ovat kirjoittaneet yhden tekstin ranskaksi ja yhden tekstin ruotsiksi. Näitä tuotoksia käyttäen tutkielman tavoitteena on selvittää, kuinka suuri sanaryypäiden osuus on opiskelijoiden tuottamissa teksteissä. Näppäilyntallennusohjelma ScriptLog mahdollistaa sekä kirjoitusprosessin että lopputuotoksen tutkimisen, joten sanaryypäiden osuutta on mahdollista tarkastella kahdella eri tasolla. Lisäksi tavoitteena on verrata sanaryypäiden osuuksia kahden eri kielen välillä. Tutkielman toisena päämääränä on tarkastella, mitä kategorioita opiskelijoiden tuottamat sanaryypäät edustavat. Tavoitteena on myös verrata, mitä kategorioita ranskassa ja ruotsissa esiintyy, ja tarkastella niiden eroavaisuuksia. Näiden päämäärien lisäksi tutkielmassa keskitytään opiskelijoiden yksilölliseen tapaan käyttää sanaryypäitä tuotoksissaan. Myös tässä tapauksessa verrataan ranskan- ja ruotsinkielistä tekstiä ja keskitytään erityisesti kirjoitusprosessin ja lopputuotoksen väliseen suhteeseen jokaisen opiskelijan kohdalla. Näiden tavoitteiden pohjalta tutkimuskysymykset on laadittu seuraavasti:

1. Mikä on sanaryypäiden osuus ranskan ja ruotsin yliopisto-opiskelijoiden kirjallisissa tuotoksissa? Onko kielten välillä eroja?

2. Mitä sanaryppäiden kategorioita opiskelijat käyttävät? Käyttävätkö he samoja kategorioita molemmissa kielissä?
3. Millaisia ovat opiskelijoiden yksilölliset tavat käyttää sanaryppäitä?

Alan aiempien tutkimusten perusteella tutkimushypoteesina on, että sanaryppäiden osuus on noin 25 % opiskelijoiden tuottamasta tekstistä (mm. Forsberg 2006; Rasmussen 2015). Lisäksi arvioidaan, että kielten välillä ei ole suurta eroa, koska opiskelijoiden kielitaito on melko samantasoinen molemmissa kielissä. Käytettyjen sanaryppäiden kategorioihin oletetaan kuuluvan ainakin leksikaaliset, diskursiiviset ja kieliopilliset sanaryppäät, ja kategorioissa voi olla eroavaisuuksia kielten välillä. Lisäksi on mahdollista, että tutkielman aineiston pohjalta syntyy uusia omia kategorioita. Viimeisen tutkimuskysymyksen kohdalla oletetaan, että kirjoitusprosessin ja lopputuotoksen avulla voidaan tunnistaa opiskelijoiden yksilölliset profiilit sanaryppäiden käytössä ja että heidän käyttötapaansa voivat vaihdella kielen mukaan.

## 2. Teoreettinen viitekehys

Sanaryppäitä käsittelevissä tutkimuksissa törmätään usein terminologiseen ongelmaan, sillä alalla käytetään useita erilaisia termejä, joiden merkitykset eivät aina vastaa toisiaan (mm. Wray 2002; Forsberg 2006; Cordier 2013; Perez-Bettan 2015). Mylesin ja Cordierin (2017, 4) mukaan ongelmana näyttäytyy se, että useat termit voivat viitata yhteen ja samaan asiaan, kun taas yhtä termiä voidaan käyttää viittaamaan moneen eri asiaan. Myles ja Cordier (2017) ovat myös tuoneet esiin, että Wrayn (2002) esittelemästä termistä *formulaic sequence* on tullut ajan saatossa kattotermi, jota alalla käytetään. Näistä syistä on tärkeää valita tutkimuksessa käytettävä termi ja sen määritelmä. Tässä tutkielmassa sanaryppäistä käytetään yleisesti termiä *séquence préfabriquée*, jonka Forsberg (2006) on esittänyt Wrayn englanninkielistä termiä mukailleen. Kuten johdannossa mainittiin, sanaryppäs koostuu vähintään kahdesta sanasta ja se käyttäytyy yhtenä kokonaisuutena. Tämän määrittelyn lisäksi tässä tutkielmassa käytetään tarkemmin Forsbergin (2006, 44) ehdottamaa määritelmää, joka käsittää sekä natiivit kielenkäyttäjät että vieraan kielen käyttäjät ja jossa oppijankieli on otettu huomioon. Lisäksi käytetään Mylesin ja Cordierin (2017, 12) laatimaa määritelmää, joka koskee yksinomaan oppijoille ominaisia sanaryppäitä (*les SP propres à l'apprenant*). Heidän mukaansa sanaryppäiksi lasketaan kaikki sellaiset ryppäät, jotka ovat oppijalle hyödyksi kognitiivisissa prosesseissa, koska ryppäät esiintyvät oppijan

muistissa kokonaisuuksina tai koska ne ovat erittäin automatisoituneita prosessien tasolla.

Automatisoituneet sanaryypäät ovat keskeisessä roolissa tässä tutkielmassa, sillä pyrkimyksenä on myös osoittaa, miten ne helpottavat opiskelijaa vieraan kielen käytössä (ks. tutkimuskysymys 3). Perez-Bettanin (2015, 21; 29; 55) mukaan sanaryypäiden käyttö lisää yleisesti ottaen kielellisen tuottamisen vaivattomuutta, ja tämä koskee niin äidinkieltä kuin vierasta kieltä. Lisäksi tämä vaikuttaa sekä puhutussa että kirjoitetussa kielessä (ibid.), mikä on tärkeä huomio tämän tutkielman kontekstissa. Tuottamisen vaivattomuuteen vaikuttavat erityisesti automatisoituneet sanaryypäät, sillä kognitiivisten prosessien automatisointi kuormittaa vähemmän työmuistia, mikä taas on hyödyksi kielenkäyttäjälle (Perez-Bettan 2015, 55). Automatisoituneiden sanaryypäiden hyödyt nousevat esiin erityisesti vieraan kielen käytössä, koska ne tekevät puhumisesta ja kirjoittamisesta sujuvampaa ja ne vapauttavat aikaa esimerkiksi tekstin suunnittelulle kirjoitusprosessissa (Perez-Bettan 2015, 22; 25).

Kuten mainittu, sanaryypäisiin liittyy moninainen terminologia. Rungas vaihtoehtojen määrä näkyy myös erilaisissa sanaryypäiden luokittelutavoissa, joita tällä tieteenalalla käytetään. Tässä tutkielmassa hyödynnetään aikaisempia kategorioita Forsbergin (2006), Rasmussenin (2015) sekä Jaworskan, Krummesin ja Ensslinin (2015) tutkimuksista. Lisäksi ruotsin kielen osalta hyödynnetään Sundmanin (2011) kategorioiden esimerkkejä. Seuraavassa listassa on esitettyä tämän tutkielman kategoriat sekä niitä vastaavat esimerkit omasta aineistosta:

|                              |   |
|------------------------------|---|
| Leksikaaliset sanaryypäät:   | <i>études en échange</i> 'vaihto-opiskelu', <i>ha tillräckligt med</i> 'olla tarpeeksi jotakin'                                 |
| Diskursiiviset sanaryypäät:  | <i>je pense que</i> 'ajattelen että', <i>till exempel</i> 'esimerkiksi'   |
| Kieliopilliset sanaryypäät:  | <i>beaucoup de</i> 'paljon jotakin', <i>någon annan</i> 'joku muu'  |
| Välikielen sanaryypäät:      | <i>tellement beaucoup*</i> 'niin paljon', <i>på sommarna*</i> 'kesäisin'  |
| Temaattiset sanaryypäät:     | <i>en français</i> 'ranskaksi', <i>i Finland</i> 'Suomessa'   |
| Individuaaliset sanaryypäät: | <i>ce pays</i> 'tämä maa', <i>man måste</i> 'täytyy'  |
| Modifioidut sanaryypäät:     | <i>dans le cas &lt; dans le meilleur cas</i> 'parhaassa tapauksessa', <i>avsikter om &lt; åsikter om</i> 'mielipiteet jostakin' |

Kaksi viimeistä kategoriaa syntyivät tämän tutkimuksen aineiston pohjalta, kun taas muut kategoriat perustuvat aiemmin mainittuihin tutkimuksiin. On myös huomioitava, että kaikki yllä olevat kategoriat sisältävät lisäksi alakategorioita, joita hyödynnetään opiskelijoiden yksilöllisten käyttötapojen tarkastelussa.

### 3. Aineisto ja metodi

Tutkimuksen korpus koostuu neljän opiskelijan kirjoittamista teksteistä ja niiden kirjoitusprosesseista, jotka on tehty ScriptLog-näppäilyntallennusohjelmalla, sekä kirjoitusprosessien verbalisoinneista ja opiskelijoiden taustatietokyselyistä. Kirjalliset tuotokset ovat pääasiallinen tutkimuskohde, kun taas verbalisointeja ja taustatietoja käytetään analyysin tukena. Opiskelijat A ja C opiskelevat ranskaa pääaineena ja pohjoismaisia kieliä sivuaineena, ja opiskelijat B ja D opiskelevat puolestaan englantia pääaineena, jolloin ranska ja pohjoismaiset kielet ovat heidän sivuaineitaan. Tutkimushenkilöiden valinnassa kriteerinä oli, että he opiskelevat kyseisiä kieliä pää- tai sivuaineena, jolloin voidaan odottaa, että heidän kielitaitonsa on riittävä, jotta he voivat kirjoittaa tekstit vaaditussa ajassa. Näin ollen esimerkiksi henkilöiden ikä tai opiskeluvuosi eivät olleet rajoittavia tekijöitä.

Opiskelijat kirjoittivat yhden tekstin ranskaksi ja yhden ruotsiksi, ja yhden tekstin kirjoittamiseen oli aikaa 30 minuuttia. Heille ei ollut annettu tiettyä sanamäärävaatimusta, vaan he kirjoittivat niin paljon, kuin tuossa ajassa ehtivät, ilman sanakirjoja tai muita apukeinoja. Ranskankielisen tekstin aiheena oli vaihto-opiskelun hyvät ja huonot puolet, ja ruotsinkielisen tekstin aiheena oli opintojen ohessa työskentelyn hyvät ja huonot puolet. Tekstit kirjoitettiin eri päivinä, ja jokainen kirjoitti tekstiään yksin tälle varatussa huoneessa. Kirjoitusajan jälkeen tutkija palasi huoneeseen, ja opiskelija verbalisoi suomeksi kirjoitusprosessiaan ScriptLogin tuottamaa videotallennetta seuraten. Videotallenne sisälsi kaiken, mitä opiskelija oli tehnyt tietokoneella kirjoituksen aikana, jolloin verbalisointi kesti myös 30 minuuttia. ScriptLog-näppäilyntallennusohjelman tarkoituksena onkin tarjota kirjoitetun lopputuotoksen lisäksi videotallenne kirjoitusprosessista, kirjoitettu teksti lineaarisessa muodossa prosessin tutkimista varten sekä tilastotietoja kirjoituksesta (mm. Cislaru & Olive 2018a; Johansson 2017; Mutta 2007).

Sanaryppäiden tunnistaminen puhutusta tai kirjoitetusta kielestä koetaan usein ongelmalliseksi, koska päätös sanaryppään identifioimisesta on aina subjektiivinen (mm. Wray 2009; Gheitasi 2017). Vieraan kielen edistyneiden opiskelijoiden kohdalla sanaryppäiden tunnistaminen voi olla vielä vaikeampaa, koska he pystyvät kielitaitonsa ansiosta muodostamaan sujuvasti erilaisia kieliopillisia rakenteita (Myles & Cordier 2017, 14). Näin ollen sanaryppäiden määritelmän hyödyntäminen korostuu niiden identifioimisessa (Wray 2009, 28), ja tätä menetelmää käytettiin myös tässä

tutkimuksessa. Lisäksi sanaryppäiden löytämisessä hyödynnettiin ScriptLogin tuottamaa lineaarista tekstiä, josta pystyi paikantamaan tauot. Tässä tutkimuksessa tauon vähimmäispituudeksi asetettiin kaksi sekuntia, sillä se on aika, jota käytetään usein alan tutkimuksissa, joissa on tarkoituksena tunnistaa yhtenäisesti tuotetut sanajonot (Cislaru & Olive 2018a, 43-44). Yli kahden sekunnin mittaisten taukojen paikantaminen on tärkeää, sillä kuten mainittu, sanaryppäät ovat yhtenäisiä kokonaisuuksia, joten niiden sisällä ei pitäisi olla taukoja.

#### **4. Tulokset**

Ensimmäinen tutkimuskysymys käsitteli sanaryppäiden osuutta kirjallisissa tuotoksissa. Kuten on aiemmin mainittu, tutkielmassa tarkasteltiin erikseen prosessin ja lopputuotoksen tasoa sekä erikseen ranskan- ja ruotsinkielisiä tuotoksia. Sanaryppäiden osuus saatiin laskemalla sanaryppäisiin kuuluvien sanojen määrä kaikkien tuotettujen sanojen määrästä. Toisin sanoen laskettiin erikseen prosessin sanaryppäiden sisältämä sanamäärä prosessin kokonaissanamäärästä ja lopputuotoksen sanaryppäiden sanamäärä lopputuotoksen kokonaissanamäärästä.

Ranskankielisissä tuotoksissa opiskelijat tuottivat prosessissa yhteensä 1588 sanaa, joista sanaryppäisiin kuuluvia sanoja oli 462 kappaletta, ja lopputuotoksessa 1315 sanaa, joista sanaryppäiden sanoja oli 393. Näin ollen sanaryppäiden prosentuaalinen osuus oli 29 % kirjoitusprosessissa ja 30 % lopputuotoksessa. Sanamäärien lasku prosessista lopputuotokseen oli odotettavaa, koska kirjoittajat tekevät muokkauksia teksteihin ja esimerkiksi poistavat kokonaan sanoja ja lauseita, joita ovat jo kirjoittaneet. Sanaryppäiden osuuden korkeampi prosentti lopputuotoksessa kertoo taas siitä, että sanaryppäiden sanamäärä on pienentynyt vähemmän kuin kokonaissanamäärä. Ruotsinkielisissä tuotoksissa opiskelijat olivat tuottaneet sanoja melko samat määrät kuin ranskassa, mutta sanaryppäiden sanamäärät olivat pienempiä. Prosessissa oli tuotettu yhteensä 1519 sanaa, joista 346 sanaa kuului sanaryppäisiin, ja lopputuotoksessa oli tuotettu 1350 sanaa, joista 321 kuului sanaryppäisiin. Sanaryppäiden prosentuaaliset osuudet olivat huomattavasti alhaisemmat kuin ranskassa, eli 23 % prosessissa ja 24 % lopputuotoksessa. Tulosten analyysi osoitti, että kielten välinen suuri ero saattaa osaltaan johtua kielten ortografisista eroista, jotka vaikuttavat sanamääriin (esimerkiksi heittomerkin käyttö ja määräisyyden ilmaiseminen), mutta toisaalta ero voi johtua siitä, että opiskelijoiden kielitaidot ovat eritasoisia näissä kielissä. Tutkimushypoteesi tämän kysymyksen kohdalla varmistui

osittain, sillä sanaryppäiden osuudet olivat lähellä 25 %, mutta kielten välinen ero oli kuitenkin suuri.

Toinen tutkimuskysymys käsitteli yleisellä tasolla opiskelijoiden käyttämien sanaryppäiden kategorioita. Sanaryppäät luokiteltiin omiin kategorioihinsa tutkimuksessa käytetyn kategorisointitavan mukaan, ja apuna käytettiin aikaisempien tutkimusten tarjoamia esimerkkejä ryppäiden luokista. Kuten on jo aiemmin ilmaistu, tässä tutkielmassa muodostettiin myös kaksi uutta kategoriaa aineiston pohjalta: individuaaliset ja modifioidut sanaryppäät. Ranskankielisissä tuotoksissa selkeästi käytetyimmät kategoriat olivat leksikaaliset sanaryppäät (38 % prosessissa ja 39 % lopputuotoksessa) ja diskursiiviset sanaryppäät (vastaavasti 37 % ja 35 %). Näiden perässä tulivat melko samankokoiset kategoriat, joiden osuudet olivat alle 10 %: kieliopilliset ja temaattiset sanaryppäät sekä välikielen sanaryppäät. Odotettavasti pienimmät kategoriat olivat uudet kategoriat, individuaaliset ja modifioidut sanaryppäät, joiden osuudet olivat alle 5 %. Ruotsinkielisissä tuotoksissa opiskelijat olivat tuottaneet huomattavasti enemmän leksikaalisia sanaryppäitä (44 % prosessissa ja 47 % lopputuotoksessa) kuin diskursiivisia (vastaavasti 32 % ja 28 %). Nämä kategoriat olivat kuitenkin selvästi isoimpia kategorioita, kuten olivat ranskassakin. Kieliopillisten sanaryppäiden ja välikielen sanaryppäiden prosenttiosuudet olivat ruotsissa noin 10 %, kun taas temaattisten, individuaalisten ja modifioitujen sanaryppäiden osuudet olivat alle 5 %. Kielten väliltä löytyi mielenkiintoisia eroavaisuuksia, ja etenkin leksikaalisten sanaryppäiden osuus oli huomattavasti suurempi ruotsissa kuin ranskassa. Toisaalta diskursiivisten sanaryppäiden osuus oli isompi ranskassa kuin ruotsissa. Nämä voivat viitata siihen, että esimerkiksi diskursiivisten sanaryppäiden omaksuminen on opiskelijoille haastavampaa ruotsissa kuin ranskassa. Lisäksi oli yllättävää, että välikielen sanaryppäitä esiintyi opiskelijoilla molemmissa kielissä melko paljon. Toista tutkimuskysymystä vastaava hypoteesi varmistui, sillä leksikaalisia, diskursiivisia ja kieliopillisiä sanaryppäitä esiintyi molemmissa kielissä, ja toisaalta kielten välillä oli suuriakin eroja.

Kolmas tutkimuskysymys koski opiskelijoiden yksilöllisiä tapoja käyttää sanaryppäitä. Tässä tarkasteltiin opiskelijakohtaisesti erityisesti kirjoitusprosessin ja lopputuotoksen välistä suhdetta ja kielten välisiä eroavaisuuksia. Prosessin ja lopputuotoksen avulla pystyttiin tunnistamaan opiskelijoille ominaiset profiilit sanaryppäiden käytössä. Opiskelija A tuotti paljon sanaryppäitä molemmissa kielissä, mutta hän myös muokkasi

tekstiä paljon ja poisti jo kirjoitettuja sanaryypäitä. Hänen käyttämissään sanaryypäiden kategorioissa näkyi hieman vaihtelua kielten välillä, mutta toisaalta hänen tuotoksissaan toistuivat tietyt alakategoriat molemmissa kielissä. Esimerkiksi mielipiteen ilmaisevia sanaryypäitä hän käytti paljon sekä ranskassa että ruotsissa. Kaiken kaikkiaan hänen tuotoksien perusteella oli selvää, että hänellä on paljon automatisoituneita sanaryypäitä käytössään. Opiskelija B tuotti myös runsaasti sanaryypäitä molemmilla kielillä, ja hänen käyttämiensä kategorioiden osuudet olivat lähes samanlaisia molemmissa kielissä. Poikkeuksena olivat välikielen sanaryypäät, joita opiskelija oli tuottanut huomattavasti enemmän ranskassa. Muutoin hän olin käyttänyt sekä ranskassa että ruotsissa toistuvasti tiettyjä alakategorioita, esimerkiksi tekstirakenteellisia sanaryypäitä (*marqueurs textuels*). Opiskelija C:n sanaryypäiden käyttöä kuvaa se, että hänellä oli paljon pitkiä taukoja kirjoitusprosessissaan. Tämä vaikutti tuotetuiden sanaryypäiden määrään, mutta toisaalta pitkät tauot osoittivat selkeästi, mitkä sanat hän oli tuottanut sanaryypäinä. Opiskelija C:n käyttämässä kategorioissa oli paljon vaihtelua kielten välillä, mutta erikoisuutena oli se, että leksikaaliset sanaryypäät olivat huomattavasti käytetyimpiä molemmissa kielissä. Opiskelija D:lle oli ominaista, että hän käytti monipuolisesti kaikkia kategorioita molemmissa kielissä, vaikka kategorioiden osuudet olivatkin erilaisia kielten välillä. Lisäksi opiskelija D tuotti runsaasti sanaryypäitä, vaikka hänen tekstinsä eivät olleet kokonaissanamääriltään kovin suuria. Kaiken kaikkiaan kolmannen tutkimuskysymyksen hypoteesi varmistui, sillä yksilölliset käyttäjäprofiilit pystyttiin tunnistamaan ja opiskelijoilla oli yksilökohtaisia eroja kielten välillä.

## **5. Johtopäätökset**

Tämän tutkielman tulokset ovat osoittaneet, että vieraiden kielten opiskelijoiden kirjallisissa tuotoksissa sanaryypäiden osuus voi olla melko korkea, jopa 30 % tuotetusta tekstistä. On kuitenkin mielenkiintoista, että sanaryypäiden käyttö voi vaihdella kielen mukaan, kun tarkastellaan niiden käyttöä yksilötasolla. Näyttää siltä, että kielitaidon taso vaikuttaa sanaryypäiden käyttämiseen, ja tämä näkyy erityisesti välikielen sanaryypäiden käytössä. Tämä tutkielma on osoittanut, että kirjoitusprosessin tutkiminen on tärkeää, kun tutkitaan sanaryypäitä, sillä se antaa erilaista tietoa sanaryypäiden käytöstä kuin pelkkä lopputuotos. Ilman prosessin tutkimista opiskelijoiden sanaryypäiden käyttö olisi vaikuttanut vähäisemmältä, kuin mitä se todellisuudessa oli. Lisäksi kirjoitusprosessin tutkiminen on välttämätöntä, kun halutaan

nähdä, miten automatisoituneet sanaryypäät vaikuttavat tekstin tuottamiseen. Yleisesti ottaen voidaan sanoa, että kaikilla tämän tutkimuksen opiskelijoilla oli automatisoituneita sanaryypäitä käytössään, ja niiden tuoma hyöty pystyttiin osoittamaan kirjoitusprosessien analyyseissa. On huomioitava, että tämän tutkimuksen osallistujamäärä oli pieni, joten näitä tuloksia ei voida yleistää koskemaan kaikkia vieraan kielen käyttäjiä. Voidaan kuitenkin todeta, että sanaryypäiden tutkiminen vieraalla kielellä kirjoitetuissa tuotoksissa on aihe, jonka tutkimusta kannattaa laajentaa. Erityisesti yksilötasolla sanaryypäiden käytön tutkiminen sekä puhutussa että kirjoitetussa kielessä voisi tarjota uudenlaista tietoa sanaryypäiden kokonaisvaltaisesta käytöstä.