

Université Paris 8 – Vincennes – Saint-Denis  
Ecole doctorale Cognition, Langage, Interaction  
U.F.R. Sciences du langage

## **THESE**

Pour obtenir le grade de Docteur en Sciences du Langage  
Discipline : Linguistique Générale

Présentée et soutenue publiquement  
par

**Annie Perez-Bettan**

Le 09 décembre 2015

# Apprentissage et utilisation du langage préfabriqué chez des apprenants de français langue étrangère

---

Volume 1

Sous la direction de

**Heather HILTON**  
**Olga THEOPHANOUS**

Composition du jury

Michel Billières	professeur, Toulouse-Jean-Jaurès
Amanda Edmonds	Maître de Conférences, Pays de l'Adour-Pau
Fanny Forsberg-Lundell	Associate Professor, Docent-Stockholm
Heather Hilton	Professeur, Lyon 2 Lumière
Olga Théophanous	Maître de Conférences, Toulouse-Jean-Jaurès
Marzena Watorek	Professeur, Paris 8-Vincennes-Saint-Denis



*A Serge, trop tôt disparu,  
A la mémoire de mes parents,*



## Remerciements

C'est après un long moment d'interrogations et de doutes sur la nécessité de cette thèse, entreprise bien tardivement, que je décidai enfin de la mener à terme. Si j'ai réussi à passer ce cap, c'est grâce d'abord et surtout à mes deux directrices de thèse, à leur confiance, leurs encouragements, leurs conseils, leur présence indéfectible et à leur patience.

Je remercie sincèrement Heather Hilton, dont j'admire la culture, le charisme et la vitalité, qui m'a consacré du temps (des jours entiers) pour me transmettre ses connaissances dans le domaine de la production de parole et des statistiques.

Un grand merci à Olga Théophanous, qui m'a initiée au domaine de la préfabrication dans le langage et avec qui j'ai partagé de beaux moments à Toulouse et ailleurs. C'est toi, Olga, qui m'as donné le goût de la recherche, et bien souvent encouragée quand les doutes m'assaillaient.

Je remercie également mes étudiants de l'Ecole des mines d'Albi et les étudiants du DEFLE à Toulouse qui ont accepté de se prêter à cette étude, ainsi que les sujets natifs qui n'ont pas hésité à me donner de leur temps.

Je dois aussi remercier mes camarades d'étude, Heglyn, Samantha et Sylvia qui m'ont acceptée et aidée par leur accueil, leur présence et leur bonne humeur et, Corinne qui m'a réconfortée bien souvent.

Mes remerciements vont aussi à mes amis les plus proches qui ne comprenaient pas toujours très bien pourquoi je m'étais engagée dans ce parcours mais qui n'ont jamais failli.

Et je n'oublie surtout pas ma famille, Bruno et Déborah, qui m'ont « supportée » pendant toutes ces années, malgré le manque de disponibilité de leur maman.

Enfin, cette thèse n'aurait pas pu être menée au bout sans l'appui de mon mari. Henri, tu m'as poussée sans relâche à poursuivre dans cette voie parce que tu as toujours su que j'y arriverai.

## **Avant-propos**

C'est à l'occasion d'un cours de « grammaire » enseigné en Master, que je fis la rencontre du langage préfabriqué (déjà pressenti lors de mes déboires en langue anglaise). En effet, j'eus la « révélation ». Je comprenais enfin ce qui m'avait manqué dans les cours de langue et pourquoi il m'était si difficile de parler avec aisance en langue étrangère. Construire des phrases mot à mot en utilisant la syntaxe ne me permettait pas toujours de parler juste et d'une manière aisée et fluide. Et j'avais toujours l'impression que cela aurait pu être dit autrement, « à la manière » des natifs. Ayant appris la langue à une époque où l'écrit semblait plus important que l'oral et où l'apprentissage se faisait encore selon la vieille méthode de la grammaire-traduction, il me manquait indubitablement des expressions courantes, des formules, facilement accessibles en temps limité me permettant de poursuivre une conversation avec aisance et confiance. Je décidai donc d'appliquer une nouvelle façon d'enseigner les cours de langue que je donnais, en évitant les listes de vocabulaire et les leçons de grammaire coupées du contexte. Passionnée par cette nouvelle façon d'appréhender le langage, je me proposai d'approfondir cette question lors de mes travaux de recherche, d'abord en Master, puis en Doctorat.

**Objects once experienced together tend to become associated in the imagination, so that when any one of them is thought of, the others are likely to be thought of also, in the same order of sequence or coexistence as before.**

James, 1890 : 561

**Le propre de la parole c'est la liberté des combinaisons. Il faut donc se demander si tous les syntagmes sont également libres.**

Saussure, 1913 : 172

**You shall know the word by the company it keeps.**

Firth, 1957

## Résumé

Cette recherche s'intéresse au langage préfabriqué (LP) dans le discours oral spontané d'apprenants de français langue étrangère (FLE). Le LP comprend des unités polylexicales qui sont souvent perçues et utilisées comme des ensembles inanalysables. En tant que phénomène langagier majeur, le LP est étudié en linguistique sous l'angle du figement (expressions idiomatiques) et de la conventionalité (pragmatèmes, routines discursives). Le LP intéresse également les psycholinguistes pour son apport dans l'acquisition du langage, le caractère d'idiomaticité et de « naturel » qu'il confère à l'utilisateur de la langue et son aide dans le traitement de la parole. Des études montrent que le LP est omniprésent dans les manifestations langagières des locuteurs natifs (Pawley et Syder, 1983 ; Erman et Warren, 1992; Wray et Perkins, 2000), sous-entendant ainsi que la maîtrise d'une langue ne passe pas seulement par l'application des règles grammaticales et la création des phrases nouvelles et originales mais également par la mémorisation et l'automatisation des bouts langagiers préconstruits (chunks). Des études en anglais langue étrangère montrent que le LP fait souvent défaut et ce, même à des niveaux avancés. A la suite de quelques rares études portant sur le LP en FLE (Forsberg, 2006 ; Cordier, 2013), nous avons entrepris cette étude avec un double objectif. D'une part, étudier l'évolution de la connaissance du LP chez des apprenants en cours d'acquisition du FLE en examinant la présence et la nature de celui-ci et, d'autre part, examiner le rôle que le LP joue dans la fluidité verbale des apprenants (Wood, 2006). Dix-sept apprenants, répartis en trois groupes de niveaux, ont réalisé à six mois d'intervalle les mêmes deux tâches narratives orales, à partir de deux supports (cartoon et film muet). Leurs narrations ont été enregistrées et transcrites en temps 1 et temps 2. Les six locuteurs natifs du groupe contrôle ont été soumis aux mêmes tâches une seule fois afin de comparer leurs performances à celles du groupe expérimental. Le LP qui a surgi dans les productions orales a été identifié selon des critères linguistiques et psycholinguistiques tels que la cohérence phonologique, la complexité syntaxique, l'opacité sémantique, la fréquence et la systématisme de l'item déviant ou idiosyncrasique. La fluence des productions a été évaluée au moyen de quatre mesures « classiques » comme la vitesse de parole, la longueur du segment de parole, la vitesse d'articulation et le temps de phonation. Elle a été ensuite mise en relation avec l'emploi effectif des chunks ou séquences préfabriquées (SP) afin de vérifier en



quel sens ces dernières améliorent le discours oral des apprenants. Nos analyses montrent qu'il existe des différences individuelles très marquées entre les sujets testés et qu'il est difficile de tirer des conclusions d'ordre général. Au vu des résultats, nous pouvons cependant affirmer que les apprenants de niveau débutant progressent dans l'emploi qu'ils font du LP, en quantité mais aussi en qualité. Les apprenants peuvent, en effet, s'appuyer sur les SP pour produire un discours plus fluide, de différentes façons (Wood, 2006). D'une manière générale, les SP améliorent le segment de parole en l'allongeant et en le complexifiant. Elles servent aussi comme stratégie de compensation mais aussi de structuration et de planification du discours.

**Mots clés :**

Langage préfabriqué, séquences préfabriquées (chunks), acquisition, oral, tâche narrative, fluence (fluidité, aisance), français langue étrangère, différences individuelles, traitement langagier

## **Abstract**

### **Learning and using formulaic language by French as a Second Language learners**

This thesis focuses on formulaic language in French as a Second Language (FSL) learners' spontaneous spoken narratives. Formulaic language is made of multi-word units which are often perceived and used as simple lexical units. ESL studies have shown that even at advanced levels the command of formulaic language can be poor. In the wake of a few studies of formulaic language in FSL (Forsberg, 2006 ; Cordier, 2013), the purpose of this thesis is twofold. First we intend to examine the evolution of formulaic language among FSL learners, and then the importance of formulaic language in learners' fluency (Wood, 2006). Seventeen learners divided into three levels were asked over a period of six months to produce two oral narratives based on a cartoon and a silent film. Their narratives were recorded and transcribed as time 1 and time 2. The formulaic language which was found in the oral productions was identified according to linguistic and psycholinguistic criteria such as phonological coherence, syntactic complexity, semantic opacity, frequency and systematicity of idiosyncratic errors and deviant items. The fluency was evaluated by means of four "classical" measurements : speech rate, length of run, rate of articulation, and phonation time. It was then related to their use of formulaic sequences (FS) in order to verify how the latter improve learners' oral production. Our analyses show that there are clear individual differences among subjects and that it is difficult to draw comprehensive conclusions. Yet, the results show a steady improvement in formulaic language among beginners, both quantitatively and qualitatively. Indeed, learners can rely on formulaic sequences to gain fluency and this in different manners (see also Wood, 2006). More generally formulaic sequences improve the speech run by making it longer and more complex. They can also be used as strategies of compensation and planification of speech.

#### **Key words :**

Formulaic language, formulaic sequences, acquisition, oral, narrative task, fluency, French as a Second Language, individual differences, language processing

# Table des matières

---

<i>Table des matières</i> .....	11
<i>Liste des abréviations</i> .....	15
<i>Liste des figures</i> .....	16
<i>Liste des tableaux</i> .....	17
<b>Introduction générale</b> .....	<b>19</b>
<b>I. Partie théorique</b> .....	<b>28</b>
<b>Chapitre 1 Évolution historique et définitions</b> .....	<b>29</b>
<b>1.1 Aperçus historiques</b> .....	<b>29</b>
1.1.1 Les précurseurs.....	29
1.1.2 Deux positions extrêmes : la grammaire générative et la grammaire de construction.....	31
1.1.3 L'école de Sinclair.....	32
<b>1.2 Universalité et ampleur du phénomène</b> .....	<b>33</b>
1.2.1 Présence dans le discours oral et écrit.....	33
1.2.2 La préfabrication : un vaste choix de termes.....	35
<b>1.3 La phraséologie en linguistique française</b> .....	<b>36</b>
1.3.1 Définitions de la phraséologie.....	37
1.3.2 Le figement.....	38
<b>1.4 Les catégories formelles</b> .....	<b>39</b>
1.4.1 Les collocations.....	39
1.4.2 Les expressions idiomatiques.....	45
1.4.3 Les proverbes.....	47
1.4.4 Les expressions conventionnelles.....	47
1.4.5 Les formules discursives.....	49
1.4.6 Les constructions.....	50
<b>1.5 Choix du terme « préfabriqué »</b> .....	<b>52</b>
<b>Chapitre 2 Traitement, Acquisition et Identification des séquences préfabriquées</b> .....	<b>55</b>
<b>2.1 Rôle des SP dans le traitement langagier</b> .....	<b>55</b>
<b>2.2 LP et approches basées sur l'usage</b> .....	<b>58</b>
<b>2.3. La fréquence et le processus de chunking</b> .....	<b>60</b>
<b>2.4 L'acquisition des SP en L1 et en L2</b> .....	<b>63</b>
2.4.1 L'acquisition des SP en L1.....	63
2.4.2 Acquisition des SP en L2.....	66
2.4.3 Influence de la L1 dans l'acquisition en L2.....	69
2.4.4 SP et apprentissage.....	69
<b>2.5 Définitions du LP</b> .....	<b>70</b>
<b>2.6 Identification des séquences préfabriquées</b> .....	<b>73</b>
<b>Chapitre 3 Rôle des SP sur la fluence, les SP dans les modèles de production</b> .....	<b>81</b>
<b>3.1 Définitions de la fluence</b> .....	<b>81</b>
<b>3.2 Le modèle de production en L1 et le « schéma du locuteur » (blueprint) de Levelt (1989, 1999)</b> .....	<b>84</b>
3.2.1 Les différentes étapes de la production de parole.....	85

3.2.2	Place des SP dans le lexique mental .....	89
a)	Les séquences préfabriquées dans le modèle de Levelt .....	89
b)	le lexique hétéromorphique .....	90
<b>3.3</b>	<b>Un modèle de production en L2 .....</b>	<b>91</b>
3.3.1	Les caractéristiques du modèle de production bilingue .....	91
3.3.2	Procédures d'encodage et de stockage des connaissances en production orale L2 .....	93
3.3.3	Développement de la compétence L2 dans le modèle bilingue .....	95
<b>3.4</b>	<b>Rôle des SP en production orale .....</b>	<b>96</b>
<b>3.5</b>	<b>Fluence et théories de l'automatisme .....</b>	<b>99</b>
3.5.2	Relation entre automatisme et modèles de production .....	99
3.5.3	Modèles basés sur les règles .....	99
3.5.4	Modèles de l'instance .....	102
<b>Chapitre 4</b>	<b>Mesures de la fluence et séquences préfabriquées .....</b>	<b>105</b>
<b>4.1</b>	<b>Les mesures de la fluence .....</b>	<b>105</b>
<b>4.2</b>	<b>Les mesures quantitatives .....</b>	<b>106</b>
4.2.1	La vitesse de parole .....	107
4.2.2	La vitesse d'articulation .....	107
4.2.3	Taux de phonation (phonation time ratio) .....	108
4.2.4	Les pauses, ou variables temporelles secondaires .....	108
4.2.5	Localisation et distribution des pauses silencieuses .....	111
<b>4.3</b>	<b>Les mesures qualitatives .....</b>	<b>112</b>
4.3.1	Longueur du segment de parole .....	112
4.3.2	Les reprises ou corrections .....	113
<b>4.4</b>	<b>Différences individuelles dans l'aisance .....</b>	<b>116</b>
<b>4.5</b>	<b>Rôle d'un séjour à l'étranger sur la fluence .....</b>	<b>118</b>
<b>Chapitre 5</b>	<b>Méthodologie .....</b>	<b>119</b>
<b>5.1</b>	<b>Choix de la méthode .....</b>	<b>120</b>
5.1.1	Description des deux dispositifs : .....	121
<b>5.2</b>	<b>Sujets testés .....</b>	<b>122</b>
5.2.1	Groupe expérimental .....	122
5.2.2	Groupe contrôle .....	126
<b>5.3</b>	<b>Administration du dispositif et collecte des données .....</b>	<b>126</b>
<b>5.4</b>	<b>Transcriptions des données et annotations des disfluences .....</b>	<b>127</b>
5.4.1	Les annotations .....	128
5.4.2	Choix du logiciel .....	131
5.4.3	Mesures et analyses .....	132
<b>5.5</b>	<b>Identification des SP .....</b>	<b>132</b>
5.5.1	Première étape .....	132
5.5.2	Nos critères d'identification .....	137
<b>5.6</b>	<b>Classement des séquences préfabriquées .....</b>	<b>139</b>
5.6.1	Les SP lexicales .....	142
5.6.2	Les SP discursives .....	143
5.6.3	Les SP grammaticales et les constructions .....	144
5.6.4	Les SP déviantes .....	145
<b>5.7</b>	<b>Notre typologie .....</b>	<b>146</b>
5.7.1	Les SP lexicales .....	146
5.7.2	Les SP discursives .....	147
5.7.3	SP grammaticales .....	148
5.7.4	Les SP déviantes .....	150
<b>5.8</b>	<b>Mesures de la préfabrication .....</b>	<b>152</b>
<b>5.9</b>	<b>Mesures de la fluence .....</b>	<b>155</b>

<b>5.10 Limites de la méthodologie</b> .....	<b>157</b>
<b>II. Présentation et analyse des résultats</b> .....	<b>162</b>
<b>Chapitre 6 Présentation des résultats Comparaison groupe natif et groupe non-natif</b> .....	<b>164</b>
<b>6.1. Valeurs brutes du groupe natif</b> .....	<b>164</b>
<b>6.2 Résultats des sujets natifs</b> .....	<b>168</b>
6.2.1 Mesures de la préfabrication .....	168
6.2.2 Mesures de la fluence.....	172
6.2.3 Corrélation entre la fluence et l'emploi du LP .....	176
<b>6.3 Comparaison entre le groupe natif et le groupe non-natif</b> .....	<b>177</b>
6.3.1 Mesures de productivité de la préfabrication .....	177
6.3.2 Mesures de productivité de la fluence des locuteurs natifs et non-natifs .....	180
<b>6.4. Résultats des tests statistiques</b> .....	<b>182</b>
6.4.1 Mesures de la préfabrication.....	182
6.4.2 Résultats statistiques des mesures temporelles de productivité.....	184
<b>6.5. Corrélations entre fluence et préfabrication</b> .....	<b>185</b>
6.5.1 Groupe natifs et groupe non-natif entier (t2).....	186
6.5.2 Groupe des locuteurs non-natifs, t1 et t2 .....	187
<b>Chapitre 7 Présentation des résultats du groupe non-natif</b> .....	<b>190</b>
<b>7.1. Résultats du groupe des non-natifs, préfabrication</b> .....	<b>190</b>
7.1.1 Répartition des SP, groupe entier et sous-groupes, t1 et t2 .....	190
7.1.2 Résultats pour les mesures de la préfabrication, groupe entier, t1 et t2 .....	198
7.1.3 Les quatre premières mesures de la préfabrication (Cordier) .....	199
7.1.4 Mesures de la préfabrication : taux de séquences préfabriquées pour cent mots .....	207
<b>7.2 Développement de la fluence</b> .....	<b>209</b>
7.2.1 Données brutes, nombre de mots produits bruts et nombre de mots nettoyés .....	209
7.2.2 Mesures de la fluence : résultats du groupe entier.....	215
7.2.3 Les variables temporelles .....	216
7.2.4 Comparaison avec les études de Towell et collègues (1996) et Cordier (2013) .....	220
7.2.5 Interprétation de l'augmentation de la fluence en relation avec le modèle de Levelt et les théories de l'automatisme.....	222
<b>Chapitre 8 Analyse qualitative</b> .....	<b>225</b>
<b>8.1 Des SP discursives en grand nombre</b> .....	<b>225</b>
<b>8.2. SP et fluence verbale</b> .....	<b>227</b>
8.2.1 Répétition de la même SP dans un même segment.....	228
8.2.2 Répétition de la même SP dans plusieurs segments.....	228
8.2.3 Succession de plusieurs SP dans le même segment.....	229
8.2.4 Amélioration de la fluence grâce à une utilisation de type méta-discursif des SP .....	231
8.2.5 Répétition du dernier mot de la SP .....	231
8.2.6 Les autocorrections .....	233
8.2.7 La réduction phonologique .....	234
8.2.8 Les marqueurs discursifs .....	234
<b>8.3 SP et structuration du discours</b> .....	<b>239</b>
<b>8.4 Quelques formes utilisées comme aide au traitement / SP ou pas ?</b> .....	<b>242</b>
<b>8.5 Les séquences interlangagières ou déviantes</b> .....	<b>244</b>
8.5.1 Séquences interlangagières lexicales .....	244
8.5.2 Évolution des SP interlangagières.....	247
<b>8.6 Complexification de la production</b> .....	<b>249</b>
<b>8.7 Comparaison corpus natif et non-natif</b> .....	<b>255</b>

<b>Chapitre 9 Étude de deux SP, « c'est » et « petit »</b> .....	<b>259</b>
<b>9.1. Rôle de <i>c'est</i> dans le discours oral</b> .....	<b>259</b>
9.1.1 <i>C'est</i> en production native .....	261
9.1.2 <i>C'est</i> dans les productions non natives.....	264
<b>9.2 Les séquences construites avec l'adjectif « petit »</b> .....	<b>273</b>
<b>Conclusion générale</b> .....	<b>285</b>
<b>Bibliographie</b> .....	<b>298</b>

# Liste des abréviations

---

L1 : langue maternelle

LM : langue maternelle

L2 : langue cible, langue étrangère

LP : langage préfabriqué

SP : séquence préfabriquée

LN : locuteurs natifs

LNN : locuteurs non-natifs

FLE : Français Langue Etrangère

MLT : mémoire à long terme

MdT : mémoire de travail

VP : vitesse de parole

VA : vitesse d'articulation

TdP : taux de phonation

SPS : syllabe par seconde

LMP : longueur moyenne des pauses

NPM : nombre de pauses par minute

SP lex : séquence préfabriquée lexicale

SP dis : séquence préfabriquée discursive

SP gram : séquence préfabriquée grammaticale

DEFLE : Département d'études du français langue étrangère

t1 : temps 1

t2 : temps 2

Nb : nombre

Bd, BD : bande dessinée

GH : groupe d'hésitation

SN : syntagme nominal

SV syntagme verbal

GN : groupe nominal

CECRL : Cadre Européen Commun de Référence pour les Langues

# Liste des figures

---

Figure 1. Schéma des processus de production orale (A Blueprint of the speaker, Levelt, 1999 : 97, cité par Hilton, 2009 : 80) .....	86
Figure 2. Distribution des SP .....	167
Figure 3. Distribution des SP groupe non-natif.....	178
Figure 4. Distribution des SP, corpus non-natif, t1 .....	191
Figure 5. Distribution des SP, corpus non-natif, t2 .....	191
Figure 6. Répartition des SP, par sous-groupe et par sujet, t2 .....	197
Figure 7. Nombre de syllabes produites par sous-groupe et par sujet, t1 et t2.....	199
Figure 8. Taux de syllabes préfabriquées par sous-groupe et par sujet, t1 et t2 .....	203
Figure 9. Taux de syllabes préfabriquées par sous-groupes, t1 et t2 .....	205
Figure 10. Taux de segments avec SP, t1 et t2 .....	205
Figure 11. Nombre moyen de syllabes préfabriquées par segment de parole.....	206
Figure 12. Nombre moyen de syllabes préfabriquées par segment avec SP .....	206
Figure 13. Nombre de mots produits par sous-groupes et par sujet, t1 et t2.....	210
Figure 14. Nombre de mots produits nettoyés, t1 et t2 .....	213
Figure 15. Nombre de mots produits bruts et nettoyés, t1 .....	214
Figure 16. Nombre de mots produits bruts et nettoyés, t2 .....	214



# Liste des tableaux

---

Tableau 1. Sujets non-natifs.....	125
Tableau 2. Sujets natifs.....	126
Tableau 3. Séquences lexicales.....	143
Tableau 4. Type d'erreurs.....	151
Tableau 5. Mesures des variables temporelles.....	156
Tableau 6. Nombre de mots par tâche, natifs et non-natifs.....	164
Tableau 7. Données brutes, locuteurs natifs.....	167
Tableau 8. Mesures de la préfabrication.....	170
Tableau 9. Mesures de productivité.....	174
Tableau 10. Mesures de productivité et préfabrication, groupe natif et non-natif.....	177
Tableau 11. Répartition des SP, groupe natif et non-natif (n=17).....	178
Tableau 12. Mesures productivité et préfabrication, natifs et sous-groupes non-natifs.....	179
Tableau 13. Mesures temporelles de productivité, groupe natif et non-natif.....	180
Tableau 14. Mesures temporelles de productivité, natifs et sous-groupes non-natifs.....	181
Tableau 15. Comparaisons Mann-Whitney, mesures de préfabrication.....	182
Tableau 16. comparaisons Mann-Whitney, groupe natif et groupe avancé non-natif.....	183
Tableau 17. Résultats des tests statistiques des mesures temporelles de productivité.....	184
Tableau 18. Résultats des tests statistiques des mesures temporelles de productivité.....	185
Tableau 19. Répartition des types de SP groupe natif entier.....	190
Tableau 20. Nombre de SP par types et par sous-groupes, t1.....	192
Tableau 21. Nombre de SP par types et par sous-groupes, t2.....	192
Tableau 22. Répartition par type de SP, sous-groupe débutant.....	192
Tableau 23. Répartition par types de SP, sous-groupe intermédiaire.....	193
Tableau 24. Répartition par types de SP, sous-groupe avancé.....	193
Tableau 25. Nb et types de SP, par sujet, niv 1, t1.....	194
Tableau 26. Nb et types de SP par sujet, niv1, t1.....	194
Tableau 27. Nb et types de SP par sujet, niv 2, t1.....	195
Tableau 28. Nb et types de SP par sujet, niv2, t2.....	195
Tableau 29. Nb et types de SP par sujet, niv3, t1.....	195
Tableau 30. Nb et types de SP par sujet, niv3, t2.....	196
Tableau 31. Résultats des mesures de la préfabrication, t1 et t2.....	199
Tableau 32. Résultats mesures de la préfabrication (Cordier) par sujet et par groupe.....	201
Tableau 33. Test de Kruskal-Wallis, t1 et t2.....	202
Tableau 34. Taux de préfabrication, test de Wilcoxon.....	207

Tableau 35. Taux de SP, SP lexicales, valeurs brutes et occurrences types .....	208
Tableau 36. Nombre de mots produits par chaque sujet, t1 et t2 .....	210
Tableau 37. Nombre de mots produits nettoyés, t1 et t2 .....	212
Tableau 38. Mesures de productivité de la fluence, t1 et t2 .....	216
Tableau 39. Mesures de la fluence par groupe et par sujet.....	218
Tableau 40. Résultats du test Kruskas-Wallis.....	219
Tableau 41. Résultats du test de Wilcoxon.....	220
Tableau 42. Comparaison entre les études de Towell et al. Cordier, et la nôtre .....	220
Tableau 43. SP discursives avec et, groupe natif et non-natif, t1 et t2.....	235
Tableau 44 SP discursives, catégorie des remplisseurs .....	238
Tableau 45. Liste des SP spécifiques au groupe natif .....	256
Tableau 46. Occurrences de "petit" chez les locuteurs natifs .....	278
Tableau 47. Distribution des syntagmes avec « petit » et « p'tit » dans le cartoon.....	280
Tableau 48. Emplois de petit dans le film .....	282

# Introduction générale

---

Pendant longtemps on a considéré que, pour maîtriser une langue, il suffisait de connaître ses règles grammaticales et son lexique. Cette conception de la maîtrise linguistique met en avant la dimension « créative » de la langue, largement soutenue par les théories générativistes de Noam Chomsky (1965). Pour ce linguiste, la créativité linguistique signifie l'aptitude du sujet parlant à créer et comprendre un nombre infini d'énoncés nouveaux qu'il n'a jamais entendus auparavant et qu'on ne peut répertorier en totalité. La grammaire générative a relégué au second plan, comme des phénomènes marginaux, tout ce qui ne pouvait pas entrer dans ses principes et « paramètres » généraux et universels. Ainsi Chafe, dès 1968, constate l'« incapacité du modèle générativiste à traiter de la non-compositionnalité et de l'agrammaticalité du déficit des propriétés transformationnelles » (Chafe 1968 cité par Legallois et Gréa, 2006 : 6).

Néanmoins, la vision hyper « créativiste » des générativistes a pendant longtemps guidé autant les travaux en linguistique qu'en didactique des langues. Pour cette dernière, qu'on pense seulement à l'importance donnée à la grammaire dans les cours de langues étrangères prônée pratiquement dans toutes les méthodologies et aux nombreux exercices sur la maîtrise des règles grammaticales qu'on fait faire encore aux apprenants pour s'en convaincre !

Si l'une des théories linguistiques dominantes affirme que les unités lexicales (les mots) sont petites et les règles de grammaire plutôt complexes et que le langage serait l'utilisation de ces unités combinées selon ces règles, il est néanmoins indéniable qu'il existe des suites de mots plus longues, qui renferment des combinaisons grammaticales complexes, qui sont utilisées telles quelles sans que le locuteur ait recours à la grammaire. Ces unités peuvent être associées à d'autres unités larges dans la même phrase ainsi qu'à des unités plus petites. En langue maternelle, elles sont le plus souvent produites d'une traite sans que l'on réfléchisse à la construction grammaticale ni au vocabulaire. On pourrait même dire qu'une expression ou une construction idiomatique est quelque chose qu'un locuteur L2 pourrait ne pas connaître alors qu'il connaît tous les autres moyens de la langue<sup>1</sup> (Fillmore *et al.* 1988 : 504). Ces « expressions », qualifiées d'exception par les grammaires pour les plus idiomatiques, sont

---

<sup>1</sup> « An idiomatic expression or construction is something a language user could fail to know while knowing everything else in the language. » Fillmore *et al.* 1988 : 504

néanmoins attestées depuis fort longtemps. Ce phénomène avait été reconnu principalement par des praticiens de la langue dès le XIX<sup>ème</sup> siècle. Ainsi, l'un des pionniers, un professeur de langue, Claude Marcel, utilise de nombreux exemples écrits dans son enseignement, pour illustrer les règles de grammaire, et les compare à la langue maternelle. Il fait répéter à ses élèves des phrases-modèles qu'ils mémorisent sans passer par la traduction. « Les élèves sont assis face à un grand tableau et le professeur explique une règle ou deux de grammaire chaque jour, comparant la langue maternelle et la langue étrangère, en illustrant les nouvelles règles avec de nombreux exemples »<sup>2</sup> (Marcel, 1820 : 50–51). Un peu plus tard, à la fin du siècle, Michel Bréal, un des fondateurs de la linguistique en France, atteste le phénomène dans son *Traité de sémantique* (1897). Dans un tout autre domaine, celui de la neurologie, lors d'une atteinte cérébrale de l'hémisphère gauche, les patients atteints d'aphasie, sont alors dans l'incapacité de créer leur langage mot à mot. John Hughlings Jackson (1835-1911) constate qu'ils ont recours à un langage stéréotypé, le seul qui subsiste. Il fait ainsi la distinction entre langage propositionnel (stéréotypé) et langage analytique (mot à mot). Il est ainsi le premier à contester l'idée d'une localisation cérébrale unique pour le langage. « Il considérait que la capacité de prononcer des mots et de courts énoncés de façon automatique, relevait de la compétence de l'hémisphère droit, l'hémisphère gauche, lésé chez ces patients, étant responsable du langage propositionnel »<sup>3</sup> (Clarke, 2008 : 31). Un peu plus tard, Saussure (1916 : 243) parlera de ces suites de mots traitées comme un long mot. « [...] quand un concept est exprimé par une suite d'unités significatives très usuelle, l'esprit, prenant pour ainsi dire le chemin de traverse, renonce à l'analyse et applique le concept en bloc sur le groupe de signes qui devient alors une unité simple. ». Ces suites peuvent donc être tout simplement des ensembles familiers et courants que nous produisons fréquemment à l'oral, l'objectif n'étant pas l'originalité ou la créativité mais l'efficacité et la rapidité de traitement. Jespersen (1924/1976) mettra l'accent sur l'idiomaticité des formules (*formula*) qui « peuvent être une phrase entière ou un groupe de mots, ou même un mot ou une partie du mot – ce n'est pas ça qui est important, mais cela doit être toujours quelque chose qui instinctivement

---

<sup>2</sup> Notre traduction de : « The pupils sit facing a large blackboard, and the teacher explains a rule or two of the grammar each day, comparing the mother tongue and the foreign language, and illustrating the new rules with numerous written examples ». (Marcel 1820 : 50–51).

<sup>3</sup> Voir Nespoulous et Lecours, 1980; Nespoulous, Code, Virbel et Lecours, 1998

semble ne pas pouvoir être analysé ou décomposé selon les règles de la combinatoire libre<sup>4</sup> ». Ces « ensembles » dont parle Jespersen renvoient au figement ou à l'irrégularité syntaxique et doivent être connus des locuteurs ; ce qui implique leur mémorisation holistique. Tous ces exemples illustrent parfaitement la question du langage préfabriqué, plaçant les séquences sur un continuum allant de l'expression la plus figée à la plus libre, l'utilisation des séquences préfabriquées étant finalement un moyen aisé de s'exprimer avec un minimum d'effort. Et c'est sans doute parce que notre capacité à nous exprimer en temps réel ne nous permet pas de combiner chaque mot, en appliquant de façon concertée une règle, que nous produisons de larges suites de mots déjà prêtes à l'emploi : elles nous permettent ainsi de parler plus vite en permettant de planifier la suite de notre discours. En effet, la tâche serait complexe si nous devions à chaque fois penser et organiser nos phrases.

Si le courant générativiste (Chomsky, 1965) a longtemps prôné la créativité du langage qui permet une infinité de combinaisons, John Sinclair (1991), linguiste travaillant sur des corpus et, dans le contexte de son travail pour le dictionnaire *Collins-Cobuild* (2003b), remet ce principe en question. Il émet à côté du principe du libre choix des théories générativistes qu'il ne conteste pas, un deuxième principe tout aussi fondamental de la langue, à savoir le principe idiomatique<sup>5</sup>. C'est suivant ce principe que l'utilisateur de la langue dispose d'un vaste stock de formules préconstruites qui constituent des choix uniques même si elles peuvent apparaître analysables en segments.

La linguiste cognitive Wong-Fillmore (1976 : 24) avait elle aussi, quelques années auparavant, mis l'accent sur la pré-construction comme dimension fondamentale de la langue : « Une vaste partie d'une langue naturelle est stéréotypée, automatique et répétitive, plutôt que générative, créatrice ou librement produite. »<sup>6</sup> À la même époque, Bolinger (1975 : 2) avait fait remarquer de manière métaphorique qu'on « ne construit pas notre langage mot à mot avec des planches, un marteau, des clous et un plan, mais qu'on parle avec un nombre incroyable d'expressions toutes faites ».<sup>7</sup>

Progresser dans une langue étrangère implique donc nécessairement la maîtrise d'expressions

---

<sup>4</sup> « [it] may be a whole sentence or a group of words, or it may be one word, or it may be only a part of a word, - that is not important, but it must always be something which to the actual speech instinct is a unit which cannot be further analyzed or decomposed in the way a free combination can ». (Jespersen, 1924/1976 : 88 cité par Wray, 2002 : 7).

<sup>5</sup> « Open-choice principle et idiom principle ». (Sinclair, 1991)

<sup>6</sup> « An enormously large amount of natural language is formulaic, automatic and rehearsed, rather than propositional, creative or freely generated. » (Wong-Fillmore, 1976 : 24), notre traduction.

<sup>7</sup> « Our language does not expect us to build everything starting with lumber, nails, and blueprint, but provides us with an incredibly large number of prefabs. » (Fillmore, 1979 : 92)

préfabriquées. Ou, comme le dit Fillmore (1979 : 92), « La capacité d'une personne à utiliser efficacement une langue consiste en grande partie à maîtriser des séquences stéréotypées ». <sup>8</sup> Ce linguiste, l'un des fondateurs de la grammaire des constructions, met justement en avant la présence d'expressions exigées telles quelles dans une situation de communication donnée (*How do you do ? Comment ça va ? Toutes mes félicitations*). Une partie importante de la langue, à l'oral comme à l'écrit, consiste en des combinaisons de mots, des bouts de langage, des phrases que chaque locuteur a à sa disposition pour chaque situation et qu'il utilise sans passer par le processus de construction grammaticale (Pawley et Syder, 1983 ; Nattinger et DeCarrico, 1992). Le rôle que jouerait le langage préfabriqué dans la communication est largement reconnu par la sociolinguistique mais aussi la psycholinguistique (Miller, 2011 : 129).

Du point de vue des psycholinguistes, en production, ces suites de mots facilitent la production d'énoncés en limitant la charge cognitive du locuteur qui peut ainsi s'exprimer plus aisément ; et on comprend d'autant mieux leur importance en langue étrangère. Dès les premiers stades de l'apprentissage, ces ensembles de mots permettent aux apprenants de pouvoir entrer en communication. A un niveau plus avancé, les locuteurs arrivent à maîtriser d'autres ensembles plus complexes qu'ils peuvent modifier en les adaptant au contexte. Ces ensembles sont des combinaisons de mots que les locuteurs natifs semblent préférer à d'autres qui auraient pu aussi bien convenir. C'est pourquoi un locuteur non-natif aura plus de mal, dans un premier temps, à se les approprier et à les employer dans la bonne situation. Il est parfois difficile de les traduire littéralement à cause des conventions lexicales propres à chaque communauté linguistique. Si elles sont très éloignées de la langue maternelle de l'apprenant, à moins de les mémoriser, il aura beaucoup de mal à les activer. Par exemple, un locuteur français dira *manger un sandwich* mais *prendre un repas* et on trouvera étrange d'entendre *\*manger un repas*, qui tout en restant compréhensible sera ressenti comme légèrement étrange et non-idiomatique. De même, on dit naturellement *deux blessés graves* et on trouvera étrange l'expression *\*deux blessés sérieux*. Ainsi, malgré toutes les combinaisons possibles que nous offre la langue, en L1 nous produisons en production sans trop réfléchir celle qui nous paraît la plus appropriée grâce à un répertoire de formules faciles à utiliser et bien pratiques à l'oral. Ainsi que le disait Bréal (1897), comme l'enfant, nous avons tendance à utiliser toujours les mêmes phrases qui permettent une meilleure production et

---

<sup>8</sup> « A large portion of a person's ability to get along in a language consists in the mastery of formulaic utterances. » (Fillmore, 1979 : 92), notre traduction.

compréhension :

C'est grâce à l'analogie que l'enfant, sans apprendre l'un après l'autre tous les mots de la langue, sans être obligé de les essayer un à un, s'en rend maître dans un temps relativement court. C'est grâce à elle que nous sommes sûrs d'être entendus, sûrs d'être compris, même. (Bréal, 1897 : 233)

En effet, bien que la capacité humaine à planifier de la parole à l'avance ou pendant la production soit limitée, les locuteurs natifs arrivent à produire aisément et rapidement des suites de mots. D'où vient cette capacité chez le natif alors qu'il est si difficile pour un apprenant L2 de trouver en production spontanée les phrases justes ? Nous trouvons des théories complètement opposées sur l'utilisation de la langue. Ainsi, pour l'ex-généraliste Pinker (1994 : 22), chaque phrase produite ou comprise est une toute nouvelle combinaison de mots qui apparaît pour la première fois à l'échelle cosmique. Cette affirmation s'oppose à celle prônée par la grammaire des constructions (*Construction Grammar*) qui dit que tout dans la langue est formulaïque (Fillmore, 1979). Entre ces deux extrêmes, on peut dire que la langue « dans la vraie vie » est un subtil mélange de processus créatifs exercés sur des unités plus ou moins grandes déjà préassemblées.

Le langage préfabriqué tel qu'il est traité dans cette thèse comprend une très large palette de phénomènes nommés de différentes façons par les chercheurs, selon le point de vue adopté. On trouve dans la littérature une terminologie très riche qui renvoie au phénomène de la préfabrication ; Wray (2002 : 9) dénombre pas moins de 50 termes pour désigner ce type de langage. Des chercheurs ont nommé ce phénomène *formulaic language* en anglais ou « langage préfabriqué » en français (Wray, 2002, 2008, 2009 ; Forsberg, 2006, 2008 ; Edmonds, 2013). Ce terme « préfabriqué » pour traduire l'anglais *formulaic* est loin de faire l'unanimité chez les linguistes et chercheurs en France<sup>9</sup>, même si certains comme Forsberg (2006) l'ont utilisé pour le français. Néanmoins, nous utiliserons ce terme de langage préfabriqué (désormais LP) ainsi que celui de « séquences préfabriquées » (désormais SP) traduction de *formulaic sequences* ou *prefabs*. Wray (2002 : 9) explique que le terme *formulaic* ou préfabriqué véhicule l'idée d'unité et d'habitude, alors que *sequence* ou séquence indique qu'il s'y trouve plus d'une seule unité<sup>10</sup>.

---

<sup>9</sup> « Le terme de « langage préfabriqué » (ou préconstruit ou préformé) n'a pas l'air très scientifique, avouons-le. Pourtant, il s'avère être un chapeau très utile qui embrasse et dépasse les notions de stéréotypes et de figement. » (Confais, 2015)

<sup>10</sup> Notre traduction de : « ...we shall use formulaic sequence. The word formulaic carries with it some associations of « unity » and of « custom » and « habit », while sequence indicates that there is more than one discernible internal unit, of whatever kind. » Wray, 2002 : 9).

On assiste depuis les années 1990 à un intérêt croissant de la recherche pour la préfabrication, grâce sans doute à l'impact des travaux de la linguistique de corpus, émanant de l'école de Birmingham autour de Sinclair et Coulthard dans les années soixante (William, 2005). Cet intérêt coïncide avec l'analyse informatisée des données issues de grands corpus. Les chercheurs ont pu ainsi regarder de manière plus aisée et efficace les relations que les mots entretiennent entre eux et la fréquence de certaines occurrences. Le contexte est donc essentiel dans la langue qui est atteinte à travers la parole et qui n'a pas d'existence propre en dehors du contexte. La linguistique de corpus est née dans le monde anglo-saxon et, en France, elle commence à être appliquée dans de nombreux domaines (traitement automatique des langues, lexicologie) et dans la recherche en acquisition des langues.

Le LP dépasse donc le cadre traditionnel de la phraséologie, domaine de la linguistique, qui a décrit les expressions du point de vue du figement ; la langue contient des séquences dont les unités apparaissent plutôt ensemble et de manière récurrente et qui ne répondent pas forcément à ce critère. Nous montrerons précisément ce que la recherche en phraséologie accepte dans ses définitions, et ce que les recherches plus récentes considèrent comme préconstruit ou préfabriqué. Nous verrons également que le terme de langage préfabriqué recouvre des phénomènes linguistiques aux frontières mal ou peu définis. En effet, pour Wray, la notion de préfabrication recouvre un grand nombre de phénomènes linguistiques et non linguistiques qui peuvent sembler, pour certains, trop généraux pour être considérés comme formulaires (Wray, 2008 : 4). Ils sont cependant considérés comme tels parce qu'ils sont à la limite, aux frontières, du modèle théorique qu'elle a établi dans un précédent ouvrage (Wray, 2002). Wray englobe dans le LP non seulement les idiomes et les suites de mots qui jouent un rôle dans la communication mais aussi les phrases mémorisées, et jusqu'à des récits oraux. Les textes longs mémorisés sont composés de « *chunks* » (blocs ou items) plus petits, susceptibles de variation. On peut dire que la préfabrication va du plus petit *item* au script le plus large.

Dans cette recherche, nous n'adopterons pas cette définition maximaliste du LP, car il ne semble pas s'appliquer au cadre de la production en L2. Nous allons, en revanche, présenter les différentes catégories ou typologies du LP que l'on trouve dans la littérature avant de donner notre typologie. Le langage préfabriqué faisant partie de la compétence native, le locuteur natif possède des dizaines de milliers de séquences en mémoire (Pawley et Syder, 1983), à la différence d'un locuteur non-natif, excepté peut-être les locuteurs très avancés



(voir Forsberg-Lundell *et al.* 2014). D'autant que l'enseignement de la langue jusqu'à récemment en France n'a consacré que peu d'intérêt à la préfabrication, à part quelques expressions idiomatiques que l'on trouve dans les manuels. Cependant, les manuels de Français Langue Etrangère (FLE) insistent sur les actes de langage (se présenter, s'excuser, etc.) dès les premières leçons. Les apprenants dans l'approche communicative ou l'approche actionnelle (une version ciblée de l'approche communicative) peuvent acquérir un répertoire de phrases qui les aident à entrer en communication dès les premiers stades de l'apprentissage. Depuis des années, l'utilisation de l'approche lexicale de Lewis (1993) en anglais langue étrangère permet de reconnaître et mémoriser des ensembles de mots, sans passer par l'enseignement dichotomique vocabulaire et syntaxe. Les didacticiens commencent à mieux prendre en compte le phénomène de la préfabrication et on trouve dans certains manuels l'amorce d'une approche lexicale, même si on trouve encore dans beaucoup d'entre eux seulement des expressions idiomatiques en bas de page. Si le langage préfabriqué comprend les expressions figées et idiomatiques, il englobe aussi un grand nombre de suites de mots pas toujours contraintes, utilisées à l'écrit comme à l'oral. Les apprenants qui n'ont pas accès à certaines expressions peu utilisées, même en L1, comme « *il pleut des cordes* » ou « *avoir maille à partir* », vont mémoriser les suites de mots entendues fréquemment. En L1, ces suites de mots jouent un rôle dans la production orale, et en L2 *a fortiori* elles se révèlent essentielles à l'apprenant pour l'aider à s'exprimer avec plus d'aisance, et lui permettre de libérer du temps pour la planification. En L2, leur utilisation facilite donc la production orale de plusieurs manières : fluidité et aisance à l'oral, mais aussi stratégie de planification.

Notre recherche se propose d'observer comment le langage préfabriqué de sujets non-natifs se développe en milieu homoglotte. Si ce sujet intéresse de plus en plus les linguistes, surtout ceux de l'école de Sinclair, et tous les linguistes qui travaillent sur des corpus oraux de la langue, il intéresse aussi dans une certaine mesure les psycholinguistes qui se posent la question de la mémorisation/activation du langage et ce qui en découle, rapidité et fluence. En effet, des chercheurs (Wray, 2002, 2008 ; Weinert, 1995 ; Edmonds, 2010a, b, 2011) reconnaissent les avantages psycholinguistiques du langage préfabriqué en L1 comme en L2. Notre recherche longitudinale

s'intéresse à ce phénomène essentiel lors de l'apprentissage d'une langue car il permet aux apprenants un usage de la langue qui se rapproche d'une façon plus fine et naturelle de celui des locuteurs natifs et facilite ainsi la production. Mais si l'utilisation du langage préfabriqué

leur permet de se rapprocher du parler « comme un natif », il leur permet tout également de pouvoir garder la parole sans s'interrompre trop longtemps ou trop souvent. Une partie de notre travail consiste à examiner et à relever le langage préfabriqué utilisé par des apprenants de français langue étrangère de niveau débutant à avancé dans des productions orales semi-spontanées. Les résultats de notre groupe expérimental seront comparés à ceux de notre groupe contrôle, afin de mieux interpréter leurs résultats.

Cette recherche soulève en effet de nombreuses questions : qu'est-ce que le langage préfabriqué ? Comment se développe-t-il en langue étrangère (L2) ? Comment peut-il aider l'apprenant d'une L2 à s'exprimer plus aisément dès les premiers stades de l'apprentissage ? Quels liens pourrait-on établir entre langage préfabriqué et automatisation des processus communicatifs ?

Dans le chapitre 1 de notre partie théorique, après un bref rappel historique du phénomène, une large place sera accordée aux définitions des principales catégories de séquences préfabriquées du point de vue de la phraséologie française. Les recherches en linguistique phraséologique ont surtout cherché à identifier et décrire les séquences sans trop s'intéresser à leur acquisition en L1 comme en L2. Notre travail de recherche, en effet, s'intéresse au développement des séquences préfabriquées en langue étrangère, sur une période de six mois, et à leur rôle en production orale. C'est pourquoi le chapitre 2 traitera du rôle des SP en production orale, de leur acquisition en L1 comme en L2 et de leur identification. Les théories et modèles de production de parole s'accordent à montrer qu'il existe un lien entre l'emploi des séquences préfabriquées et une plus ou moins grande aisance à l'oral. Aussi, le développement de la fluence dépendra en grande partie, surtout en début d'apprentissage et en milieu homoglotte, de l'acquisition d'un répertoire solide de séquences. Ces séquences ne seront pas seulement les séquences identifiées et décrites en phraséologie mais pourront être des séquences propres à un locuteur donné dans la mesure où elles répondent aux critères présentés et retenus pour cette recherche. Le chapitre 3 sera consacré aux théories et modèles de production de parole en L1 et en L2. Nous montrerons que la fluence en production orale dépend de l'automatisation des connaissances et des séquences préfabriquées. Le chapitre 4 sera consacré aux mesures de la fluence de production. Dans le chapitre 5, nous présenterons notre méthodologie. Nous présenterons les sujets et les tâches auxquelles ils ont été soumis ainsi que nos outils d'analyse. Les mesures de la préfabrication choisies permettent de mesurer l'importance du langage préfabriqué en termes quantitatifs. Les mesures de la fluence

présentées dans ce même chapitre permettront de voir s'il y a une corrélation entre l'usage des SP et une meilleure fluence

La seconde partie de cette recherche comprend la présentation des résultats, d'un point de vue quantitatif et qualitatif, ainsi que leur analyse. Dans le chapitre 6, nous présenterons dans un premier temps les résultats du groupe contrôle afin de les comparer au groupe expérimental et, dans un deuxième temps, au groupe avancé en particulier. Les résultats du groupe et des sous-groupes non-natifs seront présentés dans le chapitre 7, ainsi que les résultats individuels.

Enfin, nous terminerons par l'analyse quantitative et qualitative des SP dans le chapitre 8 afin de montrer à travers des exemples de notre corpus le rôle que jouent les SP en production orale. Le chapitre 9 analysera le rôle que jouent deux SP récurrentes dans le corpus, *c'est* et le syntagme dét + *petit* +N, en production orale.

# I. Partie théorique

---

# Chapitre 1 Évolution historique et définitions

---

Ce chapitre est consacré en partie au rappel historique des principaux courants linguistiques qui ont abouti à la reconnaissance du langage préfabriqué. Ce phénomène, constaté bien avant les années 80, n'a pas toujours reçu l'intérêt qu'il méritait. Cependant, ces dernières années, dans le domaine de la recherche en acquisition des langues, que ce soit en L1 ou en L2, certains chercheurs ont montré que la maîtrise des SP aussi bien à l'écrit qu'à l'oral permet au locuteur ou au scripteur une meilleure aisance. La terminologie très vaste utilisée pour désigner les SP, est bien souvent fonction du chercheur et de ce qu'il cherche. Néanmoins, tous les termes que nous trouvons dans la littérature désignent des suites de mots qui n'ont pas toujours la même forme ni la même fonction mais que nous appelons dans cette recherche « séquences préfabriquées ». Avant de brosser un aperçu des phénomènes phraséologiques de la tradition française, nous consacrons une section à l'évolution historique de cette branche de la linguistique qui s'est intéressée aux expressions idiomatiques. Comprendre d'un point de vue diachronique ce qui s'est passé en linguistique et notamment en linguistique française nous permettra d'aborder en synchronie les différents phénomènes traités dans d'autres champs d'étude comme la psycholinguistique, la sociolinguistique et la pragmatique. Nous ne pouvons pas faire l'impasse d'une approche linguistique, externe au locuteur, du phénomène de la préfabrication et encore fortement ancrée chez les chercheurs en France. Nous nous placerons ensuite dans une perspective plus psycholinguistique, centrée sur le locuteur et l'usage qu'il fait du langage préfabriqué.

## 1.1 Aperçus historiques

### 1.1.1 Les précurseurs

On constate depuis peu un regain d'intérêt scientifique pour le langage préfabriqué qui avait été quelque peu négligé par les linguistes depuis Bailly (1905). Cependant, bien avant Bailly, quelques linguistes avaient montré de l'intérêt pour la phraséologie. Dès le début du XIX<sup>ème</sup> siècle, Marcel, considéré comme l'« un des pionniers de la linguistique appliquée », aussi important sans doute que Sweet (1845-1912) et Palmer (1877-1949), s'y est intéressé. Bien

que son travail n'ait eu apparemment que peu d'influence sur ses contemporains, Marcel base son enseignement sur des principes et élabore une méthode d'enseignement qui semble, à première vue, le désigner comme un des premiers linguistes importants (Smith, 2009 : 172). Enseignant de français en Grande-Bretagne, il met l'accent sur la répétition d'exemples pour faire comprendre à ses apprenants les constructions sans passer par la traduction mot à mot. En France, dès 1897, Bréal constate lui aussi, dans son *Traité de Sémantique*, la présence d'expressions qu'il appelle « groupes articulés », « des mots que l'usage a réunis depuis si longtemps qu'ils n'existent plus pour notre intelligence à l'état isolé » et « on ne les reconnaît bien que quand on rapproche la langue maternelle d'une langue étrangère. » (Bréal, 1897 : 186-187). En effet, la difficulté à traduire les expressions toutes faites a attiré l'attention de ce linguiste comme elle avait attiré l'attention du professeur de langue Claude Marcel. Des expressions ou phrases toutes faites comme *à tire-larigot*, *avoir maille à partir*, *au fur et à mesure*, qui ne pouvaient, pour une grande partie, subir aucune transformation syntaxique ou sémantique, étaient à prendre telles quelles et étaient considérées et traitées comme des exceptions, à la frontière du lexique et de la grammaire.

Au XX<sup>ème</sup> siècle, des linguistes vont continuer à s'intéresser au phénomène phraséologique. Dans son *Traité de Stylistique française* (1909), et son *Précis de Stylistique* (1905), Bailly propose une définition de la phraséologie qui regroupe différents types de suites de mots, comme les expressions figées mais aussi les clichés, les exclamations, les séries verbales, etc. Il distingue les unités phraséologiques indécomposables - qui correspondent aux expressions idiomatiques qui sont des ensembles non-compositionnels, et les séries phraséologiques - que l'on peut rapprocher des collocations (*chaleur suffocante*, *omelette baveuse*, *diamétralement opposé*, *prendre une décision*, *remporter une victoire*, *rendre hommage*). Bailly avait déjà posé les jalons d'une description et une typologie des phénomènes phraséologiques. « Entre les extrêmes (groupements passagers et unités indécomposables) se placent des groupes intermédiaires appelés séries phraséologiques (p. ex. les séries d'intensité ou les périphrases verbales) » (Bailly, 1909 : 66). « Il y a série ou groupement usuel lorsque les éléments du groupe conservent leur autonomie, tout en laissant voir une affinité évidente qui les rapproche, de sorte que l'ensemble présente des contours arrêtés et donne l'impression du « déjà vu », (Bailly, 1909 : 70). Ce « déjà vu » montre bien que certains mots vont bien avec d'autres, toujours les mêmes. Il souligne l'importance des « séries » dans la maîtrise d'une

langue étrangère et introduit aussi la notion de phraséologie exclamative que l'on trouve à l'oral dans les interactions principalement.

[Cette] catégorie comprend les expressions nettement affectives qui sont caractérisées, outre l'unité d'impression, par une intonation exclamative. D'une manière générale, toute série de mots subissant l'action d'un mouvement exclamatif de prononciation tend à perdre, de ce fait, l'autonomie de ses éléments [...]. Si l'on parle d'un travail comme d'une chose difficile, je puis répondre : « La belle affaire ! » c. à d. « Rien n'est plus facile » ; voyant quelqu'un au désespoir à la suite d'un malheur, je lui dirai peut-être : « Que voulez-vous ? » [...]. (Bailly, 1909 : 274)

Saussure, dès 1913, mentionne également l'importance des groupements dans la langue, « [...] les mots offrant quelque chose de commun s'associent dans la mémoire, et il se forme ainsi des groupes au sein desquels règnent des rapports très divers. » Plus loin, il dit que « ...nous ne parlons pas par signes isolés, mais par groupes de signes, par masses organisées qui sont elles-mêmes des signes. Dans la langue tout revient à des différences, mais tout revient aussi à des groupements. » (Saussure, 1913 : 177).

### **1.1.2 Deux positions extrêmes : la grammaire générative et la grammaire de construction**

La recherche en phraséologie est passée au second plan avec l'arrivée de la grammaire générative (Chomsky, 1965) et sa préoccupation majeure, le « principe du libre choix ». Selon ce principe, l'être humain a la possibilité de produire une quantité infinie d'expressions nouvelles. Pour Chomsky (1965), la créativité linguistique par opposition à la stéréotypie signifie l'aptitude du sujet parlant à créer et comprendre un nombre infini d'énoncés nouveaux qu'il n'a jamais entendus auparavant et qu'on ne peut répertorier en totalité. Les générativistes vont s'intéresser surtout à ces possibilités combinatoires illimitées du langage humain selon les principes de la grammaire universelle (*Universal Grammar*) sans accorder trop d'attention aux phénomènes phraséologiques. Ces chercheurs se sont plus intéressés à la compétence des locuteurs et à la possibilité de produire à l'infini des combinaisons de mots au détriment de leur performance. Si certains ne nient pas la préfabrication, ils la considèrent comme un phénomène peu intéressant qui ne dit rien sur le traitement de la langue (Pinker, 1994 : 44)<sup>11</sup>. Alors que Fillmore (1979), contemporain de Chomsky, considère que tout est formulaique dans la langue et que même les constructions, comme par exemple l'anglais *let alone* (*encore moins*), véhiculent un sens. L'énoncé composé de deux parties distinctes dont

---

<sup>11</sup> « Using prefabricated chunks of language is a peripheral pursuit that tells us nothing about real language processing. »  
(Pinker, 1994 : 44)

l'une évoque un empêchement indiqué par la forme négative, prend un sens particulier du fait de la position de la séquence figée *let alone* après la virgule. L'emploi de l'expression renforce l'idée de l'impossibilité de réaliser ce qui est dit dans la première partie de l'énoncé, comme le montre l'exemple suivant :

- Tu as une nouvelle voiture ?

- Tu plaisantes ! Je ne peux me permettre un litre d'essence, *encore moins* une voiture !

Dans cet exemple, la place de l'expression est porteuse de sens et signale une situation d'empêchement (section 1.4.3). On parle alors de grammaire de construction (*construction grammar*).

### 1.1.3 L'école de Sinclair

Depuis les années 1980, grâce aux travaux des linguistes s'intéressant aux grands corpus textuels, les sciences du langage ont « redécouvert » l'importance de la préfabrication dans le langage humain. Les pères fondateurs de la linguistique de corpus, Firth à la suite de Palmer, s'étaient intéressés au lexique et aux collocations. Firth (1957) est celui qui a initié la recherche des collocations dans des corpus. Ses disciples, Halliday et Sinclair, ont développé la linguistique de corpus contextualiste. « Halliday a surtout développé l'aspect grammatical, Sinclair sera à l'origine de la partie lexicale et donc l'inventeur de l'analyse de corpus contextualiste. » (William, 2005 : 153).

C'est dans ce contexte historique que Sinclair (1991), plus tard, remet la question de la préfabrication au centre des préoccupations en linguistique. Il confronte deux principes tout aussi importants : le « principe du libre choix » (*open-choice principle*) cher aux générativistes et le principe idiomatique (*idiom principle*). Les locuteurs ont ainsi à leur disposition un ensemble de phrases préconstruites qui constituent des choix uniques même si elles peuvent être analysées en segments. Sinclair et le courant de Birmingham ont mis en évidence, dans de grands corpus numérisés, que le mot n'acquiert pas son sens de manière isolée mais en relation avec d'autres avec lesquels il forme des unités de sens. Ces items lexicaux peuvent contenir un ou plusieurs mots et parfois être des phrases entières (phrases lexicalisées et expressions idiomatiques). L'utilisation de l'outil informatique va rendre possible, dans ces larges corpus, l'extraction et l'analyse d'ensembles, qu'il serait impossible au chercheur d'embrasser de son seul regard. La linguistique de corpus tentera ainsi de « fournir de meilleures descriptions des langues naturelles et de leurs variantes » en s'intéressant aux items lexicaux et à la relation que les mots entretiennent entre eux (Teubert,



2009 : 5)<sup>12</sup>.

## 1.2 Universalité et ampleur du phénomène

### 1.2.1 Présence dans le discours oral et écrit

L'une des plus grandes contributions de la linguistique de corpus est, sans doute, d'avoir mis en avant que certains mots se trouvent fréquemment associés à d'autres, le plus souvent les mêmes. Les expériences sur des corpus de langue écrite ou orale ont permis d'extraire des suites de mots récurrentes, comme par exemple dans les discours politiques, et même dans des textes littéraires. Les domaines spécialisés, comme entre autres - la santé, l'informatique, l'ingénierie, l'aviation – regorgent d'expressions spécifiques. A l'oral, certains chercheurs ont montré aussi qu'en situation de production intense, les locuteurs font une utilisation quasi totale des séquences préfabriquées sur lesquelles ils s'appuient fortement pour parler en temps réel et parfois limité. Kuiper et Haago (1984) constatent que les commissaires-priseurs ont à leur disposition un répertoire de séquences, adaptées à chaque étape de la vente aux enchères, qui leur permet d'organiser en temps limité la vente avec un maximum d'efficacité. Les reportages sportifs (Kuiper, 1991) sont un exemple aussi très révélateur de la place des phrases toutes faites dans la langue. Les journalistes sportifs les utilisent pour commenter en temps réel et sous la contrainte du direct une partie de tennis, un match de football ou toute autre manifestation sportive.

Une large partie de la langue se trouve donc être préconstruite, c'est-à-dire composée d'ensembles de mots, aussi bien à l'oral qu'à l'écrit. C'est pourquoi des chercheurs (Pawley et Syder, 1983, Nattinger et DeCarrico, 1992 ; Hickey, 1993 ; Bardovi-Harlig, 2010 ; Wood, 2006 ; Yorio, 1989) ont reconnu l'importance des séquences « préfabriquées » dans le langage, que ce soit en langue maternelle (désormais L1) ou en langue étrangère (désormais L2). En L1, les locuteurs font un usage plus ou moins important de formules déjà prêtes à l'emploi ou d'ensembles composés de mots préférés à d'autres équivalents. Et en effet, il semblerait étrange à un natif d'entendre « *comment va l'état actuel de votre santé ?* » à la place du rituel « *comment allez-vous ?* ». En L2, pour parler comme un natif, l'apprenant devra mémoriser un certain nombre d'expressions car « une grande partie de la capacité d'une

---

12 « La linguistique de corpus : une alternative [version abrégée] », Semen [En ligne], 27 | 2009, mis en ligne le 01 juin 2009, consulté le 05 juillet 2014. URL : <http://semen.revues.org/8914>.

personne à se débrouiller avec la langue vient de la maîtrise d'expressions préfabriquées »<sup>13</sup> (Fillmore, 1979 : 92). Ces expressions ne sont pas forcément des séquences idiomatiques ou figées, mais des séquences privilégiées des natifs parmi d'autres, tout aussi correctes linguistiquement (Pawley et Syder, 1983), considérées comme préfabriquées compte-tenu surtout de leur valeur socio-pragmatique. On trouve par exemple des formules liées à une situation « *comment ça va ? joyeux anniversaire* » ou des expressions comme « *ça marche ! ça m'est égal* », etc. Pawley et Syde

r (1983 : 213) estiment notamment à plusieurs centaines de milliers le nombre de séquences préfabriquées en L1 chez un locuteur anglais adulte. Jackendoff (1995), à partir d'un petit corpus de langue orale relevé lors d'une émission de TV, arrive à la conclusion que la proportion de séquences préfabriquées est égale sinon plus importante dans le lexique que celle des mots simples. Pour Mel'çûk (1993 : 83), les locuteurs experts de français produisent naturellement des expressions, appelées phrasèmes, qu'il estime à des dizaines de milliers. Il ne relève pas moins de vingt et une séquences dans une page et demie d'un célèbre journal, ce qui tend à démontrer l'importance du langage formulaïque aussi à l'écrit. En L1, Wray et Perkins (2000 : 1-2) évaluent à 70% la part du LP à l'oral. Erman et Warren (2000 : 37) trouvent, de leur côté, 58% de langage préfabriqué à l'oral, et 52,3% à l'écrit, (avec plus de formules discursives à l'oral et plus de séquences lexicales à l'écrit). Et trente pour cent des mots consistent en « paquets lexicaux » selon Biber et Conrad (1999). La part du langage préfabriqué représenterait environ un tiers de la production orale des locuteurs natifs lors d'une tâche non planifiée, soit précisément 32,29% du discours (Foster, 2001 : 85) en L1, et seulement la moitié, soit précisément 16,83% chez les locuteurs non-natifs pour la même tâche. Foster (2001) constate aussi que les locuteurs non-natifs utilisent plus de SP (17,23% vs 16,83%) lors d'une tâche planifiée et les locuteurs natifs moins (25,08% vs 32,29%), c'est-à-dire que les natifs s'appuient dans ce cas moins sur le LP et produisent un langage plus créatif. Mais Foster ne s'intéresse qu'à un type de SP, les phrases lexicalisées identifiées par des juges natifs.

Dans certaines études, en L2, on peut trouver des proportions plus grandes encore. Oppenheim (2000 : 228) estime, à partir d'un corpus oral, entre 48% et 80% (moyenne de 66%) la présence de « séquences récurrentes » dans le discours de locuteurs non-natifs, ce qui représente les deux tiers de leur production. Si les objectifs et donc les comptages ne sont pas

---

13 « [...] I believe a very large portion of a person's ability to get along in a language consists in the mastery of formulaic utterances. » (Fillmore, 1979 : 92)

tous les mêmes selon les chercheurs, ces chiffres montrent bien que le LP représente une part importante du discours aussi bien à l'oral qu'à l'écrit, en L1 et aussi dans une proportion à peine plus réduite en L2, chez des locuteurs très avancés. On peut sans doute expliquer les écarts de chiffres par des approches différentes du langage préfabriqué : par exemple, Biber et Conrad (1999), Biber (2004) ne recherchent qu'un type de séquences, « les paquets lexicaux » ou *lexical bundles* qui sont des blocs composés de trois ou quatre unités, *quelque chose comme, comme tu sais, plus ou moins*, dans des textes avec des registres spécifiques ; alors que d'autres chercheurs considèrent le LP dans son ensemble.

Notons que la plupart des recherches citées ont été entreprises sur des corpus en langue anglaise. Cependant, si le LP semble largement attesté en anglais et en français (Mel'çûk, 1993, 2011), à un certain niveau de maîtrise, il a également été observé dans de nombreuses langues (Conklin et Schmitt, 2008). Si on ne détient pas encore de preuve chiffrée que le LP soit un trait universel des langues, néanmoins, l'existence de ce phénomène langagier dans de nombreuses langues, souvent très éloignées comme l'hébreu et l'anglais par exemple, suggère que c'est un phénomène universel. Et même en langue des signes, des formes préfabriquées sont attestées, malgré un mode de transmission totalement différent des langues vocales. En effet, en langue des signes, il existe des signes lexicaux associant une forme à un sens et qui correspondent aux expressions préfabriquées de la langue vocale (Johnston et Schembri, 2010). Un signe se lexicalise lorsque dans une unité signée la forme est régulièrement associée à un sens (Johnston et Ferrara, 2013 : 229)<sup>14</sup>. Par exemple, la séquence *maison du handicap* (mot composé) réunit en langue des signes française (LSF) deux signes *maison* et *handicap* qui sont réunis en un seul signe.<sup>15</sup>

### **1.2.2 La préfabrication : un vaste choix de termes**

Si le langage préconstruit ou préfabriqué est largement attesté, il se trouve qu'il n'est pas si simple de le décrire précisément. En effet, les chercheurs dans ce domaine ne travaillent pas

---

<sup>14</sup> « Specifically, lexicalization in SLs occurs when a signed unit acquires a clearly identifiable and replicable citation form which is regularly and strongly associated with a meaning. The meaning is unpredictable and/or more specific than that implied by the value (meaning) that each component may contribute to the overall form of the signed unit. » (Johnston et Ferrara, 2012 : 229)

<sup>15</sup> Nous remercions vivement Brigitte Garcia, professeur à l'université Paris 8, pour nous avoir expliqué les grandes lignes de la LSF et le figement en LSF.

tous sur les mêmes types de séquences et ils n'utilisent pas une même terminologie pour les désigner. Ainsi, Wray (2002 : 9) recense plus de cinquante termes pour décrire le phénomène. Provenant du domaine de la linguistique, de la sociolinguistique, de la psycholinguistique et de la didactique des langues, voici pêle-mêle quelques termes qui ont été utilisés en anglais et en français pour référer à ces suites de mots - phrasillons (Tesnière, 1965) ; *prefabricated routines* (Brown, 1973) ; holophrases (Corder, 1973) ; *prefabricated patterns* (Hakuta, 1974) ; *preassembled speech* (Bolinger, 1975) ; gambits (Keller, 1979) ; *formulaic expressions, formulaic speech* (Wong-Fillmore, 1979) ; *conventionalized forms* (Yorio, 1980) ; *memorised sentences, lexicalised stems, sentence stems* (Pawley et Syder, 1983) ; *formulae* (Raupach, 1984) ; composites (Cowie, 1988) ; *formulaic chunks* (Widdowson, 1989) ; *multiword units* (Sinclair, 1991) ; *lexical phrases* (Nattinger et DeCarrico, 1992) ; *lexical chunks* (Lewis, 1993) ; *formulas* (Ellis, 1994) ; phrasèmes (Mel'çûk, 1995) ; séquences figées (Mejri, 2004) ; *prefabs* (Erman et Warren, 2000) ; *formulaic language, formulaic sequences* (Wray, 2002) ; unités lexicales préconstruites (Grossmann et Tutin, 2003) ; séquences préfabriquées (Forsberg, 2006) ; unité de construction préformée (Schmäle, 2013).

Tous ces termes désignent des facettes du même phénomène et recouvrent un vaste domaine, parfois aux frontières du lexique et de la grammaire. Voici quelques exemples de séquences préfabriquées en français à l'oral comme à l'écrit, qui montrent la diversité et l'étendue du phénomène : *Quoi qu'il en soit ; Aussi vieux que le monde ; A prendre et à laisser ; ça marche ; bien à vous ; c'est ça ; je pense ; mais encore ; Tu connais la meilleure ? Ça jamais ! Suite à ma lettre du... Toutes mes félicitations ; Ah bon ? à votre tour ; grande tristesse ; sans foi ni loi ; roi maudit ; Qui peut le plus peut le moins !* Cette liste regroupe aussi bien des collocations, des formules conventionnelles, des expressions phatiques, que des expressions ou phrases moins figées. Si tous ces termes renvoient à la notion de préfabrication, on peut cependant distinguer différents types de séquences que l'on regroupe sous le même terme, mais qui n'ont pas toutes la même fonction ni les mêmes critères d'identification. Pendant longtemps, la phraséologie s'est surtout intéressée au figement et à l'irrégularité syntaxique ou sémantique des expressions. Nous présentons dans la section suivante la phraséologie en linguistique française et sa relation au figement.

### **1.3 La phraséologie en linguistique française**

Des colloques récents en France illustrent un renouveau d'intérêt scientifique pour la

phraséologie, ainsi qu'un élargissement de ses champs d'intérêt. La phraséologie, considérée comme une branche de la lexicologie, s'occupe spécifiquement des combinaisons de mots (Granger et Meunier, 2008 : xix). Mais, récemment, les recherches en phraséologie se sont étendues à d'autres disciplines de la linguistique comme la syntaxe, la didactique, la psycholinguistique, la sociolinguistique, le traitement automatique des langues. Longtemps considérée comme un phénomène marginal, à cause des prétendues anomalies des structures phraséologiques, la phraséologie est considérée aujourd'hui comme une branche à part entière de la linguistique. Sa présence importante dans le discours, constaté depuis Bally (1909), justifie qu'elle soit au cœur des thématiques fondamentales de la linguistique (Tutin, 2011). Si les recherches en phraséologie se sont surtout intéressées à la L1, on note malgré tout un intérêt grandissant en France pour le rôle que joue le LP dans l'acquisition d'une langue étrangère. Lors du colloque « Europhras14 » qui s'est tenu à La Sorbonne-Paris en septembre 2014, des chercheurs venus de tous horizons ont montré que les études qui s'intéressent aux séquences préfabriquées en L2 commencent à gagner du terrain en France, en acquisition comme en didactique (Théophanous, 2008, 2013).

### **1.3.1 Définitions de la phraséologie**

Dans une première définition très étroite donnée par le Centre National de Ressources Textuelles et Lexicales (CNRTL), l'idée générale de la phraséologie renvoie à un type de discours propre à une catégorie d'individus ou d'hommes politiques, qui emploieraient des phrases, toujours les mêmes, pour véhiculer des idées ou une idéologie. On parle même de la phraséologie d'un homme ou d'un parti politique. Dans « phraséologie », on reconnaît le mot « phrase » ; chaque langue naturelle possède ses phrases composées d'expressions singulières, de proverbes, ou d'idiomatismes. La phraséologie est définie dans les dictionnaires comme faisant partie du patrimoine d'une langue (Mortchev-Bouveret, 2004). La phraséologie serait donc l'« Ensemble des tournures typiques d'une langue, soit par leur fréquence, soit par leur caractère idiomatique » (*Trésor de la Langue française* <http://atilf.atilf.fr/>, consultation en ligne), ou encore un « système de particularités expressives liées aux conditions sociales dans lesquelles une langue est actualisée, c'est-à-dire à des usages » (Rey et Chanterreau, 1993 : vi). Cette dernière définition montre que la langue n'est pas fixe et qu'elle évolue dans l'emploi de ses phrases.

Une définition donnée par Paillard (2000) reconnaît le statut de ces « tournures » à côté du mot simple, et établit l'aspect mémoriel et sociolinguistique de celles-ci. La phraséologie représente l'ensemble des locutions, expressions, énoncés conventionnels, unités de niveau supérieur au mot (*multi-word units*) mémorisées et reconnues par une communauté linguistique car elles présentent un certain degré de figement. Cette définition soumet les phénomènes phraséologiques à l'idée du figement. Or, une séquence préfabriquée, telle que nous la définissons dans cette recherche, ne sera pas forcément figée. Avant d'examiner les phénomènes phraséologiques les plus connus, il nous semble important, dans un premier temps, de définir le figement.

### 1.3.2 Le figement

Les chercheurs en lexicologie ont souvent défini les phénomènes phraséologiques par les critères de figement syntaxique et sémantique (voir Granger et Paquot, 2008), dans un continuum allant des collocations (deux mots entretenant une relation privilégiée) aux locutions figées en passant par les expressions idiomatiques métaphoriques (Cowie 1981 ; Gross 1996 ; Mel'çûk 1998). Pour bien comprendre ce que l'on entend par « expression figée », le mieux est de la comparer à une « expression libre » (Mel'çûk, 2011). Selon Mel'çûk, une expression figée se distingue d'une expression libre car elle ne peut pas toujours être construite « selon les règles générales de la langue ». Il fait ainsi une distinction utile entre les syntagmes libres et les syntagmes figés ou contraints :

Un phrasème est un syntagme **NON LIBRE**<sup>16</sup>, c'est-à-dire, de façon informelle, un syntagme qui ne peut pas être construit selon les règles générales de la langue ; un syntagme « libre », par opposition, étant un syntagme non contraint et compositionnel. Les deux propriétés des signes, l'une paradigmatique (axe de sélection) et l'autre syntagmatique (axe de combinaison), prises ensemble, caractérisent un syntagme libre. (Mel'çûk, 2011 : 42)

Cependant si le figement est un « processus par lequel les éléments d'un syntagme perdent leur autonomie », on peut reconnaître des degrés plus ou moins hauts au figement. Ainsi, Mejri (2011 : 71-73) distingue le figement absolu, qui concerne les séquences ayant conservé leur forme d'origine provenant d'archaïsmes (*bayer aux corneilles, chercher noise, se tenir coi, sans mot dire, garder bouche cousue, sans coup férir, ficher le camp*, etc.), les séquences héritées du latin (*nec plus ultra, a priori, ad vitam aeternam, et cetera.*), ainsi que celles qui

---

<sup>16</sup> Souligné dans le texte par Mel'çûk (2011 : 42)

après soudure sont devenues des séquences monolexicales (*pourboire, tournevis, vinaigre, bonjour, voici, voilà, etc.*) du figement relatif, qui admet au moins une variation relevant de la combinatoire libre. Pour Gross (1988 : 21-22), il semblerait même que les phrases les plus figées possèdent des degrés de liberté, dans le sens où elles s'analysent pratiquement toutes de façon syntaxiquement régulière et qu'elles subiraient les mêmes règles que la syntaxe libre. Les locutions sont reconnues comme des ensembles figés, l'expression est prête à l'emploi - préfabriquée - elle semble fonctionner comme un "mot" avec un sens spécifique (spécialisé, lexicalisé) au même titre que les autres mots "monolexicaux" du dictionnaire. Cependant, malgré une certaine prise de conscience du champ élargi du phénomène de la préfabrication, la phraséologie en linguistique française s'intéresse encore et surtout au figement et à la non-compositionnalité des expressions. Dans chacune des catégories que nous présentons, le figement (ou le semi-figement) reste un des premiers critères pour l'identification des phrasèmes.

Ce grand ensemble phraséologique regroupe de multiples expressions, formules et suites de mots, nommées de différentes façons, classées selon des typologies variées et selon le point de vue adopté de ces chercheurs. Svensson (2004 : 14) fait une distinction utile entre les « critères » de la préfabrication (figement, opacité, fréquence, etc.) - et les « catégories » (collocations, expressions idiomatiques, expressions conventionnelles et les formules discursives, etc.). Nous présentons les catégories les plus connues dans la section suivante.

## **1.4 Les catégories formelles**

Les linguistes ont regroupé les séquences préfabriquées sous diverses étiquettes selon des critères formels. On trouve parmi elles un ensemble formé par les collocations, les expressions idiomatiques et les proverbes. Pour essayer d'y voir plus clair, nous proposons d'examiner chaque catégorie, de manière non exhaustive, afin d'en dégager les grandes lignes.

### **1.4.1 Les collocations**

#### ***a) Les collocations lexicales***

Les collocations font partie des unités phraséologiques les plus connues. Omniprésentes dans le discours écrit et oral, elles ont été largement décrites par les linguistes. Firth (1957), considéré comme le « père de la collocation » aurait été l'un des premiers à utiliser le terme.



Mais le mot est employé bien avant ; on trouve une acception du mot « collocation » dès le XVIIIème (1750)<sup>17</sup> avec plutôt le sens de colligation (Bartsch, 2004 : 28) et, au début du XXème siècle, chez Palmer (1933) cité par Legallois, 2012 : 34.

La collocation est définie le plus couramment comme la « cooccurrence lexicale privilégiée de deux éléments linguistiques entretenant une relation syntaxique » (Firth, 1957 cité par Tutin et Grossmann, 2002 : 5). Cette définition de la collocation est reprise par Sinclair en 1991, comme l'occurrence de deux ou plusieurs mots dans un espace limité et, par bien d'autres par la suite (Dubreil, 2008). Une collocation peut signifier, de manière très générale, la tendance pour certains items de se trouver ensemble dans la langue (Gregory, 1981 : 4). En linguistique française, les définitions de Tutin et Grossmann (2002 : 5) et de Mel'çûk (1995), « Association habituelle (fréquence élevée) de deux unités lexicales », privilégient une conception étroite de la collocation que l'on trouve dans l'approche dite fonctionnelle (Bolly, 2010).

Dans ces définitions, une collocation lexicale est un phénomène mettant en jeu essentiellement des lexèmes (la collocation réunit deux unités lexicales) et non des unités grammaticales ou fonctionnelles, ceci pour les différencier des collocations grammaticales ou colligations. L'approche fonctionnelle insiste sur la relation qu'entretiennent deux éléments, qu'on nommera la base et son collocatif pour reprendre la terminologie de Hausman (1989, 1997) :

Une collocation est l'association d'une lexie (mot simple ou phrasème) L et d'un constituant C (généralement une lexie, mais parfois un syntagme par exemple *à couper au couteau* dans *un brouillard à couper au couteau*) entretenant une relation syntaxique telle que : C (le collocatif) est sélectionné en production pour exprimer un sens donné en cooccurrence avec L (la base) ; Le sens de L est habituel. (Tutin et Grossmann, 2002 : 5)

Par exemple, les collocations *poser une question*, *passer un examen*, *célibataire endurci*, *le bruit court*, sont la combinaison phraséologique d'une base (*question*, *examen*, *célibataire*, *bruit*), qui garde son sens habituel et de son collocatif (*poser*, *passer*, *endurci*, *court*) qui peut avoir un sens métaphorique (Hausmann et Blumenthal, 2006 ; Mel'çûk, 1998). La base est un mot que le locuteur choisit librement, et le collocatif est un mot que le locuteur sélectionne en fonction de la base. En effet, aussi bien chez Hausmann (1989) que chez Mel'çûk, il y a collocation lorsque voulant produire une suite, le choix d'une des

---

<sup>17</sup> Entrée dans le Oxford English Dictionary (OED 2<sup>nd</sup>) qui mentionne une citation de Harris en l'an 1750 « 1750 Harris Hermes ii. Iv. Wks. (1841) 197 The accusative.in modern languages being subsequent to its verb, in the collocation of the words. »



expressions est imposé par l'autre : le locuteur pour exprimer l'intensité de sa *peur* choisira *bleue*, dans une cooccurrence restreinte. On se rend compte ici que l'adjectif *bleue* ne peut être compris dans son sens habituel de couleur et qu'associé à *peur* il prend une toute autre signification. Ce qui signifie que l'idée d'un figement relatif est tout-à-fait présente dans les collocations.

Un locuteur L2 non averti ne pourra pas comprendre le sens de cette expression à moins qu'il ne l'ait déjà rencontrée et ne pourra la produire qu'à la seule condition de l'avoir mémorisée. Son intuition pourra peut-être le guider et lui faire comprendre que l'adjectif *bleu* amplifie le sentiment de *peur* mais rien n'est moins certain.

Le caractère binaire de la collocation - que l'on trouve donc dans ces quelques exemples – ne se limite pas toujours à deux mots. Dans des collocations pluri-lexicales (voir Tutin, 2013), du type *fort comme un turc*, *un bruit à crever les tympans*, *un brouillard à couper au couteau*, un mot lexical (*fort*, *bruit*, *brouillard*) est accompagné/qualifié par un syntagme (*comme un turc*, *à crever les tympans*, *à couper au couteau*) ; on retrouve le caractère binaire dans ces expressions, mais les deux parties sont cette fois de longueur inégale. On peut aussi trouver des collocations « complexes » imbriquées l'une dans l'autre comme dans *avoir une peur bleue* ou *avoir une faim de loup*, qui sont en fait la fusion de deux collocations (*avoir peur* et *peur bleue*, *avoir faim* et *faim de loup*).

Les collocations ne sont pas un phénomène propre au français. Elles sont présentes dans toutes les langues aussi bien en langue naturelle qu'en langue spécialisée. Ainsi « en français, on fait un pas, en espagnol, on le donne (= *dar un paso*) et en anglais, on le prend (= *take a step*). La pluie est forte en français, mais lourde (= *heavy rain*) en anglais. On fait un rêve en français, mais on l'a en anglais (= *have a dream*) ; vous donnez un cours en français, mais vous l'enseignez (= *teach a course*) en anglais et vous le lisez (= *citat' kurs*) en russe (même si vous ne le lisez pas !). Toutes ces particularités d'usage sont des collocations – des expressions préfabriquées d'un certain type, qu'on ne peut pas prévoir et que la linguistique ne savait pas présenter de façon logique et facilement calculable » (Mel'çûk, 1997 : 23). Si elles sont interprétables plus ou moins correctement hors contexte, elles ne peuvent pas être produites correctement si elles ne sont pas connues de la communauté linguistique (Cruse, 2004 : 73-74). Cependant, les collocations, présentes de manière très forte dans la langue, peuvent être tout-à-fait transparentes. Calaque (2006) va jusqu'à dire que l'on fait des collocations sans le savoir comme M. Jourdain faisait de la prose !

Si une collocation est la réunion d'une base et de son collocatif, le choix de celui-ci n'est pas toujours prévisible et est fait de façon arbitraire. En effet, en français on dit une *faim de loup* et non pas un *\*appétit de loup* ; à l'inverse, on a un *appétit d'ogre* et non pas une *\*faim d'ogre*. En anglais, on dit *a powerful car* et non *\*a strong car*, *strong tea* et non *\*powerful tea*, alors que les deux adjectifs ont le même sens. Si on peut être *gravement* ou *grièvement blessé*, on ne peut être que *gravement malade* et non *\*grièvement malade* (c'est le fameux exemple de Bally) ; de *grands blessés* est possible alors qu'une blessure est *grave*, pas *\*grande*. Je peux être *très fatigué*, ou *bien reposé* mais non *\*très reposé* ; *fabuleusement riche*, et *bien connu* et non *\*bien riche*, *\*fabuleusement connu* ; la liste n'étant pas exhaustive (la plupart des exemples cités sont tirés de Mel'çûk et Polguère, 2007). Ces quelques exemples montrent que les habitudes langagières se construisent de manière arbitraire, et sans doute que la fréquence d'emploi de certaines expressions (combinée à l'histoire de la langue, à son évolution) a supplanté d'autres combinaisons tout aussi probables. Ces collocations ont sans doute aussi à voir avec le figement.

Certaines collocations, où l'un des mots a un usage restreint et ne peut être commuté comme dans *hausser les épaules* vs *\*lever les épaules*, et *prêter attention* vs *\*donner attention*, sont considérées comme figées. Elles peuvent l'être également du fait de l'usage<sup>18</sup> (Aisenstadt 1981 : 54) comme dans les exemples précédents. Si l'un des constituants est sélectionné par l'autre, on parlera de collocation restreinte ; ainsi dans *feu rouge*, *rouge* (la base) sélectionne *feu* au sens de « feu de la circulation ».

Etant donné le nombre incroyablement important des collocations dans la langue, des chercheurs ont essayé d'en proposer un classement. Par exemple, Grossman et Tutin (2003) ont classé les collocations selon leur degré d'opacité. Rappelons qu'opacité et non-compositionnalité sont synonymes, c'est-à-dire que l'on ne peut déduire le sens de la collocation de ses constituants suivant le schéma suivant :  $AB = C$ .

- i. Les collocations opaques où le collocatif est imprédictible et démotivé sémantiquement comme *peur bleue*, *nuit blanche*, *colère noire*. Dans *peur bleue*, *bleue* est synonyme d'extrême lorsqu'il est associé à *peur* (*\*angoisse* / *\*inquiétude bleue*).

---

18 « [RCs [rRestrictives collocations] are combinations of two or more words, [ ... ] in which one word at least is restricted in its commutability not only by its grammatical and semantic valency, but by usage. » (Aisenstadt 1981 : 54)

- ii. Les collocations transparentes comme *commettre un crime, faim de loup, avoir faim*, où le collocatif est motivé sémantiquement mais imprédictible.
- iii. Les collocations régulières (*rude journée, mortelle randonnée*) sont des « associations de mots ou de suites de mots dans lesquelles le sens du tout est généralement déductible et semble prédictible, bien que les règles d'association soient parfois complexes à élaborer » comme dans *célibataire endurci* ou *feuilleter un livre*.

François et Manguin (2006 : 51) constatent qu'un grand nombre de collocations ne sont pas des collocations restreintes mais se situent sur un continuum du plus au moins figé et, considèrent même que deux mots peuvent entretenir une collocation à distance comme dans *un fait historique* qui peut devenir *ce fait est historique*.

On reconnaît d'une manière générale le caractère binaire de la collocation comme la réunion de deux expressions linguistiques de même importance dont l'une peut présenter une plus ou moins grande restriction. Pour certains auteurs, elles peuvent aussi être la réunion d'un mot avec d'autres mots, sans aucune contrainte. Elles « signifient exactement ce que l'on peut comprendre en connaissant les significations lexicales individuelles et les règles syntaxiques » comme dans « *partir en voyage* », où aucun des constituants ne sélectionne l'autre et, de ce fait, aucun n'est le collocatif de l'autre (Blank, 2001 : 152). Chaque item de l'expression *partir en voyage* peut être associé à un autre « *partir à Paris, partir de chez soi ou aller en voyage*, de même dans l'expression *heureux événement* où aucun des deux items n'est contraint par l'autre. On la considère comme une collocation dans la mesure où les deux unités fonctionnent ensemble comme une unité plus large dont les items sont liés d'une certaine façon sur le plan sémantique. C'est ce que l'on reconnaît dans des expressions dont les deux constituants ont un sémantisme propre comme *conversation téléphonique* ou *long conciliabule* (François et Manguin, 2006 : 52).

En marge de la tradition française, surtout dans la recherche anglo-saxonne, des chercheurs ont privilégié la fréquence d'apparition de ces expressions dans de grands corpus pour valider le statut de collocation sans faire du figement un critère absolu. On trouve ainsi une définition plus statistique de la collocation où elle serait, en fait, l'association d'items lexicaux sur l'axe syntagmatique, quantifiables, comme la probabilité d'occurrences à intervalles d'un certain item x avec d'autres items a, b, c...<sup>19</sup> (Halliday, 1961 : 276). Dans une collocation au sens

---

19 Notre traduction de : « collocation is the syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur, at n removes (a distance of n lexical items) from an item x, the items a, b, c... ». (Halliday, 1961 : 276).

statistique, on peut se contenter d'observer que deux mots se trouvent ensemble fréquemment, on parle de « cooccurrence préférentielle » (François et Manguin 2006 : 51). Cooccurrences que Bolly (2010) caractérise en termes de collocation au sens statistique.

Les cooccurrences sont des combinaisons de deux mots graphiques (*tokens*) caractérisés par une proximité cotextuelle et qui apparaissent de manière fréquente dans l'usage. On parle alors de récurrence. Par extension, une collocation est une cooccurrence statistiquement significative, i.e. une séquence dont la co-apparition des termes ne semble pas due au hasard (Bolly, 2010 : 16)

On voit bien qu'on est loin du seul critère de figement cher à la linguistique française. Cependant si la notion de fréquence est importante, elle ne doit pas être le seul critère pour l'identification des collocations. En effet, une collocation peut n'apparaître que très peu de fois si elle appartient à un domaine spécialisé (ornithologie, mécanique, aéronautique, etc.) ou si elle appartient à un registre soutenu (*rendre hommage*). Une collocation, comme d'ailleurs toute expression idiomatique, peut être sur-employée à un moment de l'histoire et n'apparaître que très peu à une autre époque.

#### *b) Les collocations grammaticales*

Les linguistes ont tendance à distinguer les collocations lexicales (deux unités lexicales) des collocations grammaticales. Une collocation grammaticale est une combinaison courante, composée d'un mot lexical (de contenu) du type VERBE, NOM ou ADJECTIF et, d'un mot grammatical (mot-outil), le plus souvent une préposition ou une structure syntaxique (proposition infinitive) comme dans *dépendre de* qqn ou *de* qch ; *essayer de* + Vinf ; *fâché contre* ; *différent de* (Hausmann, 1998).

Benson (1985 : 61) propose une catégorisation des collocations qui distingue les collocations lexicales composées de deux mots « égaux » et les collocations grammaticales. Cette division est également mentionnée par Fontenelle (1994 : 44). Voici quelques exemples en français fournis d'après Benson (1985 : 62) : *accusé de* (verbe + préposition), *accès à* (nom + préposition), *curieux de* (adjectif + préposition).

Certains chercheurs emploient parfois le terme de colligation, « autre nom de la collocation grammaticale ou autre logique de la relation mutuelle entre syntaxe et sémantique » (Legallois, 2012 ; Schmäle, 2012 ; Granger et Paquot, 2008 : 33). Ils les considèrent comme synonymes ou quasi-synonymes. Cependant les deux phénomènes doivent être distingués. La colligation pour Sinclair (2003 : 145) est semblable à la collocation car toutes les deux

concernent la cooccurrence de traits linguistiques dans un texte. Une colligation est l'occurrence d'une classe grammaticale ou d'une structure syntaxique avec une autre, ou avec un mot ou un groupe de mots. Les catégories les plus représentées sont la négation, la possession et les modaux<sup>20</sup>. On peut sans doute distinguer entre la collocation qui est un « phénomène d'association entre mots lexicaux (*un gros buveur*), entre mots lexicaux et grammaticaux (*un jour sans*), entre mots grammaticaux (*le \* de \* - le livre de Marie*) » et la *colligation* qui est un « phénomène d'association entre un mot lexical ou grammatical et une catégorie grammaticale (partie du discours, fonction syntaxique, marqueurs aspectuels, modaux, temporels, marqueurs grammaticaux des catégories de la négation « *ne...pas / plus / jamais* », de la propriété, etc.) » (Legallois, 2012 : 39).

Rappelons que Bally intégrait déjà les « faits de syntaxe » dans les unités phraséologiques. « La négation *ne...que* est une unité phraséologique : d'une manière générale on pourrait prouver qu'une foule de faits de syntaxe ne sont que des groupements phraséologiques à éléments séparables » (Bally, 1909 : 76). De même, on notera cet autre exemple (Bally, 1909 : 77) : « la notion d'appartenance ou de propriété est exprimée en français par des groupes comme « *la maison de Paul* » où « maison » ou « Paul » peuvent être remplacés par n'importe quel substantif, la préposition étant la partie fixe du groupe : qu'est-ce là sinon une locution à éléments interchangeables ? ». Ces exemples montrent déjà que certaines constructions (voir section 1.4.3.) sont porteuses de sens (possession, négation).

#### **1.4.2 Les expressions idiomatiques**

Dans cette section, nous présentons la catégorie de SP qu'on appelle « expressions idiomatiques » et qui existent dans toutes les langues. Aucun des éléments qui composent une expression idiomatique n'a son sens habituel à la différence de la collocation qui en conserve au moins un comme dans *nuit blanche*, où la base, *nuit*, garde son sens propre.

En français, quand une collocation est figée, lexicalisée, on peut l'appeler « expression idiomatique ». D'où parfois des confusions car certains ont tendance à confondre les

---

20 Notre traduction de : « Colligation is similar to collocation in that they both concern the co-occurrence of linguistic features in a text. Colligation is the occurrence of a grammatical class or structural pattern with another one, or with a word or phrase. « Negative », « possessive » and « modal » are the kinds of largely grammatical categories that figure in colligation. (The term was first used by J.R. Firth, and has been widened a little for corpus work. » (Sinclair, 2003 : 145).

expressions idiomatiques composées de deux éléments où aucun élément n'a son sens habituel (*casser sa pipe, retourner sa veste, jeter de l'huile sur le feu*) et les collocations figées où l'un des éléments garde son sens habituel (*célibataire endurci*).

Dans les expressions suivantes *donner sa langue au chat* (demander la réponse à une question), *couper les cheveux en quatre* (se concentrer sur des détails sans importance), *mettre la charrue avant les bœufs* (ne pas faire les choses dans le bon ordre), aucun des éléments ne garde son sens propre. L'ensemble de l'expression prend un sens particulier différent du sens des unités qui la composent. « Un idiomme étant une expression plutôt figée dont le sens global ne reflète pas le sens de ses parties »<sup>21</sup> (Benson, 1985 : 66) ; il est donc non-compositionnel. Les expressions idiomatiques sont des tournures stéréotypées, lexicalisées et (souvent) imagées (Galisson, 1984) comme *jeter de l'huile sur le feu* et renvoient à la notion de locution (Mel'çûk, 2011). Confais (2005 : 3) distingue l'idiomaticité au sens large qui comprend les expressions figées et conventionnelles (idiomaticité formelle) et idiomaticité au sens étroit, où le mot *idiomatique* implique l'idée d'une image d'un sens figuré qui n'est pas toujours facile à « recalculer » à partir du sens littéral des constituants de la locution (idiomaticité sémantique). On parle aussi de lexie en français pour désigner ces expressions.

En phraséologie française, Mel'çûk (2011 : 47) appelle les expressions figées **locutions fortes**, ou complètes. Une locution est une locution FORTE<sup>22</sup> (= COMPLETE) si et seulement si elle n'inclut le sens d'aucun de ses constituants, c'est-à-dire qu'on aura le schéma suivant « AB » = C. Grant et Bauer (2004), mais aussi Nation et Webb (2011) à leur suite, distinguent dans cette catégorie trois types d'expressions idiomatiques, selon leur degré de figement : les « *core idioms* », les « *figuratives* » et les « *literals* ». Les *core idioms* sont des expressions idiomatiques, complètement figées car aucune transformation sur l'axe syntagmatique et paradigmatic ne peut être appliquée (*au fur et à mesure, en même temps, avoir les yeux plus gros que le ventre*). Les *figuratives* désignent les phrases qui peuvent avoir deux sens, comme dans *les carottes sont cuites* qui peut signifier *c'est fichu* au sens métaphorique, ou, tout simplement, *les carottes, le légume, sont cuites*. Dans leur sens métaphorique, les *figuratives* sont figées et on ne peut les actualiser (*\*tes, \*mes carottes sont*

---

21 Notre traduction de : « An idiom is a relatively frozen expression whose meaning does not reflect the meaning of its component parts. » (Benson, 1985 : 66).

22 Souligné dans le texte par l'auteur (Mel'çûk).



*cuites*,) à la différence des *literals*. Grant et Bauer (2004 : 49-51) précisent qu'on peut reconnaître l'emploi métaphorique en remplaçant l'expression dans son contexte, ce qui ne peut être fait que par des natifs ou des apprenants de niveau très avancé. Il faudra que l'apprenant d'une L2 les mémorise pour pouvoir les utiliser. Ceci étant, certaines d'entre elles sont peu utilisées même par les natifs. Mais beaucoup, comme le signale Confais (2005), sont facilement reproductibles car usuelles.

C'est surtout son caractère facilement reproductible, donc aussi le fait qu'elle soit facilement disponible en tant que concept compact dans la mémoire, qui la rend usuelle. Les raisons peuvent tenir paradoxalement à une anomalie (*etw. aus dem Effeff beherrschen, prendre la poudre d'escampette*), à une iconicité particulière – par ex. un jeu sur la forme (*in Bausch und Bogen, à tue-tête*) –, ou encore à la fonction de l'expression, par exemple quand elle est mentalement très fortement associée à un emploi dans une situation ou un contexte thématique (*Sachbereich*) bien précis, par exemple *herzliches Beileid* ou *avoir les yeux plus gros que le ventre* pour la situation, *der mündige Bürger* ou *la langue de bois* pour un domaine thématique. (Confais, 2005 : 4)

### **1.4.3 Les proverbes**

Nous évoquerons brièvement les proverbes, ces phrases entièrement lexicalisées qui sont le reflet d'une culture ou qui représentent des valeurs universelles, *Pierre qui roule n'amasse pas mousse* ; *Rien ne sert de courir il faut partir à point*, etc. On peut trouver des équivalents dans d'autres langues où parfois la même expression existe, comme dans *A attendre que l'herbe pousse, le bœuf meurt de faim* qui donne en anglais, *While the grass grows the steer starves*. Mais, le plus souvent, ces expressions doivent être apprises telles quelles, la traduction littérale étant impossible du fait de leur non-compositionnalité, *Abondance de biens ne nuit pas* donne en anglais *Plenty is no plague*. Contrairement aux phrases figées, « seuls les proverbes expriment une vérité morale, valable hors contexte » (Lamiroy, 2008, épreuve en ligne). Les proverbes peuvent être également considérés comme des expressions conventionnelles, institutionnalisées qui relèvent d'un contexte particulier.

### **1.4.4 Les expressions conventionnelles**

Les expressions conventionnelles (Yorio, 1980) ou « expressions institutionnalisées » (*institutionalized expressions*, Nattinger et DeCarrico, 1992) se distinguent des autres unités phraséologiques par leur association à une situation sociale ou à un contexte communicationnel précis. Une forme est dite conventionnelle quand elle est prévisible et

attendue dans une certaine situation par les membres de la communauté linguistique<sup>23</sup> (Yorio, 1980 : 434). Yorio met en avant la notion de compétence communicative qui comprend l'utilisation appropriée d'une expression dans une situation de communication donnée. C'est la capacité des locuteurs à dire la chose juste dans la bonne situation afin d'obtenir ce qu'ils veulent<sup>24</sup> (Yorio, 1980 :433). Galatanu (1997 :156) définit la conventionnalisation comme la « fixation dans la langue » des effets conversationnels réalisés par la performance de certains actes de langage. Ces rituels conversationnels peuvent être partagés par une même communauté mais peuvent également se réduire à un microgroupe professionnel, une famille, voire un couple.

Cette vaste catégorie des expressions conventionnelles comprend toutes sortes de formules associées aux nombreux actes de parole (saluer, se présenter, s'excuser, prendre congé, exprimer ses sentiments, féliciter, etc.) et enseignées en didactique. On trouve ainsi des formules, plus ou moins longues, allant des formules brèves du type « *Comment allez-vous ? Au revoir, à demain, tant pis, ah bon, tu parles ! Ne quittez pas ! Toutes mes félicitations, Excusez-moi* » à des phrases entières plus longues « *De quoi s'agit-il ? Je vous en prie ; Combien ça coûte ? Que voulez-vous dire ? Ravi de vous avoir rencontré* ». Galatanu (1997) inclut aussi dans les expressions conventionnelles certains proverbes ou dictons comme « *ça mange pas de pain, on n'est pas sorti de l'auberge, ne cherchons pas midi à quatorze heures, Pierre qui roule n'amasse pas mousse* ». Ces dictons très courants sont comme le résultat d'un processus de conventionnalisation, dû à une utilisation fréquente dans les conversations. « Ce qui apparaît comme « sens global », conventionnel, est une suite d'instructions permettant de construire la situation d'énonciation et la valeur communicationnelle de l'acte performé par l'énonciation. » (Galatanu, 1997 : 161).

Il faut cependant distinguer ce type d'expressions des formules de routine qui peuvent être a) des formules situationnelles qui peuvent varier en fonction des locuteurs en présence (âge, hiérarchie) « *puis-je vous aider ?* » vs « *t'as besoin d'aide ?* », b) des formules stylistiques, à l'écrit selon un certain registre (écriture académique, journal, lettre) ou à l'oral « *mesdames et messieurs*, c) des formules rituelles « *toutes nos condoléances* » (Yorio, 1980 : 436-37).

Les expressions conventionnelles (Bardovi-Harlig, 2009, 2012) sont aussi appelées *énoncés liés* (Fonagy, 1997), *holophrases* (Galatanu : 1984), *phrasèmes communicatifs* (Burger, 2010)

---

<sup>23</sup> « a certain form can be said conventional when it is predictable and expected by the members of a speech community in a certain situation. » (Yorio, 1980 : 434)

<sup>24</sup> « the ability of speakers to say the right thing, in the right situation, in order to get what they want. » (Yorio, 1980 :433)



ou *routines linguistiques* (Coulmas, 1979, 1981) ; Mel'çûk (1998 : 29) les appelle *pragmatèmes*. Ces termes techniques soulignent l'aspect pragmatique du phénomène, aspect fondamental, selon Schmäle (2013), qui va plus loin et pour qui toute unité phraséologique appelée *unité de construction préformée* a un ancrage pragmatique. Il rejoint en cela Yorio qui classait déjà les expressions conventionnelles en deux grandes catégories : les expressions idiomatiques et les formules de routines, celles-ci pouvant être aussi idiomatiques (Yorio, 1980 : 435).

Les expressions préformées ne sont du reste pas seulement frappées d'une « empreinte idiomatique », elles portent également une empreinte pragmatique, c'est-à-dire qu'elles sont rattachées pragmatiquement à des contextes d'utilisation, leur réalisation adéquate nécessitant le respect de conditions séquentielles, contextuelles, sociales, stylistiques, etc. très spécifiques. Et cet « ancrage pragmatique » ne se limite aucunement aux phrasèmes communicatifs, par définition « contraints par rapport à une situation extralinguistique » (Mel'çûk, 2011 : 44), il touche de même, dans une mesure plus ou moins marquée, d'autres types d'UCP comme le souligne I. Mel'çûk (1998) : « toutes les expressions, même si elle sont entièrement compositionnelles sémantiquement et syntaxiquement, sont des pragmatèmes : elles sont non compositionnelles pragmatiquement »<sup>25</sup> (Mel'çûk, 1998 : 29). (Schmäle, 2013 : 36)

En effet, un certain nombre de phrases idiomatiques et, *a fortiori*, toutes les suites de mots sont le plus souvent associées à une situation pragmatique (voir Confais, 2005). Ainsi pour reprendre l'exemple de Schmäle, l'expression familière *tomber dans les pommes* n'est pas totalement neutre stylistiquement puisqu'elle est produite dans une situation particulière par des locuteurs utilisant un registre de langue familier. De même, on aurait pu faire entrer les proverbes dans cette catégorie des expressions conventionnelles.

#### 1.4.5 Les formules discursives

Cette catégorie regroupe différents types de séquences qui aident à organiser et à structurer le discours. Yorio (1980 : 438) distingue des formules de routines, les *gambits*, qui ont pour fonction l'organisation, appelés aussi « organisateurs » par Raupach (1984). Les gambits peuvent être a) des stratégies pour l'interaction, (*conversational gambits*), b) des formules d'ouverture d'un tour de parole (*opening gambit*), *je pense que...* ; *excusez-moi*, ou de clôture (*closing gambit*), *je dois y aller*, c) des formules qui créent la conversation, *à votre tour maintenant*, ou organisent la conversation (*organisational gambit*), *dans un premier temps*,

---

<sup>25</sup> Notre traduction de : « All ready-made expressions, even if they are wholly compositional semantically and syntactically, are pragmatemes : they are non-compositional pragmatically. » (Mel'çûk, 1998 : 29).

*pour conclure, non seulement...mais aussi...; et d'autres comme merci beaucoup, excusez-moi* qui peuvent apparaître dans une interaction.

Cette catégorie des formules discursives comprend donc des séquences qui permettent au locuteur d'élaborer, de construire son discours, et d'autres de le maintenir. Raupach (1984 : 114-137), dans sa typologie des formules discursives, fait une distinction entre :

a) les « organisateurs » qui aident à structurer le discours : *c'est pourquoi, et comme ça, d'abord, premièrement, je veux faire, c'est, il y a*

b) les « remplisseurs » qui aident à maintenir le discours tout en évitant les pauses : *je ne sais pas, comment dire, je crois, quelque chose comme ça*. Les remplisseurs, sortes de bouée de sauvetage (Dechert, 1984), aident le locuteur à maintenir la parole et certains chercheurs les comparent à des pauses lexicales.

On trouve également, dans cette catégorie des « formules discursives », des formules propres à un genre discursif a) à l'écrit : *il était une fois*, formule consacrée de l'ouverture d'un conte (qu'on a utilisée d'abord à l'oral, le conte appartenant à l'oralité), *veuillez agréer, monsieur, l'expression de mes salutations distinguées*, pour clore une lettre formelle, b) à l'oral, lors d'une interaction *ça marche, ça m'est égal, en ce qui concerne...*

Entrent dans la catégorie des expressions discursives des séquences qui comportent des vides lexicaux, sorte de « schémas lexicaux » ou « patrons » (*patterns*) qui peuvent recevoir différents items. Nattinger et DeCarrico (1992) les appellent les contraintes phrastiques (*phrasal constraint*) : *d'une part\_\_d'autre part\_\_, ce qui\_\_c'est\_\_, non seulement\_\_mais encore, d'un côté\_\_de l'autre\_\_, plus on est X plus on Y, l'(e/a)\_\_dernier(e)*. Bolly (2010) parle aussi de « cadres collocationnels », qui sont des séquences discontinues de deux unités lexicales, séparées l'une de l'autre par un mot graphique non défini à l'avance. En français, on a aussi « *un(e) \_\_ de* » comme dans *un genre de, une espèce de, un type de* ; ou la structure *trop + adj + pour (trop grand pour)*. Mais on pourrait aussi bien les considérer comme des constructions selon les principes de la *construction grammar* (Fillmore, Kay et O'Connor, 1988). Les constructions phrastiques de Nattinger et DeCarrico (1992) et les cadres collocationnels de Bolly (2010) nous conduisent tout naturellement à évoquer la grammaire des constructions.

#### **1.4.6 Les constructions**

Deux théories, la grammaire des constructions et la *pattern grammar*, rejettent le formalisme

descriptif et tentent de regarder exactement comment la langue fonctionne, selon les mots eux-mêmes. Elles relèvent des approches « *usage-based* » propres à la linguistique cognitive (chapitre 2).

L'idée générale de la grammaire des constructions (Goldberg, 2005), dite parfois constructionnelle, est que les constructions elles-mêmes sont porteuses de sens, quels que soient les mots qui les composent (Legallois et Tutin, 2013 :15). Les *pattern grammars* qu'on pourrait traduire par « grammaire des modèles » établissent également un lien entre lexique et grammaire (Hunston et Francis, 2000). Selon la *pattern grammar*, chaque mot possède ses propres propriétés grammaticales, c'est-à-dire qu'il contraint et détermine les types de mots qui vont le précéder ou le suivre. On obtient un modèle, un schéma associé à un sens. Par exemple, le verbe va déterminer les propriétés sémantiques et grammaticales des mots qui l'entourent ; l'énoncé qui sera construit autour de ce verbe sera, en effet, assez lourdement contraint par le verbe lui-même. Un verbe peut avoir plusieurs modèles (*pattern*) et un modèle peut être associé à différents verbes. Le sens du mot ne peut pas être traité indépendamment de son « pattern » ou modèle. En réception, on pourra reconstituer le sens de la phrase même si on n'a pas entendu des mots à partir des modèles dans lesquels ils sont utilisés.

Nous sommes d'accord avec la conclusion de Sinclair pour dire que « pattern » et sens sont indissociables. Il s'ensuit que la plupart des mots n'ont aucun sens hors du contexte ou du moins sont très ambigus; ils n'ont de sens qu'au sein d'un contexte particulier. La phraséologie ou « pattern » est donc essentielle pour déterminer le sens d'un mot et ne peut pas être dissociée.<sup>26</sup> (Hunston et Francis, 2000 : 270)

Une construction peut véhiculer un sens « sémantique » mais aussi un sens pragmatique ou informationnel, qui n'est pas entièrement déductible de ses parties (Goldberg 1995 : 1). Par exemple, la construction schématique à deux compléments : *donner* N-objet à N, *écrire* N-objet à N, véhicule un sens de transfert (Croft et Cruse, 2004). Mais d'autres verbes comme *glisser* (verbe de mouvement) peuvent suivre ce schéma lorsqu'ils sont utilisés dans un sens métaphorique : *Pierre a glissé un papier à Marie*.

On peut faire entrer dans la catégorie des constructions, les phrases décrites comme des constructions figées qu'on appelle aussi constructions substantielles (*substantive constructions*) a) dans lesquelles on ne peut faire varier aucun élément comme dans *couper les ponts* (*il a coupé les ponts*) qui n'est rien d'autre qu'une phrase idiomatique, b) *tu peux*

---

26 Notre traduction de : « We agree with Sinclair's conclusion that pattern and meaning are associated. It follows from this that most words have no meaning in isolation, or at least are very ambiguous, but have meaning when they occur in a particular phraseology. The phraseology, or pattern, is therefore central to the meaning of the word and cannot be treated as separate from it. » (Hunston et Francis, 2000 : 270).

*parler !* » Une expression conventionnelle, complètement figée, c) *la mayonnaise n'a pas pris*, une sorte de vérité générale.

Les constructions se situeraient sur un continuum en fonction de leur plus ou moins grande schématicité / substantialité. Ainsi *plus* SN SV, *moins* SN SV (par ex. *plus* il parle, *moins* on l'écoute), constitue une construction en (grande) partie schématique, mais pas totalement vu qu'elle met en jeu, en plus des variables SN et SV, deux constantes, *plus* et *moins*. (cf. table 6.1 de Hoffman, 2011 : 228). Dans ces exemples, on peut une fois encore constater des recoupements avec certaines catégories, comme les *contraintes phrastiques* de Nattinger et DeCarrico (1992).

L'avantage de la notion de construction est qu'elle permet de ne pas séparer de façon arbitraire syntaxe et sémantique (distinction forme/sens) et qu'elle permet aussi de concevoir sur un même continuum les trois domaines que sont la syntaxe, la morphologie et la sémantique.

Toutes ces catégories auraient pu être présentées selon divers axes. Tantôt un axe linguistique qui comprendrait les expressions idiomatiques et les collocations, un axe pragmatique (proverbes, routines discursives) et un axe grammatical (constructions). Et on pourrait tout aussi bien considérer les proverbes comme faisant partie de l'axe linguistique, d'un point de vue purement lexical.

### **1.5 Choix du terme « préfabriqué »**

Notre étude s'intéresse aux suites de mots dans un corpus de français langue étrangère, à leur acquisition et au rôle qu'elles jouent en production orale. Il s'agit, bien évidemment, d'une toute petite palette du langage préfabriqué qui peut comprendre des séquences telles qu'elles ont été définies dans la section précédente mais aussi un grand nombre de séquences propres à la L2 et même à un seul locuteur de cette L2. Peu d'études sur l'acquisition du LP en FLE ont été menées, excepté, à notre connaissance et à ce jour, celles de Raupach, 1984; Towell, Hawkins et Bazergui, 1996 ; Towell, 2002 ; Forsberg, 2006 ; Cordier, 2013.

C'est pourquoi, parmi tous les termes pour désigner le phénomène, nous avons choisi le terme de séquence préfabriquée, traduction de *formulaic sequence* (Wray, 2002, 2008), proposée par Forsberg en 2006. La définition du terme est liée d'office à l'acquisition du langage préfabriqué en L2 plus que sur la description de celui-ci. Voici comment Forsberg explique ce choix terminologique :

Les 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. La raison est qu'une séquence préfabriquée ne suit pas les critères stricts de figement mais qu'elle peut être considérée comme préfabriquée pour d'autres raisons, comme par exemple sa complexité par rapport au reste de la production (chez des apprenants débutants) ou sa force pragmatique (des formules de politesse ou des marqueurs de discours. (Forsberg, 2006 : 3)

Le critère de figement n'est donc pas le seul critère pour identifier une séquence préfabriquée. On peut ajouter que des séquences sémantiquement transparentes et syntaxiquement régulières peuvent être potentiellement préfabriquées (Wray, 2000 : 446).

Pour conclure, nous dirons qu'il est bien difficile de bâtir une typologie des catégories de séquences. Les catégories se recoupent et les noms peuvent changer selon les chercheurs. Cependant, ces travaux qu'ils soient formels ou fonctionnels décrivent les SP d'un point de vue externe au locuteur. Ils ne nous disent pas comment elles sont acquises ni quel rôle elles jouent dans le traitement langagier en L1 et en L2. Les chercheurs examinent le langage préfabriqué non plus seulement comme une réalité linguistique mais aussi psycholinguistique dans la mesure où le LP jouerait un rôle dans le langage.



# Chapitre 2 Traitement, Acquisition et Identification des séquences préfabriquées

---

## 2.1 Rôle des SP dans le traitement langagier

De nombreuses études d'ordre psycholinguistique démontrent que les séquences préfabriquées jouent un rôle dans le traitement du langage aussi bien chez les locuteurs natifs que chez les non-natifs. Les résultats de ces études ont montré que la connaissance des séquences préfabriquées facilite la compréhension et la production (Nation et Webb, 2011 : 175) et permet le traitement du langage (réceptif et productif) de façon plus aisée, que ce soit en langue maternelle ou en langue étrangère (Kuiper et Hago, 1984 ; Wray, 2000, 2002 ; Wray et Perkins, 2000 ; Conklin et Schmitt, 2008). Même si la plupart des études entreprises à ce jour se sont plutôt intéressées à un type de SP, les expressions idiomatiques.

En effet, comme les SP semblent être stockées comme un tout dans la mémoire à long terme (Pawley et Syder, 1983 ; Bybee et Scheibman, 1999 ; Schmitt et Carter, 2004 ; Peters 1977, 1983), conformément à la définition psycholinguistique de Wray (2002 : 9), leur connaissance permet aux locuteurs une plus grande aisance à l'oral. En production, les séquences préfabriquées stockées en bloc, peuvent être récupérées avec un minimum d'effort, ce qui facilite la production en temps réel (Underwood, Schmitt et Galpin, 2004 : 153). En compréhension, le destinataire ne peut pas analyser pas chaque mot qu'il entend. Mais même s'il n'entend pas tous les éléments d'un énoncé, il peut en reconstruire le sens, grâce aux mots qui l'entourent et au contexte de l'échange<sup>27</sup> (Hunston et Francis, 2000 : 270). Comme il existe des expressions préfabriquées pour un grand nombre de concepts et d'actes de parole, le locuteur est capable de les récupérer directement de la mémoire à long terme facilitant ainsi le travail de décodage (en réception) et d'encodage (en production) - libérant du temps pour d'autres tâches comme l'interprétation du sens ou la planification d'unités plus larges<sup>28</sup> (Pawley et Syder, 1983 : 192).

Le LP présenterait donc un grand avantage, celui de limiter la charge de la mémoire de travail

---

<sup>27</sup> « [...] we can afford to mishear some words in a dialogue, for example, because we can reconstruct their meaning from the pattern they are used in. » (Hunston et Francis, 2000 : 270).

<sup>28</sup> « [in the store of familiar collocations] there are expressions for a wide range of familiar concepts and speech acts, and the speaker is able to retrieve these as wholes or as automatic chains from the long-term memory : by doing this he minimizes the amount of clause-internal encoding work to be done and frees himself to attend to other tasks in talk-exchange, including the planning of larger units of discourse. » (Pawley and Syder, 1983 : 192)

(MdT) qui ne peut traiter que quelques items à la fois (Miller, 1956 ; Baddeley, 2000). En effet, si l'on devait décoder les expressions mot à mot, la charge cognitive sur la mémoire de travail serait trop lourde.

Selon Conklin et Schmitt (2008), Jian et Nebraskova (2007), Miller (2011), les SP semblent être traitées plus aisément que le langage non préfabriqué. Les expériences de Schmitt, Grandage, et Adolphs (2004), Schmitt et Underwood (2004) et, celle de Underwood, Schmitt et Galpin (2004) que nous présentons ici, sont parmi les premières à s'intéresser à la représentation et au traitement des SP.

En compréhension, les recherches récentes étudiant les mouvements oculaires (*eye-tracking*) chez des locuteurs natifs ont montré que ceux-ci lisent plus rapidement les SP. Mais l'expérience de ces chercheurs s'est surtout focalisée sur le traitement des expressions idiomatiques, et en particulier les *figuratives* (Grant et Bauer, 2004), qui peuvent être utilisées dans un sens métaphorique (*hit the nail on the head, frapper en plein dans le mille*) ou un sens littéral (*enfoncer le clou dans la tête*).

Underwood, Schmitt et Galpin (2004) constatent que les séquences, qu'elles aient un sens littéral ou un sens métaphorique, sont lues plus rapidement que les phrases contrôles (ayant seulement un seul sens). Ils utilisent en particulier la technique des mouvements oculaires pour mesurer le nombre et la durée des pauses fixes de l'œil sur le dernier mot des vingt séquences choisies. Ils se sont aperçus que les locuteurs natifs fixent avec l'œil moins souvent et moins longtemps sur le dernier mot des SP que s'ils lisaient ces mêmes mots dans un contexte non préfabriqué. Ces résultats semblent montrer que les expressions idiomatiques sont représentées et traitées de manière holistique par les natifs (Underwood *et al.* 2004 : 167). Les séquences joueraient donc un rôle dans le traitement du langage en le rendant plus efficace. Cependant rien ne dit à quel sens, littéral ou idiomatique, les lecteurs ont accès en premier, ce qui corrobore les conclusions de Gibbs *et al.* (1997) (cité par Sinayova et Schmitt, 2008 : 433)<sup>29</sup> et celles de Conklin et Schmitt (2008) plus récemment. En d'autres termes, les séquences préfabriquées présenteraient un avantage de traitement pour les natifs comme pour les non-natifs quel que soit leur sens, idiomatique ou littéral (Conklin et Schmitt, 2008 : 85). Si le traitement holistique est avéré pour les natifs, les résultats ne semblent pas aussi nets pour les locuteurs non-natifs car rien ne permet d'affirmer, en effet, si le traitement se fait de

---

<sup>29</sup> « For example, Gibbs, Bogdanovich, Sykes, and Barr (1997) show that idiomatic expressions in story contexts (*He blew his stack*) are processed just as quickly as their literal paraphrases (*He got very angry*) but faster than control phrases (*He saw many dents*). » (Sinayova et Schmitt, 2008 : 433)



manière holistique.

Signalons également le paradigme expérimental de l'amorçage lexical (*lexical priming*), pendant lequel on peut mesurer le temps de réaction à un stimulus verbal. Dans une expérience en anglais langue maternelle (Hoey, 2005), les sujets doivent choisir entre les deux réponses proposées : *disputes* (B) ou *meanings* (B') pour compléter la phrase *the dictionary editor tried to resolve* (A). Le digramme AB ayant une plus forte probabilité d'apparition que AB', le sujet mettra plus de temps pour réagir sur B' en AB'. Le cerveau étant capable de traiter l'information de façon à estimer rapidement les probabilités lexicales d'apparition, l'œil est moins fixé lorsque le mot attendu est le bon. Hoey considère que l'emploi d'un mot est en quelque sorte déterminé, « amorcé », par son environnement qui peut être lexical, mais aussi sémantique, syntaxique, pragmatique et/ou discursif.

L'expérience en oculométrie et celle de l'amorçage lexical montrent que le traitement d'une SP en compréhension demande moins d'attention que celui d'une expression libre, créée mot à mot (Conklin et Schmitt, 2008 : 72). Cependant, elles ne s'intéressent qu'à un type de SP, les expressions idiomatiques et, seulement à l'écrit. Elles n'apportent pas la preuve que ces séquences sont stockées comme un tout dans le lexique, bien que Underwood, Schmitt et Galpin (2004) l'affirment. Elles ne nous disent pas ce qui se passe pour des locuteurs non-natifs en situation de production orale.

Si ces études se sont focalisées sur le traitement des expressions idiomatiques d'une manière générale, une autre étude, plus récente, s'est intéressée au traitement des expressions conventionnelles (Edmonds, 2013). « Une fois un acte de langage reconnu (en production ou en compréhension), toute expression conventionnelle pouvant réaliser l'acte en question serait activée, ce qui expliquerait les avantages psycholinguistiques associés à de telles séquences, avantages qui se limiteraient aux contextes déclencheurs » (Edmonds, 2013 : 130). Si celles-ci jouissent d'avantages psycholinguistiques, ils se manifesteraient différemment chez les natifs et les non-natifs d'après Edmonds (2013). En fait, Edmonds nuance la thèse d'un traitement holistique « indiscutable », tout comme Wray (2002) qui évoque seulement la « possibilité » d'un traitement holistique des SP. Edmonds donne une autre explication à la rapidité de traitement des expressions conventionnelles, qui semble confirmée par les résultats de l'expérience qu'elle a menée et qui concerne le traitement des expressions conventionnelles. En L2, par exemple, une grande partie de l'apprentissage est basée sur la répétition de séquences permettant de réaliser un grand nombre d'actes de parole (s'excuser, se présenter,

s'informer, etc.). A force de les avoir entendues et répétées, il peut se produire « une plus grande automatisation dans l'appréhension des séquences fréquentes – phénomène qui peut être tout-à-fait indépendant d'une inscription mémorielle » (Edmonds, 2013 : 133-34). Si Edmonds ne nie pas le traitement holistique des expressions idiomatiques non-compositionnelles, les expressions conventionnelles, en revanche, bénéficieraient plutôt d'un traitement automatique.

## **2.2 LP et approches basées sur l'usage**

La question qui nous intéresse est de savoir comment se fait l'acquisition des SP. Récemment, les recherches dans diverses disciplines des sciences cognitives ont montré que les acquis linguistiques émergent des formes linguistiques entendues dans des interactions sociales (Beckner *et al.* 2009 : 2). Cette position théorique est à contraster mais non pas à opposer à l'hypothèse de la période critique, chère aux générativistes. Les recherches en acquisition montrent, de manière générale, que l'input est primordial pour l'acquisition du langage et la structuration du discours et on le constate en L1 chez les enfants (voir Ellis, 2002).

Dans cette thèse, j'adopterai une perspective sur l'acquisition du langage qui est « *usage-based* » (basée sur l'utilisation du langage). Les structures du langage ainsi que le savoir se construisent à travers des milliers d'interactions. Ces processus ne seraient pas indépendants les uns des autres mais des facettes d'un « Système Complexe d'Adaptation » (*Complex Adaptive System*), (Beckner *et al.* 2009 : 1-2). Ce système nécessite qu'un certain nombre de facteurs soient réunis : a) le système comprend de multiples agents (les membres de la même communauté linguistique interagissent entre eux), b) le comportement langagier des locuteurs se nourrit des interactions passées qui, elles-mêmes, évoluent grâce aux interactions en cours, ce qui contribue à nourrir les comportements langagiers à venir, c) des facteurs aussi bien mécaniques que sociaux influent sur le comportement du locuteur, d) les structures du langage émergent des expériences personnelles, des interactions sociales et des processus cognitifs. Le langage serait donc adaptable et évolutif. Ce système s'inscrit radicalement à l'opposé des principes de la grammaire générative.

Dans les modèles de parole basés sur l'usage comme ceux de Langacker (1987, 1988, 2000) et Bybee (1985, 1995) qui sont, en réalité, très proches des grammaires de construction (Croft, 2000), les compétences linguistiques, directement liées aux situations de la vie, se renforcent grâce à l'acquisition d'un stock de structures ou unités symboliques, *structured inventory of*

*symbolic units*. En effet, à force d'entendre les mêmes expressions dans les mêmes situations, on constate un renforcement de ces compétences (Ellis, 2002). Pour chaque situation, le locuteur dispose d'un répertoire d'expressions et de concepts qui peuvent d'ailleurs prendre la forme de *scripts*, afin de rendre la communication plus aisée. Par exemple, le *script* de l'« achat de matériel » générera ses propres expressions, selon un scénario mettant en scène deux interlocuteurs, en fonction de leur statut, de leur âge et de l'ordre des énoncés de l'acte de parole, - entrée en matière, demande de prix, modalités de la garantie, délai de livraison - « *Combien coûte X ? Quel est le prix de X ?* ». Ces ensembles lexico-pragmatiques stockés en mémoire allègent fortement la charge cognitive que constituerait de mettre en place à chaque interaction de nouveaux *scripts*. Le langage se met en place, se construit dès le plus jeune âge en L1. Dans les théories basées sur l'usage, les êtres humains apprennent et utilisent des suites de mots, des unités polylexicales plus ou moins figées qui, bien qu'elles soient décomposables en unités plus petites, sont stockées dans la mémoire et récupérées comme des unités simples (Tomasello, 2000 : 62).

Des études, qui se sont intéressées à l'acquisition du langage, ont montré comment ces séquences sont acquises chez l'enfant dans sa langue maternelle. Au tout début, l'enfant veut comprendre le sens des expressions produites par les adultes. Puis, lorsqu'il se trouvera dans la situation de communiquer avec d'autres personnes, il essaiera de produire l'expression entière, même si souvent il ne peut en reproduire qu'un seul élément, par exemple « ça ! » pour dire « je veux ça ! » (Barrett, 1982). Les enfants utilisent un modèle constitué d'un mot ou d'un ensemble de mots, parfois même de morphèmes, produit en bloc comme structure d'une séquence liée à une situation particulière, à un acte de parole, qu'ils complètent un peu plus tard avec une variable qu'ils adaptent à la situation. Ces expressions-pivots sont caractéristiques dans l'apprentissage précoce de la langue, quelle que soit la langue. *Je veux X, plus de X, au revoir X, donne X*, etc. et jouent le même rôle que les holophrases (mot simple qui exprime une idée) (Tomasello, 2000 : 66).

Dans les théories grammaticales basées sur l'usage (*usage-based grammar models*), l'organisation cognitive du langage vient directement de l'expérience de chacun avec la langue. Au lieu d'être un ensemble abstrait de règles ou de structures en relation indirecte avec notre expérience, la grammaire est vue comme un réseau construit à partir des exemples (*exemplar-based*). Ces théories considèrent le langage comme un ensemble de constructions, composées d'unités qui associent une forme à un sens. On peut les ranger du plus spécifique

(niveau du lexique : mots et expressions) au plus général (niveau de la grammaire : construction passive, verbes à double objet), et des plus petites unités (mots et morphèmes) aux propositions ou même au niveau des unités du discours (Beckner et *al.* 2009 : 5). Dans ces modèles basés sur l'usage, l'acquisition de la grammaire se fait petit à petit grâce à l'apprentissage de milliers de constructions dans lesquelles on trouve des régularités (voir Clark et Kelly, 2006). Dans l'acquisition langagière en L1, l'enfant associe à ces formes régulières des fonctions. C'est à force de rencontrer les mêmes formes liées aux mêmes fonctions qu'il va acquérir ces structures, la fréquence ayant un impact déterminant pour leur acquisition. L'enfant finit par être conscient de ces régularités structurales et les analyse à tous les niveaux, de la phonologie au discours en passant par la syntaxe. La grammaire n'étant pas un système figé mais un système dynamique qui prend forme dans l'interaction humaine (Ellis, 2002 : 145), ce qui signifie que c'est en communiquant que se forment et s'établissent des régularités qui peuvent aussi se modifier et évoluer.

### **2.3. La fréquence et le processus de chunking**

Selon Bybee et McClelland (2005), Bybee (2010), Ellis (2002), la fréquence d'emploi renforce les représentations cognitives en mémoire. Répéter les mêmes séquences permet donc fluidité et mémorisation. Ces séquences, entendues et produites fréquemment, vont être traitées non pas mot à mot mais en bloc comme une simple unité. Néanmoins, une fois le modèle mémorisé, il sera plus difficile de le modifier. Depuis les années 1950, en psychologie cognitive, on parle de *chunking* - du regroupement en mémoire d'unités plus simples, un phénomène de base de la cognition humaine (Miller 1956 ; Bybee, 2010 : 34). Selon Bybee (2010 : 25), les regroupements (*chunking*) sont un processus cognitif par lequel des mots, utilisés fréquemment ensemble, finissent par entrer en relation pour former des séquences d'unités plus grandes et plus complexes. Un chunk est donc un ensemble de mots, de morphèmes ou de phonèmes organisés qui vont se souder dans la mémoire en une unité plus large. Le chunking implique aussi la capacité de construire des structures récurrentes conduisant à une organisation hiérarchique de la mémoire (Newell, 1990 : 7, cité par Ellis, 2003 : 76). Comme la mémoire de travail est limitée, seuls 7 items (plus ou moins deux selon les individus) peuvent être traités à la fois (Miller, 1956) le chunking permet alors d'augmenter la capacité de manipuler des blocs plus grands (comportant plus d'informations).

La mémoire de travail permet de mémoriser des procédures qui permettront de stocker des unités ou chunks dans la mémoire à long terme.

Lorsqu'on s'apprête à retenir un ensemble de plusieurs chiffres par exemple, il est plus facile de les grouper par deux ou trois pour les maintenir en mémoire. Par exemple, pour se rappeler un numéro de téléphone de plusieurs chiffres, dans les pays anglo-saxons on regroupe une série de neuf items 2 4 6 1 9 2 2 3 6 en trois blocs 246-19-22-36 ; en France, on a déjà des numéros « pré-chunkés » 04 93 33 35 40, ce qui facilite leur mémorisation qui correspond à une représentation ou « iconicité » c'est-à-dire « tout élément visuel ou auditif qui me permet de mieux « chunker » une série de mieux la percevoir comme une « bonne forme » » (Confais, 2015 : 33). Selon cet auteur, si on y trouve aussi des répétitions, des rimes (33 35), c'est encore plus facile. Et si on arrive à sémantiser (33= Gironde 35= Ile-et-Vilaine), on peut alors être quasiment sûr de le mémoriser. Dans la langue, on fait la même chose ; on constate que certaines expressions sont des regroupements de mots dont la mémorisation est facilitée de par leur iconicité (constructions en X et Y *bel et bien, par monts et par vaux* ; mais aussi par des groupements de sonorités (Ellis, 1996), assonances et allitérations, *au fur et à mesure, tout à coup* (voir Confais, 2015).

Selon Ellis (2002), l'apprentissage par association d'unités est central en acquisition L1 mais aussi L2. C'est à force de les rencontrer, de les entendre et de les utiliser en cours ou en immersion que l'apprenant va mémoriser les séquences de mots propres à la langue cible et essentielles à son utilisation efficace. C'est grâce aux processus de chunking que les séquences préfabriquées seraient acquises. En début d'apprentissage, par exemple, l'apprenant va tout de suite comprendre dans quelle situation il peut utiliser la séquence « *je m'appelle* » avant même de pouvoir la segmenter et l'analyser. Comme chez le petit enfant en L1, les blocs préfabriqués seront utilisés par l'apprenant de L2 comme stratégie de production. Les séquences préfabriquées ont donc à voir avec les modèles théoriques basés sur l'usage, usage et représentation étant liés.

A part le chunking, on constate le phénomène de base des regroupements, résultant de l'effet de fréquence : la réduction phonologique sur des mots employés très fréquemment. Certaines syllabes des ensembles utilisés très fréquemment sont articulées de manière plus rapide. Ce phénomène est à associer au processus du chunking. A force de produire des chunks, on fait moins attention aux unités qui composent le groupe, ce qui entraîne ces réductions. Pour Bybee et Scheibman (1999), les réductions phonologiques résultent d'un processus de

grammaticalisation : par exemple en anglais, la réduction de *going* en *gonna* comme marqueur du futur proche ou de *I don't know* en *dunno*. Mais on trouve aussi en français des formes réduites, *je ne sais pas* en *chépa*, *il y a* en *y'a*, résultat de la fréquence d'emploi de ces groupes très usités à l'oral (voir Regan, 1997 ; 2002). En L2, l'apprenant qui a entendu fréquemment ces expressions va finir par les assimiler et les répéter dans cette forme, pour se rapprocher le plus du parler natif. Et plus le niveau est avancé, plus on constate d'occurrences avec réduction phonologique (chapitre 8).

Regan (1995, 2002) s'est intéressée au rôle de la négation « ne » en français et, sa disparition dans le groupe verbal [(ne) + V + pas] en tant que variable sociolinguistique. Dans cette étude, Regan montre qu'après un séjour d'un an à l'étranger, la réduction phonologique chez les locuteurs L2 augmente sensiblement par imitation avec les locuteurs natifs et sans doute par leur exposition prolongée dans la langue cible. Elle constate cependant d'importantes variations individuelles chez ces locuteurs L2, comme on peut aussi en trouver chez les locuteurs L1. La réduction phonologique serait aussi un bon indice de la nature préfabriquée d'un ensemble lexical (Théophanous et Perez-Bettan, 2012).

Nos habitudes langagières dépendent donc de notre expérience avec le langage. Cela veut donc dire que les représentations pragmatico-linguistiques varient d'un locuteur à l'autre mais aussi selon le groupe auquel il appartient. En L2, par exemple, tout dépendra de l'exposition de l'apprenant, de l'input reçu. Il en est de même avec les SP qui varieront en fonction des parcours langagiers de chacun, de son expérience avec la langue. Dans la situation courante d'un moindre contact avec les langues étrangères, par exemple, l'apprenant va probablement développer des séquences qui lui sont propres et qui seront différentes de celles d'un autre apprenant, même s'ils partagent un grand nombre de séquences canoniques aussi. Le langage dépendrait donc, pour une grande partie, des représentations linguistiques que chaque individu se fait au cours de son existence. Des constructions - des expressions non compositionnelles, des collocations plus ou moins fixes mais aussi des séquences régulières comme *indiquer le chemin* - sont mémorisées comme des « exemples » en mémoire parce qu'elles correspondent au vécu des locuteurs. Bien sûr, le revers de la médaille, c'est que notre langage est plutôt répétitif et parfois sans grande originalité, puisqu'on répète à l'oral souvent les mêmes structures, les mêmes modèles « exemplaires ». Mais c'est justement grâce à ces répétitions que des constructions s'établissent et se fixent.

Selon Bybee (2010), le lexique et la grammaire sont complètement imbriqués, la relation forme-sens étant, on l'a vu, au centre de la grammaire des constructions. Cette conception de la langue s'oppose à celle des générativistes qui s'intéressent surtout aux règles permettant une créativité langagière infinie. Entre ces deux conceptions extrêmes, on peut dire, sans risque, que les locuteurs d'une L1, s'ils utilisent un grand nombre de séquences toutes faites, produisent aussi des énoncés en combinant les mots et en appliquant des règles.

## **2.4 L'acquisition des SP en L1 et en L2**

Les théories basées sur l'usage soutiennent que le langage se construit grâce aux interactions par un ensemble de processus communicationnels et cognitifs. De même, les processus d'acquisition des SP, en L1 et L2, sont fortement marqués par les facteurs d'exposition à la langue et son corollaire, la répétition. Ce n'est qu'après un temps d'exposition et de répétition que les apprenants analysent les séquences mémorisées et peuvent en extraire des règles qu'ils pourront utiliser de manière productive (Wong-Fillmore, 1976). Comme on constate ce même processus en L2, on peut établir une similitude entre l'acquisition des séquences chez les enfants et les locuteurs non-natifs en début d'apprentissage.

### **2.4.1 L'acquisition des SP en L1**

Peters (1983), dans une étude très importante, distingue deux types d'apprenants en L1 qui correspondent à deux styles d'apprentissage selon l'input reçu : le style gestalt ou holistique et le style analytique. Le premier fait référence à l'utilisation de suites de mots mémorisées et produites en bloc sans passer par une analyse préalable ; le second concerne la construction d'énoncés à partir d'items plus petits en procédant par analyse. Bien évidemment, les deux styles sont complémentaires. On ne peut trouver de productions orales constituées exclusivement de séquences préfabriquées ni de production n'en utilisant aucune. Toutefois, on note que la préférence pour l'un ou l'autre de ces styles dépend bien souvent de l'exposition et de l'objet. Selon Peters (1983), les enfants font appel au style analytique lorsqu'ils parlent des choses du monde et au style gestalt lorsqu'ils sont en situation de communication, en interaction. Les variations entre les enfants proviendraient de ce qu'ils entendent, du style prédominant dans l'entourage. L'enfant « analytique » a reçu plus d'*input* en termes de langage construit mot à mot, alors que l'enfant « gestalt » a sans doute reçu plus d'*input* conversationnel et de phrases construites en un bloc. Si les parents communiquent



avec des mots ou avec des suites de mots, le développement langagier de l'enfant risque d'être différent. Les enfants peuvent, bien sûr, passer d'un style à l'autre selon les situations. Mais selon le type de langage entendu, le style de l'enfant tendra vers l'un ou l'autre.

Comment les SP se mettent-elles en place en L1 ? Peters (1983 : 5) a étudié le rôle des unités de la langue parlée (*speech units*) dans la production orale des enfants. Elle constate qu'ils apprennent, en même temps que les mots, de petites phrases de façon « holistique » qu'ils arriveront peu à peu à segmenter pour pouvoir faire varier les constituants.

L'enfant n'est pas exposé à un dictionnaire de morphèmes mais plutôt à un flux discontinu de sons dont certains forment des unités (*chunks*) souvent plus longues qu'un simple mot et qui reviennent plus ou moins fréquemment ; l'enfant doit arriver à saisir dans ce flux de sens et de structures inconnues quelques fils dont il devra trouver la signification et qu'il devra conserver pour un usage futur. (Peters, 1983 : 5)

Quelques exemples permettront de mieux comprendre les processus d'acquisition chez l'enfant dans sa langue maternelle. Tout se passe entre l'enfant et son entourage. L'enfant va imiter les mots ou phrases qu'il entend. Par exemple, en L1, les enfants entendent le syntagme *wannago* (*voudrais aller*) dans certaines situations récurrentes. Après un certain temps, ils ont conscience que d'autres phrases du même type, telles que *wannaplay* (*voudrais jouer*) et *wannaget* (*voudrais avoir*), sont utilisées dans d'autres contextes : ils vont donc commencer à analyser la phrase et en extraire un élément stable et un item interchangeable, *wanna* + V (*voudrais* + V) (Peters 1983 : 25).

Ce n'est donc pas seulement un phénomène d'imitation mais plutôt un processus implicite de généralisation ou d'analogie, qui consiste à extraire des ensembles « pivots » (Tomasello, 2000 : 66 ; Lieven *et al.* 2003) où une partie de l'expression reste totalement stable et une autre partie varie. Selon Tomasello (2000) de ces ensembles syntaxiques (à structure « pivot ») émerge la grammaire, c'est-à-dire la maîtrise des règles combinatoires de la langue. D'après Peters, ces expressions ou séquences préfabriquées, correspondent à une séquence plus longue et plus complexe que celle que l'enfant est susceptible de maîtriser à cette étape de l'apprentissage de la langue.

Dès 1973, Brown, dans une étude empirique de grande envergure, observe trois enfants (Adam, Eve et Sarah) qu'il va suivre pendant plusieurs mois. Ces observations le conduiront à établir des généralisations sur les différents stades d'acquisition du langage. Un de ses sujets, Adam (2 ans et 3 mois) ne peut encore construire des phrases mot à mot à son stade



d'acquisition<sup>30</sup> (Brown distingue cinq stades d'acquisition dont un premier stade entre 15 et 30 mois) mais mémorise correctement, à force de les avoir entendues, celles que ses parents produisent. Cependant, il ne peut encore les analyser. Adam produit correctement très tôt des cadres (*patterns*) comme *it's* et *that's*, acquis avant que le système grammatical ne soit mis en place et indépendamment d'autres constructions créatives. Selon Brown, toute forme produite fréquemment par les parents sera reprise par l'enfant même s'il n'est pas capable d'en comprendre la structure. Il trouvera toujours le moyen de la produire à un moment donné. Le processus d'acquisition est toujours le même, l'enfant passe d'abord par la phase de mémorisation de formules, puis de leur adaptation au contexte et enfin par l'analyse de leurs constituants.

Dans une étude déjà ancienne mais qui fait date, Clark (1974 : 9) constate, elle aussi, que l'enfant produit des formules apprises en bloc avant toute analyse. Le langage de l'enfant devient créatif lorsqu'il commence à en analyser les structures. On peut se demander si l'enfant analyse un jour ces structures, en dehors de l'école où il le fera à un degré plus ou moins important, selon sa L1, sa culture. Selon l'hypothèse de Tomasello (2003), mais aussi de Hilton (2009) la grammaire émergerait suite à un processus implicite de catégorisation (généralisation). Pour reprendre le sujet de Brown (1973), Adam, alors âgé de trois ans (stade III), entend fréquemment *wait for it to cool* (*attends que ça refroidisse*) quand un plat chaud est servi à table. Après plusieurs semaines, Adam utilisera à bon escient l'expression *wait for it to dry* (*attends que ça sèche*) lors de l'étendage du linge. Ainsi l'enfant est capable, après un certain temps d'exposition, d'utiliser et d'adapter à d'autres situations les longues expressions ou phrases qu'il a entendues et qui deviendront des *cadres* linguistiques (voir Brown et Hanlon, 1970). Si l'enfant apprend par imitation et par renforcement des expressions, il passe aussi par une phase créative. C'est en entendant les expressions varier en fonction de l'acte de parole que l'enfant comprend qu'il peut adapter la phrase à la situation. L'expérience de Lieven, Behrens, Speares et Tomasello (2003 : 333) consiste à analyser le discours d'un enfant de deux ans lors d'interactions avec la mère, enregistré pendant cinq heures par semaine pendant six semaines. Ils constatent que l'enfant utilise des formules entendues dont il modifie au moins un des constituants. Sur les 295 expressions enregistrées, 37% d'entre

---

<sup>30</sup> stade I : 15-30 ( en mois)  
stade II : 28-36  
stade III : 36-42  
stade IV : 40-46  
stade V : 42-52 (d'après Brown, 1973)

elles étaient nouvelles au sens où elles n'étaient pas produites dans leur forme complète. Dans 74% des cas, il aura suffi aux chercheurs d'une seule opération pour voir qu'un seul mot a été substitué (un nom pour la plupart d'entre eux) et plus d'une opération pour les 26% restants. Dans la majorité des cas, ce que l'enfant disait était basé sur ce que Lieven, Behrens, Spears et Tomasello (2003) appellent des « schémas d'expression » (*utterance schema*), c'est-à-dire qu'une partie de l'expression reste fixe et l'autre varie en fonction du contexte.

Dans une autre expérience, Diessel et Tomasello (1999) examinent les expressions produites par sept enfants utilisant des formes verbales du type *je pense (I think)*. Ils constatent qu'à leur stade, ils produisent ces formes uniquement à la première personne, au présent, jamais à la forme négative ni suivies de complément et avec le sens de *peut-être*.

Ces études, liées aux « modèles basés sur l'usage », montrent que l'enfant en acquisition L1 passe d'abord par une phase de répétition et d'imitation des phrases entendues fréquemment qu'il reproduit dans une seule forme dans un premier temps puis qu'il adaptera au contexte ; cette dernière phase s'effectue après analyse et segmentation des constituants.

#### **2.4.2 Acquisition des SP en L2**

Les recherches qui concernent l'apprentissage de la langue maternelle chez les enfants permettent aussi de comprendre en partie comment fonctionne l'apprentissage en L2. Si les études en acquisition du langage montrent que les enfants en L1 utilisent des séquences et des phrases liées à une situation pour entrer en communication, les études en L2 d'enfants bilingues montrent que les enfants utilisent encore plus les SP. Comment se passe alors l'acquisition des SP chez les enfants un peu plus âgés apprenant une L2 ?

Hatch (1972) examine les productions de Paul, un enfant chinois de 4 ans, apprenant l'anglais aux États-Unis, en situation d'immersion. Paul semble utiliser les deux stratégies d'apprentissage en parallèle, ensembles plurilexicaux et phrases construites mot à mot. L'enfant s'appuie sur le langage préfabriqué pour entrer en communication, ce qui lui permet d'avoir un rôle social avec un minimum de compétence.

Hakuta (1974) rapporte le développement d'Uguisu, une Japonaise de cinq ans, apprenant l'anglais aux États-Unis. L'enfant apprend par mémorisation des suites de mots sans pour autant en connaître leur structure interne (Hakuta, 1974 : 287). Elle est capable de produire les phrases entendues dans des situations identiques sans aucune analyse préalable. En L2, lorsque l'enfant se trouve dans des situations qui demandent une interaction linguistique, à l'école, dans la cour, ou pendant les activités sportives, il utilise ces formules et ces patrons

linguistiques. Ainsi, avant même d'atteindre une compétence créative dans la langue, il a la capacité mentale de mémoriser les séquences en les associant à une situation et de les utiliser. Wagner-Gough (1978), dans une autre étude, montre que des phrases mémorisées peuvent côtoyer des structures minimales. Son sujet Homer utilise les formules uniquement en situation de communication. Ces *cadres* n'évolueraient pas directement en analyse grammaticale. On peut ainsi trouver à un certain stade de l'apprentissage les deux formes : *my name is Homer (je m'appelle Homer)* ou *he Fred (il Fred)*. En début d'apprentissage, des formules apprises en bloc correctement utilisées peuvent côtoyer des groupes créés mot à mot. Cependant, pour Krashen et Scarcella (1978), les deux chemins, langage routinisé ou langage créé mot à mot, seraient indépendants l'un de l'autre.

Vihman (1982) étudie le comportement de sa fille Virve, âgée de moins de deux ans. L'enfant entend ses parents parler uniquement l'estonien à la maison, leur langue maternelle, et elle entend l'anglais seulement au jardin d'enfants. Elle dispose alors de 260 mots isolés en L1, et de quelques suites de mots acquises en bloc en L2. Vihman constate qu'en L2, Virve utilise d'abord des expressions puis des phrases lexicalisées, alors que dans sa langue maternelle, elle semble créer son langage mot à mot. Du fait de son exposition à la langue cible, elle produit par répétition, dans les mêmes situations, les expressions entendues dans le jardin d'enfants, sans analyse car à ce stade elle n'en est pas capable. Vihman constate que les structures intonatives différentes des deux langues détermineraient pour une part le style de langage. Autrement dit, la structure intonative de la langue favoriserait le stockage des formules prêtes à l'emploi (MacWhinney, 1978 : 6-7). Une langue, où l'accent est régulier, sera beaucoup plus facile à segmenter et à mémoriser et favorisera ainsi la création de la phrase mot à mot.

Tous ces exemples semblent indiquer que l'input est essentiel pour l'acquisition du langage. Le petit enfant construit selon les modèles que produisent ses parents depuis la naissance, alors qu'au petit allophone à l'école maternelle on va parler comme à tous les enfants – en phrases et non pas en mots. L'enfant va donc extrapoler des ensembles plurilexicaux, car c'est sans doute quasiment tout ce qu'il entend.

Pour Krashen et Scarcella (1978 : 295-296), les enfants décrits par Wong-Fillmore ont acquis un répertoire de routines qui dépassent largement leur compétence linguistique, car ils se trouvaient dans des situations où ils étaient obligés de communiquer. Ces chercheurs se demandent si c'est la façon dont le langage est généralement acquis. En effet, l'*input* n'est pas

toujours constitué de formules. Tout dépend donc du contexte d'acquisition qui joue un rôle dans le processus d'apprentissage et dans l'acquisition des structures utilisées par la suite dans d'autres contextes linguistiques. L'enfant, en L1, en fonction de l'âge et de l'input, répétera des mots et/ou des expressions. Et en L2, la situation et l'input reçu en situation d'apprentissage seront d'autant plus essentiels dans l'acquisition des SP.

En situation, l'apprenant adulte comme l'enfant va « rentrer dans la L2 » en s'appuyant sur des ensembles plurilexicaux. Il apprend d'abord une formule en un bloc sans passer par la segmentation. Celle-ci se fera à un stade plus avancé lorsqu'il prendra conscience qu'une partie de la formule peut être utilisée avec d'autres unités dans d'autres situations. Pour Myles (2004), il y aurait même une phase préliminaire où les locuteurs L2 utilisent des expressions sans verbe, phase comparable à celle des petits enfants (*one-word stage*), ce n'est que dans la phase suivante qu'ils produisent des SP. Ellis (1997 : 11) ainsi que Schmidt (1992) ont remarqué que les apprenants, aux stades débutants, utilisent un grand nombre de formules (*formulae*) ou d'expressions figées sans pouvoir utiliser l'un de leurs constituants de manière productive. Ce n'est qu'à un stade plus avancé qu'ils pourront les adapter ou les segmenter. Dans les cours de langue, on l'a vu, la répétition des mêmes structures facilite leur mémorisation. En tout début d'apprentissage, dans les approches communicatives, l'accent est mis à l'oral sur la répétition de certaines structures, plus complexes que ce que l'apprenant est à même de produire et même avant d'être en mesure de les segmenter. Ce n'est que plus tard qu'interviendra la segmentation et notamment lors du passage à l'écrit. On ne se préoccupe pas des unités lexicales mais d'ensembles sémantiquement liés à une situation et qui font sens pour l'apprenant. Comme le dit Bybee (2008 : 216) « Notre bon sens nous dit que la répétition et la pratique sont essentielles pour le développement des structures cognitives qui conduisent à une production orale fluide et syntaxiquement correcte<sup>31</sup>. » De même, Tomasello (2000 : 62) parle d'ancrage de certaines expressions qui serait dû aux usages répétés de celles-ci dans certaines situations courantes. Cette idée de fréquence, déjà vue plus haut, est intéressante pour expliquer l'acquisition des séquences préfabriquées en L2 en réception mais aussi en production. Dans l'acquisition, la fréquence d'utilisation jouerait un rôle central dans le processus : répéter une unité conduit peu à peu à la fixer dans le cerveau (Bybee et Hopper,

---

<sup>31</sup> « Common sense tells us that for second language learners, repeated exposure and practice are essential to the development of the cognitive structures that lead to fluent and grammatical speech. » (Bybee, 2008 : 216)

2001, Ellis, 2002). L'utilisation fréquente de SP dans le discours favorise ainsi la routinisation de ces structures et une activation plus directe.

### **2.4.3 Influence de la L1 dans l'acquisition en L2**

Si l'acquisition des SP en L1 comme en L2 dépend chez l'enfant de l'input, il semble que chez l'adulte en L2, leur fréquence en production favorise aussi leur mémorisation. Cependant on peut se demander si le processus est identique en L1 et en L2. En effet, malgré des similitudes, la façon dont se fait l'acquisition en L2 est différente de celle de la L1. Chez les enfants, le langage se développe en même temps que le réseau conceptuel alors que, chez les adultes, ce réseau est déjà construit. L'apprentissage de la langue se fera donc de manière explicite. En L1, l'apprentissage se fait de manière naturelle ; en L2, tout un réseau linguistique (celui de la L1) est déjà présent et fortement ancré au réseau conceptuel. Les structures de la L1 vont influencer la L2 ; les représentations linguistiques déjà présentes peuvent avoir des effets positifs ou négatifs sur l'interlangue (Ellis, 2003 : 72). D'autre part, si les processus cognitifs sont semblables, les conditions (sociales, matérielles, temporelles) de l'acquisition en L2, chez un adulte, seront différentes de l'acquisition de la L1 chez les enfants, beaucoup moins de temps de contact, une moindre motivation sociale (dans beaucoup de cas), etc. La grande différence entre les enfants et les adultes est que les adultes ont déjà acquis une langue, qu'ils ont également des connaissances métalinguistiques qu'ils pourront mobiliser pour l'analyse et la mémorisation de la L2.

Pour Bybee (2008 : 232), si les « théories basées sur l'usage » (*Usage-Based Theories*) considèrent les connaissances linguistiques comme un ensemble de modèles ou cadres (*patterns*) automatisés, la première langue peut être envisagée comme une aide et une entrée pour l'acquisition de la L2. Dans la mesure où les schémas syntaxiques sont identiques, il faut seulement adapter le lexique et la phonologie. Dans le cas contraire, si les constructions sont trop éloignées de la L1 (cas des langues asiatiques), les constructions de la L2 auront tendance à être calquées sur celles de la LM. Et cela peut poser problème.

### **2.4.4 SP et apprentissage**

Une question que les chercheurs en acquisition d'une langue se posent est celle de savoir si l'acquisition d'énoncés utilisés de manière holistique peut entraîner une analyse subséquente de leurs parties. A notre connaissance, il existe peu d'études là-dessus. Une des plus célèbres est celle de Myles, Hooper et Mitchell (1998). Ces chercheurs ont étudié l'acquisition de deux

séquences en français L2 *j'aime* et *j'adore*. Sans avoir reçu un enseignement des séquences en *je + aime ; je + adore*, leurs élèves pouvaient les utiliser correctement à l'oral tant qu'il s'agissait de parler à la première personne. Ils se sont mis par la suite à dire *Monique j'aime* (pour *Monique aime*) et *Monique j'adore* (pour *Monique adore*) mais petit à petit et sans que le professeur n'explique quoi que ce soit, les élèves ont pu « casser » les séquences *j'aime* et *j'adore* en deux unités significatives (*je + aime/je + adore*). L'analyse reste donc possible même après avoir mémorisé une séquence comme une unité simple.

D'autres chercheurs reconnaissent le rôle des SP dans l'apprentissage. Pour Gaonac'h (2006 : 145), les formules servent de base à une généralisation de la fonction d'une forme grammaticale (morphèmes, ordre des mots, *etc.*). Pour Hagège (1996 : 62), les énoncés « tout faits » sont utiles du point de vue de la communication mais sont surtout source d'apprentissage. Si un grand nombre d'auteurs comme Myles, Hooper et Mitchell (1998) disent que la préfabrication peut donner lieu à l'apprentissage des règles grammaticales, à la segmentation et au transfert de la préfabrication vers la grammaire, d'autres (les générativistes) disent que les séquences préfabriquées n'ont aucune incidence sur la grammaire et donc le transfert de la préfabrication vers la grammaire n'est pas possible (Krashen et Scarcella, 1978). Pour ces auteurs, ce sont deux systèmes indépendants, avec d'un côté un ensemble de formules qui évolue effectivement en structures, et de l'autre un processus de création du langage qui se développe au même moment et de façon indépendante. Selon Wray et Perkins (2000 : 36), les SP permettent l'activation directe d'ensembles langagiers qui sont combinés avec des procédures plus créatives.

Malgré tout l'intérêt du sujet, notre recherche ne se place pas du point de vue de l'acquisition, à la manière de Myles *et al.* (1998, 1999) qui étudient le transfert et la segmentation des SP vers la grammaire et la création de phrases nouvelles, libres. Nous voulons montrer quel rôle joue le LP dans le discours oral de sujets non-natifs et quel est le lien avec la fluence de leur production.

## **2.5 Définitions du LP**

Le langage préfabriqué joue, sans aucun doute, on vient de le voir, un rôle important, en L1 comme en L2. Cependant, ce terme regroupe une grande variété d'expressions, de séquences préfabriquées que le chercheur a du mal parfois à identifier. Aussi pour identifier correctement les SP, il est nécessaire d'en avoir une définition précise (Wray, 2008 : 94).

Cette section est consacrée tout d'abord aux définitions des séquences préfabriquées, d'un point de vue fonctionnel, puis à leur identification à l'oral.

On trouve dans la littérature diverses définitions des SP, selon le point de vue adopté du chercheur. Soit on se place d'un point de vue externe au locuteur et on se contente de décrire de manière formelle les phénomènes phraséologiques à l'oral ou à l'écrit, sans lien avec ce qui se passe au niveau cognitif (chapitre 1) ; soit on s'intéresse au rôle des SP dans la production et on se place d'un point de vue interne au locuteur (Wray, 2008 ; Cordier, 2013).

Les définitions les plus connues, axées sur la fonction du LP, expliquent et définissent le phénomène d'un point de vue pragmatique ou psycholinguistique. Les psycholinguistes et les sociolinguistes abordent la préfabrication, entité qui peut regrouper divers phénomènes en plus de ceux évoqués dans le chapitre 1, du point de vue du locuteur.

La définition d'Erman et Warren (2000 : 31), souvent mise en avant, met l'accent sur la préférence native : « Une séquence préfabriquée consiste en une combinaison d'au moins deux mots que les locuteurs natifs préfèrent utiliser plutôt qu'une autre équivalente si la conventionalisation n'existait pas »<sup>32</sup>. Dans cette définition, la suite de mots « préfabriquée » (*formulaic*) est une formule conventionnelle que les natifs emploient de préférence à d'autres équivalentes. Ainsi l'expression *I'm afraid (je crains de)* utilisée pour atténuer une mauvaise nouvelle ne peut être remplacée par un équivalent littéral *I am scared (je suis effrayé)*, sans que le sens et sa fonction ne soient modifiés. En effet, les séquences produites en situation de communication sont souvent des séquences choisies par les locuteurs natifs parmi d'autres tout aussi valables linguistiquement mais non-conventionnelles. L'accent est mis dans cette définition sur l'aspect pragmatique des SP. Pawley et Syder (1983) déjà, mettaient en avant les particularités de la sélection native (*native-like selection*). Si les locuteurs natifs possèdent un répertoire de séquences adaptées à des situations particulières (*tous nos meilleurs vœux de bonheur, bonne année, joyeux Noël*) en L2, les locuteurs non-natifs doivent se les approprier pour parler comme un natif le ferait. La définition d'Erman et Warren s'intéresse aux expressions conventionnelles en L1 mais ne dit rien des séquences produites en L2. Et même s'ils reconnaissent leur fonction pragmatique

---

<sup>32</sup> Notre traduction de : « A prefab is a combination of at least two words favoured by native speakers in preference to an alternative combination which would have been equivalent had there been no conventionalisation. » (Erman et Warren, 2000 : 31).



dans le discours, Erman et Warren ne disent rien du rôle de ces expressions dans le traitement langagier.

Forsberg (2006 :44) adapte la définition pragmatique d'Erman et Warren dans sa recherche des séquences préfabriquées de locuteurs non-natifs de français langue étrangère.

Une SP est une séquence polylexicale (deux unités minimum) préférée des locuteurs natifs à d'autres moins fréquentes, tout aussi correctes linguistiquement mais non-conventionnelles. Comme le répertoire des SP chez les locuteurs non-natifs est différent de celui des natifs, une séquence préfabriquée peut être une séquence préférée par un locuteur L2, reconnue par son emploi idiosyncrasique et sa fréquence dans l'interlangue d'un locuteur L2. (Forsberg, 2006 : 44)

Cette définition élargit la notion de la préfabrication à l'utilisation du langage préfabriqué en cours d'acquisition. Elle met l'accent sur les différences entre SP natives (attestées) et non natives (idiosyncrasiques).

Cependant, la définition la plus largement reconnue est celle de Wray (2002 : 9). Selon ce chercheur, une séquence préfabriquée (SP) est une « séquence continue ou discontinue de mots ou autres éléments qui sont ou semblent être préfabriquées c'est-à-dire mémorisées et extraites de la mémoire au moment de l'usage plutôt que d'être générées ou analysées par le système grammatical de la langue »<sup>33</sup>. Depuis, Wray (2008 : 12) a complété cette définition, en développant l'idée d'une construction équivalente à un morphème, *Morpheme Equivalent Unit* (MEU) qu'elle définit comme « un mot ou une suite de mots, complète ou comportant une partie fixe et une partie variable, traité comme un morphème, sans recourir à une quelconque relation forme sens que ses unités pourraient avoir »<sup>34</sup>. Ces deux définitions (Wray, 2002, 2008) d'ordre psycholinguistique renvoient à la structuration du lexique mental. Le locuteur semble activer les séquences telles quelles sans avoir recours à l'analyse grammaticale ni à la traduction mot à mot. La séquence est donc produite d'une manière holistique en bloc et sans avoir recours au sens de chaque élément. Elles semblent être préfabriquées, dit Wray. Ce n'est donc pas une affirmation mais une possibilité. On peut alors se demander si elles ont toutes été mémorisées vraiment comme des ensembles ou construites mot par mot par le locuteur. Rien dans cette définition ne nous permet de l'affirmer car Wray ne dit pas comment on apprend ni comment on mémorise ces ensembles. Cependant, ces

---

33 Notre traduction de « A sequence continuous or discontinuous, of words or other meaning 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. » (Wray (2002 : 9).

34 Notre traduction de « a word or word string, whether complete or including gaps for inserted variable items, that is processed like a morpheme, that is, without recourse to any form-meaning matching of any subparts it may have » (Wray, 2008 : 12).



définitions privilégient la thèse de la mémorisation holistique des séquences, qui semble de toute façon évidente pour les formules non-compositionnelles. Le locuteur est ainsi capable de récupérer un ensemble de mots en bloc de la mémoire à long terme (Pawley et Syder, 1983 : 192). Ces définitions concernent les SP en L1 mais on peut les appliquer à la L2. Des chercheurs plus centrés sur la didactique en L2 proposent aussi leurs définitions.

Si Wray définit les SP en L1 comme des suites de mots présentant certains avantages pour le locuteur parce qu'elles sont ou semblent stockées en bloc et récupérées telles quelles de son lexique mental, en L2 elles peuvent suivre un autre schéma d'acquisition. En effet, elles peuvent, par exemple, avoir été automatisées grâce à un emploi fréquent en cours de langue (jeux de rôle, répétitions) ou à force d'avoir été utilisées dans certaines situations (voir les modèles basés sur l'usage, section 2.2). Les définitions psycholinguistiques du LP concernent les séquences que la linguistique traditionnelle a longuement décrites mais aussi toutes les séquences propres au locuteur parce qu'elles font partie de son répertoire. Une séquence pour être reconnue préfabriquée devrait faciliter la production orale en L2, en permettant une meilleure aisance et une plus grande fluidité. Elle présente donc un avantage de traitement pour un locuteur donné (Cordier, 2013 : 21). A l'instar de Cordier (2013 : 84), la définition dans cette thèse sera psycholinguistique, à savoir qu'« une séquence préfabriquée est une suite de mots qui présente un avantage de traitement pour un locuteur donné soit parce qu'elle a été stockée comme un tout dans son lexique soit parce qu'elle a été automatisée<sup>35</sup> ». Cette définition nous paraît mieux adaptée à un corpus d'apprenants L2 que celle de Wray, car elle met davantage l'accent sur le rôle des SP dans le traitement sans qu'elles soient nécessairement stockées et récupérées de manière holistique. Une définition d'ordre psycholinguistique, tout à fait opérationnelle, comme celle de Cordier met l'accent sur le rôle des SP dans la production, *a contrario* de celle de la tradition phraséologique française qui reconnaît une séquence préfabriquée par des critères stricts de figement dans une perspective uniquement descriptive et externe au locuteur. Notre définition des SP dans cette recherche sera celle proposée par Cordier (2013).

## 2.6 Identification des séquences préfabriquées

---

35 Notre traduction de : « A formulaic sequence (FS) is a multiword semantic/fonctional 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 automatised. » (Cordier, 2013 : 84),.

Une des plus grandes difficultés pour le chercheur est sans aucun doute l'identification des SP (Myles, 2006 : 168 ; mais aussi Schmitt, 2010 ; Wray, 2002, 2008). Identifier les SP n'est pas une tâche aisée car il faut être sûr d'avoir bien identifié toutes les séquences et seulement les bonnes. Le chercheur se trouve confronté à deux dangers : ne pas identifier toutes les séquences ou en identifier trop<sup>36</sup> (Wray 2009 : 28).

Les études présentées dans la section 2.4 sur l'acquisition, montrent qu'en L1, les enfants mémorisent des ensembles qu'ils associent à des situations avant même de pouvoir les analyser. On peut identifier ces ensembles grâce à un certain nombre de critères que certains chercheurs ont établis. Quels sont ces critères ?

Un des premiers travaux sur l'identification des SP et le plus connu en L1 est celui de Peters (1983) inspiré de Coulmas (1979). Peters (1983) distingue une expression préfabriquée d'une expression libre selon les six critères suivants :

1. emploi fréquent d'une suite idiosyncrasique, c'est-à-dire propre à un locuteur,
2. complexité par rapport à ce que l'enfant serait capable de produire à son niveau d'apprentissage, la structure de la SP est plus complexe,
3. usage souvent inapproprié : la SP peut être utilisée hors contexte ou sans rapport avec le contexte,
4. cohérence phonologique : une suite est produite d'un trait sans pause interne,
5. emploi lié à une situation donnée, une SP est produite en relation à une situation,
6. la séquence, utilisée fréquemment dans l'entourage immédiat, a été entendue souvent et associée à une situation.

Comme le souligne Peters (1983), une séquence n'a pas besoin de réunir ces six critères pour être considérée comme « préfabriquée ». Récemment, Cordier (2013 : 85) a fait remarquer très justement que si les six critères sont importants, Peters ne dit pas lesquels sont plus importants que les autres. Et en effet, on ne peut pas appliquer tous les critères à la fois. Une séquence ne peut pas être à la fois idiosyncrasique (critère 1) et en usage dans la communauté linguistique (critère 6).

---

36 « Researching formulaic language has many challenges but probably the single most persistent and unsettling one is knowing whether or not you have identified all and only the right material in your analyses. » (Wray, 2009 : 28)

Wray, dans un ouvrage récent (2008), recense tous les critères que l'on peut trouver dans la littérature pour l'identification des SP. Les onze critères ont été proposés dans différentes approches (formelle, pragmatique, statistique) et appliqués à divers types de locuteurs (natifs et non-natifs).

Nous les présentons ci-dessous (Wray, 2008 : 116-121) :

- 1) irrégularité syntaxique : la séquence ne suit pas les règles habituelles de la grammaire,
- 2) critère d'opacité : on ne peut déduire le sens de la séquence même en connaissant le sens de chacune des unités qui la composent, ex : *casser sa pipe, prendre le taureau par les cornes,*
- 3) fonction performative de la séquence, l'action est réalisée du fait même de son énonciation, *je vous déclare mari et femme,*
- 4) fonction pragmatique : la séquence est utilisée exclusivement en rapport à une situation donnée « *Sincères condoléances* » lors d'un décès ou *Joyeux Noël* lors de la fête éponyme. Elle est produite dans un registre correspondant à la situation et au statut des locuteurs en présence,
- 5) en usage dans la communauté linguistique,
- 6) utilisation idiosyncrasique d'une séquence par un locuteur donné,
- 7) cohérence phonologique et prosodique : une séquence est prononcée de manière fluide avec un contour intonatif spécifique,
- 8) emploi inapproprié : *j'ai peur de* au lieu de *je crains de* (*je crains de ne pouvoir assister à la réunion*) où *je crains* a perdu son sens lié à la « peur » physique,
- 9) complexité syntaxique ou sémantique : le locuteur n'est pas capable de créer la « phrase » à son niveau de connaissance,
- 10) même si la séquence est productive, elle semble provenir d'une SP qui a été délibérément transformée,
- 11) la séquence provient d'une suite produite par d'autres locuteurs.

Ces critères servent à identifier des séquences qui peuvent être très éloignées/différentes les unes des autres. C'est pourquoi Forsberg (2006) reconnaît l'importance et le besoin d'avoir une bonne base théorique pour expliquer de telles différences : une séquence identifiée selon le critère phonologique, peut être tout à fait transparente et se trouver à côté d'une autre identifiée selon le critère d'opacité. Des séquences du type *j'aime + (le tennis)* et *manger les pissenlits par la racine*, qui n'ont rien en commun, ne seront pas identifiées selon les mêmes

critères. Si les critères recensés par Wray permettent d'identifier toutes sortes de séquences préfabriquées, aussi bien les séquences figées et opaques que les séquences libres, nous ne savons pas, là encore, lesquels sont plus importants que les autres. Et en effet, on ne peut appliquer les onze critères à la fois.

Hickey avait déjà proposé, en 1993, son modèle d'identification des SP en L1 chez l'enfant (applicable à la L2). A la différence des autres modèles qui ne distinguent pas entre les critères, Hickey présente un système basé sur des préférences.

Elle propose une liste de critères nécessaires ou obligatoires sans lesquels une séquence ne peut être considérée préfabriquée, et de critères suffisants (typiques ou gradués/évolutifs) pour identifier une séquence préfabriquée. L'intérêt de ce modèle tient aux neuf critères qui n'ont pas tous la même importance.

Les deux premiers critères proposés, polylexicalité et cohérence prosodique, sont des critères nécessaires, obligatoires, applicables à toutes les SP sans exception.

- 1) la séquence doit comporter deux unités au minimum (**nécessaire**), deux morphèmes (L1) au minimum sont nécessaires pour remplir la première condition,
- 2) la séquence doit être prononcée de manière fluide sans pauses (cohérence phonologique de l'expression) (**nécessaire**).

Les critères 3 à 9 sont des critères typiques et gradués (évolutif). Nous les résumons ici :

- 3) les éléments individuels de la séquence ne sont pas employés dans la même forme séparément ou dans d'autres contextes, critère de non-compositionnalité,
- 4) la séquence est plus complexe que ce que l'apprenant pourrait produire à son niveau de compétence, critère de complexité grammaticale,
- 5) la séquence est partagée par la communauté,
- 6) la séquence est une suite idiosyncrasique, propre à un locuteur,
- 7) la séquence est utilisée de manière répétitive, dans la même forme sans possibilité de l'adapter,
- 8) la séquence est dépendante de la situation,
- 9) la forme de la séquence est correcte mais elle est utilisée hors contexte, de manière inappropriée.

Le modèle d'identification des SP de Hickey est basé sur celui de Peters (1983) proposé en acquisition L1 chez les enfants, repris plus tard par Wray (2008) et Bardovi-Harlig (2012) en

L2. On note cependant une grande différence : la liste des critères proposés par Hickey est plus fine car elle distingue entre critères obligatoires et critères facultatifs.

Le premier critère (polylexicalité) concernait en réalité le nombre d'unités morphémiques mais à la différence des enfants, chez les adultes, et notamment en L2, les unités sont représentées par un mot et non plus par un morphème. Le critère de polylexicalité, reconnu dans la plupart des recherches, est la condition *sine qua non* de la préfabrication. Cependant et malgré cela, certains chercheurs considèrent comme préfabriqués des formules qui, après soudure, sont devenues des séquences monolexicales, comme par exemple *pourboire, tournevis, vinaigre, bonjour, adieu, voici, voilà, pourquoi, toujours, puisque*, etc. Schmäle (2013) les nomme unités de construction préformées (UCP) ; il ne parle pas de critère de polylexicalité mais de polyfactorialité, ce qui lui permet d'accepter les formules monolexicales. En effet, les UCP peuvent faire référence aussi bien à des facteurs contextuels, situationnels, fonctionnels que non-verbaux. Perrin (2011 : 86) signale que des formules comme *hélas, bonjour, bonsoir, svp* « sont issues d'une forme de figement exclusivement diachronique... » ce sont des « formules conventionnelles » grammaticalisées par délexicalisation et qui jouent un rôle dans le discours.

La question de la polylexicalité est problématique pour le chercheur. En effet, à l'oral, les apprenants débutants n'ont pas vraiment conscience de la segmentation, qui se fait plus tard dans l'apprentissage, lorsqu'on aborde l'écrit. C'est pourquoi il peut paraître paradoxal d'utiliser un critère graphique quand on s'intéresse à la production orale. C'est toute l'ambiguïté des recherches sur la préfabrication. Cependant, le critère de polylexicalité, largement décrit dans la littérature, est dans cette thèse et conformément à Hickey (1993) un des deux critères nécessaires pour identifier une SP.

Le critère de cohérence phonologique est le deuxième critère obligatoire pour identifier clairement une SP (Hickey, 1993 : 31) mais il ne peut être le seul. En effet, une séquence dite d'un trait ne sera pas forcément préfabriquée. Peters (1983) est la première, dans la littérature sur le sujet, à avoir mis en avant ce critère. Pour un grand nombre de chercheurs qui s'intéressent à la production orale, les SP doivent être normalement prononcées avec fluidité, sans pause(s) interne(s), avec une courbe mélodique continue (Peters :1983 : 8 ; Dechert *et al.*, 1984 ; Wray, 2002 : 36). Une SP est dite de manière plus fluide qu'une expression libre avec un contour intonatif (Schmitt, 2010, Van Lancker *et al.* 1981, Peters 1983 :10) et sans hésitation dans la séquence (Kuiper 1996, Dahlman, 2009). Pour arriver à produire une SP, le

locuteur devra l'avoir mémorisée en bloc ou automatisée. Nous rappelons la définition de Wray (2002 : 9) qui considère qu'une séquence est préfabriquée lorsqu'elle semble avoir été mémorisée, et stockée d'une certaine façon, de manière holistique, et récupérée comme un tout. Des études postérieures à celle de Hickey ont développé la notion de cohérence phonologique. Lin (2010) distingue deux aspects de la notion de phonologie : l'*intonation* (contour prosodique et unité de ton) et le *tempo*, ou l'absence de disflueur, c'est-à-dire aucune hésitation dans la séquence.

Le critère de cohérence phonologique a été limité pourtant aux locuteurs experts par Bardovi-Harlig (2010) pour qui une séquence produite par un locuteur débutant peut éventuellement être hors critère phonologique. Ainsi une séquence, avec allongement ou pause interne, pourrait être préfabriquée car elle serait en cours d'acquisition. Dans ce cas, il nous paraît peu logique de la considérer comme une séquence facilitant le traitement. Rappelons que notre recherche s'intéresse au LP du point de vue du locuteur au sens où il l'aide dans le traitement langagier. Le critère de cohérence phonologique tout comme celui de polylexicalité est un critère nécessaire et obligatoire pour identifier une SP. Ces deux critères obligatoires sont applicables à la L1 mais un grand nombre d'études postérieures à Hickey les applique à la L2, quel que soit le niveau des apprenants. Concernant les autres conditions typiques ou graduées, Hickey ne dit pas si certaines ont plus d'importance que d'autres. Pour identifier une séquence préfabriquée, il faudra aussi déterminer quels critères parmi les typiques ou gradués sont les plus importants.

Les sept autres critères peuvent également être appliqués à la L2 en début d'apprentissage (Weinert, 1995 et Myles *et al.* 1998, 1999). Les apprenants débutants produisent des suites de mots souvent complexes alors qu'ils ne sont pas encore capables de créer leurs phrases en combinant les mots selon la syntaxe. Ils les emploient à côté d'unités simples. C'est d'ailleurs souvent leur complexité syntaxique ou sémantique qui permet de les identifier.

Cependant, il n'est pas toujours facile de distinguer une séquence produite de manière productive d'une séquence préfabriquée. Ainsi, les apprenants débutants produisent souvent des séquences de verbes fréquents du type *il/elle a* ou *il/elle est* qu'ils ont apprises et qu'ils ont répétées dans les cours de langue. Celles-ci semblent avoir une fonction dans le traitement de séquences plus élaborées. La séquence *il a* peut-elle être considérée comme préfabriquée ? Selon nous, cette séquence est préfabriquée lorsqu'elle est suivie d'un nom indiquant l'âge comme dans *il a 20 ans*, où un élément fixe *il a* est suivi d'un élément changeable (âge).

Cordier, en revanche, identifie *il a* comme une séquence « préfabriquée » qui permet le traitement de la forme complète du verbe au passé composé (*il a / il a mangé*). Cependant nous ne sommes pas d'accord avec ce chercheur dans la mesure où nous considérons *il a / il a mangé* plutôt comme une hésitation qui prend la forme d'une répétition.

Comment identifier les SP chez des locuteurs avancés ? Cordier (2013 : 86-88), faisant référence à Peters (1983) et à Hickey (1993), avance qu'il est relativement plus aisé d'identifier les SP chez les locuteurs débutants par leur structure grammaticale complexe ou leur emploi inapproprié. Mais les apprenants avancés, capables à leur niveau de segmentation et d'analyse syntaxique, peuvent très bien avoir construit les séquences plus complexes mot par mot. Toutefois, on pourra les considérer préfabriquées parce qu'elles présentent un avantage dans le traitement langagier à un moment donné, conformément à la définition des psycholinguistes.

Parmi les autres critères facultatifs, le critère de fréquence, généralement cité dans l'identification des SP, a été initié par Peters (1983) et Wong-Fillmore (1976). Il est possible qu'une formule, qui apparaît très fréquemment, puisse avec le temps se lexicaliser. Comme une formule peut n'apparaître que peu de fois, la fréquence n'est pas une condition obligatoire mais typique. Les enfants, peuvent produire une formule qui, peu fréquente au début, pourra, dans un deuxième temps, lorsqu'elle aura été modifiée, apparaître plus souvent. En L2, une formule peut n'être produite que peu de fois, voire une seule fois dans la production d'un locuteur donné.

Toutes les études existantes montrent un certain consensus concernant les critères d'identification. Toutes s'accordent à dire qu'un seul critère ne suffit pas. Lorsqu'on s'apprête à identifier des séquences préfabriquées dans le discours, il faut s'appuyer non pas sur un seul critère mais sur un ensemble de traits<sup>37</sup> (Wray, 2002 : 43). Certains chercheurs (Foster, 2001 ; Forsberg, 2006) ont même basé « en partie » leur identification sur leur intuition, ou celle de juges natifs, sans doute parce que leur définition des SP était plus linguistique que psycholinguistique. Cependant, l'intuition d'un natif pour certains chercheurs est incertaine par rapport à une approche scientifique (Nation et Read, 2004 : 29 ; Wray, 2002).

Après avoir évoqué le rôle des SP dans le traitement de la parole à travers quelques expériences, nous avons mis en avant le rôle de l'input dans la mémorisation et

---

37 Notre traduction de : « It may simply be that identification cannot be based on a single criterion, but rather needs to draw on a suite of features. » (Wray, 2002 : 43).

l'automatisation des séquences. Les modèles basés sur l'usage montrent que l'acquisition des SP se ferait dès la petite enfance grâce à l'input et la répétition. Il semblerait que le même processus se produise en L2. Nous avons enfin évoqué la difficulté pour le chercheur d'identifier les SP dans un corpus oral et présenté les critères d'identification applicables à un corpus L2. Dans le chapitre suivant, nous montrerons que les SP jouent un rôle important dans le traitement, la fluidité et l'aisance en production orale.



## Chapitre 3 Rôle des SP sur la fluence, les SP dans les modèles de production

---

Quoique facile en apparence, l'expression orale en L1 est un ensemble de processus d'une grande complexité. Cette compétence si « banale » a pourtant demandé des années pour se construire en L1 (Kormos, 2006 : 38). En situation de communication, un locuteur doit pouvoir gérer sa production orale en temps réel et parfois limité. En L1, les locuteurs sans problème spécifique produisent facilement un discours fluide sans conscience des procédures en jeu, même en faisant autre chose (cuisiner, conduire, écouter de la musique, *etc.*). Cette aisance si caractéristique de la production L1 est possible grâce à la nature automatique des processus d'encodage formel (activation des mots, leur agencement en énoncés, leur articulation). En L2, ces mécanismes d'encodage formel exigent souvent un effort cognitif – ils ne sont pas automatisés. La nature souvent problématique des traitements formels génère des disfluences très caractéristiques dans la production L2 comme les hésitations, les pauses, les pauses pleines du type *euh*, les corrections, les faux départs, les répétitions (mais aussi les interjections, le balbutiement et les lapsus qui ne feront pas l'objet de remarques dans cette étude). Dans ce chapitre, nous tenterons d'expliquer en quoi consiste la fluence en production orale en L2 à la lumière des théories de l'automatisme. De nombreuses études sur l'aisance à l'oral – ou la « fluence » - affirment que l'utilisation de grands ensembles lexicaux permet d'améliorer la fluidité du discours en L1 comme en L2 : les séquences préfabriquées seraient même une explication de l'automatisme dans les traitements langagiers (Kormos, 2006 ; Hilton, 2009). Nous utiliserons les termes de fluidité et fluence alternativement pour désigner la qualité de la production orale, *fluence* étant le terme anglais, souvent utilisé en français, pour désigner cette compétence.

### 3.1 Définitions de la fluence

La fluence est un terme très général fréquemment utilisé lorsqu'on enseigne ou apprend une L2. En L2, une personne est « fluente » si elle parle sans faire trop d'erreurs (prononciation, grammaire, vocabulaire, *etc.*). On dit, dans ce cas, qu'elle parle « couramment » la langue. Mais la fluence, ou la fluidité, est un phénomène plus complexe qui comprend une multitude de traits linguistiques mais aussi psycholinguistiques et sociolinguistiques (Freed, Segalowitz et Dewey, 2004 : 279) et les chercheurs ont du mal à se mettre d'accord sur sa définition.

On trouve d'ailleurs dans la littérature deux sens pour la fluence (Lennon, 1990, 2000). Un sens large qui définit la fluence comme une compétence globale de l'oral incluant la précision et la complexité de la production. La fluence serait alors une composante de la compétence orale qui associerait vitesse et correction. La fluence est même parfois considérée comme « le point culminant sur une échelle mesurant la maîtrise de la L2 » (Lennon, 1990 : 389)<sup>38</sup>. Fillmore (1979 : 93) identifie quatre types de fluence, la première étant la capacité de remplir le temps avec la parole ; les autres étant la capacité de parler de façon dense et cohérente (fluence sémantique et grammaticale), la capacité de savoir dire ce qu'il faut au bon moment (fluence pragmatique) et la capacité de jouer avec la langue en faisant par exemple des jeux de mots (fluence créative ou esthétique). En L2, seul un locuteur très avancé pourra tendre vers la maîtrise de ces quatre types de fluence. En effet, on peut se demander si, en L2, être « fluent » signifie parler comme dans sa LM. On peut être fluent sans pour autant avoir développé des compétences grammaticales ou communicatives et l'inverse est tout aussi vrai<sup>39</sup> (Sajavaara, 1987 : 62). Pour certains chercheurs, il n'y aurait pas forcément une corrélation entre une bonne fluence et une compétence linguistique, même si les disfluences (faux départs, hésitations, etc.) sont dues à des lacunes linguistiques (Sajavaara et Lehtonen, 1978). On peut en effet commettre des erreurs linguistiques sans pour autant qu'elles gênent l'interlocuteur. Schmidt (1992) fait d'ailleurs bien la distinction entre la fluence et l'exactitude.

Au sens étroit, la fluence est une compétence de l'oral parmi d'autres et elle relève de la performance du locuteur. C'est ce que Segalowitz (2010) appelle la fluence de production (*utterance fluency*) qui fait référence aux propriétés physiques de l'expression. Segalowitz (2010) distingue en réalité trois entités de fluences différentes et complémentaires qui ont chacune leurs propres mesures : la fluence cognitive, la fluence de production et la fluence perçue (ou de perception)<sup>40</sup>. La fluence cognitive fait partie de la compétence du locuteur et dépend de l'automatisation des processus d'encodage. La fluence cognitive a trait aux opérations mentales qui sous-tendent la performance comme la planification de la phrase, la recherche lexicale, l'ajustement grammatical, etc. La fluence de l'expression, en revanche, renvoie aux traits de la fluence objectivement observables et mesurables lors de la

---

38 Notre traduction.

39 « Good linguistic or communicative competence is not always realized in fluent speech : one person may master the finest nuances of literary expression and yet be incapable of good speech performance ; another person, whose command of structure and vocabulary is restricted, may carry on efficient speech communication with great skill. » (Sajavaara, 1987 : 62)

40 « cognitive fluency, utterance fluency and perceived fluency ».

performance comme les pauses, les hésitations, les réparations, le taux de phonation, etc. Elle est le résultat de la fluence cognitive. La fluence perçue est l'impression ressentie par l'auditeur qui se base sur sa perception de la fluence produite. Selon Lennon (1990 : 391), la fluence dépendrait de « l'impression de l'auditeur que les processus psycholinguistiques de la compréhension et de la production fonctionnent aisément et efficacement »<sup>41</sup>. En effet, la perception de la fluence, soumise aux « impressions » de l'auditeur, est très subjective (Fillmore, 1979 ; Freed, 1995 ; Lennon, 1990).

Néanmoins, nous ne traiterons pas de la fluence perçue, qui dépasse le cadre de cette thèse. Notre étude veut montrer en quoi consiste la fluence de production, ce qui la favorise, et s'intéresse donc à la performance des locuteurs.

Selon Schmidt (1992 : 358), la fluence en L1 est le résultat de processus automatiques qui ne demandent pas beaucoup d'attention ni d'effort. Mais en L2, les procédures (liées à l'encodage linguistique des idées) n'étant pas automatisées, en tout cas les premiers temps, les locuteurs seront « disfluents ». La fluence dépend donc de l'activation automatique et fiable du réseau de connaissances lexicales, phonologiques, prosodiques et grammaticales. Il semble, en effet, qu'il y ait une corrélation entre les mesures temporelles de la fluence et les connaissances linguistiques dont font partie, d'ailleurs, les SP (Kormos et Dénes, 2004). Un locuteur L2 serait capable de parler avec fluidité lorsque les mécanismes de la production sont largement automatisés et que des séquences sont mémorisées, récupérées et utilisées avec exactitude<sup>42</sup> (*ib.* 160).

Une définition opérationnelle de la fluence associe fluence et traitement langagier, « parler de manière fluide, c'est être capable de traduire rapidement, de façon aisée, précise et efficace sa pensée en temps réel »<sup>43</sup> (Lennon, 2000 : 26). Autrement dit, lorsque les procédures formelles d'encodage (planification et choix lexico-syntaxique) ne sont plus conscientes en L2 (comme c'est le cas en L1), un locuteur arrive à s'exprimer avec aisance – c'est-à-dire sans trop d'hésitations. D'après certains chercheurs, la fluence est possible en L2 lorsque le locuteur

---

41 Notre traduction de « An impression on the listener's part that the psycholinguistic processes of speech meaning and speech production are functioning easily and efficiently. » (Lennon, 1990 : 391).

42 « The correlations between the temporal and linguistic variables also reveals that accuracy is positively related to temporal variables that are influential in fluency judgements. In other words, it seems that those students who were fluent in terms of speed and pace also produced accurate output. In psycholinguistics terms this means that one is only able to speak fluently if speech production mechanisms are largely automatic and if automatic sequences are memorised, retrieved and used accurately. » (Kormos et Dénes, 2004 : 160)

43 Notre traduction.

réussit à regrouper des ensembles plus complexes que de simples mots<sup>44</sup>. (Rebhein, 1987 : 101). Faire le lien entre la fluence et la connaissance des SP (une des composantes de la compétence générale linguistique) est un des axes de cette recherche.

Afin de comprendre comment l'usage de ces ensembles plus complexes, c'est-à-dire des séquences préfabriquées, contribue à la fluence, il nous paraît nécessaire tout d'abord de voir quels processus sont en jeu dans la production orale en L1, puis en L2. C'est pourquoi nous présentons dans la section suivante le modèle le plus connu de la production orale en L1 (qui est aussi le plus cité en L2), celui de Levelt (1989, 1999).

### **3.2 Le modèle de production en L1 et le « schéma du locuteur » (blueprint) de Levelt (1989, 1999)**

La plupart des théories psycholinguistiques suggèrent que la production orale passe par trois étapes principales de traitement qu'on retrouve dans tous les modèles de production de parole proposés en L1 comme en L2 (Kormos, 2006) : la préparation conceptuelle du message préverbal, la formulation ou lexicalisation du message préverbal et l'articulation (voir Ferrand, 2002 : 27). Une quatrième étape, celle du contrôle qui permet au locuteur de se corriger, finalise le traitement. En production L1, la planification du message lors de la première étape demande de l'attention alors que la formulation et l'articulation, les deux étapes suivantes, sont automatiques. En L2, ces deux dernières phases demanderont souvent plus d'attention, selon le niveau du locuteur. Si les étapes 2 à 4 ont été étudiées, à notre connaissance peu de recherches ont été entreprises en L1 au niveau de la première étape, la planification conceptuelle du discours et encore moins en L2 (Hilton, 2009).

Bien qu'il existe d'autres modèles de production (Dell, 1986) comme le modèle en cascade (*spreading activation*), influencés, entre autres, par les recherches (Poulisse, 1993) dans le domaine des lapsus (*slips of the tongue*), le modèle que nous présentons est celui de Levelt (1989) révisé en 1999, élaboré pour la L1, et qui s'inscrit dans une perspective psycholinguistique. Pour la L2, d'autres modèles de production de la parole existent (De Bot, 1992), mais celui de Levelt reste le modèle plus connu. C'est, d'ailleurs, ce modèle qui a servi pour élaborer d'autres modèles de production spécifiques à la L2 dont celui proposé par Kormos (2006), que nous présenterons dans la section suivante. Le modèle de Levelt est

---

<sup>44</sup> « Fluency in a second language requires the capability of handling routinized complex speaking plans. » (Rebhein, 1987 : 101)

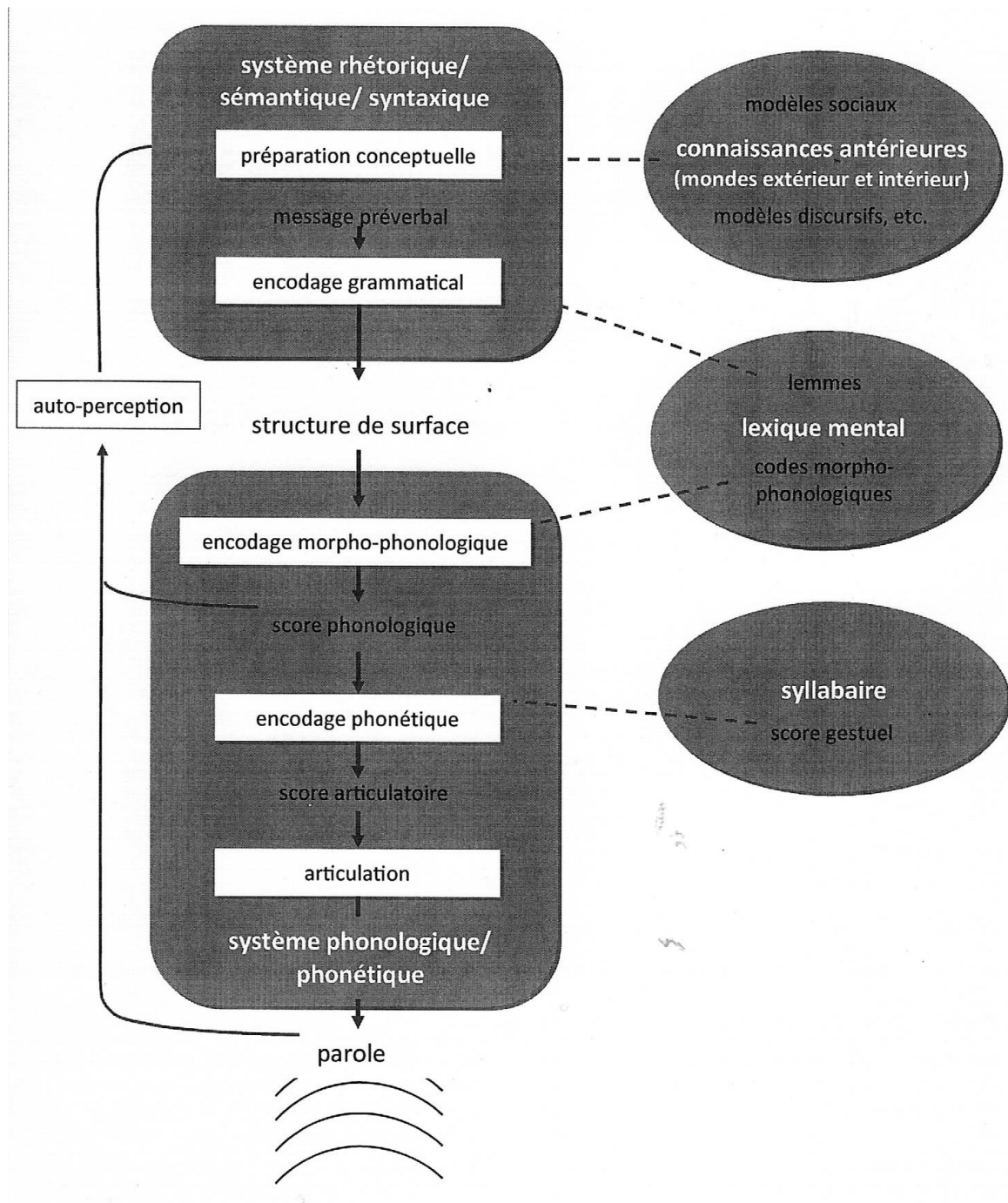
souvent cité, en littérature, notamment dans les études traitant du développement de la fluence chez des apprenants L2 (Towell, Hawkins et Bazergui, 1996 ; Towell et Dewaele, 2005).

Nous présentons dans la figure 1, un schéma plus récent du même auteur (Levelt, 1999) qui insiste sur la modularité de la production de parole. Comme on le voit sur la figure 1, le traitement se fait à travers deux systèmes cognitifs : le système rhétorique ou sémantico-syntaxique, qui s'occupe de la conceptualisation et de l'encodage grammatical du message préverbal et le système phonologique, qui comprend les mouvements articulatoires pour produire les syllabes. Dans ce modèle, trois types de connaissances (très interconnectées), (illustrées par les ellipses dans la figure 1), alimentent ces systèmes : le lexique mental, le syllabaire gestuel, et les connaissances du monde extérieur.

### **3.2.1 Les différentes étapes de la production de parole**

Nous rappelons ici brièvement, les différentes étapes du modèle de production de Levelt (1989). Le traitement se fait de manière progressive, étape par étape (en cascade), tout au long de la production. Les quatre étapes se suivent théoriquement dans cet ordre mais les processus étant imbriqués, on peut être dans la phase d'articulation à un moment donné alors qu'on planifie déjà la suite. Concernant le traitement des SP, on peut expliquer comment les SP influencent les différentes étapes du modèle de production de parole.

Figure 1. Schéma des processus de production orale (*A Blueprint of the speaker*, Levelt, 1999 : 97, cité par Hilton, 2009 : 80)



- La conceptualisation : étape pendant laquelle on planifie ce que l'on veut dire.

La première phase consiste à préparer la planification du message avant de pouvoir le générer (macro-planning). Le « message préverbal » contient les informations qui permettront de

transformer le concept en message : (représentation de la situation de communication, activation du thème, prise de perspective sur les informations à communiquer). Pour expliquer ce qui se passe pour le LP, à cette étape de la conceptualisation, un locuteur qui voudrait planifier le concept « mourir » peut déjà choisir le registre dans lequel le concept sera formulé. Ce choix sera déterminé par la situation de communication et la perspective que le locuteur veut adopter sur les événements dont il parle (*casser sa pipe* sera utilisé dans une interaction où les locuteurs se connaissent et où le langage familier peut être utilisé sans danger). Le message préverbal va ensuite passer dans ce que Levelt appelle le « formulateur ».

- La formulation :

Cette étape concerne la formulation du message préverbal et tient compte des éléments grammaticaux, lexico-sémantiques, pragmatiques et phonologiques. Le lexique mental consiste en unités lexicales composées de lemmes, contenant des informations sémantiques et syntaxiques, et des lexèmes qui sont, eux, composés d'informations morpho-phonologiques. La première étape de l'encodage est la sélection de lemmes qui correspondent le mieux à l'information sémantique de la phase préverbale. Une fois le lemme sélectionné, le locuteur activera sa forme phonologique et morphologique. Le formulateur produit alors une forme syntaxique de surface et une forme phonologique de l'expression. A cette étape, le locuteur active (selon les théories développées au chapitre 2) des ensembles de mots - des séquences préfabriquées - dans le lexique mental, qui vont (de par leur nature pluri-lexicale) permettre une plus grande rapidité dans l'encodage du message (le locuteur n'encodant pas mot à mot mais phrase par phrase). Par exemple, pour formuler « ME PRESENTER » représenté dans le conceptualisateur, je peux activer la forme *Bonjour, je m'appelle* sans passer par tous les mots et morphèmes qui composent cette SP : l'ensemble des informations (lexicales, syntaxiques et phonologiques) sera traité en bloc. Ces séquences libèrent ainsi beaucoup de ressources qui pourront être utilisées pour la planification d'un prochain message préverbal.

- L'articulateur (ou la production des sons)

Lors de la dernière phase de traitement, l'énoncé passe dans l'articulateur qui produit l'expression. L'articulateur a en mémoire, dans le syllabaire, les gestes articulatoires nécessaires pour transformer le message préverbal en parole ; selon Oppenheim (2000 : 220),

les suites articulatoires (sous formes de *chunks*) présentes dans le syllabaire faciliteraient la fluence de production.

- Contrôle ou auto-surveillance ou monitoring

L'auto-surveillance (ou « *monitoring* ») permet au locuteur de vérifier la correction de ce qui est articulé d'un point de vue lexical, morphologique, syntaxique et phonologique. « Les locuteurs contrôlent ce qu'ils ont à dire et comment ils vont le dire »<sup>45</sup> (Levelt, 1989 : 458) et ceci pour toutes les formes de discours et sans intervention de l'interlocuteur, que ce soit pour corriger leurs erreurs ou pour apporter une précision. Dans le modèle de production de Levelt (1989), trois boucles phonologiques (*phonological loop*) examinent la production. La première boucle compare le message préverbal et les intentions du locuteur. La seconde boucle concerne la surveillance au plan phonétique (*internal speech*) avant l'articulation, qui est appelée aussi *covert-monitoring*. La dernière boucle, qui prend en charge l'acoustique et la phonétique, consiste à vérifier l'expression après l'articulation. Si une erreur ou une approximation est détectée, un signal est alors émis, ce qui déclenche une nouvelle production à partir de la première étape de conceptualisation.

La correction peut servir aussi parfois à lever une ambiguïté. Au niveau conceptuel, si le locuteur s'aperçoit de son erreur avant la formulation du message, cela se traduit par une hésitation ; et après que le message a été formulé, le locuteur se reprend sous la forme d'une reformulation. « Les locuteurs peuvent se corriger presque à tous les niveaux de leur production »<sup>46</sup> (Levelt, 1989 : 463). Le signal qu'ils donnent à leurs interlocuteurs peut se traduire par une pause sonore du type *euh*, *mm* ou des formules comme *je pense, plus exactement*, etc.

Dans le modèle de Levelt, le traitement se fait de manière progressive : dès que le message préverbal passe dans le formulateur, le conceptualisateur commence son travail sur le groupe suivant même si le précédent n'a pas fini d'être traité. Ainsi l'articulation d'une phrase peut commencer bien avant que le locuteur ait terminé de planifier la totalité du message (Kormos, 2006, 7-8) et la planification s'enclenche avant que l'énoncé précédant ne soit complètement articulé. Ceci est possible parce que les activations, notamment dans la phase d'encodage,

---

45 Notre traduction de : « Speakers monitor what they are saying and how they are saying it. » (Levelt, 1989 : 458)

46 Notre traduction de : « Speakers can monitor for almost any aspect of their own speech. » (Levelt, 1989 : 463)



sont hautement automatisées. C'est sur ces automatismes que repose la grande vitesse de la parole.

En L1, les opérations formelles - et un certain nombre de traitements sociaux - sont automatiques et le locuteur n'a pas directement conscience des règles sous-entendues qui régissent une interaction (Hilton, 2009). Selon Rehbein (1987 : 99), en L1 les phases de planification et de formulation dans la production doivent être exécutées en partie simultanément par le locuteur<sup>47</sup> mais, pour être fluent, l'expression entière doit être planifiée à l'avance<sup>48</sup>. En L2, les processus de bas niveau risquent de rendre la production plus lente, voire disfluente. L'étude des variables temporelles peut, dans ce cas, éclairer la fluence cognitive, surtout dans le cas de locuteurs débutants qui ont un lexique limité et peu de pratique de la langue (Derwing, Rossiter, Munro et Thomson, 2004 : 674). Mais ce modèle de production permet de comprendre à quelle étape peuvent intervenir les difficultés rencontrées par les locuteurs d'une L2.

### 3.2.2 Place des SP dans le lexique mental

#### a) Les séquences préfabriquées dans le modèle de Levelt

Dans le modèle de Levelt (1989, 1999), le lexique joue un rôle essentiel. Il se trouve entre la conceptualisation et la formulation. D'après Levelt (1989 : 186-187), les SP (pour lui *idioms* et collocations) sont stockées dans le lexique mental au même titre que les unités simples. « Les locuteurs ont un stock de mots, de phrases ou d'expressions idiomatiques. Certains concepts vont directement chercher des phrases telles que *Dutch uncle* ou *Red tape* »<sup>49</sup>.

Dans son modèle de 1999, Levelt suppose que « probablement, le nombre des *idioms* et des collocations dans le lexique mental est au moins aussi important que celui des unités simples<sup>50</sup> (Levelt, 1999 : 95). Mais Levelt ne dit pas s'ils requièrent un traitement spécial. Il ne leur consacre en fait qu'un court paragraphe dans la partie traitant du lexique mental.

---

47 « A. Fluency of speech means that uttering and planning/controlling are to be executed at least partly simultaneously ». (Rehbein, 1987 : 99)

48 « B. Fluency depends on the activity of planning an entire utterance in advance ».

49 Notre traduction de : « Speakers have, over and above a stock of words, stocks of phrases and idioms. Certain concepts map directly onto phrases, such as Dutch uncle or red tape. » (Levelt, 1989 : 186-187).

50 Notre traduction de : « Probably, the amount of idiom and collocation in the mental lexicon is of the same order of magnitude as the number of words. » (Levelt, 1999 : 95).

## b) le lexique hétéromorphique

Quel que soit le modèle de production, les SP ne peuvent être stockées que dans le lexique (Kormos, 2006). Ce lexique serait de même nature que le lexique hétéromorphique proposé par Wray (2002) ; il comprend des ensembles de mots et des unités simples.

Selon Wray (2002), les SP seraient stockées et activées dans le lexique comme des unités, de simples mots. Dans son modèle de stockage lexical, le « lexique hétéromorphique », les unités simples et les formules préfabriquées seraient stockées dans le même lexique (Wray, 2002 : 237). Dans ce lexique, les unités sont classées selon leur taille et le registre dont elles relèvent. Le lexique hétéromorphique, composé de cinq « sous-lexiques » appelés « notionnels » (Wray, 2002), permet de classer des morphèmes, des mots polymorphémiques et des séquences préfabriquées.

- a) le lexique grammatical (*en train de, puisque, et*),
- b) le lexique référentiel (*donner quelque chose à quelqu'un, malheureux, chien*) ;
- c) le lexique interactionnel (*ça marche ! salut ! ok !*) ;
- d) le lexique de mémorisations (stratégies mnémoniques, poèmes, comptines) ;
- e) le lexique réflexif (*va te faire \_\_, merde !*). (exemples tirés de Forsberg, 2006)

On peut donc stocker un grand nombre d'unités de manières différentes. Ainsi la séquence *prends ton temps* en français (*take it slowly*) peut être stockée comme telle dans le lexique interactionnel (c) mais chaque unité de la séquence peut être également stockée dans le lexique qui lui correspond, ainsi *prends* et *temps* dans le lexique référentiel (b), *ton* dans le lexique grammatical (a) en tant que morphème cette fois. Le locuteur récupère la séquence entière lorsque c'est ce concept – dans le registre particulier de l'expression – qui encode son idée (Wray, 2002 : 262).

Cette organisation du lexique permet de ne pas dissocier les unités simples et les séquences qui trouvent leur place, selon leur rôle et leur fonction, dans chaque lexique notionnel. Selon Kormos (2006), la récupération holistique ne pose aucun problème car les séquences (des phrases idiomatiques, des suites de mots et des collocations) sont récupérées comme des mots simples. Cependant, elle fait remarquer que l'activation peut être plus difficile pour certaines phrases longues qu'on trouve dans les actes de parole (*je vous prie de bien vouloir m'excuser*). La création de ces larges unités se placerait alors au niveau conceptuel ; l'activation se ferait par l'idée complexe associée à cet ensemble : « s'excuser dans une situation socialement très formelle ». Les SP sont des exemples de l'automatisme dans le

langage ; au lieu d'assembler une phrase de façon concertée, mot par mot, on active directement, en bloc, la matière linguistique nécessaire à l'encodage d'une idée (ou d'une partie d'une idée). C'est la définition d'une activation automatique selon la théorie de l'instance de Logan (1988).

### **3.3 Un modèle de production en L2**

Afin de mieux comprendre comment se fait la production en langue étrangère, nous consacrons cette section au traitement de la parole dans les modèles de production pour la L2 comme celui de De Bot (1992) et en particulier celui proposé par Kormos (2006 : 166-173). Ce modèle reprend étape par étape, en l'adaptant à la L2, le processus de la production orale en L1.

#### **3.3.1 Les caractéristiques du modèle de production bilingue**

Kormos (2006) propose un modèle de production bilingue inspiré du modèle de production de Levelt pour la L1 (1989, 1999) présenté dans la section 3.2. Ce modèle montre des similitudes entre la production orale L1 et L2 mais surtout les spécificités du processus de la production orale en langue étrangère. La production orale en L2 serait composée des mêmes modules qu'en L1. Le traitement de la parole passe par les trois étapes principales évoquées dans la section section 3.2.1. Comme en L1, le traitement en L2 se fait de manière progressive, chaque étape du modèle se faisant l'une après l'autre. Chaque nouvel input dans le module déclenche des procédures d'encodage. Par exemple, une fois la première syllabe d'un mot activée, l'articulation peut commencer dans l'articulateur. Le traitement parallèle est théoriquement possible pour les apprenants les plus avancés en L2 ; mais tant que les procédures d'encodage demandent de l'attention, elles ne peuvent se faire que l'une après l'autre. Ce modèle de production bilingue n'est pourtant pas à proprement parler un modèle sériel dans la mesure où l'activation en cascade (*spreading activation*) est possible depuis le niveau lexical jusqu'au niveau phonologique.

Le modèle de Levelt (1999) distingue trois types de connaissances différentes situées dans la mémoire à long terme (MLT) : les connaissances du monde extérieur et intérieur, le lexique mental et le syllabaire. Kormos, qui s'appuie sur de nombreuses théories de production, propose un nouveau modèle qui comprend un seul et large magasin de stockage, pour la L1 et la L2, qui est la mémoire à long terme. Nous rappelons ici que la mémoire à long terme

comprend toutes nos connaissances, expériences et apprentissages dont le langage fait partie (les règles morphosyntaxiques, les informations sémantiques et le lexique). En situation de production, la mémoire de travail (Baddeley, 1986) permet d'assembler des éléments que le locuteur récupère de la MLT et des éléments qu'il reçoit de l'input. En langue étrangère, le locuteur peut se trouver devant de réelles difficultés s'il ne peut récupérer de la MLT des éléments d'information qui lui permettront de produire un message en temps réel. Nous rappelons que la MLT comprend plusieurs composants (résumé dans Hilton, 2009 : 42-44) : 1) la mémoire épisodique, qui contient notamment les traces de toutes nos expériences et événements vécus, 2) la mémoire sémantique qui comprend les concepts linguistiques et non linguistiques ainsi que les traces en relation avec ces concepts, les lemmes et les lexèmes, 3) le syllabaire qui comprend les gestes articulatoires automatisés indispensables pour produire les syllabes et, 4) un dernier composant, spécifique au modèle de production L2, qui est l'endroit où sont stockées les connaissances déclaratives des règles de la L2 (Kormos, 2006 : 167). Tous ces composants contiennent des informations aussi bien de la L1 que de la L2. En d'autres termes, il y aurait une mémoire sémantique et épisodique commune à la L1 et à la L2, un seul endroit pour stocker les lemmes et les lexèmes de la L1 et de la L2 ; et de même pour le stockage des sons. La nouveauté de ce modèle de production L2, c'est de supposer l'existence d'un quatrième endroit de stockage, spécifique à la L2, qui est la partie de la mémoire déclarative consacrée aux règles syntaxiques et phonologiques de la L2 (Kormos, 2006 : 167). En L1, notre connaissance des « règles » combinatoires pour la langue sont acquises implicitement et font partie des connaissances non-déclaratives (ou « procédures ») dont nous sommes à peine conscients. En L2, un grand nombre de règles grammaticales, discursives et phonologiques sont stockées comme des connaissances déclaratives (appprises explicitement en classe de langue). La recherche en neurologie, notamment chez des patients amnésiques, semble indiquer que les connaissances déclaratives pourraient se trouver dans une région du cerveau distincte de celle qui est responsable des traitements automatiques (Cohen et Squire, 1980 : 209, cité par Hilton, 2009 : 44).

Mais excepté ce quatrième magasin, le modèle proposé en production L2 par Kormos est de même nature que celui proposé en L1. Les procédures de traitement sont les mêmes, en L1 et L2, pour les apprenants exposés dès leur plus jeune âge à la L2 ainsi que pour les apprenants très avancés. En revanche, les apprenants débutants, qui ont été peu exposés à la langue, semblent activer des régions plus larges et sensiblement différentes du cerveau quand ils

parlent la L2. Le traitement chez les locuteurs bilingues repose sur des connaissances déclaratives implicites alors que le traitement chez les apprenants à des stades moins avancés se fait de manière explicite, ce qui expliquerait leurs problèmes de disfluen.

### **3.3.2 Procédures d'encodage et de stockage des connaissances en production orale L2**

En production orale, la première étape concerne la conceptualisation du message. Dans la conceptualisation en L2, on doit décider dans quelle langue les concepts sont formulés (la plupart des concepts en L1 et en L2 étant stockés ensemble dans la mémoire sémantique) – mais parfois certains concepts pourraient être plutôt associés à l'une ou l'autre langue. Dans le modèle de production proposé par Kormos, si le concept est partagé dans les deux langues, en principe seul le concept de la langue choisie est activé pour le traitement. Mais il est fort probable que la représentation conceptuelle en L2 au stade débutant sera calquée sur celle de la L1 et les locuteurs peuvent se trouver en difficulté lors de l'encodage.

Dans la phase de conceptualisation, les concepts ne sont pas toujours représentés par des unités simples mais aussi par des ensembles de mots, la plupart des expressions que nous utilisons (chapitre 1) étant des combinaisons de mots mémorisés, des phrases, des propositions qui forment, on l'a vu, le langage préfabriqué (Pawley et Syder, 1983). La création de ces groupes (*chunks*) aurait lieu dans le conceptualisateur. En d'autres termes, selon Kormos, les locuteurs natifs ont dans le conceptualisateur de larges ensembles, ayant des fonctions communicatives correspondant aux actes de parole, aux situations, au contexte, qu'ils peuvent activer comme des unités simples. Mais en L2, le choix est plus problématique. Dans la phase du message préverbal, la formulation, l'apprenant débutant qui doit se concentrer sur des processus de bas niveau (lexique, grammaire, phonologie) le fait au détriment souvent de la gestion de la situation (aspects pragmatiques et gestion du sens). Le choix de la langue dépendra de facteurs pragmatiques comme la situation de communication, la relation entre les locuteurs, le statut de la langue (dans un contexte plurilingue), etc. L'apprenant d'un niveau avancé peut mieux gérer la situation de communication et se concentrer davantage sur la planification du message.

Dans la deuxième phase, l'encodage syntaxique suit les différentes étapes proposées par Kempen et Hoenkamps's (1987) : d'abord l'activation du lemme, en fonction de la L2 pour les locuteurs les plus performants et, pour les apprenants de niveaux débutants, en fonction le plus souvent des informations de leur L1. Néanmoins, des erreurs de transfert de la L1 vers la L2 peuvent alors se produire. Un apprenant de niveau avancé pourra activer des procédures

propres à la L2 alors qu'un apprenant débutant risque fort de calquer les procédures de sa L1 sur la L2, quelle que soit l'étape. En effet, lorsque les langues sont très éloignées, les risques d'erreurs risquent d'être bien plus fréquents (Dechert, 1983). Dans le modèle bilingue, le lexique (unités simples mais aussi les larges ensembles et phrases) est conçu comme un réseau dans lequel chaque entrée crée des connections avec d'autres entrées. Les connexions peuvent exister entre les lemmes et les lexèmes de la L1 et de la L2. Signalons que les entrées, activées plus souvent, se trouvent au centre du réseau et les moins utilisées à la périphérie. La force des connections est variable selon le niveau ; en début d'apprentissage, des liens entre unités de la L1 et de la L2 peuvent être plus forts que ceux de la L2. En production orale, le locuteur pourra activer facilement des mots ou des ensembles lorsqu'il aura déjà une certaine maîtrise de la L2.

Enfin, lors de la troisième étape qui concerne la phase d'encodage phonologique, les gestes articulatoires sont enclenchés. Les syllabes de la L2 seraient stockées dans le même syllabaire que celles de la L1. Les apprenants débutants utiliseraient les gestes articulatoires de leur L1 alors que les plus avancés arrivent en principe à créer des groupes de syllabes propres à la L2 (de Bot, 1992 ; Kormos, 2006 : 173).

L'étape finale, l'auto-surveillance, fonctionnerait en L2 de la même façon qu'en L1. Elle nécessite toutefois plus d'attention en L2, les ressources attentionnelles étant plus limitées et les processus formels (lexique, syntaxe et phonologie) demandant plus d'attention (contrairement à la L1). En réalité, on peut se demander si les locuteurs L2 ne se surveillent pas plus qu'en L1, ce qui pourrait générer aussi les disfluences. S'ils se trouvent confrontés à des difficultés d'encodage, il semblerait que les locuteurs L2 privilégient d'abord le contenu (les concepts) avant la forme, le lexique avant l'application des règles, ou *vice et versa* (Kormos, 2006 : 173). Ce qui expliquerait bien des erreurs en production orale spontanée.

Finalement, la production en L2 est également caractérisée par une plus grande lenteur dans l'encodage des concepts (de Bot, 1992), due à la compétition entre les unités de la L1 et ceux de la L2, et au déploiement de connaissances déclaratives et de procédures conscientes (à la place des traitements formels plus automatiques en L1). On distingue différents types d'influence de la L1 : le transfert de la L1 sur la L2 et l'alternance de code linguistique (*code switching*). Par exemple, l'apprenant veut activer un mot de la L2 ou une règle acquise en classe mais au lieu du mot ou de la structure souhaitée, un mot ou une procédure de la L1 est activé (qui est interprété, d'ailleurs, souvent comme un défaut de performance). Ces

interférences surviennent peut-être en partie parce que les lieux de « stockage » dans la MLT sont partagés – concepts, lemmes, lexèmes, syllabes sont stockés ensemble - et qu'il y a compétition pour l'activation des formes nécessaires à l'encodage. Bien sûr, tout dépend du niveau de l'apprenant. Si l'apprenant n'a pas un niveau bilingue, l'interférence de la L1 risque d'être beaucoup plus élevée. Le transfert peut être dû à un manque de compétence en L2 qui oblige le locuteur à s'appuyer sur les connaissances de sa L1 (ou d'une autre langue étrangère).

D'une manière générale, le locuteur doit gérer l'influence de la L1 sur les procédures de traitement en L2. Cette influence se manifeste par le transfert des connaissances non-déclaratives (ou « procédures ») de la L1 : grammaticales, articulatoires, prosodiques, dans la phase d'encodage, qui peut nécessiter l'utilisation de stratégies communicatives.

Cependant, les locuteurs L2, même à des niveaux moins avancés, peuvent résoudre certaines de leurs difficultés grâce à des stratégies compensatoires (Kormos, 2006 : 175). En cas de difficulté dans l'activation lexicale (mot manquant) l'apprenant peut changer un ou plusieurs traits en utilisant une stratégie de substitution. Une stratégie semblable peut être adoptée en cas de difficulté syntaxique ou phonologique. Si le locuteur a conscience de ses manques en L2, il peut modifier un ou plusieurs aspects du message préverbal et transférer les connaissances de sa L1 ou Lx ou en sur-généralisant les règles de la L2, ce qui peut expliquer aussi certaines erreurs.

### **3.3.3 Développement de la compétence L2 dans le modèle bilingue**

Pour une production orale aisée en L2, le locuteur doit disposer de différents types de connaissances : les mots (avec leurs caractéristiques sémantiques, grammaticales, phonologiques) et les règles d'encodage grammatical et phonologique. En L2, une fois que les règles sont apprises dans leur forme déclarative, leur procéduralisation peut commencer, ce qui signifie que les procédures, contrôlées au début, deviendront automatiques. Il en est de même pour les mots, au début ils sont récupérés de manière consciente et après un temps d'exposition et plus de compétence, la récupération se fait de manière automatique.

Ces différents types de connaissances (déclaratives, non déclaratives) doivent se mettre en place en L2 (Kormos, 2006 : 177). L'acquisition d'un mot nouveau en L2 implique, initialement, l'association du réseau sémantique (existant déjà) à un ensemble formel nouveau (ensemble phonologique qui doit être également associé à un ensemble orthographique).



Selon le modèle de Kormos (et c'est vrai dans beaucoup de cas d'apprentissage scolaire des langues), un apprenant L2 va d'abord apprendre explicitement une « règle » de grammaire (ou de prononciation) qu'il va déployer de façon consciente (il doit réfléchir à la règle pour produire la forme voulue). En L2, les règles, apprises de façon explicite, font partie des connaissances déclaratives. D'abord appliquée de manière consciente, une règle de grammaire deviendra ensuite automatique, à force de répétition. Par exemple à l'écrit, le [s] du pluriel ou les terminaisons verbales du pluriel [-ent] ; à l'oral, la règle qui veut que les déterminants *lella* se transforment en *l'* devant une voyelle. Avec la pratique, une connaissance déclarative deviendra automatique et ne demandera plus aucun effort. L'acquisition d'unités grammaticales plus larges (constructions, phrases) laisserait aussi des traces au niveau du conceptualisateur et du lexique. Selon un vieux mythe de la didactique (Poulisse, 1999 ; Towell *et al.* 2006), ces règles vont « se procéduraliser » - mais il serait plus exact de dire que l'apprenant pourra activer la bonne forme directement en MLT, sans passer par une phase d'analyse attentionnée. Cette théorie de la procéduralisation des connaissances déclaratives renvoie à la théorie d'Anderson (1983) appliquée au langage (Towell *et al.* 1996, Towell, 2002). Cette thèse est fortement rejetée par Logan (1988, 1991), Paradis (2004, 2005) et Hilton (2008a)<sup>51</sup>, pour qui une connaissance est soit déclarative soit non-déclarative.

### 3.4 Rôle des SP en production orale

Pour parler en temps réel d'une manière fluide, il faut essentiellement que les processus d'encodage formel se déroulent de façon automatique (« inconsciente ») comme c'est le cas en L1. Et cette automatisation des procédures d'encodage repose très probablement sur l'utilisation de larges ensembles préfabriqués (Kormos et Dénes, 2004 : 160).

Chez l'adulte, la fluence est le résultat de la compétence dans les traitements langagiers. Exception faite des cas pathologiques (sujets atteints de bégaiement ou patients cérébraux-lésés), un adulte sain parle de manière fluide et aisée dans sa LM sans avoir aucune conscience des mécanismes en jeu. Mais l'aisance en production orale est le fruit d'un long apprentissage chez l'enfant, en L1. L'adulte apprenant une L2 se trouve dans une situation analogue à l'enfant en L1. Mais alors que l'enfant n'est pas vraiment conscient de cet apprentissage (implicite) l'adulte, lui, en est conscient. Pour que la production soit fluide, les

---

<sup>51</sup> « Dans le domaine de la L2, Paradis (2004, 2005) refuse un modèle très répandu qui veut que l'apprenant avancé en L2 ait « procéduralisé » ce qui au début a été mémorisé sous forme déclarative : « Declarative knowledge is not « proceduralized » through practice. What is proceduralized is an altogether different kind of entity (2004 : 41). » (Hilton, 2008 : 66).



processus formels doivent être automatisés (articulation, accès au lexique, processus morphosyntaxiques). Chez les locuteurs moins avancés, ces processus, moins automatisés, demanderont plus de temps, d'effort et de concentration. Aussi, le locuteur sera moins libre pour se concentrer sur des processus de haut niveau comme la conceptualisation du message, la planification de la parole et la gestion socio-pragmatique des tours de parole. Le résultat chez les locuteurs débutants est une production souvent hautement disfluente, basée sur la traduction mot à mot de la L1 en utilisant les structures syntaxiques de leur langue maternelle (Kormos, 2006 : 172). La production orale d'un apprenant débutant peut aussi alterner entre mots simples organisés laborieusement et phrases mémorisées que l'apprenant n'est pas capable de produire à son niveau. Dans une langue que l'on ne connaît pas encore (une langue étrangère, une future langue seconde), les traitements formels ne se feront pas automatiquement : le locuteur doit fournir un effort attentionnel pour « mettre en forme » ses idées : chercher un mot, le prononcer, le combiner avec d'autres mots, effectuer des opérations morphologiques, etc. La phase d'encodage se trouve donc beaucoup plus laborieuse qu'en L1, et le discours sera moins aisé, plus hésitant, moins fluide. En L2, les processus automatiques en syntaxe et en phonologie peuvent ne pas avoir été mis en place chez l'apprenant débutant mais peuvent l'être en partie chez les locuteurs avancés. Il est aussi probable que les processus de la L1 empêchent une partie de ces processus en L2 surtout s'il s'agit de langues très éloignées. En langue étrangère, l'emploi des séquences préfabriquées permet donc aux locuteurs d'augmenter le degré d'automatisme en évitant le recours à une série de « règles » morphosyntaxiques de nature déclarative nécessitant un traitement attentionné, ou une production qui suit la grammaire combinatoire de la L1. Elles faciliteraient donc la production de parole et permettrait une meilleure fluence. Ceci dit, en L2, seuls des locuteurs de niveau très avancé seraient capables d'utiliser de larges suites de mots sans conscience des règles syntaxiques et phonologiques en jeu (Forsberg Lundell *et al.* 2014).

La fluence dépendrait aussi de la tâche, selon qu'il s'agit d'un monologue, d'une narration ou d'une interaction. C'est cette dernière tâche qui permettrait aux locuteurs L2 de s'exprimer avec le plus d'aisance parce qu'ils semblent pouvoir adapter leur discours à celui de leur interlocuteur (Ejzenberg, 2000). La fluence dépendrait donc des traitements automatiques dans l'encodage formel et de l'utilisation de larges ensembles préfabriqués (Kormos, 2006 : 156). La constitution des SP peut suivre deux chemins différents :

- a) les règles apprises de façon explicite deviennent « automatiques » lorsqu'elles ne demandent plus d'efforts attentionnels au locuteur : l'apprenant arrive à activer une forme plurimorphémique ou une construction directement de la MLT, sans appliquer consciemment « une règle »
- b) les expressions et phrases construites au début mot à mot, grâce à l'application consciente de règles syntaxiques et phonologiques, peuvent être ensuite stockées comme des unités simples dans la mémoire et récupérées en bloc, de manière holistique.

Dès les années 80, des chercheurs (Pawley et Syder, 1983 ; Raupach, 1984) ont constaté que l'usage des SP en L1 facilite la production et permet de parler de manière plus fluide et plus précise. Des recherches plus récentes se sont aussi intéressées au rôle du langage préfabriqué en production orale (Lennon, 1990 ; Nattinger et DeCarrico, 1992 ; Chambers, 1996 ; Towell *et al.* 1996 ; Ejzenberg, 2000 ; Wray, 2002 ; Schmitt et Carter, 2004 ; Wood, 2010). L'utilisation de séquences préfabriquées ou de phrases, souvent les mêmes en L2, permettraient aux locuteurs non-natifs de s'exprimer même en tout début d'apprentissage. Elles peuvent les aider à communiquer mais très vite ils se trouvent démunis lorsqu'ils doivent exprimer des idées plus complexes.

Des expériences sur le terrain tentent de montrer que l'usage des SP améliore la fluence en langue étrangère. Après un enseignement spécifique des SP dans une classe de langue étrangère, Boers, Eyckmans, Kappel, Stengers et Demecheleer (2006) constatent une relation entre l'emploi des SP et une amélioration de la fluence. Leur groupe expérimental (n=17) est sensibilisé aux combinaisons de mots (phrases idiomatiques et collocations) alors que le groupe contrôle (n=15) reçoit un enseignement traditionnel d'un total de vingt-deux heures (lecture et écoute de textes) basé sur le lexique et la grammaire. Deux juges, sollicités pour juger la fluence, constatent que le groupe expérimental obtient de meilleurs résultats que le groupe contrôle. Dans une étude similaire, Altenberg (1998 : 256) pensait trouver un effet facilitateur de l'apprentissage explicite de SP mais ses résultats n'ont pas illustré un tel effet. Aucun effet de groupe n'a pu être détecté, seuls quelques sujets réussiraient mieux que d'autres.

### **3.5 Fluence et théories de l'automatlicité**

#### **3.5.2 Relation entre automatlicité et modéles de production.**

Les théories traitant de l'automatlicité considèrent que l'apprentissage de la langue est, parmi les nombreux exploits sociocognitifs humains, l'un des plus complexes. Si l'acquisition se fait de manière naturelle en L1, les traitements formels en L2 seront plus laborieux. En L2, Le traitement automatique alors peut se faire lorsque l'on n'a plus besoin d'utiliser les ressources mentales pour contrôler les processus. Le processus est identique à celui mis en place lorsqu'on apprend à conduire une voiture (Skehan, 1998 : 59). Au début de l'apprentissage, les différents processus nécessaires à la conduite sont conscients et mobilisent toute notre attention (changement des vitesses, pédales de frein, etc.) puis, peu à peu, ces différents processus demandent moins d'attention jusqu'à devenir automatiques. Il en est de même pour le langage : au début de l'apprentissage d'une L2, on doit réfléchir de façon consciente aux mots (en traduisant mentalement, en ajustant leur place et leur combinaison avec les autres mots, en faisant des efforts d'articulation, etc.) puis, petit à petit, leur activation devient automatique – en lien direct avec l'idée que nous cherchons à exprimer, sans effort attentionnel.

Depuis le milieu des années 80, deux visions de l'automatlicité ont émergé (Kormos, 2006 : 39). L'une, basée sur des règles considère l'automatlicité comme le résultat de la conversion de connaissances déclaratives en procédures (*rule-based approach*) ; et l'autre, basée sur les mots, comme le moyen d'accéder directement à un item mémorisé (*item-based approach*). Dans les approches basées sur les règles, l'attention est requise alors que dans les approches basées sur l'item, la mémoire jouerait un rôle central. Dans les approches basées sur l'item, le locuteur est capable de récupérer de manière holistique les mots en mémoire à long terme. Ces deux approches tentent de montrer comment cette transformation se fait.

#### **3.5.3 Modéles basés sur les règles**

Dans le paradigme développé par Anderson (1983, 1995) et repris par de Keyser et d'autres chercheurs (notamment dans les années 1990), tout apprentissage, dont le langage, commence par l'acquisition de connaissances déclaratives qui vont finir à force de pratique par se « procéduraliser » ou s'automatiser (Gass et Selinker, 2008 : 219 ; résumé dans Kormos, 2006 : 39-40). Dans le modèle d'Anderson (1983) ACT (*adaptive control of thought*) et son modèle

révisé ACT-R (*adaptive control of thought-revised*) (1995) basés sur les règles, le développement des processus automatiques résulte non seulement dans une plus grande rapidité mais aussi dans des modifications qualitatives dans la nature même du traitement. Dans ce modèle, cinq mécanismes de traitement (composition, procéduralisation, généralisation, discrimination et renforcement) permettent l'automatisation. La première étape (composition) se charge de la création de macro-productions (larges ensembles ou *chunks*) à partir des unités plus petites de traitement, qui permettront de récupérer l'expression en bloc. Cette action, *chunking*<sup>52</sup> (voir section 2.3), facilite le traitement dans la mémoire de travail (Kormos, 2006 : 40-41). L'étape de généralisation permet d'élargir l'application des règles aux contextes appropriés ; la discrimination, en revanche, permet de vérifier que la règle s'applique seulement aux seuls contextes appropriés. La dernière étape se charge de renforcer les règles pour que la procédure soit mieux appliquée. Mais une fois le processus de procéduralisation mis en place, on constate qu'il sera peut-être difficile de retrouver les procédures déclaratives. D'après Segalowitz (2003), si l'automatisme signifie un traitement rapide, sans effort conscient, il est également incontrôlable ou « ballistique », une fois le processus engagé.

Le modèle d'Anderson, concernait en réalité des compétences complexes comme le pilotage des avions, l'apprentissage de la programmation, *etc.* Paradis (2004) et Hilton (2009) contestent l'application de ce modèle au domaine du langage. Si ce modèle a exercé une certaine influence dans les théories de l'apprentissage de la langue (Towell *et al.* 1996 ; Towell, 2002 ; Schmidt, 1992), peu de recherches ont été entreprises pour adapter cette théorie au développement de la fluence. Pour certains chercheurs comme DeKeyser (2001), les connaissances déclaratives deviennent procédurales avec le temps. C'est aussi le point de vue de Towell, Hawkins et Bazergui (1996) qui montrent que le discours des apprenants s'est amélioré grâce à la « procéduralisation » des connaissances grammaticales et lexicales. Pour ces chercheurs, la « procéduralisation » des connaissances linguistiques et des préfabriqués, en particulier, est le plus important facteur dans le développement de la fluence chez les

---

52 La « théorie des chunks » (hierarchical chunk theory), développée dans le chapitre 2, affirme que les suites de syllabes ou de séquences récurrentes (Oppenheim, 2000 : 220), présents dans l'articulateur (et non des syllabes ni même des mots), expliquerait la rapidité de récupération des séquences en temps limité dans les contraintes de la production orale spontanée.

apprenants avancés L2. Pour eux, la procéduralisation se ferait au niveau du formulateur (deuxième étape dans le modèle de Levelt) si les trois conditions suivantes sont remplies, a) la longueur moyenne du segment de parole doit avoir augmenté dans le temps, b) la longueur moyenne des pauses peut rester stable ou diminuer, c) le taux de phonation peut rester stable ou augmenter. La longueur du segment de parole serait, même, la mesure qui montrerait le mieux la progression des apprenants dans une langue étrangère. Dans son étude, Towell (2002) constate que les deux participants dont le segment de parole a augmenté dans le temps sont ceux qui ont réussi à « procéduraliser » leurs connaissances syntaxiques. Autrement dit, c'est parce que les sujets ont pu transformer leurs connaissances linguistiques déclaratives en procédures que la production en temps réel a été possible. Pour Raupach (1984) qui émet des réserves en langue étrangère, les connaissances procédurales ne seraient pas dues systématiquement à la conversion de connaissances déclaratives. Les locuteurs L2 pourraient acquérir certains types de connaissances procédurales directement ou par transfert des connaissances procédurales de leur L1 sur la L2 et aussi par imitation. Son étude, portant sur le discours d'apprenants germanophones apprenant le français dans un programme d'études à l'étranger, montre que certains apprenants acquièrent certaines structures de la langue cible par imitation, sans analyse des unités linguistiques. Ces structures acquises se seraient alors transformées en connaissances non-déclaratives. Tous les chercheurs ne sont donc pas d'accord avec ce modèle d'Anderson. Comme on l'a vu, certains (Paradis, 2004 ; Hilton, 2008a, 2011) considèrent qu'on ne procéduralise pas une connaissance déclarative. Une connaissance est soit déclarative soit non-déclarative (fréquemment appelée « procédurale »).

Dans les modèles basés sur l'usage (*usage-based models*, chapitre 2), la pratique renforce les connexions dans les différents réseaux (la syntaxe, le lexique, la phonétique) ce qui permet une récupération plus souple et la production de phrases appropriées au contexte. Les êtres humains étant sensibles à la fréquence des événements dans leur expérience<sup>53</sup> (Ellis, 2002 : 145), l'apprentissage se ferait donc grâce aux exemples et modèles de l'input plutôt que par l'application de règles abstraites ; c'est le modèle connexionniste basé sur la fréquence et l'association de modèles. Ainsi les SP, mémorisées à force de pratique, permettent de produire plus aisément des phrases adaptées au contexte.

---

53 « humans are sensitive to the frequency of events in their experience ». (Ellis, 2002 : 145)

La théorie de restructuration, qui s'appuie sur la théorie d'Anderson, met l'accent sur l'automatisation de la production de parole qui a lieu en premier, suivie ensuite par la restructuration. Dans le cadre de l'acquisition langagière (McLaughlin, 1990), plus on aura « routinisé » les connaissances, plus elles seront activées aisément. Les nouvelles connaissances doivent être organisées et structurées et cela passe par la restructuration du système en place. Des mécanismes psychologiques peuvent expliquer ce comportement lors de l'apprentissage d'une langue étrangère. Une première phase permet aux locuteurs de produire des séquences correctes car elles ont été automatisées. Dans la deuxième phase, de nouvelles connaissances apparaissent et entraînent des changements qualitatifs dans l'application des règles. Ce n'est qu'après un certain temps que les locuteurs sont à nouveau capables de produire les formes correctes. Même si McLaughlin ne dit rien du rôle de cet apprentissage, il est fort probable que les apprenants débutants s'appuient sur un répertoire limité d'unités mémorisées pour parler à une vitesse acceptable. Quand ils commencent à analyser ces unités et à appliquer les règles qu'ils ont déduites des séquences, ils peuvent sembler régresser et parler plus lentement. Enfin, dans un dernier temps, ces règles, utilisées de manière automatique et efficace, combinées aux unités mémorisées, pourraient améliorer la fluence. On remarque, en effet, à un moment de l'apprentissage une certaine diminution de la vitesse de parole lorsque les apprenants s'appuient moins sur les séquences mémorisées et qu'ils veulent construire leur discours mot par mot. Comme il existe peu de recherches longitudinales de ce type, on ne peut généraliser cette affirmation.

#### **3.5.4 Modèles de l'instance**

Dans le « modèle de l'instance » (*instance-based models*) de Logan (1988), le traitement automatique est une activation directe, en une seule étape de la mémoire à long terme. Le traitement attentionné et laborieux de l'expression (assemblée mot par mot en début d'apprentissage) est remplacé dans le temps par l'activation directe d'une unité plus large. C'est-à-dire que l'expression est récupérée de la mémoire de manière directe grâce aux exemples stockés, sans recours aux règles ou à des concepts. Lorsqu'on apprend les tables de multiplication (par exemple  $4 \times 4 = 16$ ) c'est en passant par la formule  $4+4+4+4=16$  ; avec la pratique, on récupère la solution (16) directement de la mémoire (Kormos, 2006 :43).

Selon ces approches (Logan, 1988 ; Robinson et Ha, 1993), la fluence pourrait être le résultat d'une performance structurée qui reposerait sur la mémoire. Ce ne sont pas des règles mais

des instances (codées contextuellement) qui déterminent le traitement automatique. Par exemple (en français L2), en début d'apprentissage, on va appliquer différentes règles (auxiliaire être ou avoir ? conjugaison de ce verbe ? forme du participe, (premier groupe ?), etc). Mais avec le temps – et surtout la répétition (et donc la « mise en mémoire ») de cette forme - l'apprenant pourra activer « *il m'a dit que* », « *nous sommes partis à* » comme un ensemble préfabriqué en mémoire – activé donc directement et d'un seul trait. Dans cette optique (Peters, 1985 ; Schmidt, 1992), l'apprentissage serait le résultat de la création d'instances, et la performance (et ce qui en résulte la fluence) s'améliorerait selon l'utilisation de ces instances.

La théorie de l'instance de Logan, totalement différente de celle d'Anderson, et cela conformément à Raupach, dit bien que l'apprentissage n'est pas la conversion des connaissances déclaratives en connaissances procédurales. Logan associe l'automatisation à la mémorisation, comme Robinson et Ha (1993), les premiers à montrer que les traitements algorithmiques finissaient par produire des traitements holistiques. Ainsi « l'automatisation » serait le résultat d'un traitement holistique.

Dans un modèle qui est totalement compatible avec cette théorie de l'instance, Pawley et Syder (1983) affirment que la « fluence native » (*native like fluency*) est déterminée par la présence d'ensembles mémorisés et d'expressions préfabriquées stockées et récupérées de la mémoire comme un tout. La fluence dépendrait donc aussi de la récupération holistique d'expressions adéquates.

La question qui nous intéresse dans le traitement de la parole concerne les SP et la façon dont on peut traiter le LP en apprentissage scolaire des langues. En L1, on apprend les SP implicitement par notre contact constant avec la langue. Mais en L2, avec 2-3 heures par semaine, comment mettre en place ce langage en quantité suffisante ? Certains chercheurs affirment que leur mémorisation peut être favorisée aussi par l'attention (*noticing*), et par la conscience (*awareness*). En effet, focaliser l'attention des apprenants L2 sur certaines séquences et constructions associé à la mémorisation peut les aider à les fixer dans la mémoire et donc à parler plus rapidement (Schmidt, 1995, 2001, Lewis, 1997). Lewis préconise une démarche explicite (1997) tout comme Ellis (1997) qui insiste sur la nécessité du *noticing*. Schmidt (1992) affirme même rien ne peut être appris sans attention (*noticing hypothesis*). Même si l'attention ne permet pas forcément et à coup sûr l'acquisition, elle

serait essentielle au début. C'est en apprenant le portugais, pendant plusieurs mois au Brésil, que Schmidt se rend compte qu'il a pu mémoriser des formes linguistiques parce qu'il les avait remarquées (*explicit learning*). Pour ce chercheur, tout apprenant d'une L2 doit être attentif et conscient de son apprentissage. Le processus d'acquisition commencerait même lorsque l'apprenant prend conscience d'une forme inconnue qui semble combler une lacune (Gass, 1988). Cependant l'hypothèse de l'attention fait débat dans la communauté scientifique. Schmidt pense que l'attention est la base de tout apprentissage – mais quid de l'apprentissage des SP en L1, qui est certainement largement implicite (non-attentionné) ? Nous savons dire « *as-tu pris ton petit déjeuner ?* » et non pas *\*mangé* en français L1, mais nous n'avons pas appris cette collocation explicitement ; nous l'avons extrapolée, par l'usage de la langue. Des recherches montrent que l'apprentissage explicite des SP est utile - mais comment faire vu le nombre d'expressions à apprendre (des milliers pour Pawley et Syder, 1983 et Mel'çûk, 1993).

Nous avons montré l'importance des traitements automatiques dans les théories de production de parole. Les SP, étant des ensembles plurilexicaux stockés et récupérés de la MLT de manière holistique, expliqueraient donc une (grande) partie de cette automaticité en L1. Pour examiner un éventuel rapport entre les connaissances préfabriquées et la fluence en L2, nous devons d'abord regarder de près les différentes façons de mesurer cette fluence cognitive en production orale.



# Chapitre 4 Mesures de la fluence et séquences préfabriquées

---

Dans ce chapitre, nous présentons les mesures de la fluence. Ces mesures permettent d'évaluer la fluence de production et nous donne des informations sur les processus à l'œuvre en production orale.

## 4.1 Les mesures de la fluence

La fluence de production étant un phénomène temporel de performance lié à la rapidité, la continuité, et l'automatisme de la parole (Chambers, 1997 ; Koponen et Riggenbach, 2000 ; Segalowitz, 2000), elle peut être mesurée précisément. Il existe d'ailleurs un certain consensus autour des mesures utilisées pour évaluer la fluence de production. Les premières recherches sur la fluence autour de la L1, entreprises par Goldman-Eisler (1958, 1968), concernent les mesures de la vitesse de parole, la vitesse d'articulation et les phénomènes d'hésitation. Ces travaux ont été repris par de nombreux chercheurs en L2, selon divers angles d'approche.

Les travaux consacrés à la définition « étroite » de la fluence considèrent les variables temporelles telles que la vitesse de parole et les pauses comme les mesures les plus pertinentes (Lennon, 1990 : 39) avec la vitesse d'articulation (Henderson *et al.*, 1965). Brumfit (2000) se demande, néanmoins, si toutes les mesures temporelles de la fluence sont de même importance pour analyser les productions disfluentes de certains sujets non-natifs. Selon cet auteur, la fréquence des pauses ne serait pas une mesure fiable pour analyser la parole hautement disfluente des locuteurs non-natifs à la différence de la mesure de la vitesse de parole<sup>54</sup>(Brumfit, 2000 : 86).

Des études en L1 et en L2 (Raupach, 1984 ; Towell *et al.* 1996 ; Wood, 2006, 2010) montrent que la meilleure façon de mesurer la fluence, c'est de calculer la vitesse de parole (*speech rate*) mais aussi la longueur moyenne des segments de parole entre deux pauses (*mean length of run*). Ces mesures, vitesse de parole et longueur des segments de parole entre deux pauses,

---

54 Notre traduction de : « [...] in the analysis of highly disfluent non-native speech, speech rate may be a reliable oral fluency measure than pause frequency. » (Brumfit, 2000 : 86).

refléteraient l'automatisme [du traitement langagier] assez directement<sup>55</sup> (Schmidt, 1992 : 362).

Cependant, certains chercheurs considèrent que la fluence ne serait pas liée seulement aux phénomènes d'hésitation et à la vitesse de parole mais qu'elle dépendrait aussi d'autres critères, plus subjectifs (Freed, 1995). Ainsi, pour Freed (1995, 2000) différents critères sont nécessaires pour évaluer la fluence des locuteurs natifs (malgré le risque de subjectivité) : la vitesse de parole, les hésitations (pauses, faux départs liés à une plus ou moins grande automatisme) mais la correction sémantique et syntaxique et la prononciation doivent aussi être pris en compte. Le critère de la vitesse de parole est important mais n'est pas le seul : une production orale sera jugée fluente s'il y a peu de pauses et d'hésitations (faux départs) et si le vocabulaire et la syntaxe sont corrects. En effet, pour Freed, vitesse ne signifie pas forcément correction. En L2, on peut parler vite sans pour autant parler « bien », c'est-à-dire sans faire trop de fautes de vocabulaire et de syntaxe. A l'oral, les trois quarts des impressions négatives des auditeurs (des juges natifs et non-natifs de niveau avancé) seraient d'ordre temporel - pauses, répétitions, débit de parole – alors que les impressions d'ordre non-temporel concerneraient principalement la prononciation, la grammaire et le vocabulaire (Rossiter, 2009).

On distingue deux types de mesures pour évaluer la fluence :

1- les mesures temporelles classiques que nous appellerons « quantitatives », telles que la vitesse de parole, la vitesse d'articulation, le taux de phonation et les pauses (silencieuses et sonores).

2- les autres mesures que nous appellerons « qualitatives » telles que la longueur du segment de parole et les reprises (répétitions, reformulations, faux départs et abandons).

## **4.2 Les mesures quantitatives**

Grosjean et Deschamps (1972) proposent pour mesurer la fluence de prendre en compte des variables temporelles qu'ils appellent complexes. Ces variables sont calculées à partir des variables simples (mesures brutes).

---

55 Notre traduction de : « Such characteristics of L2 fluency as speech rate and the length of fluent runs between pauses may reflect automaticity fairly directly. »(Schmidt, 1992 : 362).

#### 4.2.1 La vitesse de parole

La vitesse de parole (VP) est le temps total que dure la production d'un individu. Ce temps de production est la somme d'éléments plus simples (variable simples) que sont le temps d'articulation et le temps de pause, conformément à Goldman-Eisler (1968) cité par Grosjean et Deschamps (1972 : 131). La VP est mesurée en mots produits par minute ou en syllabes par seconde. Ce qui donne la formule suivante :  $VP \text{ (temps total de locution)} = \text{temps d'articulation} + \text{temps de pause}$  « où le terme *pause* réfère aux pauses non sonores » (*ib.*). Cette vitesse de parole peut être différente chez des locuteurs L2 et peut parfois varier selon la langue.

Par exemple en anglais, le débit normal chez les natifs (dans un petit corpus oral, tâche narrative) est de 130 à 170 mots par minute selon Hilton (2007) et, en français, de 130 à 200 mots par minute (Hilton, 2008a : 67) ; 173 mots par minute en moyenne soit 2 à 3 mots par seconde selon Levelt<sup>56</sup> (1989 : 22). En français, on l'estime à 173,6 mots par minute en moyenne selon Grosjean et Deschamps (1972) et peut varier selon la tâche, le contexte social, *etc.*

#### 4.2.2 La vitesse d'articulation

Cette mesure complète la mesure de la vitesse de parole. Elle ne prend en compte que les syllabes ou les mots articulés. En effet, lorsque nous parlons, nous alternons temps d'articulation et temps de pause. En L2, si un locuteur pause plus qu'il n'articule, il est fort probable que cela indique des difficultés dans l'encodage des idées, et que les mécanismes de la production ne soient pas automatisés (chapitre 3). La vitesse d'articulation apporte donc des précisions sur le comportement verbal des locuteurs (plus ou moins de syllabes articulées) et complète la mesure de la vitesse de parole qui, à notre avis, est moins fine. Elle calcule le nombre de syllabes articulées par seconde sans le temps de pause. D'après Schmidt (1992), si la vitesse d'articulation augmente dans le temps (en relation avec d'autres paramètres comme le taux de phonation et la longueur moyenne des pauses), l'automatisation aurait pris place dans l'articulateur (et peut-être même déjà dans le formulateur !).

---

56 Notre traduction de « Speech is normally produced at a rate of about two or three words per second ». (Levelt, 1989 : 22).

### **4.2.3 Taux de phonation (phonation time ratio)**

Une autre variable complexe, le taux de phonation (TdP), calcule le rapport entre le temps passé à parler et le temps de pause. Cette mesure permet de mesurer le temps passé à articuler, par rapport au temps passé à hésiter, et de voir lequel des deux phénomènes joue le plus grand rôle dans l'augmentation ou la diminution de la vitesse de parole. Si ce rapport augmente dans le temps ou à défaut reste stable, le locuteur aura sans doute gagné en fluence et cela grâce à l'automatisation des processus de bas niveau. Cette mesure importante découle des deux précédentes. On peut aussi calculer, à l'inverse, le taux d'hésitation qui mesure le temps passé à pauser (par rapport au temps passer à articuler), tout dépend du point de vue adopté.

### **4.2.4 Les pauses, ou variables temporelles secondaires**

Les pauses, considérées comme des variables secondaires par Grosjean et Deschamps (1972), sont révélatrices d'une plus ou moins grande aisance du locuteur. Le phénomène des pauses est toujours associé à la fluence. Deese (1980 : 72) note le rôle que jouent les longues pauses chez les locuteurs natifs les plus expérimentés, dans un but rhétorique. En L1, les pauses longues se situeraient en début d'énoncé pour permettre la planification (Hilton, 2008a). Tout discours oral est composé de périodes de locution et de périodes d'hésitations qui peuvent être des pauses silencieuses (ayant souvent une fonction physiologique – c'est la pause respiratoire) et des pauses sonores. En L2, les hésitations seraient des indices très fiables de la plus ou moins grande aisance à l'oral (et en L1 aussi) et de l'automatisation (ou du manque d'automatisation) dans le traitement verbal. Leur diminution dans le temps pourrait signaler en effet l'automatisation d'un certain nombre de traitements formels.

Goldman-Eisler (1968) définit les pauses comme des interruptions du mouvement articulatoire. La littérature distingue trois types de pauses : les pauses liées à des phénomènes articulatoires (par exemple, lorsque deux occlusives se suivent), les pauses respiratoires nécessaires pour la prise de souffle, les pauses liées à des hésitations et les pauses liées à la planification. Ces différents types de pauses sont présentes aussi bien en L1 qu'en L2, sans doute avec une distribution différente. En L1, chez des sujets normaux, les pauses ne sont pas systématiquement des indicateurs d'un manque de fluence (ils peuvent l'être toutefois) mais sont considérées plutôt comme normales dans le discours. Les hésitations et les disfluences font partie intégrante du traitement des informations en temps réel. Elles sont nécessaires non seulement pour la production mais aussi pour la réception. Selon Hilton (2008a :67) en L1,

« environ un tiers de la production est consacré à différentes formes d'hésitation (pauses silencieuses, pauses remplies ou vocalisées et autres bruitages paralinguistiques, reprises ». Elles ont de multiples fonctions (Chafe, 1980) : elles permettent l'organisation et la planification des énoncés, la prise de souffle mais aussi la mise en relief de certains éléments du discours.

En L2, Les hésitations sont plus nombreuses et plutôt révélatrices des difficultés rencontrées par les apprenants. Elles sont particulièrement utiles pour révéler où on peut facilement avancer en production, et où cela devient difficile<sup>57</sup> (Chafe, 1980 : 171). Un grand nombre de recherches ont également montré que les hésitations à l'intérieur d'un groupe syntaxique sont caractéristiques des productions non-natives et le signe d'une parole disfluente (Hilton, 2008a ; Pawley et Syder, 2000 : 170). En L2, chez un locuteur inexpérimenté, l'hésitation sera souvent générée par un blocage (lexical, syntaxique) ou parfois due au stress (surtout dans une situation de production contrainte, comme c'est souvent le cas en classe de langue). Nous nous intéressons donc ici au rapport entre les hésitations et l'aisance en production orale en L2.

Les études entreprises par Henderson et collègues (1965) et résumées par Goldman-Eisler (1968) concernent la distribution des pauses pour la L1. Dans des passages d'entre 350 et 450 mots tirés d'interviews enregistrés, l'équipe de Henderson a constaté que chaque locuteur alterne de courtes périodes de parole avec de longues pauses et des périodes de parole plus longues mais avec des pauses plutôt courtes<sup>58</sup> (Goldman-Eisler, 1968 : 81). Ces suites sonores plutôt courtes et les longues pauses seraient liées à la phase de planification (la planification dans le modèle de Levelt, 1989). Ces pauses longues (pour la planification) permettent donc aux locuteurs de produire de plus longues périodes de parole avec de courtes pauses. Les pauses chez les locuteurs natifs peuvent donc avoir une fonction d'aide à la planification (Goldman-Eisler, 1968). Si les pauses d'hésitation aident à la planification, les pauses respiratoires en L1 tombent, toujours d'après Goldman-Eisler, à des moments clés de l'organisation du discours.

Cependant, Goldman-Eisler constate que la distribution des pauses respiratoires et d'hésitation est différente entre une tâche de « parole spontanée » et une tâche de « lecture »

---

57 « Hesitations are especially useful in showing us where it is easy to move on in speech production) and where it is difficult. » (Chafe, 1980 : 171)

58 « Periods with relatively long pauses and short utterances alternated with periods with relatively short pauses and long speech utterances. » (Goldman-Eisler, 1968 : 81)

Dans une tâche orale spontanée, il est difficile de contrôler parfaitement le traitement psycholinguistique<sup>59</sup> (1968 : 95). Elle constate que 34,1% des pauses sont des pauses respiratoires en production orale spontanée et les deux tiers restants des pauses d'hésitation ; et 68,9% de ces pauses respiratoires apparaissent à des frontières grammaticales. En production orale spontanée, 61% des pauses d'hésitation apparaissent à des jonctures grammaticales dans les échantillons les plus fluides (1968 : 95). Et dans les échantillons les moins fluides, seulement 47% des pauses sont grammaticales<sup>60</sup>. 54% des pauses d'hésitation ou de planification apparaissent à des frontières grammaticales dans les passages les plus hésitants et 77,6% à des périodes de fluence. Goldman-Eisler reconnaît que la nature de la tâche est importante dans la production. « Une tâche facile rend la production plus fluide. Cela permet aux locuteurs de regrouper des unités linguistiques de telle façon que les pauses rendent la communication plus aisée »<sup>61</sup> ; (Goldman-Eisler, 1968 : 95).

#### *a) Les pauses vides ou silencieuses*

A part les pauses respiratoires qui sont nécessaires pour la prise de souffle, les pauses vides ou silencieuses, dont les raisons sont multiples, représentent la majorité des pauses en L1 comme en L2. Plus importantes cependant en L2, elles marquent une difficulté lexicale ou syntaxique, à partir d'une certaine durée. Riggenbach (1991 : 426) distingue différents types de pauses vides en fonction de leur durée : micro-pauses lorsqu'elles sont inférieures ou égales à 200ms, pauses d'hésitations entre 300 et 400ms, et pauses vides à partir de 500ms. Cependant les études plus récentes ne font pas cette distinction et sont d'accord pour considérer une pause entre 200 et 250 ms comme une pause d'hésitation :

- i. supérieures à 200 ms selon Hilton (2011) en conformité avec Butterworth (1980 : 156), Cordier (2013).
- ii. supérieures ou égales à 250ms (Grosjean et Deschamps, 1972 ; Ejzenberg, 2000, Perez-Bettan, 2009), à partir de 280ms pour des questions pratiques (Wood, 2006).
- iii. supérieures à 400 ms (Freed, 1995, 2000 ; Riggenbach, 1991 ; Towell et al. 1996).

#### *b) Les pauses sonores*

---

59 « [...] in spontaneous speech speakers can hardly be in complete control of the psycholinguistic process ». (Goldman-Eisler, 1968 : 95)

60 « thus even in the fluent periods of the samples investigated, only 61% of the gaps occurred in grammatical places ». « even in the most hesitant parts [...] 47% of pauses occurred in grammatical places. » (Goldman-Eisler, 1968 : 95)

61 « when the cognitive task is easy, speech is fluent and attention can be paid to grouping linguistic units in such a way that pausing serves communication. » (Goldman-Eisler, 1968 : 95)

Les *pauses sonores*, au premier rang parmi les pauses d'hésitation en français d'après Grosjean et Deschamps (1972), sont des interruptions momentanées du flux langagier. Elles peuvent être décomposables en syllabes allongées qui représentent en moyenne 57 % du nombre total des pauses sonores en L1.

[les pauses sonores] réfèrent à tout prolongement anormal de syllabes (en fin de mot) ou de mots monosyllabiques (ex : « je crois *que*: je vais venir... ») et en pauses remplies qui réfèrent à tout procédé d'hésitation intervenant dans le langage, c'est-à-dire (euh, hum, en français), /ə, œ, ø, əm/ et dont /ə/ est le plus fréquent (ex : l'endroit où je suis allé /ə/ était très beau). (Grosjean et Deschamps, 1972 : 132)

36% des locuteurs natifs ne respectent pas cette tendance, soit ils font autant de syllabes allongées que de pauses remplies, soit ils en font davantage. Les allongements de syllabes sont très présents en français qui est une langue à syllabation ouverte (Grosjean, 1980) et représentent près de 30% des pauses sonores. Elles sont donc presque aussi importantes que les hésitations vocales. Elles permettent au locuteur de réfléchir à la suite de son énoncé et ont donc une fonction d'aide dans le discours au même titre que les autres hésitations. En L2, la distribution des pauses sonores entre pauses remplies et allongements peut être aussi inégale et dépendre de la langue maternelle des locuteurs.

Certains chercheurs (Raupach, 1980) comptent comme des pauses remplies ce qu'ils nomment « pauses lexicales », *j'sais pas ou chepa, je pense, des choses comme ça*. Mais ces pauses dites lexicales sont aussi considérées comme des unités à part entière appelées remplisseurs (*fillers*) pour d'autres chercheurs (Erman et Warren, 2000 ; Forsberg, 2006 ; Perez-Bettan, 2009 ; Cordier, 2013). Cependant la plus grande différence entre les productions natives et les non-natives concerne la place des pauses dans le discours.

#### **4.2.5 Localisation et distribution des pauses silencieuses**

##### *a) en L1*

En L1, Grosjean et Deschamps (1972 : 145) ont montré que les trois quarts des pauses interviennent à des jonctures grammaticales avec cependant une différence selon le type de pauses : 38% des pauses sonores sont réalisées en dehors des points de ruptures grammaticaux contre 25% des pauses non sonores.

Les pauses peuvent se trouver à l'intérieur du SN (soit entre l'article et le nom, soit entre le nom et l'adjectif) ; ou à l'intérieur du SV (entre le verbe et le SN), entre deux éléments du verbe, etc.). Elles peuvent se placer également entre le SN et le SV, après les compléments

adverbiaux, entre SN et SV et compléments adverbiaux, avant ou après des mots de coordination ou d'apposition. En L1, toujours selon ces chercheurs, trois quarts des pauses non sonores se trouvent à des jonctures grammaticales. Sur les 25% qui restent, 5% sont situées entre le SN sujet et le SV, 4% entre des mots coordonnés ou apposés. Les 16% restants se produisent à l'intérieur d'un groupe de sens (soit à l'intérieur du SN soit à l'intérieur du SV). « Nous pouvons donc dire que les  $\frac{3}{4}$  des pauses du français oral spontané se trouvent à des endroits où il pourrait y avoir une pause en lecture. » (Grosjean et Deschamps, 1972 : 147). Cette étude de l'emplacement des pauses en L1, même si elle prend en compte aussi bien les pauses respiratoires que les pauses d'hésitation, montre que la pause en fin de phrase est considérée comme « le point de rupture naturel qui permet la reprise d'air et que la pause à l'intérieur de la phrase aurait plutôt tendance à être une pause d'hésitation ou une pause stylistique ».

### c) En L2

En L2, les pauses se situent, chez les locuteurs non fluents, à l'intérieur des propositions et non aux frontières et/ou à l'intérieur des SV ou des SN (Deschamps, 1980 : 261 ; Hilton, 2008a). Ces pauses sont dites « agrammaticales » à la différence de la plupart des pauses - grammaticales - chez les natifs ou chez les locuteurs très avancés (*native-like*). Les pauses agrammaticales, qui se situent à l'intérieur d'un groupe syntaxique, renforcent l'impression de non fluence. Deux types de pauses, une pause remplie suivie d'une pause vide ou une pause remplie au milieu de deux pauses vides, forment un groupe d'hésitation (GH).

Toutes ces variables sont évaluées par des mesures quantitatives, basées sur la vitesse en production, temps passé à parler, à hésiter, etc. D'autres mesures plus récentes, comme la longueur du segment de parole calculée en nombre de syllabes ou de mots et le nombre de reprises mesurent la qualité de la production.

## 4.3 Les mesures qualitatives

### 4.3.1 Longueur du segment de parole

Cette mesure, qui n'est pas à proprement parler une variable temporelle, est considérée comme une des mesures les plus intéressantes, surtout dans les études comparant des productions orales produites à deux moments différents (Towell et *al.* 1996 ; Towell, 2002 ; Cordier, 2013). Un segment de parole est une suite sonore qui se trouve entre deux pauses vides ou un groupe d'hésitations (composé de pauses vides et de pauses remplies). La



longueur du segment de parole se mesure en nombre de mots ou de syllabes entre deux pauses. L'allongement du segment de parole dans le temps indique que la fluence de l'apprenant s'est améliorée et traduirait donc un traitement plus aisé de la langue. Pour Towell, Hawkins et Bazergui (1996), l'augmentation de la longueur moyenne du segment de parole (associé à d'autres mesures comme le taux de phonation et la durée des pauses) montrerait même que la procéduralisation aurait pris place dans le formulateur.

C'est précisément sur l'allongement du segment de parole que semble résider, du moins en partie, le développement de la fluence. Autrement dit, l'apprenant qui parle couramment ne serait pas tellement quelqu'un qui fait moins de pauses, qui hésite moins ou qui a un débit de parole plus rapide, mais c'est surtout quelqu'un qui est capable de produire, entre deux grandes pauses (généralement égales ou supérieures à 0.25 secondes), des segments de parole de plus en plus longs. (Théophanous, 2014 : 122)

La fonction des pauses silencieuses entre deux segments de parole ne fait pas toujours consensus et peut varier selon les chercheurs. Pour certains chercheurs (Riggenbach, 1991 ; Freed, 1995), on l'a vu, la durée des pauses doit être supérieure à 300ms ou même à 400ms pour être considérée comme une pause d'hésitation (plutôt qu'une pause respiratoire). Un des problèmes des recherches sur les mesures de la fluence résulte des points de vue différents des chercheurs pour la délimitation du segment de parole. En effet, certains considèrent qu'une pause remplie ou un allongement peut délimiter un segment de parole au même titre qu'une pause vide (Cordier, 2013). Il semble pourtant que ces pauses remplies soient, quand même, la manifestation d'une tentative de production.

#### **4.3.2 Les reprises ou corrections**

Certaines études ont cherché à distinguer entre les différents types d'hésitation et les causes de celles-ci. En L1, Germain et Netten (2002 : 13) distinguent les « hésitations pragmatiques » et les « hésitations langagières ». Les hésitations de type pragmatique peuvent être « fonctionnelles » : elles se caractérisent alors par une longue hésitation pour chercher ses mots afin d'exprimer sa pensée. Elles peuvent aussi être « pragmatique socioculturelle », lors d'une hésitation entre le « tu » ou le « vous » ou lorsqu'un locuteur fait une erreur de registre de langue par exemple. Les « hésitations langagières » peuvent être « grammaticales » se caractérisant par des hésitations d'accord du verbe ou d'accord en genre et en nombre. Lorsque le locuteur hésite dans l'enchaînement des phrases, on parle d'« hésitations langagières discursives ».

Ces hésitations, qui ne sont pas à proprement parler des pauses, sont aussi appelées reprises ou corrections (Levelt, 1989 : 460-63). Elles se traduisent par des répétitions, des reformulations, des abandons ou des faux départs. Les répétitions et les faux départs font partie des variables secondaires pour Grosjean et Deschamps (1972 : 132-133). D'après Goldman-Eisler, 16,80% des disfluences en L1 seraient des répétitions de mots grammaticaux (64,22%), mais aussi dans une moindre mesure des groupes de mots (24,77%), des phrases entières (3,67%) et des phonèmes ou syllabes dans la même proportion.

Dans le modèle de production de Levelt (1989) la quatrième étape, l'auto-surveillance, permet au locuteur de corriger les erreurs de production dans le discours oral. Ces corrections (*repairs*) peuvent servir à lever une ambiguïté qui pourrait gêner l'interlocuteur. La prise de conscience de l'erreur peut se faire avant ou après le début du message si l'erreur se situe au niveau conceptuel. Avant le début du message, il n'y aura pas de correction mais sûrement une hésitation (pause sonore ou silencieuse), après le message la correction prendra la forme d'une reformulation. Les corrections peuvent être d'ordre lexical, morphologique, syntaxique, phonologique ou articulatoire (Levelt, 1989 : 463). Lors d'erreurs lexicales, lorsque le mot activé a déjà été formulé et qu'il n'est pas le bon, la correction est immédiate.

Les corrections sont dépendantes du contexte, de la situation de communication. Elles seront corrigées si elles posent problème à l'interlocuteur et si elles risquent de gêner la communication. Pour Noteboom (1980, cité par Levelt, 1989 : 463), seules 75% des erreurs phonologiques et 53% des erreurs lexicales en L1 sont corrigées par les locuteurs. Même s'ils les ont toutes détectées, en production orale, ils ne se soucient pas de les corriger toutes. Si les corrections existent en L1, les productions en L2 sont souvent perçues comme disfluentes à cause justement du grand nombre de corrections qui peuvent prendre diverses formes (répétition, reformulation, abandon ou faux départ).

#### *a) Les répétitions*

Wiese (1984), dans une étude sur des natifs allemands et anglais, constate deux à trois fois plus d'hésitations en L2 qu'en L1, dues probablement au manque d'automatisation des connaissances déclaratives ; et ce malgré des variations individuelles. Les raisons pour lesquelles les débutants hésitent viennent de problèmes à retrouver et produire des unités lexicales, à choisir la forme grammaticale correcte et à corriger leurs productions. Les formes nécessaires à l'encodage formel automatique ne sont tout simplement pas disponibles en mémoire à long terme : elles doivent être laborieusement et consciemment assemblées,

exigeant un grand effort ; alors que les locuteurs très avancés en L2 seront capables de planifier à l'avance et hésiteront surtout lorsqu'ils doivent exprimer des idées complexes. Les répétitions apparaissent lorsqu'il y a un problème dans l'automatisation des unités, qui peut être un blocage lexical, syntaxique ou morphologique. Plus de la moitié des hésitations longues en L2 tirées du corpus *Parole* (2008b) « sont directement liées à la compétence lexicale. 32% des pauses de 2 à 3 secondes et 47% des pauses de plus de 3 secondes de durée sont suivies directement par une erreur lexicale » (Hilton, 2007). « Les « recherches lexicales » qui expliquent environ 20% des pauses disfluentes sont soit totalement explicites (le sujet déclarant qu'il ne trouve pas ou ne connaît pas un mot, ou demandant directement à l'investigateur de le lui fournir), soit extrapolées mais hautement probables – la pause étant suivie d'un mot correct de fréquence basse (*parade, crane*), par un hyperonyme à la place d'un mot plus précis mais moins fréquent (*animal* pour *elephant, nose* pour *trunk*), ou par une paraphrase. » (Hilton, 2007 : 76).

#### b) *Les reformulations*

Les reformulations, souvent d'ordre morphologique, sont plus nombreuses dans les productions non-natives et indiquent là encore que le processus d'automatisation n'est pas encore fait : on reformule lorsqu'on s'aperçoit de son erreur. Cependant Freed (1995) dit que certaines hésitations comme les reformulations ne sont pas toujours l'indicateur d'un manque de fluence, mais plutôt un souci de précision (*accuracy*). On les trouve aussi bien dans les productions L1 que L2. Bien souvent, les locuteurs avancés en L2 veulent être plus précis et s'appuient moins sur le langage préfabriqué que les débutants. On peut également penser que le fait de s'auto-corriger lors d'une erreur permet à l'apprenant de garder la parole qui leur garantit une certaine fluidité. « On peut penser que le locuteur procède ainsi, mu par son désir de paraître, non seulement plus précis, mais aussi (surtout ?) plus fluent, puisqu'en disant la même chose de deux façons différentes, il garde nécessairement la parole plus longtemps ». (Théophanous, Perez-Bettan et Hilton, 2014 : 199)

Parfois le locuteur abandonne une construction pour une autre ou passe à autre chose. Ces abandons ou faux départs représentent 14,17% des pauses secondaires dans les productions natives (Grosjean et Deschamps, 1972). Leur nombre important dans les productions natives peut sans doute expliquer la difficulté des locuteurs non-natifs en production. Hilton (2011a : 152) constate, en L2, que les abandons sont plus importants chez les locuteurs avancés car ils

reflètent des processus d'autocorrection de niveau discursif<sup>62</sup> alors que les locuteurs non fluent se focalisent sur des processus de bas niveau (caractérisés par des répétitions et des reformulations).

La tentation est grande de comparer les résultats obtenus en L2 dans les différentes recherches sur la fluence. Il faut faire preuve de vigilance car chaque étude suit ses propres critères pour mesurer la fluence. Comme la fluence dépend du type de tâche imposée, on ne peut pas comparer une tâche d'expression libre, une interview et une tâche narrative, par exemple.

#### **4.4 Différences individuelles dans l'aisance**

Si en L2, la vitesse de parole et les autres mesures présentées ci-dessus sont souvent liées au niveau des apprenants, l'aisance en L2 peut aussi dépendre de facteurs extralinguistiques (Dewaele, 2002) et il sera difficile de les évaluer précisément. Des facteurs inter ou intra-individuels tels que l'âge, le statut social, l'extraversion, l'anxiété communicative, la fatigue, la situation de communication, l'auditoire peuvent affecter la fluence de la production en L1 mais aussi en L2 (ou en Lx), avec cependant plus de pression dans une langue étrangère (voir Dörnyei et Skehan, 2003 et Dulay *et al.* 1982).

Des études sur le bilinguisme ont montré que des mécanismes neurophysiologiques pourraient influencer la fluence en production orale (Cook, 2002 : 223). Grosjean (1992, 1998, 2001) constate aussi, chez des locuteurs plurilingues, qu'un certain nombre de variables interagissant, sans doute de nature neurochimique, peuvent être liées à la fluence (Cook, 2002 : 233). Par exemple, un locuteur extraverti gèrera mieux la production de parole en état de stress qu'un locuteur introverti.

L'inhibition peut être un frein à l'apprentissage des langues et surtout en production orale. Elle peut réduire la part de prise de risques de l'apprenant, essentielle pour la prise de parole en L2 devant un groupe. En situation d'anxiété, la nervosité et le stress peuvent avoir des effets sur la fluence. Et c'est sûrement avec l'expérience que le locuteur arrivera à surmonter sa peur. Dewaele (2002) analyse la production de parole de cent étudiants flamands de Français L2 pour juger de l'anxiété et de ses effets sur la production orale. Il montre que l'anxiété ne serait pas due à la personnalité de l'apprenant mais aux contextes individuel et sociétal qui déterminent le niveau de l'anxiété en communication. Le fait que le français soit

---

<sup>62</sup> « restarts [...] and false starts... found in more fluent productions also appear to reflect discursive-level monitoring processes. » (Hilton, 2011a : 152)

perçu comme une « langue de prestige » en Flandre et qu'il soit un marqueur social conduit à des niveaux d'anxiété qu'on ne retrouve pas lorsque ces mêmes apprenants utilisent l'anglais.

Il existe des variations parfois très importantes entre apprenants d'un même niveau, comme il peut en exister aussi parfois chez les natifs. Cook (2002) se demande pour quelles raisons certains apprenants accèdent à une certaine maîtrise de la langue alors d'autres n'y parviennent pas. Pourquoi la fluence diffère-t-elle d'une production à l'autre et pourquoi change-t-elle d'un apprenant à l'autre alors que dans les classes de langue les apprenants ont tous reçu le même enseignement ? Pour cet auteur, les variations synchroniques de la fluence en L2 pourraient avoir des causes neurobiologiques qu'il faudrait identifier, mais qui dépassent le cadre de cette étude.

Cependant, l'anxiété peut aussi dans certaines situations et pour certains apprenants avoir des effets positifs et même faciliter la production (Spielmann et Radnofsky, 2001 cités par Lightbown et Spada, 2006 : 62). Ce ne serait probablement pas la personnalité en soi qui serait un frein à la L2 mais, combinée à d'autres facteurs, elle peut favoriser ou pas la réussite de la prise de parole.

La qualité d'une production orale peut être influencée par un certain nombre de facteurs « neuro-psycho-physiologiques » qui jouent un rôle dans la fluence. La maîtrise de la langue chez certains locuteurs « lettrés » n'est pas toujours un gage de bonne fluence ; certains écrivains très connus - Modiano en situation d'interview - ont de réelles difficultés pour produire un discours fluide dans leur propre langue maternelle. Les causes peuvent en être multiples (timidité, réserve, *etc.*).

Le type de tâche demandée peut être également un facteur supplémentaire aux difficultés des locuteurs en L2 mais que l'on peut aussi observer en L1. En effet, on ne parle pas à la même vitesse lors d'un exposé oral sur un sujet spécifique planifié, lors d'une interview libre ou avec des pairs. Des chercheurs ont montré qu'en situation de communication intense et en temps limité en L1, on peut parler très vite dans des situations spécifiques comme les reportages sportifs par exemple (Kuiper, 1991). L'aisance à l'oral dépendrait donc de plusieurs facteurs inhérents à la tâche, à la situation de communication et/ou aux acteurs participants.

[...] le comportement des variables temporelles varie considérablement selon qu'il s'agit de langage écrit-oralisé ou de langage spontané et, même, selon la nature de ce langage spontané. Nous ne présentons pas le même débit de parole, nous ne pausons pas aux mêmes endroits, la longueur des pauses est différente selon que nous participons à une discussion, que nous faisons un discours, que nous décrivons un dessin ou que nous sommes interviewés. (Grosjean et Deschamps, 1972 : 130)

#### **4.5 Rôle d'un séjour à l'étranger sur la fluence**

Les études s'accordent à dire qu'après un séjour à l'étranger dans la langue cible la fluence s'améliore dans le temps (Freed, 1995, Riggensbach, 1991, Segalowitz et Freed, 2004). Raupach (1980) et Möhle (1984) constatent que les segments de parole (suites de mots entre deux pauses) sont plus courts en L2 mais que, après trois ans en immersion dans le pays, les locuteurs non-natifs produisent des segments plus longs, jusqu'à 95% de plus (voir aussi Towell *et al.* 1996 ; DeKeyser, 1991 ; Freed, 1995, 2000 ; Cordier, 2013). Segalowitz et Freed (2004) constatent que le groupe d'apprenants en espagnol L2 produit un discours plus fluide que le groupe contrôle, après une immersion dans le pays de la langue cible. Pour Towell et Dewaele (2005), le développement de la fluence après un séjour de six mois en France, chez douze apprenants anglophones apprenant le français serait lié au développement de l'usage des séquences préfabriquées. La vitesse de parole des apprenants a augmenté de façon significative entre temps 1 (avant) et temps 2 (après) et les locuteurs ont progressé dans la façon dont ils emploient les séquences. Cependant, pour certains chercheurs, si les apprenants L2 produisent un discours plus fluide et plus riche, cela serait dû à l'emploi des séquences préfabriquées acquises après un séjour en immersion en milieu homoglotte (Cordier, 2013). Des chercheurs ont également constaté l'acquisition des expressions conventionnelles des natifs par les locuteurs L2 en situation d'immersion (Bardovi-Harlig et Bastos, 2011) mais aussi l'appropriation des déviations du parler natif (comme la réduction phonologique) (Regan, 1995).

On a vu que la fluence de production pouvait être évaluée au moyen de mesures précises quantitatives et qualitatives. Ces mesures permettent de juger du degré d'automaticité des procédures de traitement dont dépend la fluidité. Un séjour à l'étranger, dans le pays de la langue cible, ainsi que d'autres facteurs extralinguistiques peuvent favoriser la fluence – et notre étude tentera de creuser ces phénomènes.

# Chapitre 5 Méthodologie

---

Ce chapitre méthodologique a trois objectifs principaux : décrire comment notre recherche doctorale a été menée, expliquer comment les données collectées ont été analysées, et expliquer les choix méthodologiques qui ont été faits.

Comme suite aux résumés théoriques dans les quatre premiers chapitres de cette thèse, nous formulons les questions de recherche avant de présenter le protocole utilisé pour obtenir des réponses :

## 1) les SP chez des apprenants de FLE de niveaux différents

- Dans quelle proportion les apprenants de FLE utilisent-ils les SP ?
- Quels sont les types de SP produites ?
- Observe-t-on des différences entre les groupes de niveau et les sujets à l'intérieur des groupes ?

## 2) développement des SP

- Dans quelle mesure l'utilisation des SP s'est développée sur une période de six mois ?
- Observe-t-on un changement dans l'emploi des types de SP ?
- La tâche a-t-elle un effet sur l'utilisation des SP ?
- Quel groupe d'apprenants semble avoir progressé le plus ?

## 3) développement de la fluence et relation SP / fluence

- Est-ce que la fluidité des apprenants s'est développée sur une période de six mois ?
- La fluence est-elle corrélée à un usage plus ou moins important des SP ?

Dans la première partie de ce chapitre, nous décrivons en détail nos choix puis nous présenterons le public ayant fourni notre corpus. Après avoir exposé l'objet du dispositif mis en place et avant l'analyse des résultats, nous expliquerons comment nous avons transcrit les enregistrements. Nous établirons ensuite la liste des critères qui nous ont permis d'identifier les séquences préfabriquées dans les productions orales des sujets testés. Enfin nous donnerons notre typologie des SP.

Selon notre hypothèse de départ, sur une période de six mois, nos sujets testés utiliseraient des SP en plus grand nombre et/ou de nature plus riche qui faciliterait leur production en termes

de rapidité et de fluidité, en particulier la deuxième fois. Nous présentons donc tous les calculs qui nous permettront d'évaluer a) la quantité et la variété des SP ; b) la proportion du LP dans les productions ; c) la fluence de production.

## 5.1 Choix de la méthode

Conformément à d'autres recherches et à notre objectif (examiner le développement du langage préfabriqué dans le discours en L2), nous avons opté pour une tâche narrative (Lennon, 1984 ; Towell *et al.* 1996, Towell, 2002 ; Wood, 2006, 2010). D'autres chercheurs ont préféré utiliser des tâches moins contraignantes comme l'interview ou la discussion où les locuteurs, plus libres, produiraient un plus grand nombre de SP (Towell, 2002 ; Forsberg, 2006 ; Cordier, 2013). En effet, selon Goldman-Eisler (1968), Grosjean et Deschamps (1972), Towell, Hawkins et Bazergui (1996) et Cordier (2013), les locuteurs seraient moins spontanés lors d'une tâche narrative.

Consciente des difficultés que peut poser la narration - manque de spontanéité, difficultés lexicales pour rendre compte d'une histoire - nous avons malgré tout opté pour la tâche narrative parmi toutes les possibilités de sollicitation des données à l'oral. Nous avons écarté l'interview qui nous paraissait difficile à exploiter. Nous voulions en effet montrer comment le langage préfabriqué évolue dans le temps à partir d'un support identique à chaque fois. Si l'entretien ou la discussion permet d'avoir des données langagières spontanées, il est en revanche difficile d'obtenir des productions comparables parmi les sujets testés, d'une fois sur l'autre, ou d'un sujet à l'autre. Rappelons ici une de nos questions de recherche selon laquelle nous nous proposons de vérifier le développement du langage préfabriqué dans le discours des apprenants. Pour cela, il fallait nous assurer de la plus grande homogénéité dans les conditions de la sollicitation de la parole. A nos yeux, l'interview peut facilement déborder avec le risque de nous retrouver avec des productions langagières peu comparables ; alors que la narration orale permet de collecter des données comparables. Enfin, la tâche narrative permet de montrer comment les apprenants d'une L2 s'approprient certaines SP, les modifient et les utilisent dans une narration à la troisième personne. Un autre aspect intéressant de la tâche narrative est la possibilité pour les apprenants de contourner une difficulté lexicale en utilisant des séquences préfabriquées qu'ils adaptent à un contexte spécifique.

Parmi les différentes tâches narratives possibles, nous nous sommes fixée sur le compte-rendu d'une histoire à partir de deux supports différents. Là encore, étant donné qu'il nous fallait



déclencher du discours « spontané » chez l'apprenant, nous recherchions un dispositif « sans paroles », « muet », de sorte que l'imitation ou la restitution des bouts de langue déjà vus ou entendus de la part de l'apprenant ne soit pas possible. Autrement dit, il fallait que l'apprenant soit le plus libre possible dans sa parole, tout en ayant un ensemble comparable de concepts à encoder.

### **5.1.1 Description des deux dispositifs :**

#### ***5.1.1.1 Dispositif 1 : cartoon (BD)***

Nous avons choisi un ensemble de huit vignettes en noir et blanc sans paroles racontant une histoire facile à comprendre et amusante d'un cowboy à cheval qui revient d'un long voyage<sup>63</sup>. Fatigué, il décide de se reposer un peu sur son chemin mais lorsqu'il se réveille, une surprise l'attend. Un enfant qui passait par là lui fait une blague et à la place de son cheval lui accroche un petit cheval en bois.

Raconter ce qui se passe dans un cartoon (BD) nécessite du lecteur de décrire non seulement ce qu'il voit mais aussi de faire des liens d'une vignette à l'autre, ce qui aurait pu présenter une charge cognitive plus importante que la narration d'un film.

Dans cette tâche, deux types de discours sont sollicités : un discours descriptif (description des personnages : le cow-boy et l'enfant, le cheval et le jouet) et un discours narratif (« les aventures » du cowboy). Cependant d'autres types de discours (argumentatif ou explicatif) peuvent être produits. L'avantage de ce dispositif (Annexe 1) est qu'il est facilement transportable, facile à administrer, rapide à regarder et à restituer oralement : toute l'opération ne prend que quelques minutes.

#### ***5.1.1.2 Dispositif 2 : le film***

Notre deuxième dispositif (Annexe 2) est un film muet et court d'une durée de cinq minutes, tiré de la série *Les Minuscules* et intitulé *Prisonnière*. Les apprenants ont pour consigne de regarder attentivement le film et de le restituer.

Ce film raconte l'histoire d'une abeille qui, après avoir butiné des fleurs dans le jardin d'une maison, entre dans celle-ci et se retrouve prisonnière. Après avoir erré dans les différentes pièces de la maison à la recherche de fleurs, on la voit voler d'une pièce à l'autre et essayer de butiner, d'abord des fleurs artificielles disposées dans un vase (prix indiqué sur l'étiquette)

---

<sup>63</sup> Nous remercions ici vivement Jean-Luc Nespoulous, professeur à l'Université Toulouse-Jean Jaurès) qui nous a permis d'utiliser ce dispositif.

ensuite des tournesols peints sur une toile accrochée au mur mais elle s'aperçoit qu'elle ne peut plus sortir. Pendant ce temps<sup>7</sup>

g une mouche, qui l'espionnait, décide de lui venir en aide. Elle va dans la cuisine et rapporte l'un après l'autre des grains de café qu'elle dispose en forme de flèche dans le fond de la baignoire ; cette flèche indique à l'abeille la sortie par la bonde qui doit la mener à l'extérieur dans le jardin. L'abeille, qui a compris le message, suit la flèche et se retrouve enfin libre grâce à la mouche.

Lors de ces deux tâches, aucune autre consigne n'a été donnée et aucun vocabulaire fourni aux sujets. Nous présenterons les résultats des deux tâches réunies, en temps 1, ( $t_1 = bd_1 + film_1$ ) et en temps 2 ( $t_2 = bd_2 + film_2$ ). Les tests de Wilcoxon que nous avons appliqués ont montré qu'il n'y avait aucune différence significative entre les deux tâches (Une présentation des résultats de la préfabrication par tâche est cependant disponible en Annexe 13).

## **5.2 Sujets testés**

### **5.2.1 Groupe expérimental**

Le groupe d'apprenants (du français langue étrangère) comporte douze étudiants de l'École des Mines d'Albi, cinq jeunes hommes et sept jeunes femmes, âgés de 21 à 32 ans. Venus de pays divers (le Brésil (5), la Chine (3), le Japon (1), l'Espagne (1), l'Italie (1), le Canada (1)), ils étaient tous inscrits dans un programme d'études supérieures dans les métiers d'ingénieur. Dix d'entre eux étaient des doctorants en première ou seconde année. Les deux autres étaient en France dans le cadre d'échanges universitaires. Tous ces étudiants suivaient des cours de FLE. Neuf d'entre eux étaient placés au niveau débutant/intermédiaire et trois au niveau avancé, l'institution n'ayant pas les moyens matériels de financer trois classes de niveau. Tous les apprenants de cette institution ont été reçus par l'équipe enseignante qui a déterminé les niveaux suite aux entretiens avec la direction du département.

Les apprenants L2 à l'École des Mines ont trois heures de cours de français par semaine ; la plupart sont des doctorants qui travaillent dans un laboratoire de recherche scientifique avec des collègues français. Sept de ces apprenants venaient d'arriver en France et avaient suivi dans leur pays, pendant quelques semaines, des cours de français. Un autre avait pris des cours intensifs en milieu exolingue et avait appris par lui-même. Quatre autres avaient passé de deux à quatre ans en France.

Cinq autres sujets ont été recrutés au Département d'Etudes du Français Langue Etrangère de l'Université de Toulouse-Jean Jaurès, ce sont des jeunes femmes âgées de 19 à 34 ans, ayant fait des études dans leur pays (la Russie (1), les Pays Bas (3), les Etats Unis (1)) et voulant obtenir pour la plupart un niveau de français leur permettant de suivre des études supérieures en France afin de devenir professeur de français dans leur pays de naissance ou de trouver un travail en France. Elles ont toutes été placées au niveau intermédiaire ou avancé, selon l'évaluation de l'institution. Elles suivent un cursus de cours plus intensif que les sujets d'Albi. (Quatre autres sujets avaient été recrutés mais ne se sont pas présentés lors de la deuxième série d'enregistrements). Aussi bien les sujets de Toulouse que les sujets d'Albi n'ont reçu aucun enseignement spécifique des séquences, hormis les expressions que l'on trouve dans les manuels de FLE.

Par ailleurs, nos sujets non-natifs ont des camarades français avec lesquels ils communiquent quotidiennement. Le fait d'appartenir à un laboratoire de recherche et d'avoir un directeur de thèse français ou de partager la vie d'un compagnon français leur permet d'être en contact permanent avec la langue. La plupart de ces apprenants montrent une réelle motivation pour parler comme les natifs et utiliser les expressions de leurs camarades. Par conséquent, les enregistrements effectués à 6 mois d'intervalle devraient montrer une amélioration dans l'emploi qu'ils font des séquences préfabriquées. Les sujets ont été répartis par niveau, de débutant (niveau 1) à avancé (niveau 3) avec un niveau intermédiaire (niveau 2). Une grande partie des études sur le langage préfabriqué ont examiné le langage de sujets avancés et très avancés (Forsberg, 2006 ; Cordier, 2013 ; Forsberg-Lundell *et al.* 2014) ayant la même L1. Cependant les travaux de Forsberg (2006) et de Myles et collègues. (1998) se sont aussi intéressés au langage préfabriqué d'apprenants du français de niveau débutant mais ayant la même L1. Peu d'études hormis celle de Wood (2006, 2010) ou Hilton et collègues (2010) ont eu pour sujets des sujets de nationalités diverses, la difficulté étant de comparer la production de sujets de langues très éloignées ; nous pensons notamment aux sujets d'origine asiatique. Nous prenons donc le risque de ne pas avoir des résultats homogènes, sachant que nous ne pourrions pas généraliser nos résultats, compte-tenu du nombre réduit des sujets. Pour cette raison, nous privilégions les comparaisons des résultats des groupes entre eux et des sujets à l'intérieur du groupe.

Sujet	Niveau	Groupe	âge	Sexe	L1	Etudes	temps en France	Nombre d'années d'études de français
<b>AK</b>	1	Albi	26	M	Japonais	doctorant « sciences de l'ingénieur »	6 mois	2 semaines de français au japon (auto-apprentissage) un semestre à Nancy : cours débutant 2h par semaine
<b>FA</b>	3	Albi	25	M	Italien	doctorant 2ème année	18 mois	3 ans de français au collège et au lycée. un mois intensif avant Erasmus, Erasmus à Lyon (un semestre) 3h par semaine à l'EMA pendant 1 an
<b>HE</b>	2	Albi	29	M	Portugais	doctorant 2ème année	18 mois	2h par semaine au Brésil pendant 2 mois, 3h/semaine pendant 1 an à l'EMA (École des Mines Albi)
<b>JA</b>	1	Albi	26	F	Portugais	doctorant 1ère année	1 mois	2 mois de français au Brésil
<b>MA</b>	2	DEFLE	19	F	Russe	bac en Russie	18 mois	Quelques heures en Russie, vit avec un compagnon français
<b>MJ</b>	3	DEFLE	21	F	Hollandais	Master	4 mois	3 ans de français en hollandaise, projet : professeur de FLE en hollandaise
<b>MR</b>	2	Albi	25	M	Catalan	fin d'études Master 2	2 semaines	Études dans une école en suisse, grand-mère française
<b>NO</b>	2	Albi	24	F	Anglais Canada	doctorant stage en France (6 mois)	2 semaines	doctorat à NY, stage 6 mois en France 4 ans de français au collège cours à l'université
<b>NY</b>	1	Albi	23	f	Portugais	stagiaire fin d'études, puis doctorat	1 semaine et demie	4 mois de Français au Brésil, à raison de 4H/semaine à l'Alliance Française
<b>OL</b>	1	Albi	24	f	Chinois	doctorant 1ère année	12 mois	60 H de cours en Chine (grammaire) puis stage 2 mois intensif en France (Nantes) cours 3h par semaine en France
<b>PR</b>	3	DEFLE	34	f	Américain Hindi	DEFLE, ingénieur informatique	3 mois et 24 mois	Cursus de français au DEFLE, vit avec un français mais communiquent plutôt en anglais.

<b>QI</b>	3	Albi	32	f	Chinois	doctorant 2ème année	6 ans	Français en Chine (grammaire) 2 mois intensifs à Vichy puis Nancy stage en entreprise
<b>SU</b>	1	Albi	27	f	Portugais	doctorant 1ère année	2 mois	4 mois de français en 2008 puis cours de français intensif au Brésil
<b>VI</b>	2	Albi	21	m	Portugais	4ème année ingénieur	2 mois	2H / semaine au Brésil pendant 2 ans
<b>VO</b>	3	DEFLE	23	f	Russe	DEFLE, niveau Master	2 ans et demi	Motivation : après un voyage touristique en France à 15 ans apprendra le français, français à l'université en 2ème année
<b>XI</b>	1	Albi	22	f	Chinoise	doctorant 1ère année	2 ans	2 mois intensif à Nantes puis pendant 2 ans 3h/semaine, arrivée en 3ème année en France
<b>ZI</b>	2	DEFLE	19	f	Hollandaise	DEFLE, IUFM en Hollande	2 semaines	3 ans au Lycée, veut intégrer l'IUFM. Projet : devenir enseignante de français en Hollande

Tableau 1. Sujets non-natifs

### 5.2.2 Groupe contrôle

Nous avons également constitué un groupe contrôle de six locuteurs natifs (LN), composé de jeunes adultes volontaires, de formation universitaire supérieure en cours ou récente. Nous avons fait le choix de tester ces LN une seule fois même si l'idée de les soumettre deux fois au même test aurait pu être envisagée. Partant du principe que la production des LN est à mettre en relation avec la deuxième production de nos apprenants nous n'avons pas choisi cette option. Notre objectif (pour ce corpus natif) était simplement d'avoir des valeurs de référence (normes de fluidité et un inventaire des SP utilisées) pour nos considérations du développement du langage préfabriqué et son éventuel rapport avec l'aisance en L2.

<b>sujets</b>	<b>âge</b>	<b>Sexe</b>	<b>études</b>	<b>Langues parlées</b>
<b>AN</b>	30	F	Maîtrise de géographie, Dulco en arabe, Master 2, production audiovisuelle, travaille comme assistante de production	Anglais, arabe LV3 (LO), espagnol LV2
<b>CA</b>	29	F	Directrice artistique, graphiste, école de communication visuelle (Master)	Anglais, espagnol avancé
<b>JE</b>	24	M	Master de droit, école du barreau	Anglais
<b>MG</b>	21	F	Études de commerce	Anglais bilingue (famille anglaise), italien (9 ans), un peu danois
<b>TH</b>	20	M	3ème année de droit à Paris X Nanterre	Anglais courant, notions en espagnol
<b>VC</b>	21	F	4ème année école de commerce en stage alterné	Bilingue anglais (mère anglaise), italien courant

Tableau 2. Sujets natifs

### 5.3 Administration du dispositif et collecte des données

Nous avons rencontré chaque sujet individuellement à deux reprises et ce, à six mois d'intervalle environ. Nous ne pouvons pas étendre nos enregistrements sur une plus longue période, tous nos sujets ne restaient pas trois années en France, certains étant seulement en résidence pour un semestre. Cependant six mois devraient rester une période suffisante pour vérifier que nos sujets ont progressé dans l'acquisition des séquences en milieu naturel sans enseignement spécifique de celles-ci. D'autres études, notamment celle de Wood (2006), ont choisi une période plus courte de trois mois. Les premiers enregistrements ont eu lieu en début de semestre, en même temps que débutaient les cours de FLE c'est-à-dire au début du mois de novembre, les doctorants arrivant assez tard dans l'institution. La deuxième série d'enregistrement a eu lieu d'avril à juin 2010. Nous avons soumis les sujets à une interview

avant chaque série d'enregistrements. Ces rencontres ont permis d'avoir des informations personnelles mais aussi de comprendre comment les apprenants avaient appris le français, leur motivation, et leurs projets d'avenir.

Chaque rencontre a donné lieu à un nouvel enregistrement, effectué avec un appareil numérique facilement transportable et très fiable « Ediol », que nous avons par la suite minutieusement transcrit. Tous les enregistrements ont eu lieu le mercredi matin avant ou après le cours de français dans une salle de l'École des Mines d'Albi entre novembre 2009 et juin 2010. Pour les apprenants de Toulouse, nous les avons enregistrés dans le laboratoire de recherche de l'Université Toulouse-Jean-Jaurès, dans une salle insonorisée dédiée à cet effet. A chaque rencontre, nous avons demandé aux sujets testés de regarder attentivement le dispositif proposé et de nous raconter par la suite, oralement et le plus spontanément possible, l'histoire qu'ils venaient de visionner. La prise de notes n'était pas permise et le dispositif leur était retiré après le visionnement, cela dans le but d'éviter qu'ils s'efforcent de nous raconter, dans le cas de la BD, l'histoire vignette par vignette, et afin de produire ainsi le discours le plus libre et spontané possible. Il n'y avait aucune restriction de temps et chaque compte-rendu a duré entre une demi-minute et deux minutes environ pour la BD, et jusqu'à huit minutes pour le film. Le visionnement du film s'est déroulé dans une salle au calme sans la présence du chercheur pour éviter aux apprenants de se sentir contraints ou mal à l'aise. Aucune prise de note n'était autorisée. La restitution à l'oral, en revanche, s'est faite en présence du chercheur.

Le déroulement des activités s'est fait dans la convivialité, les apprenants ont participé sans gêne et sont venus de leur propre gré en sachant aussi qu'il s'agissait de contribuer à une recherche académique qui n'allait avoir aucune incidence sur leur note de fin d'année. Malheureusement, nous n'avons pas pu rémunérer les sujets pour leur participation.

Notre corpus de productions non-natives comprend 19017 mots (9034 mots en temps 1 et 9986 mots en temps 2), et pour les productions natives 3683 mots, BD et film réunis.

#### **5.4 Transcriptions des données et annotations des disfluences**

Les paroles enregistrées ont été transcrites en prenant bien soin de noter toutes les diverses manifestations des hésitations ou disfluences. Nous avons ainsi calculé et relevé systématiquement et de manière très précise les pauses (pauses silencieuses, pauses remplies) ainsi que toutes les disfluences (bruits de gorge, rires et fragments de mot). Nous n'avons pas transcrit certaines informations acoustiques, qui ne nous paraissaient pas pertinentes pour cette recherche (pauses respiratoires, spectre acoustique de la parole). Nous n'avons gardé que

la transcription (énoncés et pauses remplies) ainsi que les pauses notées en millisecondes et *groupes d'hésitation* (désormais GH), comportant au minimum deux types de pauses, une pause silencieuse et une pause remplie).

Les transcriptions ont été faites de façon à faire apparaître les énoncés sous forme de segments de parole [transcriptions des locuteurs non-natifs (Annexe 5) ; transcriptions des locuteurs natifs (Annexe 6)] et la durée des hésitations (pauses vides mais aussi les autres hésitations dépassant 500ms). Nous rappelons ici qu'un segment de parole consiste en une suite de mots encadrés par deux grandes pauses supérieures ou égales à 250ms. Les données transcrites sous Word l'ont été ensuite sous Excel, ce qui nous a permis d'effectuer les calculs de manière aisée (un exemple de fichier Excel se trouve en Annexe 15).

#### **5.4.1 Les annotations**

##### ***5.4.1.1 Les hésitations***

Elles comprennent les pauses remplies, les syllabes allongées ainsi que d'autres bruits comme le raclement de gorge, le claquement de langue, les rires, et les fragments de mots. Rappelons que nous cherchons à identifier les séquences préfabriqués qui ne peuvent en aucun cas être interrompues par une quelconque disflueance selon le critère nécessaire de cohérence prosodique (voir chapitre 2 et ce chapitre, section 5.5.2). Dans cette recherche, les pauses qu'elles soient silencieuses ou remplies sont considérées comme les signes d'un traitement non automatique. C'est pourquoi une SP ne peut être interrompue par une pause, et ce même au niveau débutant.

##### *a) Les pauses silencieuses*

Si l'on considère les études déjà entreprises, le minimum de temps de silence pris en compte pour délimiter le segment de parole varie selon les recherches. Nous rappelons que Grosjean et Deschamps (1972 :133) considèrent une pause non sonore à partir de 0.25 sec « Seules les pauses dépassant 0,25 sec furent comptées, afin de ne pas prendre en considération les coupures dues à l'articulation (ex : les plosives) ». Selon les chercheurs, la durée peut être plus ou moins grande. Riegenbach (1991) propose de mesurer les micro-pauses à partir de 200ms et les pauses d'hésitation à partir de 400ms comme Freed et *al.* (2004). Pour des questions pratiques de reprographie, Towell et *al.* (1996) et Towell (2002) ont déterminé la pause à partir de 280ms, sinon la limite aurait été de 250ms et Wood (2010) choisit une durée de 300ms. Cordier (2013), Hilton (2008b) ont choisi la limite basse de 200ms comme indiqué déjà par Raupach (1984). Cependant, dans cette recherche, nous avons choisi de compter les



pauses silencieuses à partir de 250ms, durée qui semble mieux correspondre à une pause d'hésitation et non à une pause respiratoire.

#### *b) Les pauses sonores*

Les pauses sonores ou remplies, prises en compte dans cette recherche, sont les pauses qui ne contiennent aucune information sémantique (*euuh, hein, mm*) et qui, de ce fait, ne sont pas des mots, alors que des séquences comme *j'sais pas* ou *je pense* sont traitées par certains chercheurs comme des mots ne véhiculant aucune information sémantique (Riggenbach, 1991). Néanmoins, dans cette recherche, nous traiterons ces dernières comme des séquences discursives car elles contiennent des informations sémantiques et ont une fonction de « remplisseurs ». Lorsque l'apprenant n'arrive pas à trouver un mot ou une information lexicale, il s'appuie sur ces séquences qui lui permettent de garder la parole et qui font partie des stratégies de l'oral.

Pour certains chercheurs (Cordier, 2013), les pauses remplies ou sonores de plus de 400ms indiquent un changement de segment. Contrairement à Cordier, nous n'avons pas fait ce choix ; en effet, les sujets ayant des langues maternelles différentes, il n'était pas possible de les prendre en compte, certains apprenants faisant un usage de pauses sonores plus important que d'autres. En revanche, une séquence préfabriquée ne peut être interrompue par une pause remplie.

#### *c) Les groupes d'hésitation (GH)*

Un segment de parole peut être délimité par une grande pause mais aussi par un groupe d'hésitation, composé d'au moins deux types de pauses (pause silencieuse et pause sonore). Nous avons décidé, après avoir longuement hésité, de prendre en compte la durée totale du groupe d'hésitation (pauses vides et remplies). Nous avons pris ce parti, peut-être contestable, de noter la durée de l'ensemble de la pause alors que certains sujets produisent plus de GH que d'autres, ce qui peut être dû à leur LM ou tout simplement à leur « style » oral. Cependant, nous trouvons les mêmes différences dans notre groupe contrôle.

#### **5.4.1.2 Les syllabes allongées**

Nous trouvons dans les productions une grande quantité de syllabes allongées en finale de mot mais aussi dans le mot. Elles ont été signalées par deux points, ou plus selon la durée de l'allongement : [:] entre 0.25 et 0.40s, [::] entre 0.40 et 0.70s, [:::] plus de 0.70s). Seules les syllabes allongées à partir de 0.40s, sont prises en compte pour les mettre en rapport avec la durée de la pause silencieuse choisie (0.25s). D'autres chercheurs (Cordier, 2013) ont choisi

également 0.40s, la durée d'une syllabe en production orale étant autour de 200 à 300 ms, d'après Zellner (1994 : 47) cité par Cordier (2013 : 119). Déterminer la durée de l'allongement est important si l'on considère qu'une séquence est un groupe mémorisé et produit en bloc. Ceci signifie qu'une séquence identifiée comme préfabriquée ne peut contenir de syllabe allongée d'une durée supérieure ou égale à 0.40s. Si l'allongement intervient à l'intérieur d'une SP, nous ne pouvons l'identifier comme une SP. Ce qui n'est pas l'avis d'autres chercheurs, notamment Bardovi-Harlig (2010) qui accepte une hésitation à l'intérieur d'une SP qu'elle considère en cours d'acquisition. Nous avons cependant accepté que la dernière syllabe de la SP soit allongée. A notre avis, après avoir produit la séquence, le locuteur peut s'appuyer sur la syllabe finale, ce qui ne serait pas dû à un manque d'automatisation de la SP mais plutôt considéré comme une stratégie de planification de la suite à venir. En revanche, aucune des unités qui composent la séquence ne doit contenir de syllabe allongée excepté la dernière syllabe du dernier mot. L'allongement est souvent d'ailleurs accompagné d'une pause remplie et/ou une pause vide.

#### **5.4.1.3 Les reprises (ou corrections) : répétitions, reformulations, abandons ou faux départs**

Les disfluences comprennent non seulement les pauses silencieuses, les pauses remplies (*euh*, *hein*, *hum*, *mm*), les allongements, mais aussi toutes les corrections (répétitions, reformulations, abandons). Toutes les corrections ont été notées par une, deux ou trois barres obliques : (répétitions [/], reformulations [//], abandons [///]). Les répétitions n'entraînent pas de changement de segment contrairement au choix fait par Cordier (2013). On pourra donc trouver des répétitions à l'intérieur du segment de parole.

#### **5.4.1.4 Les syllabes finales**

A l'intérieur du segment, nous avons également transcrit les syllabes finales accentuées. Lorsqu'une syllabe finale ouverte se termine par un e muet /un schwa/ et qu'elle est appuyée, elle est suivie de deux points [:], et comptée comme syllabe pleine, *l'abeille*: (3 syllabes) ≠ *l'abeille* (2 syllabes). Si le mot est allongé par un /schwa/ marqué, nous avons transcrit *euh* et l'avons compté comme une pause remplie. En cas d'allongement de la consonne finale comme par exemple *donc* (1 syllabe) suivie d'un /schwa/ appuyé, nous l'avons transcrit *donc euh* et nous l'avons compté comme une seule syllabe suivie d'une pause remplie. Chaque pause remplie d'une durée de 400ms ou plus est noté *eu:h*. Certains apprenants utilisant un grand nombre de *euh*, il n'était pas envisageable de les considérer comme des syllabes

articulées. Ceci étant, le « e » d'appui pour reprendre le terme de Candea (2002) peut être systématique chez certains locuteurs en L1, et notamment chez les locuteurs parisiens.

#### **5.4.1.5 Autres annotations**

Les autres bruits ont été transcrits de la manière suivante et selon les instructions du corpus

Parole (Hilton, 2008b) :

[&=bche] pour tout bruit de bouche (claquement de langue)

[&=ahem] pour raclement de gorge

[&=rire] pour les rires

Les fragments de mots ont été notés comme suit : [&che] pour che-(val) et n'ont pas été comptées comme des répétitions.

Enfin pour repérer et calculer de manière aisée le nombre de ces disfluences, le texte est transcrit sous format Word de manière linéaire sans la durée des pauses, seul (.) indique qu'il y a une pause. Cela permet de rendre la lecture plus facile et de visualiser la place des pauses sans gêne pour le lecteur.

#### **5.4.2 Choix du logiciel**

Un des objectifs de cette recherche est d'examiner le développement des SP dans une perspective longitudinale en relation avec le développement de la fluence. La transcription de toutes les productions orales de nos sujets devait se faire en ayant à l'esprit qu'il nous fallait aussi identifier les segments de parole fluides, c'est-à-dire ininterrompus par une pause silencieuse ou par un groupe d'hésitation (contenant au moins une pause silencieuse et une pause remplie). Nous devons pouvoir également compter le nombre de syllabes des segments de parole, identifier les SP à l'intérieur de ces segments et compter le nombre de syllabes préfabriquées à l'intérieur de chaque segment.

Aussi, nous avons choisi le logiciel PRAAT, créé par Paul Boersma et David Weenink de l'institut de phonétique de l'université d'Amsterdam (2008), pour transcrire des données audio. Toutes ses fonctionnalités n'ont pas été exploitées parce que nous ne travaillons pas en phonologie. Il nous a permis cependant de visualiser le spectre de parole afin de détecter passages vocaux et pauses. Ce logiciel propose d'utiliser différentes rangées ou *tier* pour les annotations. Autant d'annotations différentes autant de *tier* possibles. Nous avons soigneusement transcrit le texte dans une première rangée (*tier*) et dans un deuxième noté les pauses en millisecondes.

### 5.4.3 Mesures et analyses

Dans un premier temps, après avoir enregistré le travail de transcription, fait sur les *tier*, en fichier *textgrid* (Annexe 14), nous l'avons transformé en fichier « Word », en éliminant toutes les informations qui ne nous étaient pas directement utiles (durée de chaque segment de parole, par exemple). Puis nous avons rapatrié les données dans un fichier Excel pour faciliter les calculs. Toutes ces manipulations ont été faites manuellement.

Toutes les transcriptions ont été rapportées sous Word sous forme de tableau mettant en regard les deux tâches, BD (t1) et (t2) Film (t1) et (t2) et dans un format Excel. Nous avons utilisé le format Excel pour transcrire les segments de parole (colonne B), la durée des pauses notée en ms (colonne C), le nombre de mots de chaque segment (colonne D), le nombre de syllabes du segment (colonne E) le nombre de SP (colonne F) et le nombre de syllabes contenues dans les SP (colonne G) avec le total sous chaque colonne. Nous présentons (Annexe 15) un exemple de fichier Excel avec les données brutes d'un apprenant (FA) qui permet de comprendre comment le travail a été fait. Un tableau complémentaire dans la même Annexe 15 indique les résultats obtenus à partir des données brutes.

## 5.5 Identification des SP

Comme nous l'avons vu au chapitre 2, une des plus grandes difficultés pour le chercheur est l'identification des SP (Myles, 2007 : 168 ; mais aussi Schmitt, 2010 ; Wray, 2002, 2008). Comme le fait remarquer Wray (2009 : 28), ce n'est pas une tâche aisée<sup>64</sup> car il faut être sûr d'avoir bien identifié toutes les séquences et seulement les bonnes. Le chercheur se trouve confronté à deux dangers : ne pas identifier toutes les séquences ou en identifier trop.

### 5.5.1 Première étape

Comme aucune méthode d'identification existante ne nous satisfaisait, nous avons choisi d'appliquer non pas une seule méthode mais de croiser plusieurs méthodes d'identification comme Wray le recommande (2008 : 110). Nous avons relevé ce qui nous paraissait préfabriqué selon des critères applicables à un corpus de productions L2. S'il est difficile d'identifier clairement les SP chez les natifs car le langage préfabriqué se mêle au langage créatif, il est plus facile, en revanche, de les reconnaître chez les locuteurs non-natifs grâce à leurs erreurs (Wray, 2002 : 39) comme du reste chez les enfants et les sujets pathologiques.

---

<sup>64</sup> « researching formulaic language has many challenges but probably the single most persistent and unsettling one is knowing whether or not you have identified all and only the right material in your analyses. » (Wray, 2009 : 28)

Nous nous sommes trouvée confrontée à quelques doutes, dus certainement à notre première analyse du corpus, trop formelle. De ce fait, nous avons appliqué plusieurs méthodes pour certifier notre relevé des SP. Mais notre démarche d'authentification s'est peu à peu transformée au cours de ces années de recherche.

Tout d'abord, nous avons essayé de vérifier si les séquences identifiées pouvaient être authentifiées 1) par des juges natifs et 2) par leur présence dans des corpus oraux ou des dictionnaires.

#### 1. les juges natifs

Un grand nombre d'études ont eu recours à l'expertise de juges natifs (Foster, 2001 ; mais aussi Forsberg, 2006 ; Forsberg-Lundell *et al.* 2014). Mais, dans toutes ces études, les SP ont été identifiées selon leur plus ou moins grand degré de figement. Nous rappelons brièvement la méthode d'identification qu'a choisie Forsberg (2006). Elle utilise d'abord le moteur de recherche Google pour l'identification des séquences. Une première recherche est effectuée pour identifier une SP. Par exemple, elle trouve 5240 occurrences du syntagme verbal *déposer un dossier*. Ensuite elle lance la recherche pour la même séquence mais modifiée selon le principe de la commutation restreinte. Pour vérifier la nature préfabriquée d'un syntagme, Forsberg (2006 : 52) applique aussi d'autres techniques :

- « a) Substitution d'un mot de la SP contre un synonyme
- b) Changement d'article (du défini à l'indéfini ou suppression d'article)
- c) Changement du nombre (du singulier au pluriel ou vice-versa)
- d) Changement d'ordre des mots (par exemple *égalité femmes / hommes* au lieu de *égalité hommes / femmes*)
- e) Changement contre un antonyme (par ex. *ça marche mal* au lieu de *ça marche bien*) »

La même recherche est effectuée pour une séquence synonyme, *laisser un dossier*. Le nombre très faible d'occurrences : 25 (vs 5240) obtenu sur Google certifie que le syntagme *déposer un dossier* est un bien une SP préférée des natifs et donc conventionnelle. Mais afin d'être plus sûre du statut préfabriqué des séquences, Forsberg a également demandé à « deux locuteurs natifs de choisir la séquence qui leur paraissait la plus naturelle et la plus courante » et qui ont confirmé le résultat des recherches sur Google.

Nous avons été dans un premier temps intéressée par la méthode de ce chercheur, mais très vite nous avons préféré choisir plusieurs modes d'identification Nous avons donc demandé à deux juges natifs, une étudiante en linguistique déjà initiée à la phraséologie et une didacticienne du français langue étrangère de regarder notre corpus de productions non-natives et natives. Leur contribution à cette recherche s'est faite spontanément, toutes deux

étant intéressées par le sujet. Les deux juges ont eu accès aux transcriptions complètes (texte et pauses) et avaient pour consigne de relever tout ce qui leur paraissait produit en bloc, sans que les séquences relèvent obligatoirement du figement. Dans le travail que les juges natifs ont accompli sur 20% du corpus natif et non-natif, les séquences ont été relevées selon des critères plus linguistiques que psycholinguistiques, et cela malgré nos recommandations. Si nous avons tenu compte des remarques des juges, nous avons vite compris que le jugement des natifs n'était pas d'un grand intérêt pour cette recherche qui considère la préfabrication selon une perspective plus psycholinguistique que linguistique. L'identification des SP par les juges natifs nous a permis néanmoins de confirmer ou infirmer nos choix de départ, et nous l'avons utilisée plutôt comme un garde-fou. Cette identification a permis et légitimé l'élimination d'expressions (*dans la maison, dans la cuisine*) pressenties comme des préfabriqués (avec toutefois une certaine incertitude) selon les critères de cohérence phonologique et de fréquence mais que les deux juges n'ont pas confirmé.

Dans cette étude, plusieurs modes d'identification ont été croisées, d'abord les critères d'identification des chercheurs (voir chapitre 2 : Hickey, 1993, Wray, 2008), le jugement des natifs (1), et le recours aux dictionnaires et corpus en ligne (2).

2. Voici les outils que nous avons consultés :

- Corpus en ligne : « *lexutor* », (un corpus de langage oral de 700 000 mots), qui nous a permis de trouver en quantité des formes récurrentes de certaines séquences comme la séquence *et après* employée 101 fois dans notre corpus L2 et les formes avec l'adjectif *petit*.

- Dictionnaires en ligne, le *Trésor de la langue française* (ATILF et CNRTL) ainsi que le dictionnaire *Médiadico* qui nous a permis de vérifier que des séquences du type *indiquer le chemin/la sortie*, étaient bien attestées comme des expressions. Les juges natifs ne les considéraient pas comme des SP à cause de leur transparence. L'identification des juges était basée sur des critères formels essentiellement. Nous avons trouvé dans le dictionnaire *Mediadico* des expressions construites avec *sortie* et *chemin*, que nous considérons comme des collocations verbales non contraintes. Notre « intuition » nous poussait à accepter des séquences comme « *indiquer le chemin/la sortie* » *prendre/chercher la sortie*, qui n'étaient pas reconnues comme des collocations par un des juges natifs (pour qui une collocation relevait forcément du figement). Lors de notre recherche sur le site *Médiadico*, ces expressions étaient indiquées dans la rubrique expressions avec « *sortie* ». Nous avons donc validé notre choix d'autant plus aisément que ces expressions répondaient aux deux critères obligatoires (chapitre2).

Nous avons aussi identifié certaines séquences automatisées à force de répétition et d'input, par exemple les séquences apprises en cours de langue du type *c'est, il y a, il/elle est + adjectif, petit + N* ou des séquences verbales. Ainsi, les formes réflexives (pronominales) des verbes à la troisième personne relèvent, à notre avis, de la préfabrication. Ces formes verbales à la troisième personne (*se + V*) comportent deux unités, un pronom réfléchi et un verbe, et doivent être automatisées pour être produites correctement. Pour appuyer notre propos, nous avons trouvé une forme incorrecte dans notre corpus, *on \*t'inquiète pas pour cette mouche*. Cette forme erronée montre bien que la forme verbale du pronominal à la troisième personne doit passer par la mémorisation et l'automatisation pour être produite correctement. Ici l'apprenante qui appartient au groupe des plus avancés a sans doute mixé les deux formes du verbe : l'injonction « t'inquiète pas » et la troisième personne du verbe avec « on ». Une séquence est préfabriquée lorsqu'elle a été apprise en bloc ou automatisée, c'est-à-dire lorsque l'expression vient d'elle-même sans qu'on y pense. Si l'on s'en tient strictement à cette définition, on pourrait aussi inclure dans les SP les formes verbales composées d'un auxiliaire et d'un participe passé. Nous avons cependant exclu les formes composées du verbe et cela malgré une utilisation importante du passé composé dans les productions de nos sujets. Les apprenants passent le plus souvent par une étape d'erreur dans l'utilisation des auxiliaires de conjugaison en français (souvent selon leur LM) et ce n'est que dans une ultime étape de l'acquisition que les formes composées sont produites de manière automatique par le locuteur. Comme Wood (2006), nous les avons éliminées de notre relevé considérant que cela fait partie davantage de l'acquisition de la morphologie des verbes que de la préfabrication. Nous acceptons comme séquences les verbes pronominaux (pronom réfléchi et verbe), les verbes à construction prépositionnelle du type « obéir à » et les verbes à double construction « dire N à N », (dire, donner qch à qqn) avec un *open-slot*<sup>65</sup> (Erman et Warren, 2000 : 35). Ces verbes à régime prépositionnel sont d'ailleurs souvent suivis d'une hésitation dans notre corpus, qui peut prendre la forme d'une pause vide ou sonore ou la forme d'une répétition, ce qui démontre bien qu'elles sont produites en bloc [*il commence à / à courir*].

Les formes verbales suivies d'un groupe prépositionnel tels que « *sortir de* (la maison) » qualifiées de colligation par certains (Legallois, 2012 ; Firth, 1957) ont été relevées et comptées car nous considérons qu'elles sont produites comme un tout. En effet, il arrive souvent que la préposition soit suivie d'une hésitation comme si l'apprenant sortait le bloc avant de pouvoir planifier la suite. Dans ce cas, le critère phonologique a facilité

---

<sup>65</sup> « [...], open slots are represented in the dictionary by sth or sb (something or somebody), as in sth depends on sb. » (Erman et Warren, 2000 : 35)



l'identification, ainsi que l'impossibilité de changer la préposition dans le groupe « *sortir de la maison* » \**sortir à* vs *sortir de*, où *sortir à* (*sortir à Paris*) n'a pas le même sens.

Les séquences relevées peuvent être classées selon différentes catégories ou fonctions. Elles peuvent être discursives, lexicales ou grammaticales, avoir une fonction référentielle, métalinguistique (marqueurs de discours, marqueurs organisationnels) ou constructionnelle comme [**de** (fleur) **en** (fleur)] qui indique le mouvement. Elles peuvent être des groupes verbaux (V+ préposition), des mots qui relient (*parce que*, *pendant que*), des groupes qui introduisent une proposition, (*je pense que*, *qu'est-ce que*).

Parmi les autres séquences produites de manière récurrente par les apprenants, nous avons relevé pas moins de 145 occurrences de *petit* + N sous forme de syntagmes dont voici les principaux : *petit garçon*, *petit enfant*, *petit cheval mais aussi petit banc*, *petit village*, *petite abeille*, *petit insecte*, *petits grains de café*, etc.

On trouve aussi un ensemble de syntagmes construits avec des adjectifs utilisés fréquemment en début d'apprentissage, ce que Peteers (2012) appelle les formes « primitives ». On les trouve dans tous les lexiques des langues sous forme de couple d'antonymes comme *petit/grand* et *vrai/faux* très présents dans notre corpus : *un **grand** cheval*, *le **vrai** cheval*, *le **faux** cheval*, *des **vraies** fleurs*, *des **fausses** fleurs*... L'emploi récurrent de ces adjectifs chez certains apprenants en début d'apprentissage mériterait que l'on s'y attarde. Il serait par ailleurs intéressant d'approfondir l'étude de ces formes primitives que l'on trouve en grande quantité dans les productions non-natives au niveau débutant. Cependant, nous avons décidé de compter comme des SP seulement les syntagmes composés de l'adjectif *petit* +N. En effet, les productions natives comportent également de nombreuses occurrences de l'adjectif *petit*. Comme rien n'indique que *petit* est utilisé dans son acception de « diminutif », nous avons décidé de compter toutes les occurrences, sans distinguer entre sens diminutif ou autres. En français, l'adjectif *petit* serait en cours de figement dans certains cas comme *un p'tit ami*, *un p'tit café*, *une petite annonce*. Nous développerons ce point dans un chapitre consacré aux formes sur-employées dans les productions narratives (chapitre 9).

Parfois les sujets utilisent des expressions typiques qu'ils répètent tout au long de leur narration mais qui ne sont pas toujours considérées comme des SP. Nous avons accepté comme SP des syntagmes verbaux composés de la forme très simple du verbe *être* au présent à la troisième personne suivi d'un adjectif « *il est* » + adjectif. On peut en effet s'interroger sur la forme *il est* + adjectif, appris dès les premiers stades de l'apprentissage (mais aussi *il est* + N métier : *il est professeur*, *il est médecin*). Ce modèle, mémorisé à force d'input, permet à l'apprenant de qualifier l'objet (ici le personnage de la BD) dès les premiers stades : *il est*



*fatigué, il est surpris, il est mignon*. La forme verbale très simple, *il est*, facilite la production, il suffit d'ajouter un adjectif pour former un syntagme complet. On peut donc la considérer comme un cadre collocationnel (Bolly, 2010). De même, la forme *il a* + âge, enseignée dès les premiers stades de l'apprentissage, permet la production de syntagmes complets. Mais nous n'avons aucune SP de ce type dans notre corpus. D'autres formes apprises et mémorisées semblent jouer un rôle dans la production de syntagmes, nous pensons notamment à la forme du futur périphrastique « *il va* + V infinitif » produit très souvent par les locuteurs. Nous avons compté *il va* comme préfabriqué seulement lorsqu'il était suivi du verbe à l'infinitif. Il peut être également suivi d'un groupe prépositionnel indiquant un lieu : *il va* + lieu (*il va à la maison*).

### 5.5.2 Nos critères d'identification

Dans notre partie théorique, nous avons vu que le langage préfabriqué embrasse de nombreux faits de langue et que plusieurs critères peuvent être pris en compte dans sa définition. Les critères retenus sont à mettre en relation avec les modèles de Hickey (1993) et Peters (1983) adaptés à la L2 et aux critères propres à la L2 tels que Forsberg (2006) les a définis :

- a) polylexicalité
- b) cohérence prosodique
- c) combinatoire conventionnelle
- d) irrégularité syntaxique et/ou opacité sémantique
- e) fréquence
- f) complexité syntaxique
- g) déviation interlangagière

Une suite de deux mots ou plus sera identifiée comme préfabriquée si elle est dite d'une traite sans pauses internes et si elle répond à au moins un des critères suivants (c, d, e, f, g).

Voici nos critères d'identification

#### a) *Polylexicalité*

Dans cette recherche, une SP devait comporter au moins deux items (Erman et Warren, 2000, Forsberg, 2006 ; Cordier, 2013 ; Wood, 2006, 2010). Autrement dit, nous ne prendrons pas en considération les préfabriqués d'autrefois qui, aujourd'hui, après fusion, sont des unités monolexicales comme *bonjour*. La polylexicalité est un critère obligatoire et concerne même deux items séparés par un tiret (*peut-être*) ou une apostrophe (*c'est*).

#### b) *Cohérence prosodique*

La SP devait être prononcée d'une traite sans pauses internes ni hésitations avec une courbe mélodique continue (Wray, 2002 : 42-43). La cohérence prosodique est ainsi considérée dans cette étude comme la preuve que la séquence en question est mémorisée et récupérée en bloc sans passer par le système grammatical de la langue. Une réduction comme une fusion phonologique, réduction de syllabes ou omission d'un schwa peut être présente dans la SP. Ces traits sont présents dans les productions de bouts de langue très communs comme par exemple *chepa* pour *je ne sais pas*. Inversement, les productions des bouts langue, moins courants ou créatifs, ne comportent pas ce trait. Ce critère est obligatoire.

*c) Combinatoire conventionnelle*

Nous avons vu qu'une SP est une suite qui semble relever de la mémorisation ou de la récupération holistique à partir du lexique mental plutôt que de la génération ou création suite à l'application d'une règle grammaticale. Dans ce sens, une SP peut être une collocation, un syntagme attesté plus ou moins figé, une phrase lexicalisée ayant une fonction pragmatique (ça va ?). Ce critère est typique ou gradué.

*d) Irrégularité syntaxique et/ou sémantique*

Une SP peut être une expression idiomatique ou une métaphore et comme telle, elle est apprise de façon holistique, puisque ni la génération ni la traduction mot à mot d'une langue à l'autre n'est possible (*avoir l'air, être en train de, avoir faim*). Ce critère est typique ou gradué

*e) Fréquence*

Les séquences produites avec une certaine fréquence dans le corpus sont vraisemblablement dues au contexte d'énonciation et aux dispositifs de narration utilisés (*et après*). Ce critère est typique ou gradué.

Le critère de fréquence doit être appliqué avec une grande prudence. En effet, dans le corpus non-natif on trouve de nombreuses suites de mots que les locuteurs emploient à profusion. Par exemple, les apprenants produisent le groupe nominal « *dans la maison* » de nombreuses fois, 28 fois en t1 et 26 en t2, ainsi que *dans la cuisine* et *dans le salon*. Cependant même si les deux critères obligatoires sont applicables ici (polylexicalité et cohérence phonologique) et que le critère facultatif de fréquence pourrait être justifié, nous n'avons pas considéré cette suite comme un préfabriqué. Ce choix a été fait après une longue réflexion. Si l'on se place du point de vue de la grammaire des constructions, on pourrait identifier *dans la maison* comme une construction formée d'une préposition indiquant un lieu *dans* suivi d'un item interchangeable, ici *maison*. Du point de vue des grammaires de constructions elle serait

considérée comme une construction véhiculant un sens locatif. Cela étant, nous avons décidé que ces constructions relevaient plus de la syntaxe.

#### *f) Complexité syntaxique*

Nous incluons ici des SP qui n'auraient pas pu être générées par l'apprenant via l'application d'une règle de grammaire étant donné leur niveau de compétence en L2.

En début d'apprentissage, si une séquence est plus complexe syntaxiquement que le reste des productions d'un apprenant et utilisée de manière inappropriée, son statut préfabriqué est très probable. Les apprenants produisent souvent des expressions et des phrases entières sans les adapter car ils les ont mémorisées telles quelles en bloc sans passer par l'analyse trop complexe, semble-t-il, à leur niveau. Par exemple, on trouve dans notre corpus de nombreuses phrases interrogatives directes comme « *qu'est-ce qui s'est passé* » ou « *qu'est-ce qui se passe* » produites telles quelles sans aucune modification alors qu'elles sont introduites par une proposition principale. L'apprenant les a mémorisées en bloc et ne peut les transformer, à son niveau, en interrogatives indirectes. Ce critère est typique ou gradué.

#### *g) Déviation interlangagière*

L'erreur n'est jamais sans importance en acquisition. Rien n'empêche par exemple qu'une séquence mal perçue de la part de l'apprenant puisse être mémorisée et récupérée de façon holistique par ce dernier (*comment dire* perçu comme *\*comment dit* ; *qu'est-ce qui se passe* perçu comme *\*qu'est-ce qu'il passe*). Dans une perspective psycholinguistique et acquisitionniste comme la nôtre, le langage préfabriqué pourrait inclure des séquences qui traditionnellement échappent à la phraséologie en linguistique (cf. aussi Cordier, 2013; Wood, 2010; Forsberg, 2008; Wray, 2002). Ce critère est typique ou gradué.

Il est important d'attirer ici l'attention sur le fait que, nos critères nécessaires ou obligatoires sont les critères de polylexicalité et le critère phonologique. Les autres critères sont facultatifs et si aucun d'eux n'a été considéré comme plus important qu'un autre, l'un de ces critères au moins doit être présent pour confirmer le statut préfabriqué de la séquence.

## **5.6 Classement des séquences préfabriquées**

Une fois les SP identifiées, il est nécessaire de les classer. Là encore, la tâche est ardue pour le chercheur. Quelle typologie adopter et selon quels critères ? Dans le chapitre 1, nous avons distingué plusieurs catégories de SP. Une SP peut être une collocation (*beurre rance*), un syntagme attesté ou plus ou moins figé (*de temps en temps*), une phrase lexicale ayant une fonction pragmatique (*ça va ? qu'est ce qui se passe ?*) (Nattinger et DeCarrico, 1992 : 36), un patron qui admet des variations *<plus...plus>*, *<non seulement...mais encore>*, *<de...*

en...>. Une SP peut être une expression idiomatique conventionnelle ou pas, ou une métaphore apprise et récupérée de façon holistique puisque ni la génération ni la traduction mot à mot d'une langue à l'autre n'est possible (*avoir l'air, être en train de, avoir sommeil, tant pis*). Une séquence peut être une phrase ou une expression récurrente que le locuteur a mémorisée. Elle semble donc relever de la mémorisation et donc de la récupération holistique à partir du lexique mental plutôt que de la génération ou création ou fabrication suite à l'application d'une règle grammaticale. Alors comment classer ces séquences qui correspondent finalement à des situations d'emploi diverses et qui ont des fonctions différentes dans le discours ?

En grammaire, le classement des mots ou unités simples se fait normalement et facilement en classes grammaticales. Il n'en est pas de même lorsqu'on parle de SP. Les séquences peuvent être a) des collocations construites autour d'un verbe (*faire le pont*), d'un nom (*moteur à essence*), b) des phrases (*les bons comptes font les bons amis*), c) des phrases toutes faites non verbales (*à bientôt, sous peu*). Les expressions peuvent être plus ou moins opaques et plus ou moins figées. On pourrait donc les classer en fonction de leur degré d'opacité sémantique. D'une part, les expressions sémantiquement transparentes et, d'autre part, les expressions sémantiquement opaques ; ou en fonction de leur degré de figement, les expressions soient totalement contraintes, soit semi-contraintes, soit libres.

Il serait possible aussi de les distinguer selon leur catégorie, c'est-à-dire comme nous l'avons fait dans notre partie théorique, en collocations, expressions conventionnelles et discursives, constructions, phrases idiomatiques, etc. Lorsqu'on étudie le langage préfabriqué en acquisition L2 et qu'on a des sujets de niveau débutant, on ne doit pas s'attendre à trouver des phrases idiomatiques ou figées comme, par exemple *casser sa pipe* ou *une hirondelle ne fait pas le printemps*. Il serait donc improbable de classer les séquences selon les catégories de la phraséologie (chapitre1), car en L2 on trouve très peu de séquences de ce type. Si on devait prendre en compte tous ces paramètres, il ne fait pas de doute que le classement serait complexe. S'il fallait classer les séquences polylexicales, il faudrait, en effet, avoir recours à des analyses linguistiques pointues (Pecman, 2004). Cependant nous avons précisé que notre recherche se situe en-dehors du cadre phraséologique traditionnel. Nous nous situons plutôt du côté des psycholinguistiques pour qui une séquence préfabriquée est une séquence qui répond à certains critères mais dont le rôle est de faciliter le traitement langagier.

Il ne faut donc pas négliger l'importance des recherches en psycholinguistique qui mettent l'accent sur le rôle des séquences pour un locuteur. L'on peut trouver chez un seul locuteur un

type de séquence qu'il pourra ou pas sur-employer et qui ne répondra pas toujours, et loin s'en faut, aux seuls critères de figement ou d'idiomaticité.

Doit-on alors classer les séquences selon leur forme (approche formelle) ou selon leur fonction (approche fonctionnelle) dans le discours ? Il n'est pas facile de choisir entre l'une ou l'autre approche car dans chaque cas on risque de trouver des SP qu'on ne peut ranger dans telle ou telle catégorie.

Nous avons donc identifié les SP selon une approche mixte, influencée par les recherches antérieures en L2 (Erman et Warren, 2000 ; Forsberg, 2006) et une approche plus récente (Cordier, 2013). Nous avons relevé dans le corpus les SP selon un certain nombre de critères tout en ayant à l'esprit le rôle qu'elles jouent dans le langage (facilitation du traitement langagier et meilleure fluidité). Aussi, pour notre corpus de productions orales L2, notre typologie sera à la fois formelle et fonctionnelle. Nous avons classé les séquences préfabriquées en trois grandes catégories (Annexes 3 et 4).

A. les *séquences lexicales*, SP lexicales (Forsberg, 2006) ou SP « référentielles » (Cordier, 2013) qui servent à nommer (objets, actions), à identifier les objets du monde, à exprimer le temps, l'espace ;

B. les *séquences discursives* qui se répartissent en deux sous-catégories.

i. les séquences qui servent a) à construire le discours, les organisateurs (*au premier plan, au début, à la fin*), b) à relier des phrases (*parce que*) ou c) des parties du discours (*et après*),

ii. les séquences qu'on appelle « remplisseurs », utilisées comme stratégie pour gagner du temps (*j'sais pas*) et d'appui (*islands of reliability*, Dechert *et al.*, 1984) pour avancer dans la production.

C. les *séquences grammaticales* qui comprennent des constructions de phrase (*sentence-builders*, au sens de Nattinger et DeCarrico, 1992), composées d'une partie fixe et d'une partie libre (*je crois que* + proposition, ou un groupe du type, *à la fin de* + GN, mais aussi des outils qui servent à relier les mots entre eux (*à cause de*).

Aussi, pour classer les séquences préfabriquées identifiées, nous avons adapté les typologies de Forsberg (2006, 2008) et de Cordier (2013). Forsberg s'est largement inspirée de la classification d'Erman et Warren (2000) en l'adaptant au français. Cordier ne suit pas exactement la même typologie même si l'on peut rapprocher la catégorie des SP référentielles à celle des SP lexicales de Forsberg, la catégorie des SP métadiscursives à celle des discursives. En revanche, la catégorie des SP grammaticales, beaucoup moins étendue chez Forsberg (2006), comprend la catégorie des constructions de phrases « *sentence-builders* »

chez Cordier (2013). Malgré quelques sous-catégories en commun, elles présentent des différences. L'ensemble des *sentence-builders*, terme emprunté à Nattinger et DeCarrico (1992) ou « cadres collocationnels » selon Bolly (2010), comprend toutes les constructions de phrases composées d'une partie fixe suivie d'une partie variable. Cordier se place, nous semble-t-il, dans une perspective plus constructionnelle. Nous pensons aussi que la catégorie grammaticale telle qu'elle apparaît chez Forsberg est ambiguë car elle comprend exclusivement les mots-outils reliant des mots entre eux, des adverbes de quantité ou des marqueurs d'aspect. Leur identification reste strictement grammaticale. Si nous avons choisi la terminologie de Forsberg, pour des raisons strictement pratiques, nous avons en revanche adopté en partie le classement de Cordier que nous avons adapté à notre corpus. Nous avons relevé aussi un certain nombre de séquences qui pourraient correspondre aux « paquets lexicaux » (*lexical bundles*) dont parle Biber et collègues (2004). Ces « paquets lexicaux » seront classés selon notre typologie en séquences lexicales *il comprend pas pourquoi, ça s'appelle*, en séquences grammaticales “*venir de + verbe*” (*je viens de voir*), *je pense que, je crois que, il peut pas*, et en séquences discursives *je sais pas quoi, ça y est*.

Ce classement permet de ne pas entrer dans les distinctions purement descriptives de figement et d'idiomaticité, de collocations lexicales ou grammaticales - le point de vue adopté étant moins linguistique que psycholinguistique.

Nous présentons notre typologie qui comprend les trois grandes catégories de SP, lexicales, grammaticales et discursives annoncées.

### 5.6.1 Les SP lexicales

Conformément à Erman et Warren (2000) et Forsberg (2006), les séquences lexicales (SP lex) sont des séquences qui ont un contenu référentiel. Elles renvoient à une idée ou quelque chose que l'on a identifiée et que l'on peut se représenter. Elles ont un contenu extralinguistique. Nous avons classé les différentes SP lexicales d'un point de vue syntagmatique. On peut reconnaître des collocations lexicales classées en différents types de syntagmes. Nous donnons dans le tableau 3 suivant quelques exemples extraits de notre corpus.

Syntagmes nominaux	<i>journée de voyage, maison de campagne</i>
Syntagmes adjectivaux	<i>étonné de, divisé en</i>
Syntagmes verbaux	<i>faire la sieste, commencer à, se réveiller</i>
Syntagmes adverbiaux	<i>en bois, à ce moment-là</i>
Syntagmes prépositionnels	<i>à la maison, à l'intérieur</i>
Syntagmes propositionnels	<i>c'est drôle, c'est la surprise, qu'est-ce qui</i>

Tableau 3. Séquences lexicales

### 5.6.2 Les SP discursives

Les SP discursives (SP disc) n'ont pas de contenu propositionnel ni extralinguistique. En revanche, elles permettent d'organiser le discours et sont donc liées aux mots de contenu contrairement aux SP grammaticales (SP gram) que l'on trouve à l'intérieur des constructions syntaxiques et qui sont dépendantes de la structure dont elles participent. C'est ce que disent en substance Erman et Warren (2000 : 43) qui distinguent la préfabrication grammaticale de la préfabrication pragmatique, les séquences de type pragmatique apparaissant en-dehors de la structure grammaticale.<sup>66</sup>

En ce qui concerne les deux catégories proposées par ces auteurs, les remplisseurs et modificateurs et les organisateurs, la distinction n'est pas toujours facile à établir dans certains cas ; et il se peut que certaines formules puissent être classées dans les deux catégories selon le contexte où elles apparaissent. Les SP pragmatiques ou discursives organisent le discours : on trouve dans cette catégorie des connecteurs qui « sont des items qui servent à lier les propositions entre elles mais qui de ce fait ne font pas corps avec elles » (Riegel *et al.* 2004 : 617). On peut d'ailleurs à l'écrit les remplacer par la ponctuation (*parce que* par deux points [:] par exemple).

Des catégories discursives de la typologie de Forsberg (2006), nous avons gardé les marqueurs textuels et les marqueurs autorégulateurs, les marqueurs interactifs étant sans pertinence pour nous étant donné qu'il n'y avait pas de tâche d'interaction. La séquence *c'est*, qui apparaît très fréquemment dans notre corpus, peut être classée dans les trois catégories, tout dépend du contexte d'emploi. *C'est*, séquence discursive est une séquence produite dans une intention discursive, soit répétée (*c'est c'est*) afin de trouver la bonne séquence, *c'est c'est drôle la conclusion*, soit comme appui pour permettre la planification. Nous développerons une section sur la séquence *c'est* dans le chapitre 9 consacré aux séquences sur-employées de notre corpus.

Les marqueurs textuels permettent de structurer l'énoncé, de marquer des liens entre les idées et de lier les énoncés entre eux. Les marqueurs textuels correspondent à la fonction d'organisation chez Raupach (1984). Fernandez (1994 : 31) les définit comme des particules qui « relie entre elles des parties du texte » par exemple des connecteurs de conjonction, des connecteurs adverbes.

<sup>66</sup> Notre traduction de « Pragmatic prefabs [...] differ from grammatical prefabs in that they may occur outside the syntactic structure. » (Erman et Warren, 2000 : 43).



Les discursives « autorégulateurs » (*je pense, j'sais pas, etc.*) n'ont pas de fonction dans la structuration du discours mais dans la planification. Elles permettent au locuteur de s'accorder une pause lexicale pour planifier ou pallier un manque lexical. Elles peuvent aussi être des îlots de sécurité (*islands of reliability*, Dechert, 1984) qui permettent de soutenir une certaine fluidité.

Ces séquences utilisées souvent à l'oral peuvent parfois devenir très présentes dans le discours d'un apprenant qui les utilisera de façon récurrente et idiosyncrasique, selon son niveau en L2. Elles servent à planifier le discours à la manière d'une pause remplie (*eah, hein, hum*) et/ou à remplir un espace qui serait un vide autrement.

### 5.6.3 Les SP grammaticales et les constructions

Elles n'ont pas toutes un contenu extralinguistique à la différence des SP lexicales. Ce sont des mots fonctionnels qui *a priori* n'ont pas de sens de contenu, mais qui peuvent avoir un sens indéfini. Elles servent à organiser les mots dans le groupe syntagmatique. Dans cette catégorie sont classés les pronoms indéfinis (*n'importe où, quelque chose, les uns les autres*), des expressions de quantité ou de degré (*un □ tout petit □ peu, pas du tout, pas vraiment, très très*) des marqueurs d'aspect (*vient de, être en train de, va + Vinf*). Contrairement aux SP discursives qui sont liées à des mots de contenu, les SP grammaticales apparaissent à l'intérieur des constructions syntaxiques et sont dépendantes de la structure à laquelle elles participent. C'est ainsi que Erman et Warren (2000 : 43) distinguent la préfabrication grammaticale de la préfabrication discursive ou pragmatique.

Nous avons aussi classé des séquences qui peuvent, au regard de la grammaire constructionnelle, être considérées comme des constructions telles que « *de fleur en fleur* » où la construction *de...en* marque le déplacement d'un lieu à un autre, on peut remplacer *fleur* par d'autres unités (pays, ville, village, maison, etc.). Cependant, cette construction n'apparaît qu'une seule fois et dans le corpus natif.

Les formes du futur périphrastique « *aller* au présent + Infinitif » sont des SP grammaticales. Le verbe *aller* (se déplacer), qui a perdu son sens de mouvement et s'est grammaticalisé (Bolly, 2010), est considéré comme une construction véhiculant un sens, ici le futur proche. De nombreuses occurrences de cette construction à la troisième personne (il) *va + V* sont présentes dans le corpus. Bolly (2010) définit la grammaticalisation comme un phénomène diachronique où certaines unités lexicales vont, dans le temps, acquérir une fonction grammaticale. Par exemple, le mot latin « homo », *être humain*, s'est grammaticalisé en



pronom impersonnel « *on* » et certaines unités, comme le verbe de mouvement « aller », vont acquérir une fonction davantage grammaticale. Il forme ainsi un « cadre collocationnel » avec le verbe infinitif qui le suit et nous le considérerons donc comme une SP.

Les séquences du type « *la fin de* », « *le début de* » font partie de cette catégorie des constructions qui structurent la narration. *C'est* et *il y a* suivis d'un nom ont été classés dans cette catégorie des constructions à l'instar de Cordier (2013). Les structures qui introduisent des propositions, *je pense que*, et souvent adaptées au récit à la troisième personne, *il pense que*, *il sait que*, *il trouve que*, ont été classées également dans la catégorie des constructions.

#### 5.6.4 Les SP déviantes

Enfin, nous avons relevé à l'intérieur de ces trois catégories les « séquences déviantes » (Annexe 8), c'est-à-dire comportant divers types d'erreurs. Ces séquences faisant partie de l'interlangue des apprenants (Selinker, 1972), seront considérées comme des SP, et ce malgré leur forme déviante. Un même locuteur peut les produire une ou plusieurs fois dans sa production. D'après Forsberg (2006 : 60-61), on trouve deux types de SP déviantes ou interlangagières. Ce sont soit des combinaisons natives « utilisées de manière non native (*la Monique j'adore le tennis* pour *Monique adore le tennis*) soit elles ont une combinaison non-native et sont répétées dans cette même forme à plusieurs reprises. » Les séquences déviantes peuvent aussi provenir d'une erreur morphologique, lexicale, ou phonologique. Elles auraient été mémorisées sous une forme erronée.

Toutes les séquences idiosyncrasiques ou déviantes ont été relevées. Nous avons supposé que les productions erronées qui sont conformes à un ou plusieurs de nos critères étaient pertinentes pour notre recherche. Par exemple, rien n'empêche qu'une SP mal perçue de la part de l'apprenant puisse être mémorisée et récupérée de façon holistique par ce dernier (*comment dire* perçu comme *\*comment dit* ; *qu'est-ce qu'il se passe* perçu comme *\*qu'est-ce qu'il passe*). Dans une étude très récente, Forsberg-Lundell et collègues (2014) utilisent le critère de « déviations morphosyntaxiques » qui, associé à d'autres critères comme la longueur du segment de parole et l'utilisation de séquences préfabriquées, permet d'évaluer un type de locuteurs très avancés pouvant passer « pour des natifs » (*near-native speakers*).

Les SP déviantes ont été classées dans chaque catégorie (lexicale, discursive ou grammaticale) et n'ont pas fait l'objet d'une catégorie à part. Bien évidemment, une SP erronée peut aussi être due au stress causé par la situation du compte-rendu qui reste malgré tout artificiel, où l'apprenant devait parler devant un enregistreur tout en s'efforçant de

restituer l'histoire du mieux qu'il pouvait. Comme pour les autres SP, le facteur de la fréquence n'a pas été pris en compte dans le but de qualifier telle ou telle séquence comme déviante. Dans cette recherche même un *hapax* a été compté. Pour cela, il suffisait que la SP erronée rappelle sans difficulté la SP correcte. D'autre part, nous avons relevé chez nos sujets un suremploi de certaines séquences. On peut sans doute considérer que quelques séquences seront typiques des apprenants en fonction de leur L1 mais aussi de la façon dont ils ont acquis la langue, de l'enseignement qu'ils ont reçu, ou tout simplement des situations où ils les ont entendues. Nous signalons les séquences déviantes par un astérisque.

## 5.7 Notre typologie

### 5.7.1 Les SP lexicales

a) Une phrase entière ou « syntagme propositionnel »

Les SP lexicales peuvent être des phrases entières exprimant une opinion ou une idée d'ordre général. Elles peuvent être utilisées dans de nombreuses situations et peuvent donc être un recours quand on est amené à produire une parole orale.

i. Une phrase entière exprimant une opinion ou une idée commune

Constructions personnelles (utilisation de la 3ème personne dans la narration)

*il ne comprend pas pourquoi, 2 ; il comprend pas, 2 ;*

Constructions impersonnelles

*ça s'appelle, 2 ; ça se passe, 2 ; ça a l'air, 1 ;*

Constructions interrogatives

*qu'est-ce qui s'est passé, 2 ; qu'est-ce qui se passe, 6 ; \*qu'est-ce qui passe, 1 ; qu'est-ce qu'il fait (va faire, a fait, faisait), 5 ; qu'est-ce que c'est, 2.*

ii. phrase entière exprimant une opinion du type *c'est/c'était* + adjectif. Nous n'avons pas classé cette construction dans la catégorie grammaticale car on considère l'ensemble et non la partie fixe *c'est* suivie d'une partie libre. L'ensemble de l'expression forme un tout.

*c'est amusant, 1 ; c'est drôle, 4 ; c'est différent, 1 ; c'est difficile, 1.*

b) SP indiquant une circonstance : temps, lieu, espace

i. Compléments de lieu, syntagmes prépositionnels

*là-bas, 14 ; chez + pronom (forme tonique), 2 ; à la maison, 4 ; à l'intérieur, 3 ; à l'extérieur, 15 (ces deux dernières formes en construction absolue).*

ii. Compléments de temps (phrases nominales et syntagmes prépositionnels)

*à ce moment-là, 3 ; en ce moment, 7.*

iii. Compléments exprimant la matière, la manière, syntagmes adverbiaux avec préposition

- *en* + matière : *en bois*, 8 ; *en plastique*, 6 ; *en verre*, 1 ; *en jouet*, 1

- *en continu*, 1 ; *en galope*, 1 ; \**en poing*, 1 ; *en recherche*, 1 ;

- *sans succès*, 1 ; *pour les enfants*, 2 ;

c) syntagmes nominaux

i. Une suite de mots exprimant une simple entité :

*salle de bains*, 34 ; *grain(s) de café*, 21 ; \**pois de café*, 6 ; \**haricots du café*, 1 ; \**morceaux du café*, 1 ; \**quiz du café*, 1 ; *maison de campagne*, 1 ; *coup de vent*, 1 ; *système de conduite*, 1 ; *fleurs d'soleil*, 1 ;

ii. Syntagmes nominaux avec *petit*, 141 ; *p'tit*, 27 : *petit garçon*, *petit cheval*, *petit bout*, *petit village*, *petit insecte*, *petit jaune*, \**petit jupe*, *p'tite sortie*, *p'tit garçon*...

d) Syntagmes verbaux

Les syntagmes verbaux peuvent être des locutions verbales comme « *avoir l'air* », 7 ; « *faire une randonnée* » mais aussi des constructions verbales comme « *commencer à*, *essayer de*, *changer N pour/avec N* » produites comme un tout et dont les constructions sont porteuses de sens. Les syntagmes *commencer à*, 20 et *commencer par*, 1, qui n'ont pas tout à fait le même sens, peuvent être utilisés comme marqueurs discursifs.

i. Locutions verbales (avec verbes courants) ou collocations verbales

*faire du café*, 1 ; *faire un signe*, 6 ; *avoir peur*, 4 ; *avoir l'air*, 7 ; *prendre part*, 1 ; *faire la sieste*, 4 ; *faire une blague*, 1 ; *faire une randonnée*, 1 ; *faire une bêtise*, 1 ; *avoir sommeil*, 6 ; *avoir envie de*, 4 ; *être fatigué*, 27, *se sentir fatigué*, 2 ; *sembler fatigué*, 2 ; *être libre*, 5 ;

ii. Constructions verbales (Cordier les classe dans la catégorie constructions grammaticales), verbe + préposition

*Commencer à*, 23 ; *commencer par*, 1 ; *sortir de*, 43 ; *essayer de*, 20 ; *continuer à*, 5 ;

iii. Verbes pronominaux ou réflexifs, 101 occurrences de cette forme à la troisième personne « *se* » +V

*Se réveille*, 12 ; *s'est réveillé*, 6 ; *s'est rendue compte*, 1 ; *se lève*, 4 ; *se regarde*, 6 ; *se ferme*, 3 ; *se promène*, 8 ; *se rencontrent*, 1 ; *s'appelle*, 6 ; *s'arrête*, 15 ; *s'ajoute*, 1 ; *s'endormir*, 3 ; *s'approche*, 5 ; *s'assoit*, 5 ; *s'est assis*, 1.

### 5.7.2 Les SP discursives

Ces SP ont en commun de faciliter le discours oral sans apporter d'information au contenu du message. On les appelle également méta-discursives (Cordier, 2013).

a) Les remplisseurs ou *fillers*

*je sais pas*, 25 ; *j'sais pas*, 6 ; *je sais pas quoi*, 3 ; *je pense*, 10 ; *je crois*, 10 ; *peut-être*, 38 ; *je veux dire*, 1 (VI) ; *en fait*, 21 ; *comment dire*, 14 ; \**comment dit*, 5 ; *c'est c'est*, 5 (l'emploi redoublé de *c'est* peut être suivi d'un abandon ou de la structure correcte).

Rôle métadiscursif : *je sais pas pourquoi*, 4 ; *j'sais pas comment*, 3 ; *j'sais pas si*, 2.

b) Structuration du discours

i. Les organisateurs du discours

avec pronom indéfini : *on voit / on voit pas*, 15 ; *on aperçoit*, 4 ; *on observe*, 1 ; *on a*, 2 ; *on peut*, 5 ; *on entend*, 4 ; *on dit*, 3, *on dit ça*, 3, *on sait pas*, 2 ; *on a déjà vu*, 1 ; *on commence à*, 1 ;

à la première personne, *j'ai vu*, 8 ; *je vois*, 1 (VO) ;

à la troisième personne, *le film commence par*, 2 ;

à la fin de la narration ou d'un épisode narratif, *c'est ça*, 12 ; *comme ça*, 12, *ça y est*, 2 ; *c'est comme ça*, 4, *quelque chose comme ça*, 1 ; *tout ça*, 1

ii. Les connecteurs

Expression de la chronologie : *et après*, 109 ; *et quand*, 10 ; *et puis*, 4 ; *et là*, 3 ; *et pis* ; *et aussi*, 2 ; *et maintenant*, 5 ; *après ça*, 7 ;

*un jour*, 6 ; *au début*, 3 ; *à la fin*, 1 ;

Expression de la conséquence : *et donc*, 27 ; *et voilà*, 7 ; *et finalement*, 7 ; *et alors*, 1 (VI) ; *du coup*, 6 (XI) ; *et soudain*, 2.

Expression de l'opposition : *mais finalement*, 1 (XI), *mais maintenant*, 1 ;

iii. Les constructions qui relient les propositions entre elles

*parce que*, 49 ; *pour que*, 6 ; *alors que*, 4 ; *dès que*, 1 ; *sans que*, 1 ; *avant que*, 1 ; *après que*, 6 ; *au lieu que*, 1 (AK) ; *tout le temps que*, 1.

### 5.7.3 SP grammaticales

a) Les constructions

i. pour exprimer une opinion :

À l'instar de Cordier, nous avons classé ces SP dans la catégorie grammaticale par leur structure, c'est-à-dire une construction composée d'une partie fixe et d'une partie libre. Mais on aurait pu aussi choisir de les classer dans les SP discursives car elles expriment une opinion et jouent un rôle dans le discours

*je crois que*, 11 ; *je pense que*, 7 ; *on se dit que*, 1 ; *on s'imagine que*, 1.

Contrairement à Cordier, on ne trouve pas d'occurrence des constructions suivantes : *je sais que, je trouve que* mais des formes à la troisième personne, spécifiques à la narration à la troisième personne, que nous avons considéré préfabriquées et classées dans les SP grammaticales.

- ii. *il dit que, 7 ; il sait (pas) que, 5 ; il pense que, 3 ; il croit que, 2 ; il sait que, 2 ; il trouve que, 3 ; dire que, 2 ; il a vu que, 8 ; il voit que, 7*
- iii. constructions pour décrire avec pronom sujet courant (indéfini) : *on voit que, 8.*

b) les présentatifs

*il y a, 116 ; il y avait, 5 ; il y a eu, 2 (HE), la forme au présent est la plus représentée, c'était, 22 ; ce sont, 1 ; c'est, 193 ;*

*c'est + N*

*c'est + GN + qui/que, emploi en clivée (c'est l'histoire \*où)*

*il y a + GN + qui*

*il est/elle est + adjectif*

*il est surpris, il est désolé, il est curieux, il est malin, il est méchant, il est navré*

c) les locutions prépositionnelles + GN

*avant de, 1 ; en fonction de, 1 ; en haut de, 1 ; grâce à, 1 ; à cause de, 2 ; à côté de, 21 ; à l'intérieur de, 1 ; à l'extérieur de, 3 ; autour de, 1 ; en dehors de, 1 ; de l'autre côté de, 2 ; en train de, 3 ; à travers, 2 ; autour de, 1 ;*

d) les quantifieurs

Les syntagmes nominaux marqueurs de quantité + N : *une bouteille de, 1 ;* mais aussi SN ayant d'autres valeurs (espace, manière) : *l'intérieur de, 1 ; le côté de, 1 ; la présence de, 4 ; une façon de, 1.*

Les adverbes de quantité + N : *beaucoup de, 13 ; un peu de, 3 ; \*tellement de, 1 (Qi) ; + Adj. ou en construction absolue, un peu, 44.*

e) les déterminatifs : *tout le / tous/toutes les / toute la, 29 ; le, la, les même(s), 3 ; l'autre/un autre/les autres, 51.*

f) les pronoms indéfinis : *un autre, 2 ; quelque chose, 14*

g) les adverbes négatifs : *pas trop, 2 ; pas vraiment, 4 ; plus rien, 2 ; pas très, 1*

h) les verbes fréquents marqueurs d'aspect,

*il vient de, 4 ; il va + V, 22 ; être en train de + V, 8 ;*

i) les constructions

*il va + espace, lieu : il va à la maison/ à la cuisine, 3 ;*

les modaux + inf. : *il peut*, 11 ; *il peut pas*, 5 ; *elle voudrait*, 1 ;  
les constructions impersonnelles : *il semble que*, 1 ; *il faut que*, 2 ;  
les constructions interrogatives : *est-ce que*, 2 ; *qu'est-ce que*, 8.

#### 5.7.4 Les SP déviantes.

Voici les principales catégories relevées dans notre corpus (tableau des SP déviantes en Annexe 8)

##### a) les SP déviantes lexicales

###### i. Syntagmes verbaux :

\*(il) *reste étonné* forme idiosyncrasique directement traduite de l'italien,

\*(il) *a été sommeil* qui est utilisé avec *être* au passé composé au lieu de avoir au présent,

\**descendre en* cheval, pour *descendre de* cheval : erreur morphologique d'auxiliaire verbal

###### ii. Syntagmes nominaux :

\**en* cheval, 2 au lieu de *de* cheval ou *à* cheval

###### iii. Syntagmes prépositionnels :

\* *sauf un* pour *à part un*

\**à un moment* ou \**en ce moment* pour *à ce moment*

###### iv. Syntagmes propositionnels

\**qu'est-ce qui passe* pour *qu'est-ce qui se passe* (dû sans doute à une mauvaise perception de l'ensemble).

Cas de certains syntagmes propositionnels

*Qu'est-ce qui s'est passé/ se passe* pour *il ne comprend pas/ sans comprendre ce qui s'est passé*, la phrase interrogative directe a été apprise et produite telle quelle et n'a pu être modifiée en interrogative indirecte. On a ici un exemple typique d'une phrase produite en bloc sans transformation. La forme correcte aurait été : *il ne comprend pas / sans comprendre ce qui se passe*. Ceci étant, nous ne les avons pas comptées dans les déviantes parce qu'elles sont produites correctement. Elles sont la preuve d'une mémorisation en bloc de la suite de mots.

##### b) Les SP déviantes discursives

La production de \**comment dit*, (3) pour *comment dire* est le résultat d'une erreur phonologique, parfois rectifiée lors de la deuxième narration, dû à une mauvaise réception auditive.

##### c) les SP déviantes grammaticales

Comme dans le groupe déterminant quantitatif *\*un bouteille de pour une bouteille de*, erreur morphologique, accord du déterminant avec le nom.

Les séquences préfabriquées déviantes sont pour la plupart des SP lexicales non encore maîtrisées ou non segmentés : dans *il \*pas compris* le sujet calque la troisième personne sur le modèle de *j'ai pas compris* où *j'ai* est mémorisé en bloc comme une seule unité. L'apprenant n'a pas encore la connaissance grammaticale suffisante à son niveau d'apprentissage, soit qu'il n'ait pas encore visualisé les formes écrites de la séquence verbale, soit que la forme déviante ait été fossilisée à l'oral. Ces séquences peuvent être idiosyncrasiques comme chez FA ou HE (propres à plusieurs apprenants ou à un seul) (Forsberg, 2006)

Certaines peuvent avoir été mal perçues : *\*comment dit* pour *comment dire*, *\*sans fait attention* pour *sans faire attention*. Elles peuvent comporter une erreur de type morphosyntaxique. Erreur sur le genre *une \*petit abeille*, *\*tous les choses*.

ERREURS	SEQUENCES
Morphosyntaxique	<i>*tous les choses, *petit abeille, *sans compris</i>
Lexicale	<i>*pois de café, *reste étonné</i>
Phonologique	<i>*comment dit, *qu'est ce qui passe</i>

Tableau 4. Types d'erreurs

Nous avons seulement noté les SP déviantes par un astérisque [\*], et les avons classées par types d'erreurs (Annexe 9). Certaines d'entre elles sont des SP propres à l'apprenant et à son niveau ; d'autres sont le fait d'une mauvaise réception phonologique (*\*comment dit* pour *comment dire*) ou encore d'une erreur de l'apprenant (*\*sans compris* pour *sans comprendre*). Il peut aussi arriver qu'un sujet utilise une SP correcte cette fois pour pallier un manque lexical. Le sujet Akira, ne trouvant pas le mot adéquat, utilise trois fois la SP « *tant pis* » là où il aurait fallu manifestement un verbe. La séquence a été mémorisée et employée telle quelle dans un énoncé où elle exprime la consternation et l'impuissance du personnage, le cowboy, constatant que son cheval s'est échappé. La répétition trois fois de suite de la même séquence renforce le sentiment de désappointement du cowboy. A ce stade de l'apprentissage, la séquence « *tant pis* » est une séquence figée non-compositionnelle trop complexe pour avoir été construite mot à mot.



## 5.8 Mesures de la préfabrication

Il n'a pas été aisé d'emblée de déterminer le choix de mesures qui rendent compte de la préfabrication. Comment compter les séquences préfabriquées après les avoir identifiées, d'une fois sur l'autre ; fallait-il compter les séquences, les mots à l'intérieur de la séquence, les syllabes de la séquence ?

Par exemple, Forsberg (2006) classe et comptabilise les SP selon des types (lexical, grammatical, discursif). Mais ce classement par type ne dit rien de l'évolution du LP dans le temps. Dans cette recherche, nous voulons choisir un mode de calcul qui montre que le type des SP (lexical, grammatical ou discursif) change dans le temps, en quantité mais aussi en qualité. Pour mesurer le taux de langage préfabriqué, de nombreuses mesures ont été appliquées par les chercheurs. Dans une recherche récente, Forsberg-Lundell *et al.* (2014), calculent le taux de préfabrication en utilisant un mode de calcul pour 100 mots produits. Ce calcul permet d'évaluer l'évolution quantitative des SP dans le temps par type de SP. On peut alors vérifier si les apprenants s'appuient davantage sur des SP de type discursif et/ou grammatical en t1, et si leur répertoire de SP s'enrichit de séquences lexicales, après six mois. Pour ce calcul, il faut avoir identifié et compté les SP par type (lexical, discursif, grammatical), en valeur brute, c'est-à-dire toutes les SP par type même répétées plusieurs fois dans la même narration. Ensuite, il faut compter ces SP une seule fois, on obtient un nombre de SP produites en occurrences types.

Cordier dans sa thèse (2013) adopte un autre mode de calcul. Elle ne compte pas le nombre de SP qui, selon elle, ne dit rien du LP en L2 ; les SP produites par des locuteurs L2 n'étant pas les mêmes que celles de la L1, décrites en linguistique. Elle avance que les séquences en L2 sont celles qui facilitent le traitement d'un locuteur donné ce qui explique, en partie, qu'on ne trouve pas ou très peu de séquences idiomatiques dans leurs productions. cela niveau des apprenants. Cordier mesure le rapport en pourcentage entre le nombre des syllabes des SP et le nombre total de syllabes produites par un même apprenant. Les calculs des deux recherches citées sont très différents : prenons le groupe « *c'est c'est intéressant* » Forsberg identifiera deux SP, *c'est* et *c'est intéressant*, alors que Cordier va relever seulement la séquence « *c'est intéressant* » qu'elle comptera en syllabes (5 syllabes). Selon le mode de calcul de Cordier, cet autre exemple, *il a un peu sommeil*, comporte six syllabes si l'ensemble est produit d'une traite sans aucune micro pause entre le sujet et le verbe, mais si le locuteur n'a produit que *a été un peu sommeil* d'un trait, on compte seulement 5 syllabes. Selon le mode de calcul de Forsberg, on identifie deux SP (*avoir sommeil*, *un peu*) dans tous les cas.



Comme Cordier, nous avons vérifié chaque enregistrement pour être sûre que la suite de mots était bien prononcée sans interruption. Par exemple, dans la suite *elle n'arrive pas à sortir de*, toutes les syllabes du groupe sont dites sans pause et nous comptons alors 8 syllabes. Ce calcul en syllabes permet de montrer le rapport entre syllabes préfabriquées et syllabes de la production, ce qui permet d'évaluer la proportion de syllabes préfabriquées et l'amélioration en pourcentage de syllabes préfabriquées dans le temps. Ce type de calcul est intéressant pour évaluer la proportion de syllabes préfabriquées mais cela ne nous dit rien de la qualité des SP. Il faut donc choisir un autre mode de calcul qui va compléter celui-ci. C'est pourquoi, nous avons aussi appliqué le classement de Forsberg et son mode de calcul, pour voir ce qui se passe au niveau du répertoire des SP. Cette méthode permet de montrer l'évolution des catégories de SP dans le temps, notre question de recherche étant de voir si les SP évoluent dans le temps et si les locuteurs non-natifs emploient plus de SP lexicales après un temps en immersion. Dans un premier temps, la suite *elle n'arrive pas à sortir de* est composée, selon la méthode de Forsberg, de deux SP (*ne pas*) *arriver à* et *sortir de*. Puis, nous avons appliqué un taux pour déterminer le nombre de SP produites tous les 100 mots. Afin d'avoir un meilleur aperçu de l'évolution en termes de diversité, nous avons aussi mesuré le nombre de SP pour cent mots en occurrences types ; c'est-à-dire si *sortir de* est produit plusieurs fois par le même sujet dans la même narration, cette SP ne sera comptée qu'une seule fois.

A notre avis, les deux modes de calculs sont intéressants et complémentaires. Le mode de calcul selon Cordier nous permet de constater l'évolution du langage préfabriqué dans le temps en termes de quantité. Le calcul du nombre de SP pour cent mots nous permettra de mesurer l'évolution des types de SP utilisées et de voir si des changements s'opèrent dans le temps. Rappelons une de nos questions de recherche qui se propose de voir si le type de SP évolue dans le temps.

#### ***a) Calculs selon Cordier (2013)***

Nous avons calculé le nombre de syllabes de chaque SP ainsi que le nombre de syllabes de chaque segment de parole. Ces calculs ont été faits pour chaque sujet pour chacune des tâches, la première et la deuxième fois. Nous avons ensuite rassemblé les données en temps 1 puis les données en temps 2 afin de comparer les résultats d'une fois sur l'autre. Toutes les données ont été calculées pour le groupe entier (17 sujets) et par groupes de niveau (3 groupes), ainsi que pour le groupe contrôle (6 sujets).

Données brutes

- Nombre total de syllabes

- Nombre de syllabes préfabriquées
- Nombre total de segments de parole
- Nombre de segments contenant des préfabriqués

Ces données brutes ont permis de calculer le développement des SP dans le temps, selon les quatre mesures suivantes :

- Pourcentage de syllabes préfabriquées par rapport au nombre total de syllabes de la production
  - Pourcentage de segments de parole contenant des préfabriqués par rapport au nombre total de segments d'une production. Cordier (2013 : 128) reconnaît que cette mesure est moins intéressante que la précédente car elle ne dit rien de la quantité de langage préfabriqué. Elle offre cependant une vision plus dynamique de la répartition des SP dans les productions des sujets.
    - Longueur moyenne des syllabes préfabriquées par segment. On obtient ce nombre en divisant le nombre total de syllabes préfabriquées par le nombre total de segments de parole. Cette mesure permet de relier la préfabrication (nombre de syllabes préfabriquées) à la fluence (longueur du segment de parole).
    - Longueur moyenne des syllabes préfabriquées dans des segments de parole contenant une ou plusieurs SP. On l'obtient en divisant le nombre total de syllabes préfabriquées par le nombre de segments contenant au moins une SP. Cette mesure permet voir de manière précise quelle est la proportion de LP dans le segment. L'impact du nombre de syllabes préfabriquées sur la longueur moyenne des segments de parole contenant une ou plusieurs SP est d'importance pour juger de la fluence.

#### ***b) Pourcentage de SP pour 100 mots***

Nous avons choisi d'appliquer un autre mode de calcul, utilisé dans les recherches (Forsberg, 2010 ; Forsberg-Lundell *et al.* 2014) et aussi en linguistique de corpus, en particulier dans la recherche en *n-gramm*. On extrait les occurrences d'une unité ou d'une suite de mots par 1000 mots, voire par million de mots (Ellis *et al.* 2008). Nous avons préféré comme Forsberg-Lundell et ses collaborateurs (2014) ramené ce taux à 100, compte-tenu de la taille de notre corpus.

Forsberg (2006) dans sa thèse calcule le nombre total de SP par production, par sujet et par niveau. Nous adoptons dans un premier temps cette mesure pour classer les SP dans les trois grandes catégories (section identification des SP), ce qui nous donne deux mesures : le

nombre total de SP dans les trois catégories de SP (lexicales, grammaticales et discursives) et le nombre d'occurrences types :

- Nombre total des SP : lexicales, discursives et grammaticales

Cette mesure rend compte du nombre des séquences par types utilisées par un sujet (Annexe 3). Ce classement permet de voir si des changements ont eu lieu la deuxième fois. A la différence de Forsberg qui examine le langage préfabriqué dans un corpus déjà établi, le corpus Interfra, nous avons constitué notre corpus à partir des enregistrements des apprenants en t1 et en t2 car nous nous intéressons au développement du LP dans le temps. Nous avons classé les SP par types comme Forsberg (2006) et Cordier (2013).

- Nombre de SP pour 100 mots

Cette mesure intéressante veut montrer l'importance des SP dans la production et calcule le nombre de SP produites tous les cent mots. On l'obtient en divisant le nombre de SP par le nombre total de mots de la production et on multiplie par 100 le résultat.

- Nombre de SP en occurrences types.

Un sujet peut utiliser plusieurs fois la même séquence. Cette mesure vérifie la plus ou moins grande variété des SP. Elle donne une image plus juste des habitudes du sujet qui peut ou non s'appuyer sur les mêmes séquences.

Ces calculs seront faits pour les SP en valeur absolue et en occurrences types afin de montrer la diversité des SP dans leur ensemble et, en particulier, le nombre de SP lexicales.

Évidemment ces mesures ne nous disent rien de la longueur de la SP.

## **5.9 Mesures de la fluence**

Un autre ensemble de mesures permettent de calculer la fluence. Nous présentons dans cette section les variables retenues ainsi que nos modes de calcul.

### ***a) Les variables temporelles***

Les données entrées dans des fichiers Excel pour chaque tâche et chaque sujet ont permis de faire différents calculs afin de mesurer les variables temporelles.

Données brutes pour les mesures de la fluence

- Durée totale de la production= temps de locution
- Durée de la production orale sans les pauses = temps d'articulation

- Nombre total de pauses
- Durée des pauses pour chaque tâche
- Nombre total des syllabes de la production

Ces données ont permis de calculer les variables temporelles (Grosjean et Deschamps 1972 ; Kormos et Dénes, 2004 ; Towell *et al.*, 1996) présentées ci-dessous et dans le tableau 5.

Vitesse de parole : nombre de syllabes produites par seconde, divisé par le temps de production en secondes multiplié par soixante.

Vitesse d’articulation exprimée en nombre de syllabes produites par seconde : nombre total de syllabes produites divisé par le temps d’articulation (sans les pauses).

Taux de phonation (*phonation time ratio*) le pourcentage de temps passé à parler par rapport au temps de la production, rapport TA-TL, temps d’articulation (sec) divisé par temps de locution (sec) multiplié par 100.

Longueur moyenne des pauses : temps total des pauses divisé par le nombre total de pauses.

Nombre de pauses silencieuses par minute : on divise le nombre total des pauses silencieuses par le temps de production en secondes et on multiplie par soixante.

MESURES	CALCULS
Vitesse de parole (SPS/VP)	Nb de syllabes/ tps de parole
Vitesse d’articulation (VA)	Nb de syllabes/tps sans les pauses
Taux de phonation/ratio (TdP)	Tps sans les pauses/ tps total de parole
Longueur moyenne des pauses (LMP)	Longueur moyenne pauses/ Nb total pauses
nombre de pauses silencieuses par minute (NPM)	Nb total pauses vides/ Tps total en secondes

Tableau 5. Mesures des variables temporelles

### ***b) Longueur du segment de parole***

La longueur du segment de parole mesurée en mots ou en syllabes est une des mesures classiques de la fluence. Contrairement à d’autres chercheurs, nous avons considéré dans cette recherche que seules les pauses vides  $\geq 250$ ms délimiteraient les segments de parole. Cette mesure, qui n’est pas une variable temporelle, permet de constater s’il y a évolution ou pas dans le temps en termes de longueur d’un énoncé entre deux grandes pauses silencieuses. Nous avons choisi de prendre en compte seulement les pauses silencieuses ou les groupes d’hésitation contrairement à d’autres chercheurs. Pour Cordier (2013 :119-120), une répétition, une pause remplie ou un allongement de voyelle d’une durée supérieure à 400ms

délimitent un nouveau segment. Nos résultats ne peuvent être comparés à ceux de Cordier dans la mesure où nous n'avons pas utilisé les mêmes critères.

Pour calculer la longueur moyenne du segment de parole en syllabes, on divise le nombre de syllabes par le nombre de segments pour trouver le nombre moyen de syllabes produites entre deux pauses supérieures ou égales à 250 ms (Towell *et al.* 1996).

### *c) Les reprises ou corrections*

Nous avons choisi de relever un autre type d'hésitations ou corrections que sont les répétitions, les faux départs et les abandons (quatrième étape du modèle de Levelt, auto-surveillance, chapitre 3). Grosjean et Deschamps préconisent de calculer la moyenne des corrections de la production en divisant le nombre total de corrections par le nombre de segments (1972 : 133). Nous utilisons une autre mesure qui calcule le rapport entre le nombre de corrections et nombre total de mots de la production. On divise le nombre de corrections par le nombre de mots que l'on multiplie par cent. On obtient un taux pour cent mots.

## **5.10 Limites de la méthodologie**

Le petit nombre de sujets (n=17) ne nous permettra pas de tirer des conclusions d'ordre général. Dans le cas de notre recherche, il est préférable d'observer ce qui se passe à titre individuel, sujet par sujet.

- La tâche

Le choix d'un type de tâche ne rend pas compte du langage préfabriqué dans son ensemble. En effet, la tâche de narration choisie limite le choix des séquences que l'on pourrait trouver dans d'autres tâches orales comme l'interaction, la discussion sur un sujet d'actualité ou des tâches qui nécessitent de produire des discours autres que narratifs tels que le discours descriptif, argumentatif, explicatif. Le type de tâche conditionne le type de séquences produites, en particulier les séquences lexicales. On ne peut trouver que des séquences liées au sujet, ce qui limite forcément le répertoire. Cela étant, on ne peut rendre compte de l'immense variété des séquences. La comparaison entre les deux tâches peut néanmoins être intéressante pour le chercheur. Cependant, dans une étude très intéressante, Forsberg et Fant (2010 : 64) montrent qu'une interaction peut être une tâche difficile même pour des utilisateurs de la L2 très avancés. Ces chercheurs constatent que les sujets non-natifs testés produisent moins de SP et sont moins fluents dans la tâche de dialogue (*Boss*) que les natifs alors que les résultats dans les deux groupes, à part une légère divergence dans les SP lexicales, sont similaires dans la tâche narrative (*Charlie*).

Comme le dispositif choisi implique la production d'un certain type de SP, nous n'avons aucune séquence discursive situationnelle comme chez Forsberg (2006) du type *je m'appelle, j'ai X ans*, etc. Cordier (2013) dans sa thèse soumet ses sujets à quatre tâches différentes : l'interview, le dialogue entre deux sujets sur un sujet d'actualité, une tâche où l'apprenant donne son opinion sur deux sujets d'actualité, ainsi qu'une tâche narrative. Ses cinq sujets, un groupe homogène, tous de niveau avancé, ont été soumis à ces quatre tâches avant et après un séjour de six mois en immersion dans le pays de la langue cible (la France). Elle ne fait appel à aucun groupe contrôle. Comme Cordier le fait remarquer, la tâche narrative est la tâche la moins productive en termes de séquences préfabriquées de type référentiel, comparée aux autres tâches où les apprenants ne sont pas contraints par un support et se sentent plus libres. Si effectivement le choix de la tâche influe sur la production et la liberté d'utiliser une grande variété de SP, notre objectif est plutôt de montrer comment les sujets utilisent les séquences dans une narration, par exemple *je pense que* qu'ils transforment à la troisième personne *il pense que*. D'autre part, ils mettent en place des stratégies d'évitement ou de remplacement lorsque le mot n'est pas connu, et emploient souvent des groupes « préfabriqués » qui compensent leur méconnaissance du lexique. Par exemple, le sujet OL qui ne connaît pas le mot « abeille » le remplace par la SP « *petit jaune* » sur le modèle de *petit garçon*. De même chez AK, un de nos sujets de niveau 1, l'utilisation de « *tant pis* » sert à exprimer un sentiment de déception. Ne connaissant sans doute pas le vocabulaire approprié, il utilise trois fois cette même SP.

- Hétérogénéité du groupe

### Langues diverses

Un groupe homogène, c'est-à-dire ayant une même L1 facilite la tâche du chercheur. La diversité des langues maternelles de nos sujets ne nous permet pas d'analyser les SP déviantes en regard avec la L1. Par exemple, le sujet italophone produit *\*reste étonné* pour *être étonné*, sans doute parce qu'en italien on dit *rester étonné*. Il aurait été cependant intéressant de comparer de façon systématique les erreurs de tous les sujets selon les expressions de leur L1.

### Niveau d'étude

Nos sujets n'ont pas tous le même parcours universitaire ou professionnel. Nos sujets d'Albi sont doctorants ou en stage pour une année. Les sujets de Toulouse, s'ils ont terminé leurs études au lycée n'ont pas tous un diplôme universitaire, et certains qui ont terminé leurs études (PR) travaillent ou sont en recherche d'emploi. En revanche, ils sont tous inscrits au DEFLE et suivent des cours de français.

- Jugement des natifs

Le recours au jugement des natifs aurait pu être fait de manière systématique ; nous avons fait appel au jugement de deux juges natifs avertis qui avaient pour tâche d'examiner les productions et les séquences relevées. Notre travail n'étant pas un travail strictement linguistique nous avons préféré croisé, comme nous l'avons dit par ailleurs, un certain nombre de traits d'identification.

- La motivation

Si nous n'avons eu aucun mal à tester nos sujets, il se trouve que nous constatons pour certains d'entre eux une baisse de la motivation lors du deuxième enregistrement. Certains sujets ont produit des narrations très courtes qui ne nous permettent pas de généraliser les résultats. Cependant, on peut se demander s'il n'est pas normal que quelques sujets se trouvent dans cette situation. Deux explications sont possibles : cela peut être effectivement le résultat d'une baisse de la motivation pour certains sujets ou alors la volonté de mieux faire en étant plus créatif (précis) pour d'autres.

Comme nous l'avons expliqué au début de ce chapitre, notre objectif était de voir ce qui se passe chez des apprenants n'ayant eu aucun enseignement spécifique des SP, alors qu'ils sont soumis à une tâche narrative et qu'ils avaient d'ailleurs toute liberté de donner leur opinion. Si nous sommes consciente qu'ils ne peuvent pas produire certaines SP, nous trouvons intéressant de montrer comment les apprenants arrivent à pallier le manque lexical face à une narration imposée et quelles stratégies ils utilisent.

- Les types de séquences

Les SP lexicales produites sont en rapport avec la tâche. Nous trouvons donc un certain type de séquences (lexicales) en rapport avec le sujet de la narration, comme les séquences construites avec « petit » ou la séquence lexicale « en bois » puisque dans la BD, les « personnages » sont un *petit enfant (petit garçon)* et un *petit cheval « en bois »*. On note l'utilisation récurrente des syntagmes avec *petit* aussi bien chez les LNN que chez les LN. « *petit* » n'est pas toujours utilisé avec le sens de diminutif et même quand il pourrait l'être rien ne nous l'indique clairement. Aussi avons-nous considéré tous les syntagmes avec cet adjectif comme des séquences préfabriquées. De même nous constatons chez nos sujets non-natifs un type de séquences discursives sur-employé « *et après* » pour structurer les différentes étapes chronologiques du récit. Ils s'appuient fortement sur cette SP discursive à la

différence des natifs où l'on n'en trouve qu'une seule occurrence (CA, film). Les natifs utilisent plus de séquences idiomatiques et plus variées, ce qui semble normal compte-tenu du statut de nos sujets.

Ce chapitre consacré à la présentation de notre méthodologie, clôt notre première partie théorique et annonce la partie dédiée à la présentation des résultats ainsi qu'à leur analyse.





## **II. Présentation et analyse des résultats**

---

## **Introduction aux chapitres 6 et 7**

Dans les chapitres 6 et 7, nous présentons les résultats des différentes mesures appliquées dans cette recherche aux productions L2. Nous répondrons ainsi aux questions de recherche que nous rappelons ici :

- 1) Usage et développement du LP chez les apprenants de français langue étrangère des niveaux débutant à avancé :
  - quelle est la proportion de LP dans le discours oral des sujets testés ?
  - quels types de SP produisent-ils ?
  - constate-t-on une évolution dans le temps dans l'emploi des SP et les types de SP produites ?
- 2) les SP et le développement de la fluence
  - comment la fluence en L2 se développe-t-elle après une période de six mois dans le pays de la langue cible ?
  - y a-t-il une corrélation entre l'usage des SP et la fluence ?

Dans cette étude, nous voulons tester l'hypothèse que l'usage des SP aide les locuteurs L2 à s'exprimer avec plus d'aisance. Mais les effets risquent d'être probablement différents selon le niveau des locuteurs, répartis dans trois groupes de niveaux. Il sera donc intéressant de regarder de façon détaillée (et qualitative), les différents types de SP utilisés selon le niveau, un éventuel développement dans les fonctions discursives de ces entités dans le discours L2.

# Chapitre 6 Présentation des résultats

## Comparaison groupe natif et groupe non-natif

---

Avant de présenter les résultats du groupe de locuteurs natifs, nous récapitulons (tableau 6) les tâches auxquelles ont été soumis nos sujets et les abréviations que nous utiliserons par la suite ainsi que le nombre de mots produits.

Tâche narrative	Description	Corpus non-natifs (NN) en mots	Corpus Natifs (N)
Bd1	cartoon, temps 1 ou t1	2717	-----
F1	film muet, temps 1 ou t1	6314	-----
Bd2	cartoon, temps 2 ou t2	2986	1065
F2	film muet, temps 2 ou t2	6999	2588

Tableau 6. Nombre de mots par tâche, natifs et non-natifs

La taille totale du corpus L2, obtenu suite à la réalisation de ces tâches, est de 19017 mots pour tous les sujets, pour les deux tâches (Bd et Film) aux deux moments (temps1 et temps2). Le corpus L1 comprend 3653 mots, le groupe des natifs ayant effectué les mêmes tâches, une fois seulement. Les résultats des natifs seront donc comparés avec ceux des non-natifs uniquement lors du deuxième enregistrement (temps2).

### 6.1. Valeurs brutes du groupe natif

Afin de mieux comprendre ce qui se passe en L2, il nous a paru important de nous attarder sur le groupe des locuteurs natifs et de comparer leurs résultats à ceux des non-natifs. Nous utiliserons le test (U) de Mann Whitney pour comparer les résultats des deux groupes entiers, le groupe des natifs (n=6) et le groupe des non-natifs (n=17), ainsi que ceux du groupe des natifs (n=6) et du groupe avancé seul (n=6). Les résultats du groupe natif nous donnent des indices de la production orale « fluente », qui nous serviront de base ou de point de départ pour nos analyses de la production en L2 (sans que les valeurs du corpus natif servent de « but à atteindre », bien sûr).

Cette approche qui peut sembler étrange de prime abord, nous paraît utile pour comprendre les particularités de la production en L2. Nous pourrons ainsi déterminer les mesures les plus pertinentes pour rendre compte du phénomène de la préfabrication et de la fluence liée à celle-ci. Mais tout d'abord, il sera important de voir la place du langage préfabriqué dans la

réalisation de nos deux tâches par des locuteurs natifs : les points communs et la part de la variance.

Les calculs, pour les locuteurs natifs, ont été faits pour les deux tâches, Bd et Film confondus, car une série de tests de rang de Wilcoxon appliquée aux mesures temporelles de productivité ne montre aucune différence significative entre ces deux tâches pour les natifs.

Le tableau 7 suivant présente l'ensemble des données brutes des natifs qui serviront pour le calcul des mesures de la préfabrication et de la fluence : chaque individu d'abord (colonnes grises), ensuite les moyennes descriptives pour le groupe des locuteurs natifs obtenues pour les 18 mesures figurant dans chaque ligne (liste en colonne 2). La première série de mesures (lignes 2 à 5) permet de calculer le taux de préfabrication des productions natives selon la méthode de Cordier (2013). Une deuxième série de mesures plus classiques, taux pour 100 mots (lignes 1 et 6 à 11) indique le rapport entre le nombre de SP et le nombre de mots de la production. Cette mesure, utilisée pour 1000 mots ou par million de mots dans les recherches dans de larges corpus (Ellis *et al.* 2008), a été appliquée à notre typologie adaptée de Forsberg (2006), pour 100 mots. Les données brutes serviront aussi aux divers tests khi-deux que nous utiliserons pour comparer les caractéristiques de productions natives et non-natives.

Nous complétons le tableau (lignes 12 à 15) avec les mesures temporelles brutes que sont le temps de production en secondes, la durée totale des pauses ainsi que le nombre de pauses et de groupes d'hésitation. Ces données serviront aux calculs des mesures de la fluence. Les dernières valeurs, lignes 16 et 17, concernent les diverses hésitations sonores, comme les pauses remplies et les allongements. La dernière ligne, 19, fait état du nombre de reprises de la production (répétitions, reformulations, faux départs et abandons). Ces dernières données fournissent des précisions pour mieux comprendre les spécificités de la production orale.

	Données t2	AN	CA	JE	MG	TH	VC	Moyenne	Ecart-type	min	max
1	Nb mots bd et film (t2)	679	694	512	922	342	504	608,8	200,9	342	922
2	Nb syllabes bd et film (t2)	865	883	672	1183	460	646	784.8	249.8	460	1183
3	Nb de segments	42	64	66	80	28	42	53.7	19.4	28	80
4	Nb de segments avec SP	38	53	40	59	20	32	40.33	14.12	20	59
5	Nb de syllabes préfabriquées	317	335	227	459	151	256	290.8	105.6	151	459
6	Nb de SP brutes bd et film (t2)	108	109	68	148	44	78	92.5	36.7	44	148
7	Nb SP types Bd et film	82	77	54	110	40	67	71.7	24.2	40	110
8	Nb de SP lexicales	62	60	46	73	32	39	52	15.5	32	73
9	Nb de SP grammaticales valeur brute	29	14	11	44	8	20	22,66	13,08	8	44
10	Nb de SP discursives valeur brute	17	25	11	31	4	19	17,83	9,64	4	31
11	Nb de SP lexicales occurrences types	49	45	36	62	30	32	42.3	12.1	30	62
12	Temps production (s)	200,65	260,72	229,36	290,47	124,38	156,33	210,32	62,8	124,38	290,47
13	Durée des pauses (s)	29,94	49,66	61,87	55,69	17,86	29,97	40,33	17,77	17,86	61,87
14	Nb pauses vides	40	62	64	78	26	40	51,67	19,41	26	78
15	Nb de groupes d'hésitation	4	5	7	6	5	4	5,17	1,17	4	7
16	Nb de pauses remplies	26	33	23	56	35	21	32,33	12,83	21	56
17	Allongements	26	39	36	46	10	12	28,17	14,78	10	46
18	Nb de reprises	30	36	39	68	20	10	33,83	19,86	10,00	68,00

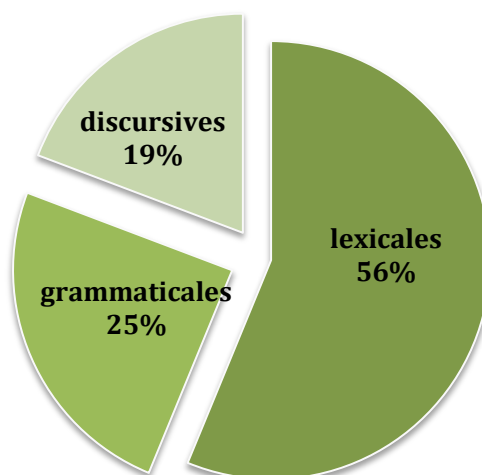
Tableau 7. Données brutes, locuteurs natifs

*a) Premiers constats sur les moyennes de groupe*

Le groupe natif produit en moyenne 608,8 mots ou 784,8 syllabes, et ces locuteurs parlent en moyenne 210,32 secondes soit trois minutes et demie, Bd et Film confondus. A 173 mots par minute, la moyenne de mots produits à la minute de ce groupe correspond à celle constatée par d'autres chercheurs (Levelt, 1989 : 22). La durée totale des pauses vides est de 40 secondes en moyenne. Les natifs produisent en moyenne 53,7 segments de parole, 52 pauses vides (nombre de segments moins un) et 32,33 pauses remplies. On rappelle ici que les segments de parole sont délimités par deux pauses d'une durée  $\geq 250$ ms. Sur les 784,8 syllabes produites, 290,8 syllabes appartiennent à des SP soit un taux moyen de 37 % de syllabes préfabriquées. La moyenne des SP, Bd et Film confondus, par sujet est de 92,5 SP. La répartition par type est la suivante : 52 SP lexicales, 22,66 SP grammaticales, 17,83 SP discursives. Nous complétons ce tableau avec les valeurs en occurrences types des SP.

Figure 2. Distribution des SP

**Distribution des types de SP,  
corpus natif**



La figure 2 ci-dessus montre la distribution des types de SP produites par le groupe natif. On constate un pourcentage plus important de SP lexicales, plus de la moitié des SP (56%), un

quart de SP grammaticales (25%) et 19% de SP discursives. Pour voir s'il y avait une variation significative dans les types de SP produits d'un individu à l'autre, nous avons utilisé le test khi-deux pour comparer cette distribution ; elle est non significative : ( $\chi^2(10, 555) = 15,8$  ( $p=0,11$ , ns)). Les six natifs utilisent donc en proportions équivalentes, les différents types de SP : lexicales, grammaticales et discursives.

### *b) Valeurs individuelles*

Les valeurs de groupe sont intéressantes mais ne sont pas suffisantes quand on s'intéresse aux individus. En effet, les valeurs individuelles, très contrastées (colonnes grisées), donnent une indication des différences entre les sujets. Les écarts-types (colonne 10), parfois élevés, signalent que, pour un même groupe en apparence homogène, les caractéristiques de la production verbale sont hautement idiosyncrasiques. On constate, notamment, que la quantité de discours produit en mots présente un écart-type de 200,9 mots sur une moyenne de 608,8 mots ; et en syllabes, un écart-type de 249,8 syllabes sur une moyenne de 784,8 syllabes, ce qui est considérable. Certains sujets comme TH, JE et VC sont très concis dans leur narration alors que d'autres comme MG et CA parlent beaucoup plus (342 et 922 mots et 460 et 1183 syllabes pour les valeurs extrêmes). Le nombre de SP produites montre également des écarts importants entre les sujets : une moyenne de 92,5 SP avec un écart-type de 36,7. Le nombre de pauses silencieuses, variant du simple au triple, est aussi à mettre en relation avec la longueur de la production et le nombre de segments. Les écarts-types des diverses manifestations des disfluences - les groupes d'hésitation (pause remplie et pause vide), les pauses remplies et les allongements - montrent des différences très importantes entre les sujets natifs. De même, l'étendue du nombre des reprises est de 10 à 68. Pour conclure sur ces données brutes, on peut dire qu'il n'y a pas, dans un groupe d'individus, de « norme » à proprement parler. Au vu des résultats, les paramètres de la production orale semblent être hautement individualisés. Ces données brutes, nous ont permis de calculer les taux de préfabrication du groupe natif. Les résultats obtenus par les LN (section 6.2) seront la base d'une comparaison des résultats du groupe contrôle avec ceux du groupe expérimental.

## **6.2 Résultats des sujets natifs**

### **6.2.1 Mesures de la préfabrication**

Nous rappelons ici les deux méthodes utilisées pour le calcul de la préfabrication qui ont été commentées dans le chapitre 5. La première méthode (Cordier, 2013) propose quatre mesures



pour calculer le taux de préfabrication dans les productions : taux de syllabes préfabriquées (pourcentage des syllabes des SP), taux de segments de parole contenant au moins une SP (pourcentage des segments contenant au moins une SP), longueur moyenne des syllabes préfabriquées par segment, longueur moyenne de syllabes préfabriquées par segment contenant au moins une SP. La deuxième méthode utilisée en linguistique de corpus calcule le nombre total de préfabriqués pour cent mots produits.

Le tableau 8 présente les taux de préfabrication pour les productions du groupe natif. Les lignes 1 à 4 du tableau concernent les quatre mesures de la préfabrication selon Cordier (2013). Les lignes 5 à 8 présentent le taux de préfabrication pour 100 mots, toutes séquences confondues et le taux des séquences de type lexical, en valeur brute et en types. Les lignes 9 à 12 concernent deux mesures intéressantes, le taux de SP grammaticales et le taux de SP discursives, en valeur brute et en occurrences types. Nous rappelons ces mesures qui permettent de calculer le taux de SP en valeur brute, c'est-à-dire qu'on calcule le pourcentage de toutes les SP pour 100 mots, même si elles sont répétées plusieurs fois dans la production (toutes les occurrences) ; et le taux de SP en types, c'est-à-dire qu'on considère non plus toutes les occurrences mais les types de SP (la SP ne sera comptée qu'une seule fois). Ces taux en valeur brute et en occurrences types sont appliqués à la totalité des SP et aux types de SP, lexicales, grammaticales et discursives, selon la terminologie empruntée à Forsberg (2006).

	Mesures de la préfabrication	AN	CA	JE	MG	TH	VC	Moyenne	Ecart-type	min	max
1	Taux syllabes préfabriquées	36,65	37,94	33,78	38,80	32,83	39,63	36,6	2,7	32,83	39,63
2	Taux segments avec SP	90,48	82,81	60,61	72,50	71,43	76,19	75,9	10,2	60,61	90,48
3	Nb syllabes préfabriquées par segment	7,55	5,25	3,44	5,66	5,57	6,10	5,6	1,3	3,44	7,55
4	Nb syllabes préfabriquées par segment avec SP	8,34	6,34	5,68	7,81	7,80	8,00	7,3	1,0	5,68	8,34
5	Nb SP_100mots	15,91	15,71	13,28	16,05	12,87	15,48	14,88	1,4	12,87	16,05
6	Nb SP_100mots_occurrences types	12,08	11,10	10,55	11,93	11,70	13,29	11,77	0,9	10,55	13,29
7	SP lexicales_100mots	9,13	8,65	8,98	7,92	9,36	7,74	8,6	0,7	7,74	9,36
8	SP lexicales_100mots_occurrences types	7,22	6,48	7,03	6,72	8,77	6,35	7,1	0,9	6,35	8,77
9	SP grammaticales_100mots	4,27	3,46	2,15	4,77	2,34	3,97	3,49	1,06	2,15	4,77
10	SP grammaticales_100mots_occurrences types	2,65	2,02	1,95	2,82	2,05	3,77	2,54	0,70	2,02	3,77
11	SP discursives_100mots	2,50	3,60	2,15	3,36	1,17	3,77	2,76	1,01	1,17	3,77
12	SP discursives_100mots_occurrences types	2,21	2,59	1,56	2,39	0,88	3,17	2,13	0,81	0,88	3,17

Tableau 8. Mesures de la préfabrication

On constate que ces différents taux sont plutôt homogènes ; cependant, quelques mesures comme le pourcentage de segments de parole contenant au moins une SP (écart-type de 10,2 pour une moyenne de 75,9) et la longueur moyenne des syllabes préfabriquées par segment (écart-type de 1,3 pour une moyenne de 5,6) présentent des contrastes. Sur ces deux mesures, le sujet AN obtient les résultats maximum (90,48 et 7,55) et le sujet JE les résultats minimum (60,61 et 3,44). Les résultats des séquences préfabriquées produites pour 100 mots dans le groupe entier natif montrent des différences entre les individus aussi bien en valeur absolue qu'en occurrences types. Les sujets s'appuient davantage sur les mêmes SP dans leurs narrations. On passe d'une moyenne de 14,9 SP pour 100 mots en valeur absolue à 11,8 en occurrences types. 8,6 %, des 14,9 %, appartiennent à des SP de type lexical en valeur brute (toutes les SP lexicales) ; sur les 11,8 % des SP en occurrences types, 7,1% sont des SP lexicales. Malgré l'homogénéité du groupe - des locuteurs natifs, d'un même âge, d'un même niveau d'études - on trouve des différences. Il est donc intéressant de regarder de plus près les résultats individuels par rapport à la moyenne du groupe.

Les résultats de MG sont proches de la moyenne pour les quatre premières mesures. Mais si MG utilise le taux de SP le plus élevé du groupe, il passe de 16,05 SP en valeur brute pour cent mots à 11,93 SP en occurrences types. Ce qui signifie que ce sujet va s'appuyer sur un certain nombre de SP qu'elle répète. On trouve chez un autre sujet, le même rapport : CA ne produit que 11,10 SP pour cent mots en occurrences types alors qu'en valeur brute elle en produit 15,71. On peut sans doute conclure que ces sujets utilisent fréquemment les mêmes SP de manière stratégique, peut-être pour parler plus vite.

JE, le sujet au plus bas taux de préfabrication (taux de syllabes préfabriquées et de segments contenant des SP), a un profil très différent de MG. JE semble moins s'appuyer sur le langage préfabriqué que les autres natifs. Cependant, les autres mesures de la préfabrication montrent que le sujet JE s'appuie plutôt sur des SP lexicales variées. Alors que le taux de SP produites tous les cent mots par JE est un des plus bas du groupe (13,28) avec celui de TH (12,87), on constate à l'inverse que le taux de production des SP lexicales, pour ces deux sujets, en valeur brute (8,98 et 9,36), est plus élevé que celui de la moyenne du groupe (8,6). D'autre part, TH qui a le taux de SP pour cent mots le plus bas (12,87), a en revanche un taux de SP en occurrences types très proche (11,70), avec un taux de SP lexicales très haut, ce qui signifie qu'il utilise des SP variées et surtout des SP lexicales. TH comme JE se prêtent peut-être plus à un exercice conventionnel. Selon les sujets, le taux de SP lexicales peut varier. Par exemple, VC et MG produisent moins de séquences lexicales que JE et TH, et plus de séquences

grammaticales et discursives. Le sujet TH produit de nombreuses SP lexicales (9,36) et moins de SP grammaticales et discursives (2,15) que les autres sujets.

En conclusion, les locuteurs natifs, dans l'ensemble, produisent les séquences grammaticales et discursives dans une moindre proportion que les SP lexicales. En effet, la tâche narrative semble favoriser l'emploi de séquences lexicales ou référentielles pour décrire les personnages, les actions, alors que d'autres types de tâches orales - l'interview ou la discussion sur un sujet d'actualité - génèrent une parole moins contrainte, plus variable (Cordier, 2013 ; Grosjean et Deschamps, 1972). Les sujets natifs disposent, sans aucun doute, d'un répertoire de séquences lexicales riche et diversifié qu'ils utilisent pour nommer et décrire. Ils ont un moindre besoin, compte tenu de la tâche, d'employer des SP discursives.

### **6.2.2 Mesures de la fluence**

Sept mesures ont été retenues pour la fluence. Les quatre variables temporelles les plus significatives et utilisées dans la recherche sur la fluence sont la vitesse de parole, la vitesse d'articulation, le taux de phonation, et la longueur du segment de parole en syllabes. Deux autres mesures, la longueur moyenne des pauses et le nombre de pauses par minute, complètent les informations sur les sujets et leur comportement langagier. Enfin une dernière mesure, le nombre de corrections pour cent mots, apporte des informations sur les habitudes des sujets, s'ils se corrigent ou pas, et sur la qualité de la production ; une production orale sans cesse corrigée donne à l'auditeur une impression de disfluence.

La longueur du segment de parole en syllabes (mesure 1) est l'une des mesures classiques de la fluence. Cette mesure associée aux variables temporelles donne un aperçu des habitudes orales des sujets dans leur L1. La longueur moyenne des pauses, variant de 673,5 à 966,70ms, est aussi un indicateur des différences entre les locuteurs. La dernière mesure utilisée, le nombre des reprises pour cent mots, indique la plus ou moins grande fréquence de ce type de disfluence dans un corpus L1. Elle pourrait se trouver en corrélation avec d'autres aspects de la fluence ou correspondre à un type de sujet selon la formation reçue, la plus ou moins grande aisance à l'oral, ou la recherche de la précision dans le discours.

On constate dans le tableau 9 ci-dessous une grande variabilité entre les locuteurs natifs qui ont réalisé les tâches narratives proposées. La mesure 1 (longueur du segment de parole) montre des résultats extrêmes allant du simple au double (AN et JE : 20,6 et 10,18 respectivement), à mettre en relation avec la longueur moyenne du segment de parole du groupe, 15,2 syllabes. Trois sujets produisent des segments d'une longueur à peu près équivalente autour de 15 syllabes (15,4 ; 16,4 ; 14,8), et un peu moins pour un autre (13,8).

Ces résultats posent la question de l'interprétation de cette mesure. Sur les six sujets natifs, de même niveau universitaire, au moins quatre d'entre eux ont des résultats à peu près équivalents, illustrant une sorte de norme de traitement pour ce type de tâche. Cependant deux sujets ont des résultats différents. Cela indique qu'il faudra être vigilant en interprétant les valeurs observées chez les sujets non-natifs, et mettre cette mesure en relation avec d'autres mesures comme le taux de préfabrication.

Mesures	AN	CA	JE	MG	TH	VC	Moyenne	Ecart-type	mini	maxi
1 Long moy du segment (sylls)	20,60	13,80	10,18	14,79	16,43	15,38	15,19	3,4	10,18	20,59
2 Vitesse de parole	4,24	3,41	3,12	4,07	3,77	4,10	3,78	0,44	3,12	4,23
3 Vitesse d'articulation	4,98	4,18	4,01	5,04	4,32	5,11	4,61	0,49	4,01	5,11
4 Taux de phonation	86,57	80,95	73,03	80,83	85,64	80,83	81,31	4,81	73,02	86,57
5 Longueur moy pause	673,48	800,98	966,70	713,96	687,00	749,25	765,2	108,9	673,48	966,70
6 Nb pauses vides minute	11,96	14,27	16,74	16,11	12,54	15,35	14,50	1,93	11,96	16,74
7 Taux de correction	4,42	5,19	7,61	7,37	5,84	1,98	5,40	2,08	1,98	7,61

Tableau 9. Mesures de productivité

A première vue, le sujet AN serait plus « rapide » que les autres sujets natifs : elle produit les plus longs segments de parole du groupe, et les trois mesures - la vitesse de parole, la vitesse d'articulation, le taux de phonation - sont également plus élevées que la moyenne du groupe. AN produit aussi les pauses les plus courtes du groupe. On voit ici un profil particulier de locuteur : des pauses plutôt courtes, peut-être par un moindre besoin de planifier son discours, contrairement à JE dont les pauses nombreuses sont aussi les plus longues. Ces deux sujets illustrent deux comportements langagiers différents.

Ces données confirment les conclusions d'autres chercheurs (voir chapitre 4) : soit on produit des segments de parole plutôt longs avec des pauses plutôt courtes soit c'est l'inverse, on produit des périodes de parole courtes avec de longues pauses (Goldman-Eisler, 1968 : 81). Les comportements langagiers décrits par Goldman-Eisler correspondent aux comportements des deux sujets de mon groupe. AN produit des périodes de parole plutôt longues (les plus longues du groupe) avec de courtes pauses (les plus courtes du groupe). JE produit des segments plutôt courts (les plus courts du groupe) avec de longues pauses qui lui permettent de planifier la suite de son discours.

TH semble avoir un comportement langagier comparable à celui de AN. Il produit des segments de parole plus longs que la moyenne (16,4) avec des pauses plutôt courtes (687ms). AN et TH sont les sujets natifs qui pausent le moins souvent (11,96 et 12,54 pauses par minute) contrairement à JE qui pause plus souvent (16,74 pauses par minute). En revanche, le taux de pauses remplies de TH est élevé, contrairement à celui de JE.

Cette différence entre les sujets réside, peut-être, dans la portée de la planification. On peut se demander à quel moment les sujets qui pausent peu planifient leur discours.

Nous verrons, dans les analyses qui suivent, une éventuelle contribution des SP à ces phénomènes temporels. Les SP permettent de libérer du temps pour la planification qui peut aussi se faire pendant les hésitations (pauses remplies, allongements, reprises).

On se rend compte également que les mesures des variables temporelles ne reflètent pas tout : elles doivent être associées aux mesures de la longueur de segment de parole et des pauses, pour une interprétation plus juste.

La dernière mesure, le taux de reprises ou corrections (pour 100 mots), montre également des écarts extrêmement importants entre les sujets : de 1,98 à 7,61 corrections pour cent mots entre les sujets VC et JE (moyenne du groupe : 5,4). JE pause plus longtemps que les autres et se corrige aussi beaucoup plus. Deux raisons possibles peuvent expliquer ces résultats : JE

cherche à produire un discours précis, d'où une parole moins fluide et des corrections plus nombreuses. En général, si quelqu'un hésite beaucoup, comme JE, c'est qu'il ne veut peut-être pas prendre de risque ; il parle plus lentement pour trouver la formulation la plus juste. Le comportement de JE est sans doute lié à sa façon de formuler les messages qui peut être influencée par sa formation (études de droit) et, peut-être, à la nature de la tâche. Une de nos questions de recherche est de savoir si la fluence est liée à l'emploi de séquences préfabriquées.

### **6.2.3 Corrélation entre la fluence et l'emploi du LP**

Les études sur le langage préfabriqué mettent toutes en avant que l'usage du LP serait un des facteurs déterminants de la fluence des locuteurs à l'oral (chapitre 3). En L1, les locuteurs ont un répertoire étendu et diversifié de séquences qui leur permettent de s'exprimer aisément. Ils libèrent ainsi des ressources attentionnelles (chapitre 2) pour la gestion et la co-construction du sens. En L2, les études révèlent que l'emploi des SP favorise une parole plus fluide non seulement en termes de rapidité mais aussi d'aisance (chapitre 3).

Pouvons-nous trouver dans notre groupe contrôle une corrélation entre la fluence et l'usage des SP ? Les sujets les plus rapides en termes de fluence, c'est-à-dire AN, VC et MG sont, en effet, les sujets qui obtiennent des résultats plus élevés sur les quatre mesures de la préfabrication selon Cordier. Ils produisent plus de syllabes préfabriquées et de segments avec SP. En outre, la longueur moyenne des syllabes préfabriquées par segment avec ou sans SP est plus longue que la moyenne du groupe. Et ils sont les plus « fluents ». En revanche, TH et surtout JE, les deux sujets les moins fluents, produisent le moins de syllabes préfabriquées du groupe. Enfin, JE a les résultats les plus bas sur les quatre premières mesures de la préfabrication.

On retrouve les mêmes écarts dans les mesures des SP pour cent mots produits. Là encore, ce sont JE et TH qui produisent le moins de SP du groupe (13,3 et 12,9 en valeur brute pour une moyenne du groupe de 14,9). Mais on a vu aussi que ces deux sujets produisent un taux de SP lexicales pour cent mots plus élevé que la moyenne du groupe et que TH produit un nombre élevé de ces SP en occurrences types (section 6.2.1), ce qui semble confirmer la tendance de ces deux sujets à vouloir produire un discours plus précis et plus riche en termes de variété lexicale puisqu'ils emploient surtout des séquences de ce type et plus diversifiées.



## 6.3 Comparaison entre le groupe natif et le groupe non-natif

### 6.3.1 Mesures de productivité de la préfabrication

Le tableau 10 présente les données du groupe des locuteurs natifs et du groupe des non-natifs en temps 2.

Mesures (t2)	LN N =6	LNN (tous) N =17
Nb mots, BD + Film (t2)	3653	9983
Nb syllabes, BD + film (t2)	4709	12607
Nb SP	550	1346
Taux de syllabes préfabriquées	0,37	0,33 (-10,81%)
Taux de segments avec SP	0,76	0,50 (-34,21%)
Nb syllabes préfabriquées par segment	5,57	2,6 (-53,32%)
Nb syllabes préfabriquées par segment avec SP	7,28	5 (-31,32%)
Nombre SP_100mots	14,9	13,1
Nombre SP_100mots_occurrences types	11,8	9,2 (-29,77%)
SPlexicales_100mots	8,6	5,9
SPlexicales_100mots_occurrences types	7,1	4,3

Tableau 10. Mesures de productivité et préfabrication, groupe natif et non-natif

Les résultats ci-dessus indiquent des différences entre le groupe natif et le groupe non-natif en entier, sur toutes les mesures que ce soit le nombre de syllabes préfabriquées, le nombre de segments contenant des SP, la longueur des syllabes préfabriquées par segment de parole avec ou sans SP ou le taux de préfabrication pour cent mots en valeur brute ou en occurrences types, toutes SP confondues ou les SP de type lexical en particulier. Ces écarts entre les deux groupes étaient prévisibles car nous avons un corpus de productions natives d'un côté et de l'autre un corpus de productions non-natives de niveau débutant à avancé. Les chiffres entre parenthèses indiquent le différentiel en pourcentage des résultats du groupe expérimental par rapport au groupe contrôle.

N – NN-tous	SP lex	SP gramm	SP disc
Non-natifs (Bd2+Film2)	658	418	345
Natifs (Bd+Film)	312	136	107

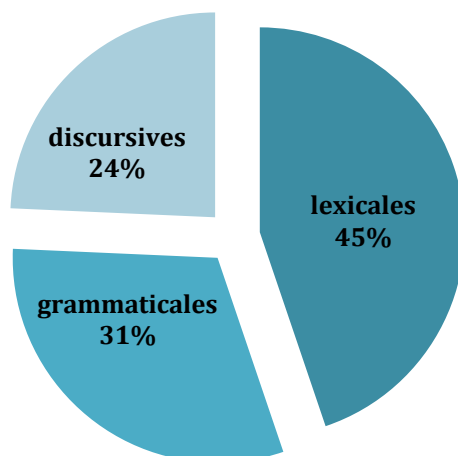
Tableau 11. Répartition des SP, groupe natif et non-natif (n=17)

La répartition des types de SP entre les deux groupes illustrée dans le tableau 11 devrait également nous donner une indication intéressante ; le test khi-deux, en effet, montre une différence significative dans la distribution par type de SP,  $\chi^2(2, 1976) = 15,78$  ( $p < 0,001$ )

La figure 3 ci-dessous montre la distribution des SP en t2 dans le groupe non-natif. On constate une distribution différente de celle des natifs (page 168), moins de SP lexicales et plus de SP grammaticales et discursives mais la proportion entre les types reste sensiblement la même.

Figure 3. Distribution des SP groupe non-natif

### Distribution des types de SP, corpus non natif t2



On remarque une forte proportion de SP lexicales (45% vs 56% chez les natifs) ; de un tiers à un quart de SP grammaticales (31% vs 25% chez les natifs) et des discursives dans une moindre proportion (24% vs 19% chez les natifs), les discursives étant les moins fréquentes (comme chez les locuteurs natifs).

Nous pouvons aussi comparer les résultats du groupe natif à ceux des trois sous-groupes non-natifs. Ces résultats sont donnés dans le tableau 12 suivant.

<b>Mesures (t2)</b>	LN N =6	LNN groupe 1 N =6	LNN groupe 2 N =5	LNN groupe 3 N =6
Nb mots, BD + Film (t2)	3653	3695	2491	3797
Nb syllabes, BD + film (t2)	4709	4736	3107	4764
Nb syllabes des SP	550	423	346	577
Taux de syllabes préfabriquées	0,37	0,29	0,32	0,38
Taux de segments avec SP	0,76	0,40	0,49	0,60
Nb syllabes préfabriquées par segment	5,57	1,9	2,4	3,6
Nb syllabes préfabriquées par segment avec SP	7,28	4,5	4,8	5,6
Nb SP_100mots	14,9	11,7	12,7	15,1
Nb SP_100mots occurrences types	11,8 (-20,13%)	7,25 (-38,03%)	9,7 (-23,62%)	10,8 (-28,48%)
SPlexicales_100mots	8,6	5,5	5	7
SPlexicales_100mots, occurrences types	7,1	3,32	4,1	5,3

Tableau 12. Mesures productivité et préfabrication, natifs et sous-groupes non-natifs

On constate dans ce tableau que les résultats du groupe avancé (groupe 3) tendent vers ceux du groupe natif. Les écarts ne sont pas aussi importants, quelques résultats sont même légèrement supérieurs dans le groupe 3 comme le pourcentage de syllabes préfabriquées (38%) et le nombre de SP pour cent mots produits (15,1 vs 14,9). Ce qui diffère en revanche, ce sont les taux de SP pour cent mots en occurrences types. Il apparaît que les natifs produisent une plus grande variété de SP alors que les non-natifs répètent plus souvent les mêmes. Les chiffres entre parenthèses indiquent une perte en pourcentage entre toutes les SP produites (même plusieurs fois) et les SP produites une seule fois. 20,13% des SP sont produites plusieurs fois par les natifs et près de 30% par le groupe non-natif entier. L'écart est d'autant plus grand quand le sous-groupe est débutant (38% pour le groupe débutant). D'autre part, les résultats des autres mesures comme le pourcentage des segments avec SP et la longueur des syllabes préfabriquées par segment restent inférieurs à ceux du groupe natif ainsi que la mesure du taux des SP en occurrences types, les non-natifs utilisant souvent les mêmes SP dans une même production. Le taux des SP lexicales, aussi bien en valeur brute (7) qu'en occurrences types (5,3), du groupe avancé non-natif reste inférieur à celui des natifs (-18,60%). Certains sujets du groupe avancé obtiennent aussi des résultats très proches de ceux du groupe natif. Les chiffres entre parenthèses indiquent en pourcentage le différentiel entre le

nombre de SP pour cent mots et le nombre de SP en types. La question de la variabilité constatée chez les natifs se pose aussi pour les locuteurs L2. Il faudra rester vigilant dans nos interprétations, car cette variabilité influencera certainement les productions et le développement longitudinal de nos apprenants.

### 6.3.2 Mesures de productivité de la fluence des locuteurs natifs et non-natifs

Le tableau 13 présente les mesures temporelles de productivité. Il donne les moyennes du groupe natif et du groupe entier des non-natifs.

<b>Mesures (t2)</b>	LN (n=6)	LNN (n=17)
Nombre de syllabes du segment	15,19	7,65
Vitesse de parole en secondes	3,78	2,68
Vitesse d'articulation en secondes	4,61	3,7
Taux de phonation	81,31	72,17
Longueur moyenne des pauses (ms)	765,23	780,09
Taux d'hésitation	18,69	27,83
Nombre de pauses par minute	14,49	21,60
Taux de correction	5,40	9,23

Tableau 13. Mesures temporelles de productivité, groupe natif et non-natif

Si l'on prend les moyennes natives comme « étalon », on constate, comme on pouvait s'y attendre, que les résultats du groupe entier non-natif sont inférieurs à ceux du groupe natif sur toutes les mesures, excepté la mesure de la longueur moyenne de la pause qui ne montre pas de différence significative. Cette mesure devra donc être traitée avec prudence. Toutes les autres mesures montrent des écarts entre locuteurs natifs et non-natifs.

<b>Mesures (t2)</b>	LN (n=6)	LNN-1 (n=6)	LNN-2 (n=5)	LNN-3 (n=6)
Nombre syllabes du segment	15,19	6,25	7,53	9,13
vitesse de parole en secondes	3,78	2,43	2,82	2,81
Vitesse d'articulation en secondes	4,61	3,53	3,96	3,65
Taux de phonation en %	81,31	68,50	71,23	76,62
Longueur moyenne pauses (ms)	765,23	829,97	800,15	713,48
Taux d'hésitation	18,69	31,50	28,77	23,38
Nombre de pauses vides par minute	14,49	23,17	22,17	19,57
Taux de correction	5,40	12,04	7,84	7,75

#### Tableau 14. Mesures temporelles de productivité, natifs et sous-groupes non-natifs

Le tableau 14 ci-dessus présente les résultats du groupe natif et des sous-groupes non-natifs. Les résultats des mesures de la fluence montrent des différences entre le groupe natif et les sous-groupes non-natifs, quel que soit le niveau. On aurait pu s'attendre à ce que le groupe avancé ait des résultats très proches de ceux du groupe natif, mais ce n'est pas le cas. La moyenne de la longueur du segment de parole (15,19) des natifs est bien supérieure à celle du groupe avancé (9,13). Les locuteurs du groupe avancé produisent des segments de parole nettement moins longs, près de 40% en moins (-39,90% très précisément). Les résultats de notre groupe avancé sont d'ailleurs en concordance avec ceux de Forsberg-Lundell *et al.* (2014), qui montrent que les locuteurs de niveau avancé (niveau 6 sur les huit niveaux établis par ces chercheurs) n'atteignent pas le niveau natif sur certaines mesures, notamment la longueur moyenne des segments de parole. Cette mesure, considérée comme la mesure phare de la fluence, serait donc la plus représentative des différences entre locuteurs natifs et non-natifs. Les trois mesures suivantes, vitesse de parole, vitesse d'articulation et taux de phonation, sont également inférieures à celles du groupe natif. Seule la mesure de la longueur moyenne de la pause, dans ce groupe des locuteurs non-natifs, est légèrement inférieure à celle du groupe natif (713,48 vs 765,23). Cependant cette mesure, on l'a souligné précédemment, doit être interprétée avec prudence. Chaque locuteur natif a sa propre vitesse de parole et certains pausent plus longtemps que d'autres, ce qui devrait être aussi le cas des locuteurs L2. Cependant, malgré la durée moyenne assez élevée de la pause, le groupe natif passe plus de temps à parler qu'à pauser, comme on pouvait s'y attendre. On constate sans surprise que chaque groupe de locuteurs non-natifs a un taux de phonation correspondant à son niveau (68,50 ; 71,23 ; 76,62 respectivement).

La mesure du nombre de corrections pour cent mots indique que les non-natifs se corrigent plus que les natifs pour toutes les raisons déjà évoquées : hésitations dans les traitements lexicaux et morphosyntaxiques. Au chapitre 3, nous avons vu que l'automatisme des encodages formels permet la fluidité en L1 ; en L2, le manque d'automatisme est certainement l'une des causes de ce manque de fluidité. A un niveau avancé, les hésitations et corrections peuvent aussi être le signe d'un souci de précision. On constate des différences entre les deux groupes (natif et non-natif entier) mais aussi entre le groupe natif et le groupe non-natif avancé dont les résultats auraient dû être proches de ceux du groupe natif. Cependant il apparaît nécessaire de vérifier au niveau individuel si les mêmes écarts existent dans ce groupe avancé.

## 6.4. Résultats des tests statistiques

### 6.4.1 Mesures de la préfabrication

Un premier test : Mann-Whitney

Afin de déterminer si ces différences entre les groupes sont significatives, nous avons appliqué le test U de Mann-Whitney pour les mesures de la préfabrication a) au groupe natif et au groupe non-natif entier, b) au groupe natif et au groupe avancé.

#### a) groupe natif et groupe non-natif entier

Nous présentons les comparaisons des résultats des deux groupes entiers pour les mesures de la préfabrication dans le tableau 15 suivant.

	Mesures	Comparaisons natifs et non-natifs groupe entier		
		U	Valeur de p	
1	Taux syll préfabriquées	U(6,17)=0	<0,001**	N>NN
2	Taux segments avec SP	U(6,17)=0	<0,001**	N>NN
3	Nb syll pref par segment	U(6,17)=8	<0,01*	N>NN
4	Nb syll préf par segment avec SP	U(6,17)=8	<0,01*	N>NN
5	Nb SP_100mots	U(6,17)=25	>0,05	N=NN
6	Nb SP_100mots occurrences types	U(6,17)=13	<0,01*	N>NN
7	SPlexicales_100mots	U (6,17)=3	<0,001**	N>NN
8	SPlexicales_100mots occurrences types	U(6,17)=13	<0,01*	N>NN

Tableau 15. Comparaisons Mann-Whitney, mesures de préfabrication

Ici nous constatons que toutes les différences entre les groupes natif et non-natif dans les taux de production sont significatives (tous les apprenants pris ensemble) pour toutes les mesures de la préfabrication, sauf pour la mesure du taux de SP pour cent mots (ligne 5), où la différence est non-significative. En revanche, la mesure du taux de SP pour cent mots en occurrences types présente une différence significative ( $p < 0,01^*$ ). La mesure du taux de SP lexicales pour cent mots présente une différence très significative ( $< 0,001^{**}$ ) en valeur brute et significative en occurrences types ( $< 0,01^*$ ). Ces mesures seraient donc les plus pertinentes pour nos analyses. Elles montrent que les locuteurs non-natifs répètent les mêmes SP dans leurs productions, peut-être par manque de connaissances lexicales mais peut-être aussi qu'ils emploient celles qui leur sont facilement accessibles dans une tâche non planifiée et en temps réel pour éviter un discours disfluent. Les locuteurs natifs, en revanche, disposent d'un répertoire de séquences plus diversifié.

b) groupe natif et groupe non-natif avancé

Mesures		Comparaisons natifs et groupe non-natif avancé		
		U	Valeur de p	
1	Taux syllabes préfabriquées	U(6,6)=0	<0,01*	N>NN
2	Taux segments avec SP	U(6,6)=0	<0,01*	N>NN
3	Nb syllabes préfabriquées par segment	U(6,6)=7	=0,08	NS
4	Nb syllabes préfabriquées par segment avec SP	U(6,6)=7	=0,08	NS
5	Nb SP_100mots	U(6,6)=18	=0,99	NS
6	Nb SP_100mots_occurrences types	U(6,6)=8	=0,11	NS
7	SPLexicales_100mots	U(6,6)=3	=0,02	(m)
8	SPLexicales_100mots_occurrences types	U(6,6)=1	<0,01*	N>NN

Tableau 16. comparaisons Mann-Whitney, groupe natif et groupe avancé non-natif

Le tableau 16 indique les résultats des tests statistiques entre le groupe natif et le groupe avancé non-natif. On s'aperçoit que les deux premières mesures de la préfabrication présentent des différences significatives ( $p < 0,01^*$ ) (lignes 1 et 2, taux de syllabes préfabriquées et taux de segments de parole contenant des SP) ; la différence entre le groupe natif et le groupe avancé ( $p < 0,01^*$ ) est cependant moins marquée qu'entre le groupe natif et le groupe entier ( $p < 0,001^{**}$ ). Concernant les mesures 3 et 4, le nombre de syllabes préfabriquées par segment et le nombre de syllabes préfabriquées par segment contenant des SP, on ne constate aucune différence significative entre le groupe natif et le sous-groupe non-natif avancé. Il en est de même pour les mesures 5 et 6, taux de préfabrication pour 100 mots produits en valeur brute ou en occurrences types. Seule la mesure du nombre de SP lexicales pour 100 mots produits, en valeur absolue et *a fortiori* en occurrences types, présente une différence significative entre les deux groupes. Il semble, en effet, que la différence est plus marquée dans ce corpus de productions orales. Si les locuteurs du groupe L2 avancé tendent à se rapprocher des natifs sur quelques mesures (longueur des syllabes préfabriquées dans le segment avec ou sans SP), ils n'arrivent pas à un taux équivalent de syllabes préfabriquées ni de SP lexicales. Sur la dernière mesure, le nombre de SP lexicales en occurrences types, le groupe avancé tout comme le groupe entier reste en dehors des valeurs natives.

### 6.4.2 Résultats statistiques des mesures temporelles de productivité

Les tableaux 17 et 18 complètent statistiquement les comparaisons des mesures temporelles de productivité du groupe des natifs avec celles du groupe entier non-natif.

a) groupe natif et groupe L2 entier :

		<b>Comparaison natifs et non-natifs, groupe entier</b>	
	Mesures	U	valeur de p
1	Longueur du segment	U(6,17)=5	<0,01*
2	Vitesse de parole	U(6,17)=4	=0,001**
3	Vitesse d'articulation	U(6,17)=9	<0,01*
4	Taux de phonation	U(6,17)=9	<0,01*
5	Longueur moyenne pause	U(6,17)=48	>0,8 (NS)
6	Taux d'hésitation	U(6,17)=9	<0,01*
7	Nombre de pauses minute	U(6,17)=6	=0,001**
8	Taux de correction	U(6,17)= 17	=0,02(m)

Tableau 17. Résultats des tests statistiques des mesures temporelles de productivité

Les tests statistiques réalisés (groupe natif et groupe non-natif entier) montrent des différences significatives pour la mesure de la longueur du segment de parole ainsi que pour les variables temporelles telles que la vitesse de parole, la vitesse d'articulation, le taux de phonation et le nombre de pauses par minute. Nous ne relevons aucune différence significative en revanche sur la longueur moyenne des pauses, ce qui semble confirmer notre commentaire (section 6.3.2) sur les pauses dans un corpus natif et non-natif. La principale différence entre les natifs et les non-natifs se situerait surtout au niveau de l'emplacement / localisation des pauses dans les productions orales (Goldman-Eisler, 1968 : 95 ; Grosjean et Deschamps, 1972 ; Hilton, 2008a).

b) groupe natif et groupe non-natif avancé :

		<b>Comparaison natifs et groupe avancé</b>	
	Mesures	U	valeur de p
1	Longueur du segment	U(6,6)=5	<0,05 (m)
2	Vitesse de parole	U(6,6)=2	=0,01*
3	Vitesse d'articulation	U(6,6)=3	<0,02 (m)



4	Taux de phonation	U(6,6)=9	=0,1495 (NS)
5	Long moyenne pause	U(6,6)=15	=0,6310 (NS)
6	Taux d'hésitation	U(6,6)=9	=0,1495 (NS)
7	Nombre de pauses minute	U(6,6)=7	=0,05 (NS)
8	Taux de correction	U(6,6)=7	>0,05 (NS)

Tableau 18. Résultats des tests statistiques des mesures temporelles de productivité

De toutes les mesures de productivité de la fluence, seule la vitesse de parole exprimée en syllabes par seconde présente une différence significative ( $p=0,01^*$ ) ; la vitesse d'articulation et la mesure du segment de parole en syllabes présentent une différence seulement marginale ( $p<0,02$ ). Les autres mesures temporelles ne présentent aucune différence significative. La différence entre locuteurs natifs et non-natifs de niveau avancé (groupe contrôle et groupe avancé) se situe donc au niveau de la vitesse de parole (nombre de syllabes produites par seconde) et dans une moindre mesure dans la vitesse d'articulation. Ces résultats statistiques indiquent que le groupe avancé tend vers le niveau natif, en termes d'aisance (fluidité du discours).

## 6.5. Corrélations entre fluence et préfabrication

Dans cette section, nous tenterons de répondre à l'une de nos questions de recherche :

- y a-t-il une corrélation entre l'usage des SP et la fluence ?

Toutes les recherches entreprises sur le langage préfabriqué ont montré que l'usage des SP améliore la fluence des locuteurs, que ce soit en L1 ou en L2 (chapitre 2). Foster et Skehan (1996) ont trouvé que des locuteurs non-natifs étaient capables de produire un discours plus fluide dans une tâche de discussion, après un temps de planification. Ils sont en outre capables, en le planifiant, de produire un discours plus riche et plus complexe. Foster (2001) montre cependant que les locuteurs natifs sont plus fluent que les non-natifs, que la tâche soit planifiée ou pas. En L2, une autre étude (Ellis et Fangyuan, 2004) montre que dans une tâche planifiée (tâche narrative en l'occurrence) les locuteurs L2 sont non seulement plus « fluent » mais aussi plus précis, ce qui corrobore les résultats de l'étude de Foster et Skehan (1996). L'étude de Cordier (2013) arrive aux mêmes conclusions, après que cinq sujets anglophones apprenant le français ont été soumis à plusieurs types de tâches : les locuteurs non-natifs produisent des segments de parole plus longs en t2 grâce à l'emploi des SP, qui se développe après un séjour dans le pays de la langue cible.

Nous avons cherché à savoir s'il existe une corrélation entre la préfabrication et la fluence en appliquant le test non paramétrique de Spearman, d'abord au groupe entier des sujets de cette étude (locuteurs natifs et non-natifs) puis au groupe des locuteurs non-natifs entier. Les corrélations par sous-groupes se sont avérées non significatives (sans doute le nombre des sujets est-il trop petit pour qu'un éventuel effet soit mesurable).

Nous avons choisi les quatre mesures de fluence les plus significatives : la longueur du segment de parole, la vitesse de parole, la vitesse d'articulation et le taux de phonation. Chacune de ces mesures a été corrélée à une mesure de la préfabrication : le taux de syllabes préfabriquées, le nombre de syllabes préfabriquées par segment de parole, le taux de SP pour cent mots et le taux de SP lexicales pour cent mots.

### **6.5.1 Groupe natif et groupe non-natif entier (t2)**

- a) Il existe une corrélation forte à très forte entre la longueur du segment de parole et
- le taux de syllabes préfabriquées  $r = ,677$  ( $p < 0,0015$ )
  - le nombre moyen de syllabes préfabriquées par segment  $r = ,831$  ( $p < 0,0001$ )
  - le nombre moyen de syllabes préfabriquées par segment contenant au moins une SP  $r = ,865$  ( $p < 0,0001$ )
  - le taux de SP ( $r = ,544$   $p = 0,01$ )
  - le taux de SP lexicales  $r = ,746$  ( $p < 0,001$ )
- a) il existe une corrélation forte entre la vitesse de parole et
- le taux de syllabes préfabriquées :  $r = ,665$  ( $p = 0,0018$ )
  - le nombre moyen de syllabes préfabriquées par segment  $r = ,717$  ( $p < 0,001$ )
  - le nombre de syllabes préfabriquées par segment contenant au moins une SP ( $r = ,788$   $p = 0,0002$ )
  - le taux de SP lexicales  $r = ,687$  ( $p = 0,0013$ )
- b) il existe une corrélation entre la vitesse d'articulation et
- le taux de syllabes préfabriquées  $r = ,622$  ( $p = 0,0035$ )
  - le taux de segments avec SP  $r = ,706$  ( $p < 0,001$ )
  - le taux de SP  $r = ,456$  ( $p = 0,0324$ ), effet marginal seulement pour cette corrélation
- c) il existe une corrélation entre le taux de phonation et
- le taux de syllabes préfabriquées  $r = ,692$  ( $p = 0,0012$ )
  - le nombre de syllabes préfabriquées par segment  $r = ,811$  ( $p < 0,0001$ )
  - le nombre de syllabes préfabriquées par segment avec SP  $r = ,706$  ( $p < 0,001$ )

Les résultats de ces tests montrent qu'il existe une corrélation forte à très forte entre l'usage des SP et la fluence pour l'ensemble des sujets (natifs et non-natifs)

### 6.5.2 Groupe des locuteurs non-natifs, t1 et t2

Nous avons appliqué le test de Spearman au groupe des locuteurs non-natifs (n=17) et avons trouvé une corrélation forte entre l'usage des SP et la fluence en t1, mais en t2 les résultats sont pour une grande partie marginaux ou non significatifs.

Regardons ce qui se passe dans le groupe des locuteurs non-natifs en entier en t1 et en t2

a) entre la longueur moyenne des segments de parole en syllabes et

- le taux de syllabes préfabriquées t1 ( $r=,635$   $p=0,01$ ) et en t2 ( $r=,297$   $p>0,5$ )
- le nombre de syllabes par segment t1 ( $r=,909$   $p<0,001$ ) et en t2 ( $r=,598$   $p=0,0167$ )
- le nombre de syllabes par segment avec SP t1 ( $r=,914$   $p<0,001$ ) et en t2 ( $r=,683$   $p<0,01$ )
- le taux de SP pour cent mots t1 ( $r=,640$   $p=0,01$ ) et en t2 ( $r=,365$   $p>0,5$ )

b) entre la vitesse de parole et

- le taux de syllabes préfabriquées, t1 ( $r=,669$   $p<0,01$ ) et en t2 ( $r=,297$   $p>0,5$ )
- le nombre de syllabes préfabriquées par segment t1 ( $r=,904$   $p<0,001$ ) et en t2 ( $r=,402$   $p>0,5$ )
- le nombre de syllabes préfabriquées par segment avec SP t1 ( $r=,877$   $p<0,001$ ) et en t2 ( $r=,578$   $p=0,02$ )
- le taux de SP pour cent mots en t1 ( $r=,642$   $p=0,01$ ) et t2 ( $r=,289$   $p>0,5$ )
- le taux de SP lexicales pour cent mots en t1 ( $r=,664$   $p<0,01$ ) et t2 ( $r=,414$   $p>0,05$ )

c) entre la vitesse d'articulation et

- le taux de syllabes préfabriquées t1 ( $r=,499$   $p<0,05$ ) et t2 ( $r=,271$   $p>0,5$ )
- le nombre de syllabes préfabriquées par segment t1 ( $r=,656$   $p<0,01$ ) et t2 ( $r=,286$   $p>0,25$ )

d) entre le taux de phonation et

- le taux de syllabes préfabriquées t1 ( $r=,564$   $p=0,024$ ) et t2 ( $r=,392$   $p>0,5$ )
- le nombre de syllabes préfabriquées par segment t1 ( $r=,860$   $p<0,001$ ) et t2 ( $r=,598$   $p<0,02$ )
- le taux de SP pour cent mots en t1 ( $r=,627$   $p=0,01$ ) et t2 ( $r=,515$   $p<0,04$ )
- le taux de SP lexicales pour cent mots en t1 ( $r=,694$   $p<0,01$ ) et t2 ( $r=,549$   $p<0,03$ )

Comment interpréter ces résultats du groupe non-natif ? En t1, on constate des corrélations que l'on ne retrouve pas en t2. Doit-on expliquer ce changement en t2 par la nature de la tâche ? La tâche narrative pourrait avoir été vécue la deuxième comme une épreuve de précision. En t2, les sujets ont voulu mieux faire dans l'ensemble en utilisant un vocabulaire plus précis, plus juste et de ce fait ont produit une quantité de SP moins importante que celle à laquelle nous nous attendions. On sait que lorsqu'une tâche est planifiée, les locuteurs s'appuient moins sur le LP, ce qui pourrait être une première tentative d'explication (Foster, 2001).

Nous allons donc examiner les résultats du groupe non-natif pour tenter de comprendre ces résultats.



# Chapitre 7 Présentation des résultats du groupe non-natif

---

Dans ce chapitre, nous présentons les résultats du groupe non-natif entier et des sous-groupes, ainsi que les résultats individuels à l'intérieur de chaque sous-groupe. Ensuite, nous examinerons attentivement les résultats de la préfabrication et de la fluence pour voir s'il y a effectivement une progression entre t1 et t2, et sur quelles mesures en particulier.

## 7.1. Résultats du groupe des non-natifs, préfabrication

### 7.1.1 Répartition des SP, groupe entier et sous-groupes, t1 et t2

Avant toute analyse des mesures de la préfabrication, nous présentons dans le tableau 19, comme nous l'avons fait pour le groupe natif, les données brutes concernant les SP par types produites par le groupe entier, puis par les sous-groupes. Nous appliquerons le test khi-deux pour comparer la distribution des SP dans les deux tâches du corpus non-natif. Ensuite, nous présenterons les analyses statistiques qui devraient indiquer s'il y a eu progression entre le premier enregistrement et le second, six mois plus tard.

#### a) Groupe entier

corpus NN-tous	SP lexicales	SP grammaticales	SP discursives
NN-tous t1 (Bd + Film)	432	373	338
NN-tous t2 (Bd + Film)	658	418	328

Tableau 19. Répartition des types de SP du groupe non-natif entier

Dans le tableau 9, le test khi-deux indique que la distribution des différents types de SP est différente aux deux moments de production (novembre à juin) :  $\chi^2(2, 2564) = 19,6$  ( $p < 0,0001$ ). Le nombre de SP lexicales a fortement augmenté alors que le nombre des SP discursives et celui des SP grammaticales a diminué en proportion après six mois, ce que les figures 4 et 5 montrent de façon très nette, en t1 (Figure 4) et en t2 (Figure 5).

Figure 4. Distribution des SP, corpus non-natif, t1

### Distribution des SP corpus non natif (t1)

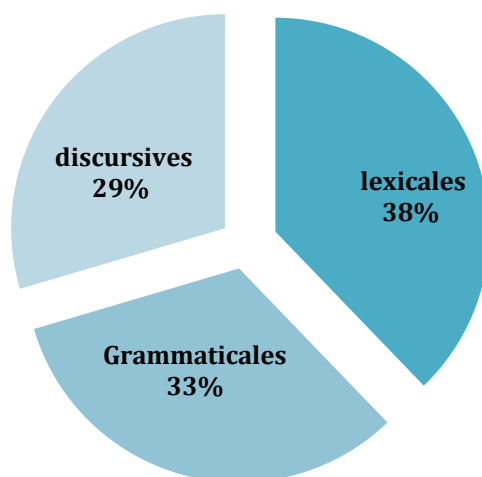
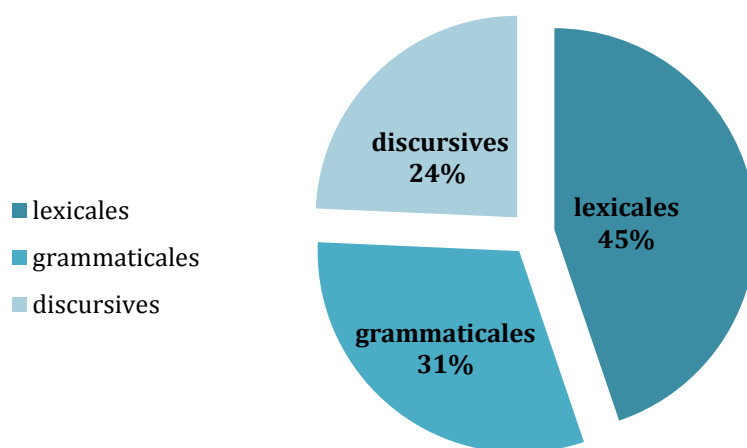


Figure 5. Distribution des SP, corpus non-natif, t2

### Distribution des types de SP, corpus non natif (t2)



La distribution des SP entre t1 et t2 évolue vers une plus grande quantité de SP lexicales (+7%) indiquant une amélioration du lexique référentiel. La légère baisse des SP grammaticales (-2%) et surtout celle des SP discursives (-5%) semblent confirmer que les apprenants ont

moins besoin de s'appuyer sur un certain nombre d'entre elles. Les résultats semblent donc en conformité avec une de nos hypothèses, à savoir une distribution des SP qui évolue dans le temps.

*b) Entre les groupes*

Cp sous-groupes t1	SP lexicales	SP grammaticales	SP discursives
NN_déb	106	132	110
NN_int	97	92	97
NN_av	229	149	131

Tableau 20. Nombre de SP par types et par sous-groupes, t1

Le tableau 20 ci-dessus présente le nombre de SP par types et pour chaque groupe en t1. La différence dans la distribution des trois types des SP entre les trois groupes d'apprenants est très significative, en t1 :  $\chi^2(4, 1143) = 22,7$  ( $p < 0,0001$ ).

Cp sous-groupes t2	SP lexicales	SP grammaticales	SP discursives
NN_déb	197	121	105
NN_int	126	102	85
NN_av	273	166	138

Tableau 21. Nombre de SP par types et par sous-groupes, t2

Le tableau 21 présente la distribution des SP produites par types et par sous-groupe en juin (t2). Cette fois, le test khi-deux ne montre pas, cette fois, de différence en t2 :  $\chi^2(4, 1313) = 4,50$  ( $p = 0,34$ ). Il y a donc une répartition plus homogène des types de SP entre les différents groupes de niveau, la deuxième fois. C'est-à-dire qu'après six mois, les apprenants semblent faire une utilisation semblable du langage préfabriqué (ce que nous regarderons de plus près dans le chapitre 8).

*c) Sous-groupe, niveau débutant*

Groupe déb	SP lexicales	SP grammaticales	SP discursives
NN déb t1	106	132	110
NN déb t2	197	121	105

Tableau 22. Répartition par type de SP, sous-groupe débutant

Si on compare les différents types de SP produites aux deux moments de production par les membres d'un même groupe, on constate un net changement chez les apprenants « débutants », entre novembre et juin :  $\chi^2(2, 771) = 20,8$  ( $p < 0,0001$ ). Les analyses *post hoc*



(taux de liaison) montrent que les débutants produisent plus de SP grammaticales et discursives en novembre, et plus de SP lexicales en juin.

*d) Sous-groupe, niveau intermédiaire*

Groupe int	SP lexicales	SP grammaticales	SP discursives
NN int t1	97	92	97
NN int t2	126	102	85

Tableau 23. Répartition par types de SP, sous-groupe intermédiaire

La proportion des SP produites par les apprenants du groupe intermédiaire (tableau 23) montre en t1 une distribution quasi-identique des trois types de SP. En t2, malgré une augmentation des SP lexicales et grammaticales, cette différence n'est pas significative :  $\chi^2(2, 599) = 3,87$  ( $p=0,14$  ns).

*e) Sous-groupe, niveau avancé*

Groupe NN	SP lexicales	SP grammaticales	SP discursives
NN av t1	229	149	131
NN av t2	273	166	138

Tableau 24. Répartition par types de SP, sous-groupe avancé

La répartition des trois types SP produites par les apprenants du groupe avancé (tableau 24) ne change pas entre t1 et t2 =  $\chi^2(2, 1086) = 0,701$  ( $p= 0,70$  ns).

Ces tests montrent que la différence, dans la répartition des types de SP (lexicales, grammaticales et discursives) aux deux moments de production, est seulement significative dans le groupe débutant. En t1, les apprenants débutants produisent plus de SP grammaticales et discursives, ils semblent s'appuyer sur celles-ci ; par exemple, ils produisent un grand nombre de constructions avec *c'est* et des SP discursives comme *et après, c'est*, mais en t2 la tendance s'inverse, leur répertoire de SP lexicales a augmenté et s'est enrichi et le nombre de SP discursives a diminué.

*e) Répartition des SP en t2 par sujet et par sous-groupe*

Des tests khi-deux, appliqués aux types de SP par sujet dans les sous-groupes, nous indiqueront si la répartition de ces types varie d'un individu à l'autre. Ces tests nous montrent que :

- i. l'utilisation des trois types de SP varie fortement pour les membres du groupe débutant en t1 :  $\chi^2(10, 348) = 33,2$  ( $p < 0,0001$ ) comme en t2 :  $\chi^2(10, 423) = 37,1$  ( $p < 0,0001$ ).

<b>NN deb t1</b>	<b>SP lexicales</b>	<b>SP grammaticales</b>	<b>SP discursives</b>
AK	12	24	37
JA	14	28	16
NA	8	13	10
OL	28	41	27
SU	15	12	7
XI	29	14	13

Tableau 25. Nb et types de SP, par sujet , niv 1, t1

En t1, les apprenants produisent plus de SP grammaticales et discursives dans l'ensemble que les SP lexicales. En t2, le rapport s'inverse mais la répartition entre les individus varie fortement.

<b>NN deb t2</b>	<b>SP lexicales</b>	<b>SP grammaticales</b>	<b>SP discursives</b>
AK	25	22	40
JA	36	16	14
NA	31	25	9
OL	34	26	18
SU	34	11	8
XI	37	21	16

Tableau 26. Nb et types de SP par sujet, niv1, t1

- ii. L'utilisation des trois types de SP ne varie pas entre les individus du groupe intermédiaire en t1 :  $\chi^2(8, 286) = 10,69$  ( $p=0,22$ ) comme en t2 :  $\chi^2(8, 322) = 6,21$  ( $p=0,62$  ns). Cette différence n'est pas significative.

<b>NN int t1</b>	<b>SP lexicales</b>	<b>SP grammaticales</b>	<b>SP discursives</b>
HE	20	17	17
MA	26	22	38
MR	18	17	7

NO	17	21	19
ZI	16	15	16

Tableau 27. Nb et types de SP par sujet, niv 2, t1

Que ce soit en t1 comme en t2, la répartition des trois types de SP ne varie pas entre les individus du groupe intermédiaire. Pourtant on constate que pour NO et ZI, cette répartition (en proportion) se rapproche de celle des natifs au t2.

<b>NN int t2</b>	<b>SP lexicales</b>	<b>SP grammaticales</b>	<b>SP discursives</b>
HE	26	16	21
MA	17	14	16
MR	13	19	12
NO	29	24	16
ZI	41	29	20

Tableau 28. Nb et types de SP par sujet, niv2, t2

- iii. L'utilisation des trois types de SP ne varie pas entre les individus du groupe avancé en t1 :  $\chi^2(10, 509) = 9,31$  ( $p=0,50$ ) ; et la différence est marginale en t2 :  $\chi^2(10, 577) = 20,1$  ( $p=0,028$ )

<b>NN av t1</b>	<b>SP lexicales</b>	<b>SP grammaticales</b>	<b>SP discursives</b>
FA	46	20	26
MJ	18	9	10
PR	48	31	32
QI	57	44	23
VI	20	17	17
VO	40	28	23

Tableau 29. Nb et types de SP par sujet, niv3, t1

Le groupe avancé est celui qui produit le plus de SP lexicales en t1 comme en t2. La différence marginale constatée en t2 révèle quelques petites différences entre les sujets. Tous les sujets du groupe avancé ont utilisé plus de SP lexicales en t2 (excepté PR) mais dans des proportions différentes. FA et surtout VI sont les deux sujets qui en produisent le plus, passant

de 46 à 64 et de 20 à 49 SP lexicales respectivement ; FA produit aussi plus de séquences discursives en t2.

<b>NN av t2</b>	<b>SP lexicales</b>	<b>SP grammaticales</b>	<b>SP discursives</b>
FA	64	26	38
MJ	26	20	12
PR	22	16	18
QI	68	48	20
VI	49	32	19
VO	44	24	31

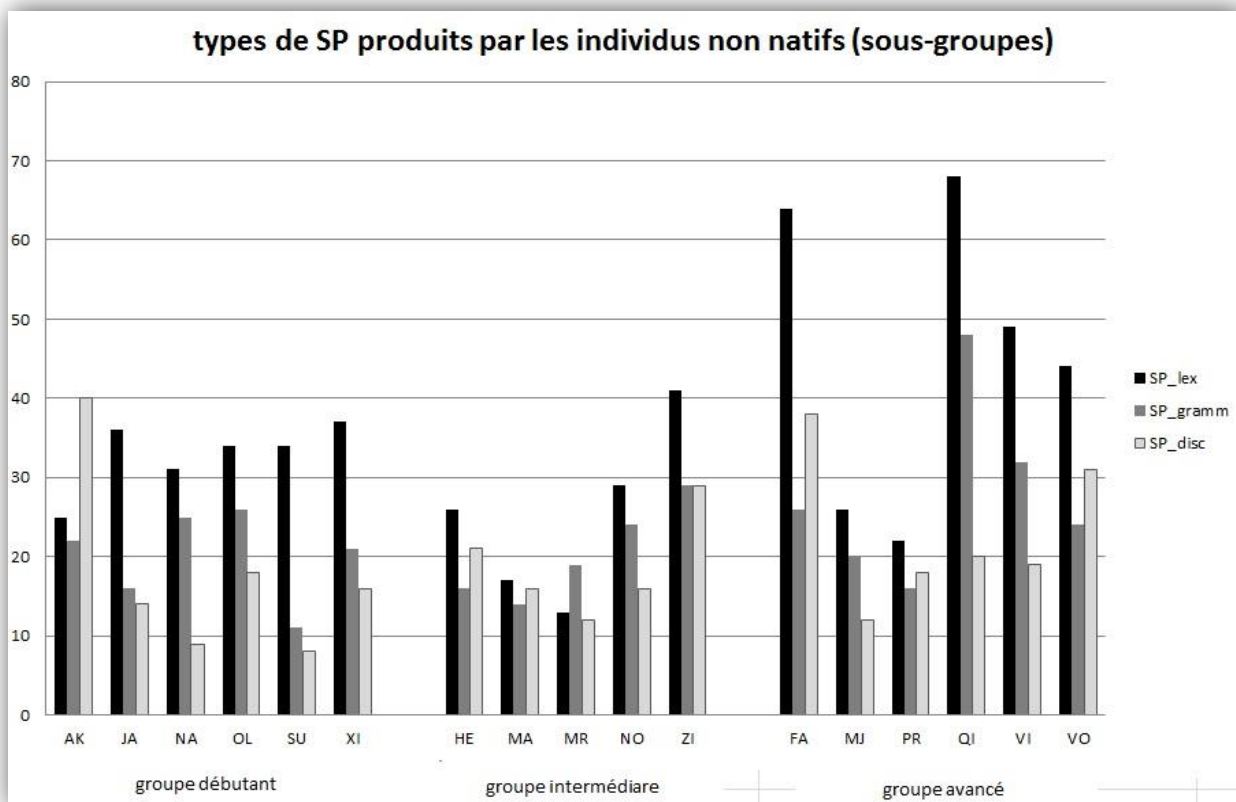
Tableau 30. Nb et types de SP par sujet, niv3, t2

En conclusion, les changements dans la répartition des SP du groupe débutant indiquent que l'utilisation des types de SP a évolué dans les productions de ce groupe, et en particulier dans l'augmentation des séquences de type lexical. Les autres groupes ne présentent pas une même évolution dans le temps si ce n'est une légère variation interindividuelle en t2 pour le groupe avancé.

La figure 6 suivante illustre la répartition des SP lexicales, grammaticales, discursives par sous-groupes et par sujet en t2. On voit que les différences individuelles dans les sous-groupes sont très marquées, notamment dans le groupe 3. Comme chez les sujets natifs, c'est au niveau individuel que l'on perçoit des différences dans l'emploi des SP non seulement en quantité mais aussi en types de SP produites. La variabilité entre les individus est aussi soulignée par les chercheurs (Forsberg-Lundell *et al.* 2014 : 269). Afin de vérifier si les séquences produites sont plutôt les mêmes ou si elles se diversifient, nous allons comparer le taux de SP produites pour cent mots en valeur brute mais aussi en occurrences types.

Cinq des six sujets du groupe débutant produisent des SP lexicales en plus grand nombre en t2. Seul AK produit plus de SP discursives, et moins de SP lexicales que les autres apprenants du groupe. La répartition des SP en types est plutôt homogène entre les cinq sujets. Il en va autrement dans le groupe intermédiaire : la figure 6 illustre parfaitement ce que les tests khi-deux avaient mis en avant, c'est-à-dire que la répartition des SP n'est pas homogène. Seule NO produit des types de SP réparties de façon homogène comme celles du groupe contrôle. Dans le groupe avancé, deux sujets FA et QI produisent nettement plus de SP lexicales que les autres apprenants. Mais si FA produit de nombreuses SP lexicales, il produit aussi plus de SP discursives que de SP grammaticales, contrairement à QI et aux sujets natifs.

Figure 6. Répartition des SP, par sous-groupe et par sujet, t2



FA et QI se détachent cependant des autres apprenants et la figure le montre bien : le nombre de SP lexicales est bien supérieur à l'ensemble du groupe. Quelles en sont les raisons ? Grâce aux informations tirées des interviews que nous avons menées, nous pouvons tenter de l'expliquer. Comme le font remarquer Forsberg et Fant (2010), certains sujets sont plus des utilisateurs de la langue (*users*) que des apprenants. C'est sans doute le cas de QI et FA. QI vit en France depuis plusieurs années et FA depuis plus d'un an. De ce fait, ils ont acquis plus d'expérience de la langue que les autres sujets de leur groupe. QI adore la langue française, elle vit avec son compagnon chinois dans une ville du Sud et ne conçoit pas sa vie ailleurs qu'en France. Elle désire y travailler et n'a pas envie de retourner vivre en Chine. FA a été embauché par une entreprise basée à Lyon et son employeur finance ses études. Il est, lui aussi, extrêmement motivé et sa carrière professionnelle devrait s'effectuer en partie en France. Il est investi dans l'institution (Ecoles des Mines) et participe à de nombreux événements, notamment à une conférence qui fera date à l'Ecole des Mines d'Albi, sur les origines du chocolat, en collaboration avec un chocolatier de renom dans la région. Un autre sujet du groupe avancé, VI,

arrivé en France en début d'année, produit moins de séquences préfabriquées que FA et QI mais leur répartition ressemble plus à celle du groupe natif. Ce sujet a appris le français pendant deux ans au Brésil pendant ses études universitaires et sa progression est surprenante au bout de six mois dans le pays de la langue cible. Il désire venir vivre en France et y travailler par la suite. A la différence des autres sujets, il est en France pour une seule année.

### 7.1.2 Résultats pour les mesures de la préfabrication, groupe entier, t1 et t2

Le tableau 31 ci-après présente les moyennes des mesures de la préfabrication aux deux moments de prise d'information (t1 et t2) pour tous les sujets non-natifs ainsi que les résultats statistiques du Test de Wilcoxon comparant ces mesures en novembre et en juin. Le test comparatif montre une légère progression dans l'utilisation de la préfabrication entre t1 et t2 (différences marginales pour les mesures 1, 2, 3, 6).

	Mesures	Moyenne	Ecart-type	Mini	Maxi	Wilcoxon
1	Taux syllabes préfabriquées t1	,312	,059	,182	,409	p=0,04 (m)
	Taux syllabes Préfabriquées t2	,332	,059	,213	,428	
2	Taux segments avec SP t1	,463	,116	,285	,696	p=0,04 (m)
	Taux segments avec SP t2	,502	,135	,308	,893	
3	nb syll pref par segment t1	2,28	1,38	,93	6,52	=0,03 (m)
	nb syll pref par segment t2	2,63	1,46	1,14	7,71	
4	nb syll préf par seg avec SP t1	4,65	1,42	3,26	9,37	>0,05 NS
	nb syll préf par seg avec SP t2	5,00	1,11	3,69	8,64	
5	nb SP_100mots t1	12,61	2,28	7,04	16,31	>0,05 NS
	nb SP_100mots t2	13,19	2,16	8,75	16,62	
6	Nb_SP_100mots occurrences types t1	9,12	2,19	5,28	13,47	=0,02 (m)
	Nb_SP_100mots occurrences types t2	9,22	1,99	5,67	12,32	
7	SP lexicales_100mots t1	4,75	1,72	2,15	8,16	<0,01*
	SP lexicales_100mots t2	5,89	1,42	3,17	8,31	
8	SPlexicales_100mots occurrences types t1	3,61	1,51	1,24	6,56	=0,02 (m)
	SPlexicales_100mots occurrences types t2	4,28	1,18	2,32	6,40	

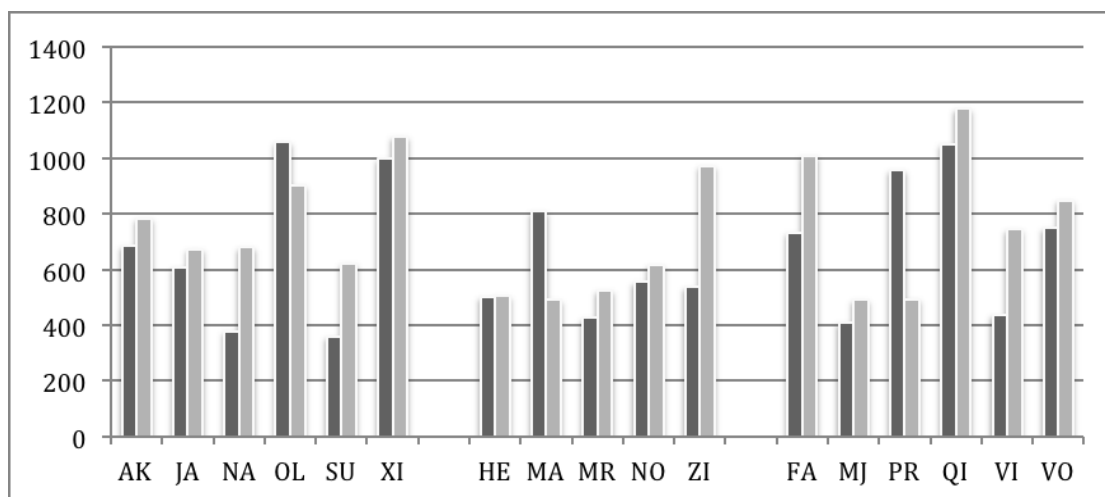
Tableau 31. Résultats des mesures de la préfabrication, t1 et t2

Malgré une légère augmentation des moyennes de toutes les mesures, la différence entre t1 et t2 dans les taux de préfabrication (mesures 1 à 3, et 6) reste modeste car les tests statistiques sont proches du seuil de significativité ( $p=,05$ ). Seul le taux des SP lexicales pour cent mots, en valeur brute, montre une forte différence ( $p<0,01$ ) mais mesuré en occurrences types l'effet est moins prononcé ( $p=,02$ ). Pour les autres mesures, on ne constate qu'un léger changement entre t1 et t2 (différences marginales - à nuancer selon les groupes certainement), et aucune évolution en ce qui concerne le taux de SP pour cent mots. Ces résultats indiquent que le changement a lieu dans l'utilisation que les apprenants font des SP lexicales, plus nombreuses et plus diversifiées en t2.

### 7.1.3 Les quatre premières mesures de la préfabrication (Cordier)

Avant d'effectuer des calculs de la préfabrication selon les mesures développées par Cordier (2013), nous présentons dans la figure 7 le nombre de syllabes produites par les sujets en t1 et en t2.

Figure 7. Nombre de syllabes produites par sous-groupe et par sujet, t1 et t2



On constate dans la figure 7 des différences dans le nombre de syllabes produites entre t1 et t2, et qui sont liés, bien sûr, à des comportements langagiers individuels. Il est intéressant de constater que les sujets qui parlent beaucoup, XI et OL, et qui produisent beaucoup de syllabes en termes de quantité sont également ceux qui hésitent beaucoup : en les écoutant, leur discours paraît *disfluent* (voir tableau 39, en particulier le nombre de pauses par minute). Lors de l'interview, XI s'est plainte d'être bloquée en production alors que la compréhension ne lui pose aucun problème. Elle a étudié en France pendant deux années, en troisième et

quatrième année d'un cursus scientifique (l'École des Mines, en partenariat avec l'université de Shangai, recrute tous les ans des étudiants qui viennent préparer un double diplôme). XI a terminé les deux années de Master et entame sa troisième année en France, cette fois en doctorat ; mais elle éprouve toujours des difficultés à l'oral, qui se traduisent par un discours disfluent avec de nombreuses pauses, notamment des pauses remplies (allongements) et des répétitions. OL est en France depuis peu de temps mais éprouve aussi de grandes difficultés en production orale. Ces deux sujets, d'origine asiatique, semblaient moins investies dans le groupe et avaient tendance à rester avec les membres de leur communauté linguistique. En t2, XI et OL produisent moins de syllabes en quantité qu'en t1. Il se pourrait que ce soit par manque d'idée ou par lassitude, mais plus probablement parce qu'elles semblent avoir plus confiance en elles et qu'elles veulent mieux faire, être plus précises.

Deux autres sujets du groupe débutant, SU et NY, parlent peu en t1 et plus en t2. Elles ont acquis plus de connaissances lexicales, des unités simples et des séquences, et semblent plus à l'aise dans leur narration. Enfin, on constate que ceux qui parlent moins en t2, comme MA (niv2) et PR (niv3), le font pour d'autres raisons que OL. OL, qui hésitait beaucoup en t1, se reprend moins en t2 et fait moins de répétitions. MA et PR qui, en t1, avaient produit des narrations plutôt riches lexicalement parlent moins en t2. Elles semblent moins motivées et peut-être lassées d'accomplir la même tâche. PR semblait pressée d'achever la narration pour des raisons personnelles d'emploi du temps. Deux autres sujets du groupe avancé, FA et VI, parlent plus parce qu'ils sont à l'aise et qu'ils ont certainement plus de choses à dire la deuxième fois.

Le tableau 32 suivant présente les résultats individuels des apprenants L2 pour les quatre premières mesures de la préfabrication selon Cordier. Les six premiers sujets font partie du groupe 1 débutant, les cinq suivants du groupe 2 intermédiaire et les six derniers du groupe 3 avancé (zones grisées). Nous donnons les moyennes, les écarts types, les résultats minimum et maximum pour chaque sous-groupe aux deux moments (t1 et t2).

Sous-groupes	Sujets	taux de syllabes préfabriquées		taux de segment avec SP		nb syll préfab par segment		nb syll préfab par segment avec SP	
		T1	T2	T1	T2	T1	T2	T1	T2
groupe 1 débutant	AK	0,30	0,37	0,44	0,49	1,65	2,26	3,75	4,63
	JA	0,28	0,32	0,42	0,44	1,86	2,08	4,38	4,74
	NY	0,26	0,30	0,34	0,40	1,25	2,14	3,70	5,29
	OL	0,25	0,27	0,28	0,32	0,93	1,21	3,26	3,80
	SU	0,30	0,31	0,48	0,47	2,04	2,41	4,24	5,14



	XI	0,18	0,21	0,29	0,31	1,04	1,14	3,55	3,69
	moyenne	0,26	0,30	0,37	0,41	1,46	1,89	3,81	4,56
	Ecart-type	0,04	0,05	0,08	0,08	0,45	0,56	0,42	0,67
	Min	0,18	0,21	0,28	0,31	0,93,	1,14	3,26,	3,69
	max	0,30	0,37	0,48	0,49	2,04	2,41	4,38	5,29
groupe 2 intermédiaire	HE	0,35	0,37	0,42	0,60	1,95	3,35	4,66	5,58
	MA	0,36	0,32	0,53	0,46	2,65	2,33	4,98	5,03
	MR	0,31	0,26	0,47	0,49	2,26	1,96	4,85	4,03
	NO	0,29	0,37	0,41	0,48	1,43	2,25	3,46	4,69
	ZI	0,23	0,26	0,34	0,44	1,26	2,01	3,68	4,55
	Moyenne	0,31	0,32	0,43	0,49	1,91	2,40	4,32	4,81
	Ecart-type	0,05	0,06	0,07	0,06	0,58	0,61	0,71	0,64
	Min	0,23	0,26	0,34	0,44	1,26	1,96	3,46	4,03
	Max	0,36	0,38	0,53	0,60	2,67	3,46	5,02	5,76
groupe 3 avancé	FA	0,41	0,42	0,70	0,88	6,52	7,59	9,38	8,67
	MJ	0,34	0,35	0,52	0,54	2,40	2,44	4,63	4,50
	PR	0,34	0,33	0,59	0,43	3,46	2,29	5,88	5,32
	QI	0,37	0,36	0,56	0,53	2,81	2,56	5,04	4,86
	VI	0,36	0,43	0,64	0,67	3,36	3,87	5,27	5,76
	VO	0,38	0,37	0,45	0,56	1,92	2,50	4,32	4,47
	Moyenne	0,37	0,38	0,58	0,60	3,42	3,56	5,76	5,59
	Ecart-type	0,03	0,04	0,09	0,16	1,63	2,11	1,86	1,57
	Min	0,34	0,33	0,45	0,43	1,92	2,29	4,32	4,47
	Max	0,41	0,43	0,70	0,89	6,52	7,71	9,38	8,64

Tableau 32. Résultats des mesures de la préfabrication (Cordier) par sujet et par groupe

Les résultats présentés dans ce tableau montrent des différences individuelles importantes entre les groupes, mais aussi à l'intérieur des groupes. Dans le groupe 1, XI produit 18% de syllabes préfabriquées en t1 alors qu'AK et SU en produisent presque deux fois plus (30%), le rapport reste le même en t2 malgré une augmentation sensible (21% vs 37%). On trouve ces écarts dans tous les groupes et pour toutes les mesures. Ces différences, on l'a vu, existent aussi chez les locuteurs natifs et semblent être « normales » en production langagière (en L1 en comme en L2). Chaque apprenant a son propre rythme d'acquisition/utilisation des SP de la langue cible en fonction de l'input reçu, de sa L1, de sa motivation (Dörnyei, 2001). Nous constatons que tous les sujets du groupe débutant sans exception produisent plus de syllabes préfabriquées en t2. Mais les résultats de certains sujets des groupes 2 et 3 ne suivent pas la

même progression. Dans le groupe intermédiaire, deux sujets, MA et MR, produisent moins de syllabes préfabriquées en t2 alors que NO en produit beaucoup plus. MA et MR ne semblent pas avoir utilisé plus de LP en t2, pour des raisons différentes, nous semble-t-il. MA, on l'a déjà évoqué, ne paraît plus aussi motivée et MR semble avoir moins besoin de s'appuyer sur les SP pour parler vite car il calque sa vitesse de parole en L2 sur les habitudes de sa L1.

A ce stade, nous voulons également vérifier s'il existe des différences entre les groupes, en t1 et en t2. Pour cela, nous avons utilisé le test statistique non paramétrique de Kruskal-Wallis. Les résultats de ces tests se trouvent dans le tableau 33.

	Mesures	T1		T2	
		H=ddl	Valeur de p	H=ddl	Valeur de p
1	Taux syllabes préfabriquées	H(2)=10,42	<0,01*	H(2)=4,89	=0,08 NS
2	Taux segment avec SP	H(2)=9,04	=0,01*	H(2)=6,63	=0,03 (m)
3	Nb syll pré par segment	H(2)=9,70	<0,01*	H(2)=7,92	=0,02 (m)
4	Nb syll préf par segment avec SP	H(2)=8,35	=0,02 (m)	H(2)=1,62	=0,4 NS
5	Nb SP_100mots	H(2)=8,21	<0,02 (m)	H(2)=7,76	=0,02 (m)
6	Nb SP_100mots occurrences types	H(2)=9,28	<0,01*	H(2)=8,57	=0,01
7	Splexicales_100mots	H(2)=10,29	<0,01*	H(2)=5,64	=0,05 NS
8	Splexicales_100mots occurrences types	H(2)=11,81	<0,01*	H(2)=9,98	<0,01

Tableau 33. Test de Kruskal-Wallis, t1 et t2

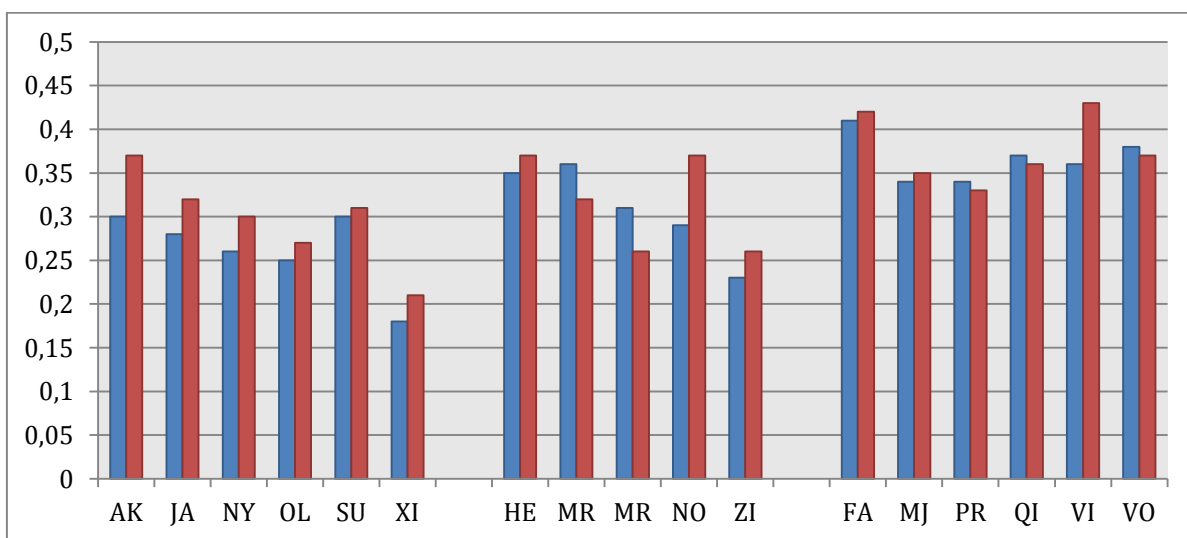
Nous voyons ici des différences significatives en t1 entre les groupes sur les premières mesures de la préfabrication préconisées par Cordier, le taux de syllabes préfabriquées, le taux de segments avec SP, et le nombre de syllabes préfabriquées par segment, (lignes 1 à 3) ; et un effet marginal, mesure 4 (nombre de syllabes préfabriquées par segment contenant des SP). On note en revanche des différences marginales ou aucune différence en t2, sauf pour deux mesures (lignes 6 et 8, les taux en occurrences types). Ces résultats peuvent être interprétés comme la réduction de l'écart entre les groupes en t2, c'est-à-dire une progression des groupes moins avancés entre t1 et t2. Le pourcentage de syllabes préfabriquées ne présente aucune différence significative au t2, contrairement au t1. Il en est de même pour

toutes les autres mesures, puisque les différences entre les groupes se réduisent. On note cependant encore une différence statistique significative en t2, pour le nombre de SP lexicales (effet très prononcé en t1) qui montre une augmentation de celles-ci, même si les écarts entre les groupes restent encore importants. Les groupes moins avancés arrivent en t2 à des résultats qui tendent vers ceux du groupe avancé.

*a) mesure 1, taux de préfabrication (% de syllabes préfabriquées)*

La figure 8 suivante présente les résultats en pourcentage du nombre de syllabes préfabriquées de la production en temps 1 et en temps 2. Il montre donc l'évolution d'une fois sur l'autre des apprenants. Les résultats d'AK, NO et VI, sont en progression de manière très nette la deuxième fois, ceux de JA, NY et ZI dans une moindre mesure. On constate que les apprenants du groupe 1, c'est-à-dire les six premiers dans la figure progressent d'une fois sur l'autre. Dans le groupe 2, deux sujets présentent des résultats en baisse (MA et MR). Les résultats du groupe 3 montrent également des différences entre les sujets, trois sujets produisent moins de syllabes préfabriquées la deuxième fois (PR, QI et VO).

Figure 8. Taux de syllabes préfabriquées par sous-groupe et par sujet, t1 et t2



La figure 8 montre très clairement les différences entre sujets en t1 et t2, à l'intérieur des sous-groupes et dans le groupe entier. Certains sujets produisent plus de syllabes préfabriquées en t2 (AK, NO et

VI) et d'autres moins (MA, MR, PR, QI et VO). Mais tous les apprenants débutants sans exception produisent plus de syllabes préfabriquées en t2 avec, cependant, une certaine variabilité constatée entre les sujets.

- Dans le groupe 1, AK produit 30% de syllabes préfabriquées en t1 et 37 % en t2, soit une progression de 7% alors que le pourcentage d'augmentation chez XI est très faible, à peine 1%. Si le pourcentage de syllabes préfabriquées augmente chez AK et XI, il se fait dans des proportions différentes. On a constaté précédemment que XI en t2 a produit moins de syllabes, ce qui peut expliquer ces résultats.

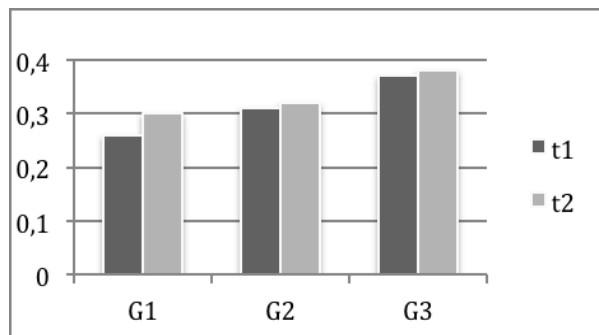
- Dans le groupe 2, NO produit plus de syllabes préfabriquées en t2 alors que MR et MA en produisent moins. Ces deux sujets n'ont pas progressé en termes d'emploi du LP, pour des raisons différentes, on l'a vu.

- Deux sujets du groupe avancé, ont des comportements langagiers différents, PR produit moins de syllabes préfabriquées en t2 et VI en produit beaucoup plus. Un possible manque de motivation de PR déjà évoqué (diminution très nette de la quantité de la production orale en mots et nombre de syllabes) peut expliquer ces résultats en baisse. En revanche VI, on l'a dit, est un apprenant très motivé qui semble avoir des facilités avec le français et qui a progressé très vite. VO a des résultats légèrement en baisse dus certainement à son désir d'être plus précise, de mieux faire.

L'apprentissage n'étant pas linéaire, chaque sujet a un développement langagier qui lui est propre selon son niveau mais aussi sa motivation, ses objectifs, son intérêt. Il existe aussi des paliers dans l'acquisition, des moments de progression et des plateaux de stagnation (McLaughlin, 1990 ; voir chapitre 3).

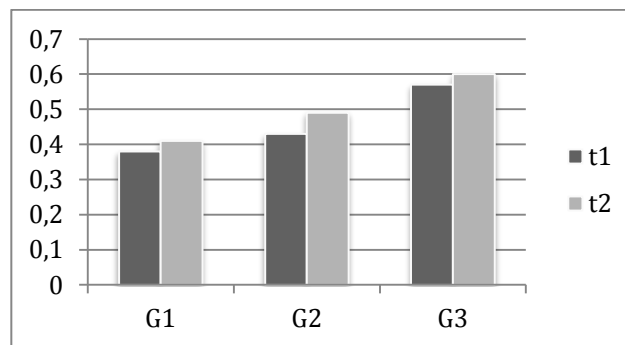
La figure 9 résume les taux de préfabrication par groupe, en t1 et t2. On constate à nouveau que le groupe 1 est le seul où le taux progresse réellement même si cette progression reste modeste. Les deux autres groupes ne progressent pas autant, ce qui confirme les résultats statistiques du test Kruskal-Wallis.

Figure 9. Taux de syllabes préfabriquées par sous-groupes, t1 et t2



*b) Mesure 2, taux de segments contenant des syllabes préfabriquées*

Figure 10. Taux de segments avec SP, t1 et t2

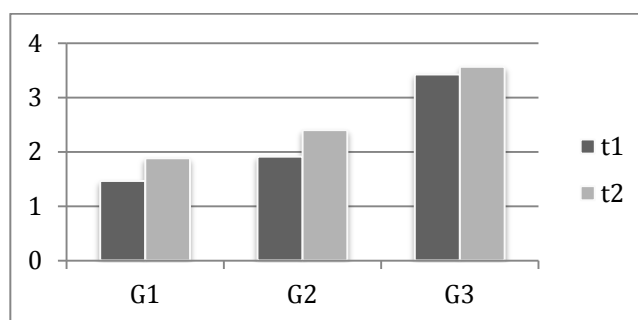


L'histogramme 10 présente la deuxième mesure, le pourcentage de segments contenant des syllabes préfabriquées. L'écart entre t1 et t2 pour le groupe 1 ne permet pas à ce groupe de se rapprocher du groupe 3 avancé mais on note cependant une légère progression en t2. Le groupe 2 produit plus de segments avec SP en t2 et tend vers les résultats du groupe avancé. Cependant les résultats statistiques ne montrent pas de différence significative entre les groupes en t1 comme en t2.

*c) mesure 3, nombre de syllabes préfabriquées par segment*

La figure 11 présente la troisième mesure concernant le nombre moyen de syllabes préfabriquées par segment de parole, par sous-groupes et en t1 et en t2.

Figure 11. Nombre moyen de syllabes préfabriquées par segment de parole

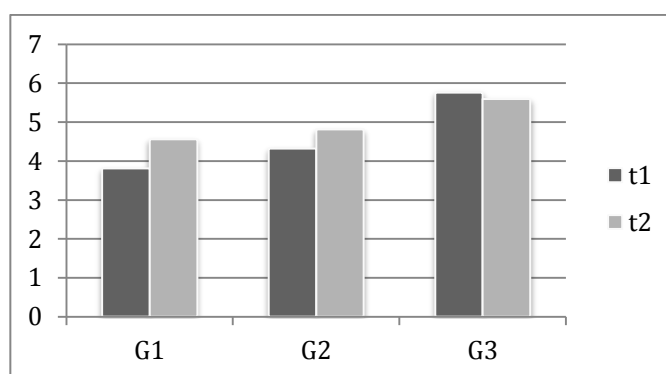


On constate encore une fois un écart entre le groupe 1 débutant et le groupe 3 avancé sur cette mesure. Même si les groupes 1 et 2 produisent plus de syllabes préfabriquées en t2, tendant vers les valeurs du groupe des natifs, la différence subsiste en t2 entre les groupes de niveau. Le groupe 3 produit à peine plus de syllabes préfabriquées par segment en t2 mais c'est quand même le groupe avancé qui produit le plus de syllabes préfabriquées par segment de parole en t1 comme en t2, dépassant en cela largement les deux autres groupes.

*d) mesure 4, nombre de syllabes par segment contenant des SP*

La figure 12 présente les moyennes des groupes pour le nombre moyen de syllabes par segment contenant des SP.

Figure 12. Nombre moyen de syllabes préfabriquées par segment avec SP



On constate dans cette figure des écarts moins marqués entre les niveaux que dans la mesure précédente. Cette mesure plus « généreuse » permet de mieux rendre compte de la quantité de syllabes préfabriquées par segment de parole avec SP. Le groupe débutant produit plus de syllabes en t2 alors que le groupe avancé en produit un peu moins, pour les différentes raisons que nous avons déjà évoquées (manque de motivation en t2, souci de précision, etc.). Le

nombre moyen de syllabes préfabriquées dans les segments contenant des SP montre que le LP est bien présent dans les productions orales des apprenants, même si le résultat moyen n'est pas comparable à la moyenne des natifs (7,3 syllabes préfabriquées par segment). La longueur moyenne produite par le groupe avancé (5,59) peut être mise en relation avec les résultats individuels de certains natifs.

#### 7.1.4 Mesures de la préfabrication : taux de séquences préfabriquées pour cent mots

##### a) Comparaisons, t1 et t2, (tous les groupes), taux pour 100 mots

Taux de SP pour 100 mots		H	T1 Valeur de p	T2 Valeur de p
5	Nb SP_100mots valeur brute	2	=0,01*	=0,02 (m)
6	Nb SP_100mots occurrences types	2	<0,01*	=0,02 (m)
7	SP lexicales_100mots valeur brute	2	<0,01*	=0,06 NS
8	SPlexicales_100mots occurrences types	2	<0,01*	<0,01*

Tableau 34. Taux de préfabrication, test de Wilcoxon

Le tableau 34 ci-dessus présente les résultats statistiques du test de Wilcoxon, comparant les mesures de la préfabrication pour cent mots en t1 et en t2 : le taux des SP produites (tous types confondus) en valeur brute, et le taux des SP en occurrences types, comptées une seule fois, c'est-à-dire sans tenir compte des répétitions. Nous avons appliqué le test de rang de Wilcoxon au groupe entier, pour voir s'il y a un effet de progression en t2. On constate une évolution entre t1 et t2 ; une différence significative en t1 sur toutes les mesures et une différence marginale en t2, ce qui signifie qu'il y a eu progression et que les écarts entre les sujets se sont réduits en t2. Le résultat statistique de la mesure 7 (taux de SP lexicales en valeur brute) montre que les écarts se sont réduits entre les groupes entre t1 et t2. En revanche, la mesure 8 (taux de SP lexicales en occurrences types) montre un écart entre les groupes même en t2. On peut conclure d'une manière générale que les apprenants évoluent dans l'emploi qu'ils font des SP entre t1 et t2. La seule différence qui subsiste entre les groupes en t2 concerne la diversité des SP lexicales, leur taux en occurrences types.

Nous présentons les résultats par groupe de niveau pour ces quatre dernières mesures de la préfabrication.

##### b) Résultats individuels, par sous-groupes, taux pour 100 mots, t1 et t2

Nous présentons dans le tableau 35 les taux de SP pour cent mots produits et le taux de SP lexicales par sujet et par groupe de niveau. Une de nos hypothèses étant que les sujets, la deuxième fois, produiront plus de SP et qu'ils devraient moins s'appuyer sur les mêmes en t2 et produire des SP lexicales plus variées.

Le tableau suivant regroupe les taux de SP par apprenant en t1 et en t2, ainsi que les taux de SP lexicales, en valeur brute et en occurrences types. Nous rappelons également les moyennes, les écarts types et les valeurs extrêmes pour chaque groupe.

Sujets		SP pour 100 mots valeur brute		SP pour 100 mots types		SPlex pour 100 mots valeur brute		SPlex pour 100 mots types	
		t1	t2	t1	t2	t1	t2	t1	t2
groupe 1 débutant	AK	13,08	14,51	8,06	7,79	2,15	4,15	1,79	2,32
	JA	11,49	11,66	7,72	8,46	2,77	6,77	2,38	4,32
	NY	9,60	12,15	6,19	8,30	2,48	5,72	1,24	3,87
	OL	9,91	12,09	5,81	6,59	3,19	4,77	1,71	2,95
	SU	12,75	11,50	10,11	6,94	5,62	7,38	3,00	3,69
	XI	7,38	8,99	5,28	5,67	3,65	4,37	3,02	2,84
	Moyenne	10,81	11,67	7,20	7,25	3,31	5,52	2,19	3,33
	Écart-type	2,27	1,86	1,80	1,12	1,25	1,33	0,73	0,75
	Min	7,04	8,75	5,28	5,67	2,15	4,15	1,24	2,32
	max	13,08	14,43	10,11	8,46	5,62	7,37	3,02	4,32
groupe 2 Intermédiaire	HE	12,81	15,63	9,83	11,22	4,80	6,34	3,60	5,12
	MA	13,33	11,74	10,35	10,77	3,96	4,36	3,65	4,87
	MR	13,08	11,65	9,69	8,05	5,63	3,17	4,38	3,17
	N0	12,89	13,19	7,66	11,11	3,83	5,97	2,70	4,53
	ZI	11,48	11,07	7,93	7,30	3,73	5,16	2,80	3,02
	Moyenne	12,50	12,73	9,09	9,69	4,39	4,30	3,43	4,14
	Écart-type	0,90	1,97	1,21	1,87	0,81	1,28	0,69	0,98
	Min	10,96	10,73	7,66	7,30	3,73	3,17	2,70	3,02
	max	13,13	15,37	10,35	11,22	5,63	6,34	4,38	5,12
groupe 3 avancé	FA	16,82	17,30	13,48	11,30	8,16	8,31	6,56	5,45
	MJ	12,96	14,38	10,19	10,79	5,56	6,24	4,94	5,28
	PR	14,97	14,63	9,76	10,18	6,33	5,74	4,09	3,92
	QI	15,86	15,77	10,13	9,84	6,96	7,27	5,25	5,78
	VI	14,65	15,62	11,11	12,32	5,42	7,84	4,34	6,40
	VO	15,66	13,16	11,75	10,19	6,62	6,60	5,96	5,25
	Moyenne	14,51	15,09	11,07	10,79	6,51	7,00	5,19	5,35
	Écart-type	1,64	1,02	1,39	0,90	1,01	0,98	0,95	0,82
	Min	11,42	13,91	9,76	9,95	5,42	5,74	4,09	3,92
	max	16,31	16,62	13,48	12,32	8,16	8,31	6,56	6,40

Tableau 35. Taux des SP, SP lexicales, valeurs brutes et occurrences types

Alors que le taux de SP devrait avoir augmenté en t2 (hypothèse), on note une baisse chez certains sujets, moins productifs en t2 en termes de nombre de SP. Les sujets du groupe débutant ont dans l'ensemble progressé en quantité mais aussi en variété, en diversité. Si le taux de SP n'a pas progressé de manière très nette, le taux de SP lexicales pour cent mots, en



revanche, présente une nette augmentation pour la majorité des sujets. Les apprenants produisent dans l'ensemble plus de SP lexicales et s'appuient moins sur les mêmes séquences en t2, surtout chez les débutants contrairement aux sujets avancés. FA, par exemple, a tendance à répéter les mêmes SP lexicales (en occurrences types, les valeurs baissent en t2). On remarque que certains locuteurs avancés, comme d'ailleurs certains natifs, ont tendance à répéter les mêmes SP, sans doute pour la planification et afin d'éviter les pauses. Ce serait donc plus une stratégie de l'oral qu'ils mettent en place.

PR et VO s'appuient moins sur les SP en t2, pour les raisons que nous avons déjà mentionnées. Les mauvaises performances de PR sont le résultat du manque de motivation. En revanche, VO semble moins s'appuyer sur les SP en t2 car elle cherche à produire une narration plus juste et plus précise. C'est pour cette raison sans doute qu'il sera difficile voire impossible de généraliser les résultats, la production orale étant propre à chaque individu. Nous présentons les résultats détaillés par tâche, Bd et film en t1 et en t2 (Annexe 15).

## 7.2 Développement de la fluence

### 7.2.1 Données brutes, nombre de mots produits bruts et nombre de mots nettoyés

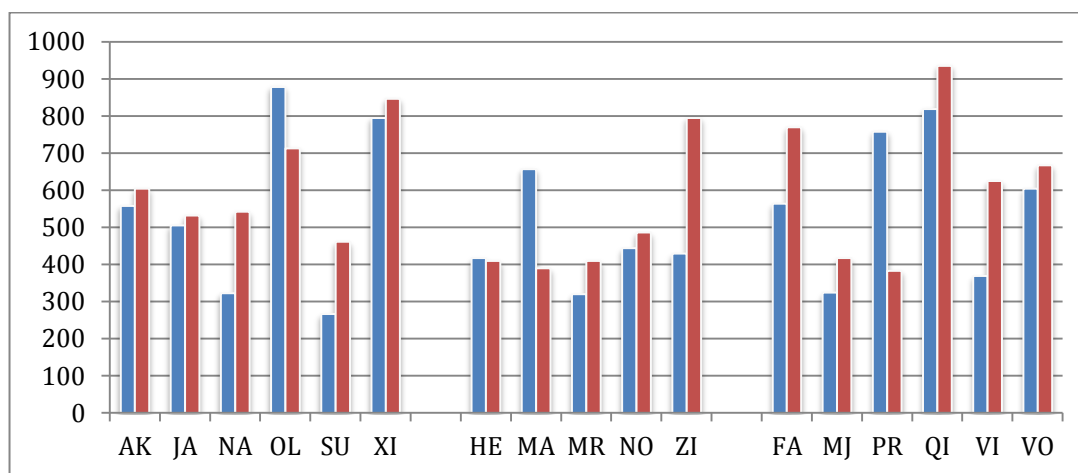
Avant de présenter les mesures temporelles de la productivité, il nous paraît utile de présenter pour chaque sujet (par groupe de niveau) le nombre de mots produits en t1 et en t2. Le tableau 36 présente le nombre de mots produits par chaque sujet, bd et film, en t1 et en t2. Les apprenants ont, dans l'ensemble, produit une moyenne de 587,41 mots en t2 contre 531,24 en t1. Une des raisons du faible différentiel (56 mots) est due à la faible performance de deux sujets, MA et PR qui produisent moitié moins de mots en t2, de 657 à 390 mots et de 758 à 383 en t2.

Sujets	t1	t2	total
AK	558	603	1161
JA	505	532	1037
NY	332	542	865
OL	878	713	1591
SU	267	461	728
XI	795	846	1641
HE	417	410	827
MA	657	390	1047
MR	320	410	730
NO	444	486	930
ZI	429	795	1224
FA	564	770	1334
MJ	324	417	741
PR	758	383	1141
QI	819	935	1754
VI	369	625	994
VO	604	667	1271
Total	9031	9986	19017
Moyenne	531,24	587,41	1118,55

Tableau 36. Nombre de mots produits par chaque sujet, t1 et t2

On constate dans le tableau 36 que certains apprenants ont parlé beaucoup plus la deuxième fois alors que, pour d'autres, c'est l'inverse. Ces résultats sont certainement un indicateur de la motivation de certains, ou du manque d'intérêt pour d'autres (une tâche narrative répétée après 6 mois). Cordier (2013 :195-196) constate chez ses sujets que le nombre de mots produits la deuxième fois a considérablement augmenté. Deux facteurs semblent l'expliquer : la nature des tâches et la durée du séjour dans le pays de la langue cible. Notons également que l'homogénéité de son groupe, groupe de niveau avancé constitué de cinq apprenants anglophones apprenant le français dans la même classe, peut vraisemblablement expliquer ces « bons » résultats. Dans cette recherche, les dix-sept apprenants n'ont pas tous le même niveau de français. Certains sujets sont testés en t1 dès leur arrivée dans l'institution en France (novembre-décembre), d'autres sont déjà en France depuis un certain temps, de quelques semaines à plusieurs années (chapitre 5, tableau 5.2). Cela peut expliquer des comportements et des résultats très différents.

Figure 13. Nombre de mots produits par sous-groupes et par sujet, t1 et t2



Les trois sujets qui parlent moins la deuxième fois sont MA, PR et OL et dans une moindre mesure, HE (7 mots en moins la deuxième fois). Ces chiffres sont à mettre en regard des mots produits nettoyés (voir tableau 37 et figure 15) afin de vérifier que les mots produits (en plus) ne sont pas des répétitions ou d'autres types de reprises.

Les sujets qui parlent le plus en t2 sont FA (+36,5%), NY (+67,8%), SU (+72,7%), VI (+69,4%) et ZI (+85,3%). Notons que les sujets qui parlent moins en t2 ont déjà fait l'objet d'une interprétation, baisse de la motivation (MA, PR) ou amélioration, plus de précision et moins de disfluences (OL).

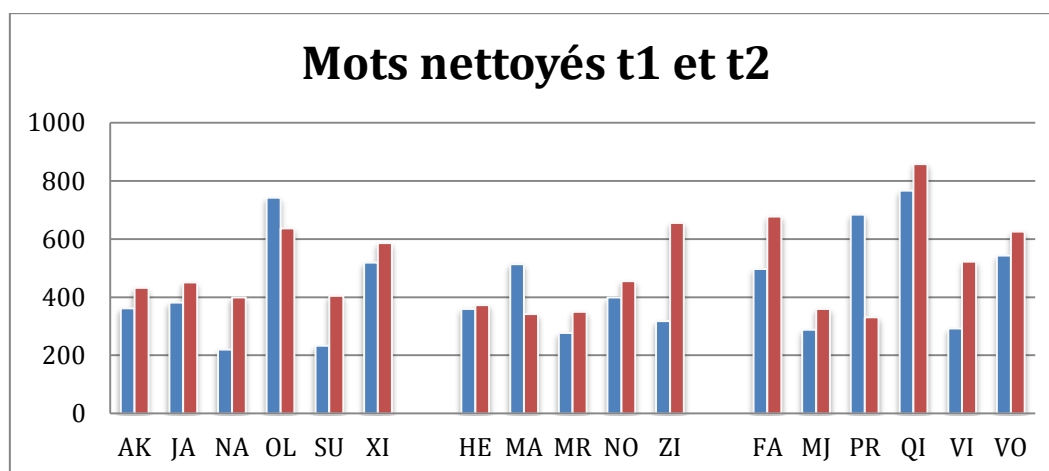
Les chiffres (bruts) montrent de grandes différences entre les sujets. On constate notamment que trois apprenants, OL, XI (groupe 1) et ZI (groupe 2) produisent beaucoup plus de mots que FA et VI du groupe 3. Ce sont des productions longues mais beaucoup moins précises que celles de sujets de niveau avancé. Il nous a alors paru intéressant de montrer ce qui se passe une fois que les productions ont été nettoyées des reprises et répétitions. Nous présentons sous la forme de tableau les résultats des mots produits nettoyés (*pruned words*) pour bien comprendre ce qui se passe pour certains. Nous reviendrons sur ces problèmes dans la section « disfluences ».

Tableau 37. Nombre de mots produits nettoyés, t1 et t2

	AK	JA	NY	OL	SU	XI	HE	MA	MR	NO	ZI	FA	MJ	PR	QI	VI	VO
t1	362	382	220	742	233	519	360	514	278	399	318	497	288	684	766	293	543
t2	432	451	399	637	405	586	373	342	350	456	656	678	360	331	858	523	626
total	794	833	619	1379	638	1105	733	856	628	855	974	1175	648	1015	1624	816	1169
baisse	-40	-18	-36	-12	-14	-44	-10	-14	-17	-7	-21	-14	-16	-16	-9	-20	-7

Le tableau 37 présente le nombre de mots nettoyés en t1 et t2 (lignes 1 et 2), le total (ligne 3), et le pourcentage de mots en moins par rapport à la production brute (ligne 4). Dans l'ensemble, les productions qui présentent peu d'écart entre productions brutes et nettoyées (moins de 10%) sont celles des sujets des niveaux 2 et 3 (NO, QI, VO), sans doute parce qu'ils se corrigent moins. Au niveau débutant, les productions nettoyées d'AK et XI présentent 40% de mots en moins une fois nettoyées (AK, -40% ; XI -44%). Quelles peuvent en être les raisons ? Les locuteurs L2 produisent une narration avec beaucoup de répétitions et de reformulations illustrant des difficultés avec les processus de bas niveau, comme l'avait déjà constaté Hilton chez des locuteurs natifs<sup>67</sup> (Hilton, 2011a : 152). Ces difficultés semblent limiter, en outre, la possibilité des sujets à planifier la suite de leur discours<sup>68</sup> (Hilton, 2011a : 155). Nous formulons cependant l'hypothèse que certains sujets avancés produisent moins de répétitions ou de reformulations parce qu'ils s'appuient davantage sur les pauses, plus longues et/ou plus nombreuses pour planifier leur discours (processus de haut niveau) ou pour des raisons lexicales (trouver le mot juste). Une autre hypothèse est émise : certains s'appuieraient davantage sur le langage préfabriqué.

Figure 14. Nombre de mots produits nettoyés, t1 et t2

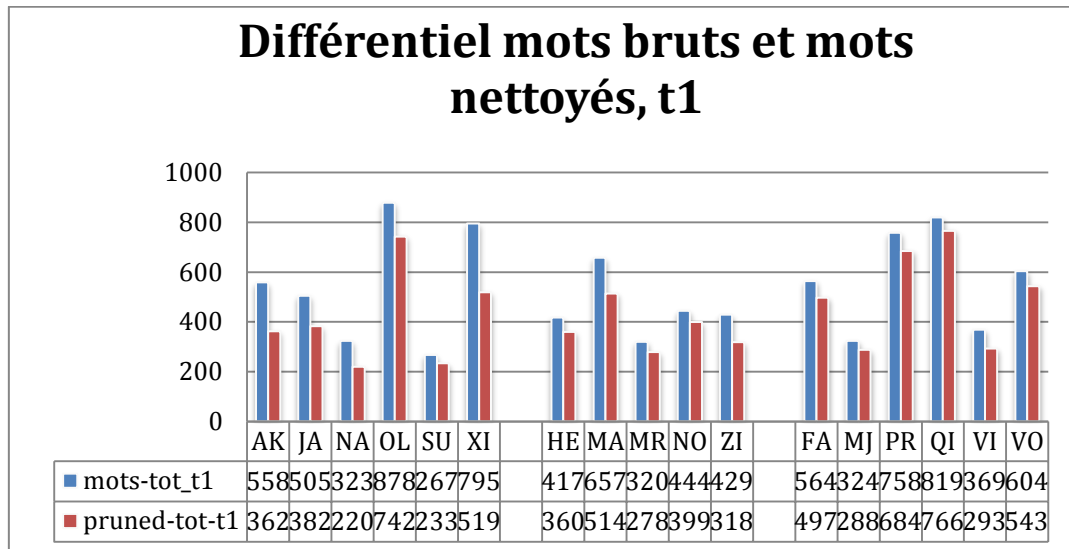


La figure 14 met en regard la quantité de mots nettoyés par sujet t1 et t2. On peut mieux ainsi se rendre compte des progrès dans la quantité de mots produits entre les deux moments. On constate alors que les sujets dans l'ensemble produisent plus de mots la deuxième fois, excepté MA, PR et OL pour les raisons déjà évoquées plus haut.

<sup>67</sup> « The higher proportion of restarts and false starts found in our NS corpus may reflect the fact that speakers explicitly focused on lower-level processes have less short-term capacity available to monitor discourse construction in this fashion. » (Hilton, 2011a : 152)

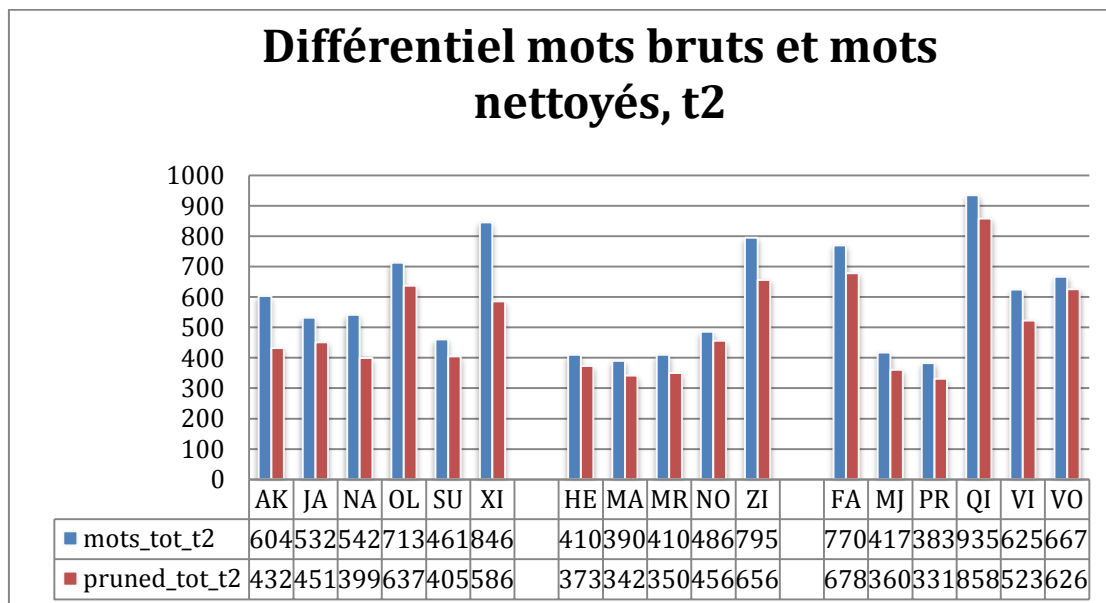
<sup>68</sup> « The attentional effort devoted to lower-level encoding in disfluent speech also appears to limit discursive planning to the serial enumeration of events. » (*ib.* 155)

Figure 15. Nombre de mots produits bruts et nettoyés, t1



La figure 15 montre, en t1, par sujet, la quantité de mots produits bruts ainsi que la quantité de mots nettoyés. On peut se rendre compte des écarts existant entre les sujets. Par exemple, ce sont les sujets qui produisent le plus de mots en quantité brute qui en perdent le plus dans le groupe débutant, c'est le cas de AK, OL et XI, alors que SU qui parle peu en perd moins et semble moins se répéter.

Figure 16. Nombre de mots produits bruts et nettoyés, t2



La figure 16 présente les résultats en t2. Elle montre les disparités entre les sujets, ceux qui parlent moins en t2, MA et PR qui ne se sont pas investies dans la tâche ; OL parle moins en quantité de mots mais de façon plus juste la deuxième fois. Ces résultats sont à mettre en relation avec le nombre de séquences produites et la fluence. Il semblerait que plus la production est longue, moins on emploie de SP. Les productions de certains sujets sont plutôt longues car ils ont du mal à trouver les mots ou les expressions, dans une tâche narrative contrainte qui exigeait un certain vocabulaire. C'est pour cette raison que les sujets débutants comme OL et XI parlent beaucoup, en utilisant des périphrases et aussi en hésitant beaucoup.

### **7.2.2 Mesures de la fluence : résultats du groupe entier**

Afin de montrer comment se développe la fluence, nous présentons les résultats en t1 et en t2 des mesures de la fluence pour tous les locuteurs non-natifs : longueur du segment de parole, vitesse de parole en syllabes par seconde, vitesse d'articulation, taux de phonation, longueur moyenne des pauses, taux d'hésitation, et taux de reprises.

Le tableau 38 regroupe toutes les mesures de productivité de la fluence en t1 et en t2. Il présente les sept mesures de la fluence que nous avons retenues pour le groupe entier en t1 et en t2, la moyenne, l'écart-type et les chiffres minima et maxima pour chaque mesure. La dernière colonne du tableau indique les résultats du test de rang de Wilcoxon appliqué au groupe entier, pour chaque mesure entre t1 et t2. Ces mesures donnent un aperçu de la relative fluence ou disfluence des locuteurs en L2.

Les résultats statistiques des mesures 2 et 7 (vitesse de parole et taux de correction) indiquent des différences significatives entre t1 et t2, représentant les écarts entre les deux moments. Cependant malgré le résultat non significatif de la mesure 1, le segment de parole s'est allongé en t2 pour le groupe entier, et l'écart - entre les extrêmes - reste important. Le taux de reprises, plus bas en t2, indique que les sujets se corrigent moins. Le test de Wilcoxon appliqué au taux de phonation, une des mesures les plus significatives pour Towell, ne montre dans ce groupe aucune différence significative d'une fois sur l'autre. Ces statistiques doivent être interprétées avec prudence car nous constatons entre les chiffres extrêmes des écarts allant parfois du simple au double. C'est donc au niveau des groupes et à l'intérieur des groupes que l'on devrait constater des différences, ce que nos analyses de départ (groupe natif et non-natif) avaient effectivement montré.

	Mesures	moyenne	Écart-type	Mini	Maxi	Valeur de p
1	long moyenne du segment t1	7,00	2,82	3,77	15,93	=0,09 NS
	long moyenne du segment t2	7,65	2,90	4,93	18,02	
2	vitesse de parole t1	2,45	,468	1,65	3,55	<0,01*
	vitesse de parole t2	2,68	,449	1,79	3,67	
3	Vitesse d'articulation t1	3,45	,399	2,54	4,12	=0,140* NS
	Vitesse d'articulation t2	3,7	,490	2,83	4,79	
4	taux de phonation t1	70,60	7,18	61,84	86,30	=0,3318 NS
	taux de phonation t2	72,17	5,29	63,26	87,48	
5	Longueur moy pauses t1	809,04	150,48	594,09	1233,09	=0,3560 NS
	Longueur moy pauses t2	780,09	143,12	622,54	1110,73	
6	Nombre pauses minute t1	21,81	4,18	12,80	30,39	>0,8 NS
	Nombre pauses minute t2	21,60	3,76	11,77	28,66	
7	taux de reprises t1	11,17	4,84	5,37	20,07	<0,01* (=0,0031)
	taux de reprises t2	9,23	3,80	4,12	16,72	

Tableau 38. Mesures de productivité de la fluence, t1 et t2

### 7.2.3 Les variables temporelles

Nous présentons dans le tableau 39 les mesures de la fluence par sujet et par sous-groupe (zones grisées) : vitesse de parole (VP = nombre de syllabes produites par seconde), vitesse d'articulation (VA = nombre de syllabes produites par seconde de parole sans les pauses), taux de phonation (TdP = temps passé à parler), nombre de pauses par minute (NPM), longueur moyenne des pauses vides à partir de 250ms (LMP), et longueur du segment de parole



groupes	sujets	Vitesse de Parole		Vitesse d'Articulation		Taux de Phonation		Nombre de Pauses/Minute		Longueur Moyenne/Pauses		Longueur segment parole	
		t1	t2	t1	t2	t1	t2	t1	t2	t1	t2	t1	t2
Groupe 1, débutant	<b>AK</b>	2,17	2,34	3,41	3,45	63,79	67,70	23,40	22,37	928,53	866,33	5,48	6,17
	<b>JA</b>	2,61	2,64	3,86	4,04	67,67	64,85	23,15	24,24	837,94	870,14	6,62	6,41
	<b>NY</b>	1,87	2,87	3,03	3,95	61,84	72,76	23,12	23,36	990,32	699,75	4,74	7,22
	<b>OL</b>	1,92	2,17	3,04	3,10	63,14	69,95	30,39	28,66	727,79	628,91	3,76	4,49
	<b>SU</b>	2,53	2,76	3,63	3,81	69,54	72,46	21,10	20,55	866,06	803,99	6,90	7,86
	<b>XI</b>	1,65	1,79	2,54	2,83	64,90	63,26	17,08	19,84	1233,09	1110,79	5,73	5,36
	moyenne	2,13	2,43	3,25	3,53	65,15	68,50	23,04	23,17	930,62	829,97	5,54	6,25
	Ecart-type	0,38	0,41	0,48	0,49	2,92	3,93	4,32	3,16	172,60	167,23	1,72	1,22
	Min	1,65	1,79	2,54	2,83	61,84	63,26	17,08	19,84	727,79	628,91	3,76	4,93
	max	2,61	2,87	3,86	4,07	69,54	72,76	30,39	28,66	1233,09	1110,73	6,90	7,86
Groupe 2 Int	<b>HE</b>	2,32	2,89	3,68	4,25	63,18	68,08	24,73	18,25	893,30	1049,45	5,52	9,16
	<b>MA</b>	2,56	2,52	3,38	3,52	75,85	71,62	20,36	20,09	711,77	847,65	7,41	7,31
	<b>MR</b>	2,66	3,45	3,72	4,79	71,47	72,01	20,99	26,80	815,50	626,63	7,34	7,50
	<b>NO</b>	2,24	2,61	3,25	3,63	68,82	72,07	26,35	25,45	709,93	658,66	5,00	6,04
	<b>ZI</b>	2,23	2,62	3,27	3,63	68,18	72,38	24,16	20,25	790,06	818,34	5,42	7,65

	moyenne	2,40	2,82	3,46	3,96	69,50	71,23	23,32	22,17	784,11	800,15	6,14	7,53
	Ecart-type	0,20	0,38	0,23	0,54	4,65	1,78	2,55	3,73	76,94	169,46	1,45	1,11
	Min	2,23	2,52	3,25	3,52	63,18	68,08	20,36	18,25	709,93	626,63	5,00	6,04
	max	2,66	3,45	3,72	4,79	75,85	72,38	26,35	26,80	893,30	1049,45	7,41	9,16
Groupe 3 Avancé	<b>FA</b>	3,55	3,67	4,12	4,19	86,30	87,48	12,80	11,77	642,48	638,26	15,93	18,02
	<b>MJ</b>	2,77	2,72	3,57	3,57	77,48	76,22	22,75	22,92	594,09	622,54	7,05	7,01
	<b>PR</b>	3,03	2,56	3,81	3,52	79,45	72,78	17,65	21,80	698,78	749,31	10,07	6,86
	<b>QI</b>	2,36	2,37	3,03	3,27	78,05	72,54	18,36	19,95	717,29	825,70	7,62	7,05
	<b>VI</b>	2,98	3,10	3,89	4,13	76,74	75,04	18,39	19,97	758,98	749,86	9,32	9,07
	<b>VO</b>	2,25	2,42	3,53	3,19	63,75	75,67	25,96	21,00	837,81	695,23	5,13	6,79
	moyenne	2,82	2,81	3,66	3,65	76,96	76,62	19,32	19,59	708,24	713,48	9,19	9,13
	Ecart-type	0,48	0,50	0,38	0,43	7,34	5,53	4,54	3,99	85,89	76,76	3,74	4,44
	Min	2,25	2,37	3,03	3,19	63,75	72,54	12,80	11,77	594,09	622,54	5,13	6,79
	max	3,55	3,67	4,12	4,19	86,30	87,48	25,96	22,92	837,81	825,70	15,93	18,02

Tableau 39. Mesures de la fluence par groupe et par sujet

Si la fluence s'améliore entre t1 et t2 pour le groupe non-natif entier, on constate des différences importantes à l'intérieur des sous-groupes. Quatre sujets sur six (groupe 1) : AK, NY, OL et SU ont progressé, entre t1 et t2, sur les quatre mesures de la fluence : l'augmentation du segment de parole, du taux de phonation, de la vitesse de parole et d'articulation et la diminution de la longueur moyenne des pauses. HE, MR, NO et ZI, quatre sujets sur cinq du groupe 2, ont aussi progressé entre t1 et t2. Dans le groupe 3, seuls FA et VO ont progressé en t2, sur toutes les mesures de la fluence. Les quatre autres sujets ont des taux de phonation en baisse en t2, pour toutes les différentes raisons déjà évoquées. Cependant et bien que leurs résultats soient en baisse, ils semblent avoir progressé sur les mesures de la préfabrication : c'est le cas de VI qui produit un pourcentage de syllabes préfabriquées en hausse en t2, et dont les résultats présentent un taux de SP et de SP lexicales pour cent mots en progression. FA parle plus vite en t2 mais ne produit pas beaucoup plus de préfabriqués. Les sujets ont donc des profils différents et ne suivent pas forcément le profil que Towell et collègues (1996) ont brossé dans leur étude.

Pour interpréter les résultats entre les groupes, nous avons aussi appliqué le test statistique Kruskal-Wallis.

	mesures	T1		T2	
		H=ddl	Valeur de p	H=ddl	Valeur de p
1	Vitesse de parole	H(2)=9.60	<0,05(m)	H(2)=1.77	>0,5 (NS)
2	Vitesse d'articulation	H(2)=2,43	>0,5(NS)	H(2)=2,16	>0,5 (NS)
3	Taux de phonation	H(2)=7,57	<0,03(m)	H(2)=10,52	<b>&lt;0,01*</b>
4	Nombre de pauses Minute	H(2)=3,07	>0,5(NS)	H(2)=2,05	>0,5 (NS)
5	Longueur moyenne Des pauses	H(2)= 7.95	<0,02(m)	H(2)= 2,25	>0,5 (NS)
6	Longueur du Segment de parole	H(2)=5.77	<0,05(m)	H(2)= 3.32	>0,5 (NS)

Tableau 40. Résultats du test Kruskas-Wallis

Le tableau 40 montre en t1 une différence marginale pour les mesures de la vitesse de parole, du taux de phonation, de la longueur moyenne des pauses ainsi que de la longueur du segment de parole. Ces résultats indiquent qu'il y a peu de différences entre les groupes en t1, seulement des différences marginales voire aucune différence entre les groupes pour les mesures de la vitesse d'articulation et du nombre de pauses par minute. En temps 2, il n'y a aucune différence significative entre les groupes sauf pour la mesure

du temps de phonation ( $p < 0,01^*$ ). Ce résultat indique que le taux de phonation, dans cette étude, est la mesure qui indique le mieux les différences de niveau entre les groupes. Cependant, pour un grand nombre de chercheurs, la longueur du segment de parole serait une des mesures les plus représentatives de l'amélioration dans le temps. Bien que cette mesure ne montre aucune différence significative entre les groupes, on note tout de même un écart entre les groupes (6,25 syllabes en moyenne au niveau 1 et 9,13 au niveau 3). Précisons cependant que les écarts se situent surtout au niveau individuel.

Nous avons appliqué au groupe débutant le test de Wilcoxon pour voir s'il y a des différences significatives entre t1 et t2.

	<b>mesures</b>	<b>valeur de p</b>
1	Vitesse de parole	=,0277
2	Vitesse d'articulation	=,0277
3	Taux de phonation	=,1159
4	Nombre de pauses minute	=,9167
5	Longueur moyenne des pauses	=,0464
6	Longueur du segment de parole	=,1159

Tableau 41. Résultats du test de Wilcoxon

Bien que les résultats présentés dans le tableau 41 montrent une progression de la vitesse de parole et d'articulation du groupe débutant, la différence est seulement marginale pour la mesure 5 (longueur moyenne des pauses). Le résultat de la mesure de la longueur du segment de parole ne montre aucune différence significative entre t1 et t2.

Les comparaisons statistiques (t1/t2) pour les mesures temporelles ne montrent pas de différences au fil du temps pour les groupes 2 et 3.

#### **7.2.4 Comparaison avec les études de Towell et collègues (1996) et Cordier (2013)**

Cordier, dans son chapitre 9 traitant de la relation entre l'usage des SP et le développement de la fluence, compare les résultats de sa recherche avec ceux de Towell *et al.* (1996) sur les quatre mesures de la fluence. Nous complétons son tableau en y ajoutant les résultats de notre étude (tableau 42). Ces résultats méritent certainement un commentaire.

	Cette étude	Towell	Cordier	Cette étude	Towell	Cordier
	T1	T1	T1	T2	T2	T2
Vitesse parole	2,50	2,28	2,15	2,70	2,61	2,93
Vitesse articulation	3,37	3,85	3,40	3,61	4,17	4,11
Segment de Parole	7,00	4,94	4,30	7,65	6,06	6,00
Taux de phonation	70,60	57,74	62,80	72,17	62,51	70,95

Tableau 42. Comparaison entre les études de Towell *et al.* Cordier, et la nôtre

Les résultats de notre étude avec celles de Towell *et al.* (1996) et de Cordier (2013) présentent de grandes différences. Rappelons que l'étude de Towell, Hawkins et Bazergui (1996), montre qu'entre t1 et T2, le taux de phonation et la longueur moyenne des segments de parole s'améliorent. Les résultats chez Cordier, la deuxième fois, marquent une nette augmentation, pour diverses raisons. Cordier propose cinq tâches dont une tâche narrative, une interview et trois discussions sur un sujet d'actualité (alcool chez les jeunes, obésité et un sujet sur l'apprentissage du français). Elle constate que la tâche narrative est la tâche la moins réussie parce que les sujets, plus contraints, parlent moins. La tâche narrative, demande plus d'attention alors qu'une discussion permet plus de liberté. La deuxième fois les sujets, plus à l'aise après un séjour dans le pays de la langue cible, parlent davantage. Ils ne sont contraints ni par le support ni par le sujet. Ainsi, ils peuvent avoir réfléchi à la question, en avoir discuté auparavant. La nature de la tâche joue sans aucun doute un rôle dans l'aisance. Si les apprenants se sentent plus libres, ils parleront davantage et plus vite sans la contrainte du support. Mais il sera plus difficile de comparer deux productions de ce type. Nous pensons qu'effectivement la tâche narrative est plus contraignante car les apprenants font a) un effort de mémoire pour rapporter le plus précisément possible ce qu'ils ont vu b) un effort de conceptualisation pour produire un discours clair et juste c) un effort lexical pour rendre compte des différentes situations de la narration. Il paraît donc important de le préciser pour comprendre les résultats différents d'une étude à l'autre.

Towell *et al.* et Cordier ont des sujets anglophones, étudiants d'un même groupe à l'université dans un programme de français, c'est donc un groupe homogène. Les sujets (Cordier, 2013) sont enregistrés en temps 2 après un séjour de sept mois dans le pays de la langue cible. Towell et ses collègues, ont prévu un troisième enregistrement dans la LM des sujets, en l'occurrence l'anglais, 2 ans plus tard.

Contrairement aux deux études citées, notre groupe d'apprenants n'est pas un groupe homogène dans la mesure où :

i. ils ne suivent pas tous le même cursus, le groupe d'Albi comprend des doctorants en sciences de l'ingénieur, certains ayant obtenu une bourse pour préparer leur thèse en France, d'autres étant venus en stage pour une année seulement ; le groupe de Toulouse est composé de sujets venus en France pour diverses raisons, études, ou compagnon français, ils suivent à l'université de Toulouse-Jean-Jaurès, au département de français langue étrangère (DEFLE) un programme de cours d'une vingtaine d'heures par semaine, Ils visent l'obtention d'un diplôme qui leur permettra de préparer un diplôme français ou de trouver un emploi.

ii. Nos apprenants sont de nationalités diverses et ont des L1 parfois très éloignées du français comme les étudiants asiatiques (trois chinois et un japonais), mais la plupart d'entre eux maîtrisent une autre langue occidentale. Les doctorants, par exemple, utilisent l'anglais comme langue véhiculaire et parfois même communiquent en anglais dans leur laboratoire. Cependant, ils ont tous pour objectif leur thèse, et parlent français aussi avec les camarades français, leur directeur de thèse et au quotidien. Le groupe d'Albi reçoit seulement trois heures de cours de FLE par semaine dont le but est de leur permettre de se débrouiller dans la vie de tous les jours.

iii. Les études de Towell *et al.* (1996) et Cordier (2013) ont pour sujets des apprenants de niveau avancé. Nos apprenants constituent un groupe hétérogène ; ils n'ont pas tous le même niveau à leur arrivée à l'école. Certains n'ont eu qu'une première approche de la langue quelques semaines seulement avant leur arrivée en France.

### **7.2.5 Interprétation de l'augmentation de la fluence en relation avec le modèle de Levelt et les théories de l'automatisme.**

Les mesures des variables temporelles sont un moyen de mieux appréhender les processus cognitifs en jeu dans la production orale. Selon Towell *et al.* (1996), l'augmentation de la vitesse de parole est un indicateur de l'automatisation des processus de la production de parole. Selon ces chercheurs, il faut examiner attentivement les mesures de la fluence pour voir quels types de traitement s'automatisent. L'augmentation du segment de parole sans qu'il y ait une diminution du taux de phonation et une diminution de la longueur moyenne des pauses peut indiquer que la procéduralisation a eu lieu. Towell *et al.* (1996) considèrent cette mesure comme la plus importante, et l'interprètent comme la preuve de traitements automatiques, au niveau du formulateur, plutôt que dans le conceptualiseur ou l'articulateur. Ils l'expliquent par l'automatisation des traitements formels et non par la façon dont les sujets conceptualisent et articulent en L2.

#### *Conclusion*

Dans cette étude, les deux tâches narratives choisies – un cartoon (BD) composé de huit vignettes et un petit film muet - sont plus contraignantes qu'une tâche orale plus libre, interview ou discussion. Pour comparer finement les productions dans le temps, il nous paraissait essentiel d'utiliser le même support, qui permet de comparer des objets

comparables. Cette tâche demande une certaine concentration de la part des locuteurs : ils doivent décrire ce qu'ils ont vu en respectant la trame narrative et la chronologie - elle demande donc un certain effort cognitif (mémorisation, explicitation, description). Comme les apprenants n'ont pas tout le vocabulaire disponible en MLT, certains peuvent employer des SP ou des expressions pour pallier une difficulté lexicale. Même si les résultats d'ensemble montrent une progression en t2, le groupe débutant est celui dont les résultats augmentent le plus, en quantité et en qualité de SP, en valeur absolue comme en occurrences types. A contrario, les résultats du groupe avancé ne sont pas meilleurs en t2. Dans ce groupe, le sujet FA est déjà arrivé à un stade de maîtrise avancé et, de ce fait, n'a pas profité autant que VI qui a progressé très vite après quelques mois passés en France. Néanmoins, les résultats d'ensemble du groupe entier montrent que l'usage des SP s'est développé en temps 2 et on peut relier ces résultats à ceux de la fluence.

Voici nos réflexions

Les apprenants en début d'apprentissage s'appuient davantage sur le LP car ils ont peu de vocabulaire. Ils disposent, grâce au cours de langue, d'un répertoire de séquences discursives qui leur permet de pouvoir s'exprimer avec un minimum de vocabulaire. Ils s'appuient largement sur les SP qu'ils répètent plusieurs fois dans la même narration ; et on constate plus de SP en t2 même en occurrences types.

Les apprenants avancés, qui ont un répertoire de SP lexicales plus riche que celui des débutants, n'en produisent pas plus la deuxième fois, peut-être par souci de précision, désir de mieux faire, ou pour certains par lassitude ou perte de motivation. Les variables temporelles montrent cependant que le segment de parole est plus long la deuxième fois, la vitesse de parole également mais que le temps de pause est aussi plus long. Le comportement des sujets avancés est surtout variable, certains sujets désirant se montrer plus précis et créatifs semblent moins s'appuyer sur les SP. Le souci de précision l'emporte donc sur la vitesse de parole, d'où un taux de phonation en baisse pour certains. Pour d'autres, FA mais surtout VO il augmente (de 63,95 à 75,67), ce qui peut être interprété comme effectivement une plus grande automatisation des processus langagiers.

Le manque de motivation de certains sujets la deuxième fois peut expliquer aussi la « stagnation » des moyennes des groupes avancé et intermédiaire.

MA et PR ont beaucoup parlé en t1 et très peu en t2, et leurs performances expliquent les résultats équivalents (t1= t2) obtenus par le groupe entier. La longueur des segments de parole

en syllabes passe de 9,19 à 9,13 alors que les minima et maxima sont plus élevés en t2 (5,13-15,93 la première fois et 6,79-18,02 la deuxième). PR a des résultats en baisse sur toutes les mesures de la fluence (TdP -7.8%, segments de parole -31.9%, VP -15.5% et VA -8.3%). MJ, malgré un taux de phonation plus élevé en t2, ne produit pas de segments plus longs ; la vitesse de parole est en très légère baisse ainsi que la vitesse d'articulation. La baisse du taux de phonation et de la longueur du segment de parole constatés, chez QI et VI, peuvent aussi être attribués au désir d'être plus précis, peut-être plus créatifs la deuxième fois. FA et VO progressent le plus, ils produisent des segments de parole plus longs et, si le taux de phonation de FA reste stable, celui de VO augmente de plus de 9%.

En conclusion, nous dirons qu'il est difficile d'émettre des généralisations au niveau des groupes. Seuls les sujets du groupe débutant progressent très nettement dans l'emploi des SP qui, à leur stade, ont le grand avantage de faciliter le traitement tout en leur permettant d'être plus fluent. L'usage des SP permettrait de pouvoir entrer en communication très vite, même si elles deviennent rapidement insuffisantes pour avancer en communication. A des stades plus avancés, l'usage des SP permettrait de libérer du temps pour un usage plus créatif de la langue, il permettrait ainsi aux sujets les plus avancés d'être plus précis dans une tâche de narration

En conclusion, les changements entre t1 et t2 sont confirmés pour le groupe entier des non-natifs ; néanmoins à l'intérieur des sous-groupes, force est de constater des variations au niveau individuel.



## Chapitre 8 Analyse qualitative

---

Les résultats présentés dans les chapitres 6 et 7 ainsi que les tests statistiques ont confirmé la réduction des écarts entre les groupes entre t1 et t2 et, la progression des locuteurs de niveau débutant. Nous constatons, dans l'ensemble, une évolution dans le nombre de SP produites pour le groupe entier des locuteurs non-natifs et une augmentation des séquences préfabriquées lexicales, en termes quantitatifs. Le discours narratif des sujets s'est diversifié, entre novembre et juin en termes de qualité grâce à un emploi plus diversifié des SP (augmentation du nombre des SP en occurrences types) et des SP lexicales en particulier. On note, parallèlement à l'augmentation des SP lexicales une diminution des séquences discursives en t2. La répartition entre SP lexicales, grammaticales et discursives est semblable en proportion en t2 à celle des natifs. Nous avons constaté également un lien évident entre l'usage des SP et la fluence. En effet, les SP permettent aux locuteurs de parler avec plus de facilité et d'aisance.

Dans ce chapitre, nous verrons que les SP jouent plusieurs rôles dans le discours. En effet, si utiliser des SP permet de parler avec plus d'aisance en évitant de construire son discours mot par mot, le plus souvent elles servent à allonger le segment de parole mais aussi à structurer la narration et à libérer du temps pour la recherche lexicale et/ou la planification. L'utilisation des SP serait donc une stratégie personnelle et « efficace » des locuteurs L2 pour parler mieux et plus vite (Rebhein, 1987 :103 ; Raupach, 1984).

D'autre part, l'analyse des autocorrections (Théophanous, Perez-Bettan et Hilton, 2014), et notamment celles qui concernent les SP, montre que les erreurs peuvent être aussi révélatrices d'une stratégie des apprenants pour parler davantage. L'autocorrection permettrait à l'apprenant non seulement de réparer ses erreurs mais aussi d'améliorer sa fluence verbale.

L'amélioration de la fluence verbale ne serait alors pas tellement (ou pas seulement) une question quantitative au sens de réduction de pauses ou de rapidité de la production mais plutôt une question qualitative au sens d'allongement et de complexification du discours produit. (Théophanous, Perez-Bettan et Hilton 2014 : 203)

### 8.1 Des SP discursives en grand nombre

En comparant les résultats t1 et t2, nous constatons qu'après six mois, les sujets produisent plus de SP, autant en nombre absolu qu'en occurrences types. Chez les locuteurs débutants, ce sont surtout le nombre de SP Lexicales qui a augmenté le plus en t2, la parole de nos

apprenants étant ainsi plus riche et plus variée après six mois. Les études reconnaissent que les apprenants ont tendance à faire un emploi plus large de SP discursives, surtout en début d'apprentissage (Fitzpatrick et Wray, 2006 ; Forsberg, 2008), lorsque les SP lexicales sont encore peu connues.

La répartition des SP, comparable en proportion à celle des natifs, montre que les SP grammaticales et discursives sont moins présentes en t2 (Figure 5. page 191), et en particulier les SP discursives. Il semblerait donc que l'acquisition du langage préfabriqué évolue de la préfabrication discursive vers la préfabrication lexicale, ce que Forsberg (2006) avait déjà observé lors d'entretiens semi-dirigés d'apprenants suédophones de FLE. Cependant, il paraît tout à fait normal que les SP discursives soient plus nombreuses (en proportion) en t1 que les SP lexicales. En effet, les SP discursives constituent une classe fermée et présentent un plus grand degré de figement. Peu nombreuses, elles peuvent être acquises rapidement. Elles sont d'ailleurs entendues et utilisées de façon répétée en cours de langue et en milieu naturel ; il est donc normal que les apprenants s'appuient sur elles en production orale. Contrairement aux SP discursives, les SP lexicales, représentant une classe ouverte, tendent à une plus grande variabilité d'un point de vue structurel, dans l'actualisation (*toutes mes/nos félicitations, de mon/ton/son point de vue*), l'extension (*il est [déjà] 8 heures [précises, pile] ; [tout à fait d'accord]*), et le complément libre ou *open-slot* (défense de [*se pencher/fumer/stationner*]), (Erman et Warren, 2000 : 40-41). De ce fait, les SP discursives seraient maîtrisées plus rapidement, leur rôle étant de permettre aux apprenants de maintenir leur fluidité verbale.

Dans les exemples 1 à 9, (SP discursives en caractères gras), on note le suremploi de certaines séquences telles que *et après, comment dire (\*comment dit), peut-être* ainsi que des formes verbales avec pronom indéfini, *on voit, on a, on a (déjà) vu, on peut* (imaginer), ou à la première personne chez les débutants, *j'ai (juste) vu*, qui servent à organiser la narration tout en maintenant la parole. Certaines SP discursives, comme celles de la catégorie des remplisseurs (*fillers*), sont également très présentes (*je sais pas, j'sais pas, j'suis pas sûr, je pense*). La formule *c'est* redoublée, *c'est c'est* (ex. 9), appartient aussi à la catégorie des SP discursives. Notons que Forsberg (2006) classe toutes les occurrences de *c'est* (dont *c'est + N* est la plus courante) dans la catégorie des discursives alors que Cordier (2013) les classe dans la catégorie des grammaticales, celle des *sentence-builders*, qui servent à construire. Nous développerons dans le chapitre 9 une section consacrée aux différents emplois de la séquence *c'est*.

(ex. 1) **et après** qu' il y: (.) et les autres situations (.) et mais **je ne sais pas comment faire** pour écrire pour décrire la situation (.) e:::t (.) hum je pense que **c' est ça** (JA, bd1)

(ex. 2) mais alors qu' il dort eu:h (.) il vien:t euh un un p'tit garçon (.) avec un: p'tit cheval un p'tit animal **j' sais pas** (.) c' est un chien euh je (.) **j' suis pas sûr** mais (.) (VI, bd1)

(ex. 3) **comme ça** le le l'abeille sort par par \*le baignoire (.) euh on et (.) **et voilà** elle est elle est hors de de la maison (VI, F1)

(ex. 4) donc **je vois** un dessin (.) qui raconte **peut être** un une petite blague (.) e:::t **il y a** (.) sur **le premier plan** il y a: (.) un un homme qui (.) (VO, bd1)

(ex. 5) euh **j' ai vu un j' ai vu** un homme qui est euh (.) sur un cheval euh (.) il il veu:t il veu:t se (.) se:: &reposer (.) sur eu :h (.) sur un:: un (xxx) (.) **j'sais pas** (XI, bd1)

(ex. 6) euh **j' ai vu** un film **je pense** c' était une animation **il y a** une\* (ZI, F1)

(ex. 7) **et après** eu::h (.) hum le petit enfant il veut échanger le (.) petit petit maquette de cheval et cheval euh un cheval &re réel (.) **et après** euh \***comment dit** (.) \***comment dit** le &che un un cheval réel il sort (.) e:::t petits enfants ils (.) il est sort aussi (.) **et après** euh

(ex. 8) **voilà donc** euh c' est une série euh comme **on a déjà vu** l' année dernière **on a** différentes scènes (.) et dans une: contexte un peu : rural donc eu:h **on voit** une: (.) un chevalier (.) un cowboy en anglais **je sais** (.) qui : qui : semble fatigué et il arrive : à côté d' une maison **on peut** imaginer que ça soit sa maison (.) (FA, Bd2)

(ex. 9) soudain eu:h **comment dire** (.) le monsieur\* (.) il **peut-être** cowboy (.) cowboy et (.) il a recommencé juste: (.) il avec euh petit maquette de cheval e:t (.) **c'est c' est** (.) pro produit un peu triste **peut-être** (.) oui &=rire (.) c'est marrant (AK, Bd2)

## 8.2. SP et fluence verbale

Dans cette section, nous examinerons comment les SP (soulignées en caractères gras) contribuent à l'allongement du segment de parole. Tout d'abord le segment de parole se trouve allongé par la simple répétition consécutive d'une même SP dans le segment (ex. 10 à 14). Ensuite une même SP peut être répétée tout au long de la narration, elle sert alors d'appui au locuteur (ex 15 et 16). Enfin, les multiples SP dans un même segment (ex. 18 à 20) rendent le discours plus fluide en évitant de marquer des pauses. Les SP servent aussi à faire des commentaires, ce qui contribue également à l'allongement du segment de parole. Elles jouent un rôle métadiscursif (ex. 21 et 22) tout en permettant à l'apprenant de parler plus.

### 8.2.1 Répétition de la même SP dans un même segment

L'emploi répété de la même SP dans un même segment permet d'allonger le segment de parole tout en réduisant le nombre ou/et la longueur des pauses. Notons ici que la répétition d'une même SP dans le même segment permet de gagner du temps afin de trouver le mot recherché (ex. 10 à 14). Parfois une pause est nécessaire (ex. 10).

(ex. 10) il a **tant pis tant pis** non il a **tant pis** non il est (.) il est /il est désolé il est navré (AK, t1, niv1)

(ex. 11) il est **un peu un peu peut-être un peu** impressionné par (.) (FA 2)

(ex. 12) **il y a** un cavalier j'sais pas s'**il y a** un mot en français pour dire jinet (.) et qu'il **est très fatigué très très fatigué ...** (MR,( t1), niv2)

(ex. 13) il y a un garçon **un p'tit un p'tit** garçon et il a un **\*une p'tit** cheval... (ZI, (t1), niv2)

(ex. 14) **à côté de** la maison mais **c' est** pas la la fleur de la place **à côté de** la maison **c' est** la fleur de de une image (.) (JA, (t1), niv)

AK (ex. 10) répète trois fois la SP **tant pis** qui lui laisse le temps de trouver la bonne formulation et d'allonger le segment de parole. AK ne trouve pas le mot juste pour décrire l'état du personnage qui se retrouve avec un jouet à la place de son cheval. La SP *tant pis* qu'il répète trois fois est un moyen/une stratégie lexicale qui lui permet de décrire la consternation du cowboy et son renoncement à vouloir récupérer son cheval mais sert aussi à trouver la bonne formulation (*il est désolé, il est navré*). De même, dans l'exemple suivant (ex. 11), la répétition de la SP (*un peu*) permet à FA de trouver le mot juste (*impressionné par*).

JA, niveau débutant, (ex. 14) répète deux SP dans le même segment, *c'est* et *à côté*, qui lui permettent ainsi de chercher le mot juste tout en gardant la parole. Ces temps de planification (ex. 10 à 14) servent aussi à formuler un énoncé plus long dans le même segment de parole.

### 8.2.2 Répétition de la même SP dans plusieurs segments

Une même SP répétée tout au long de la narration allonge également le segment de parole (ex. 15). La SP *comment dire* fonctionne ici comme îlot de sécurité (Dechert, 1984), sur laquelle s'appuie l'apprenant et qui lui permet de poursuivre sa narration tout en planifiant la suite. Les îlots de sécurité sont des expressions, des phrases que l'apprenant garde en mémoire et qui permettent une meilleure fluence.

(ex. 15) (AK, (t2), niv1)

c'est une salle de bains il y a **comment dire**...

**comment dire** (.)

**comment dire** c'est &trou \*trouv   trouv   (.)

il est sorti sur euh le : : **comment dire**...

On trouve onze occurrences de la SP *comment dire* en t2 dans les productions d'AK (5 occurrences dans la narration du cartoon dont 2 d  viantes \**comment dit*, et 6 occurrences dans la narration du film), alors qu'en t1, on n'en trouve aucune. La s  quence *comment dire* semble   tre une s  quence utilis  e apr  s un temps d'exposition    la langue. Elle permet de prendre le temps de trouver le mot juste, la bonne formulation et sert d'appui pour ne pas « bloquer » la narration. Chez AK, elle devient presque un tic de langage.

(ex. 16) (QI, (t2), niv3)

hop    la fin de l'apr  s-midi il arrive dans un petit village **il dit que** (.)

un banc sur contre un mur il me &f **il dit que** ah c'est le meilleur (.)

dans sa t  te il **il dit que** ah je peux rigoler un peu je je vais faire une b  tise (.)

QI (ex. 16) r  p  te    plusieurs reprises la s  quence *il dit que*, qui rapporte les paroles imaginaires du personnage du cartoon. Elle s'appuie sur cette s  quence pour garder la parole et faire le lien entre son r  cit et les paroles et pens  es du personnage.

### 8.2.3 Succession de plusieurs SP dans le m  me segment

La fluence s'am  liore lorsque l'apprenant est capable d'encha  ner les SP l'une apr  s l'autre en cr  ant ainsi des segments de parole plus longs (ex. 17    22). NY qui, la premi  re fois (ex. 17), n'arrivait pas    parler et elle en avait conscience (*je ne sais pas parler*) - elle ne pouvait que produire quelques syllabes - est capable la deuxi  me fois (ex. 18) de produire un long segment de parole sans pauser en encha  nant plusieurs SP (*trouver le chemin, sortir de, et apr  s, elle est* (heureuse), *tr  s tr  s, parce que, elle est, dans la nature, une autre fois*). Les six SP produites en F2 allongent le segment de parole de mani  re cons  quente (25 syllabes). On constate aussi un   nonc   plus riche gr  ce aux deux SP qui marquent une circonstance (temps, cause) et qui complexifient l'  nonc   (trois propositions, dont une coordonn  e et une subordonn  e). Cet apprenant d  butant r  ussit    produire un   nonc   complet sans aucune disfluence, et cela gr  ce    l'utilisation qu'elle fait des SP. Notons que lorsque le locuteur ne peut produire que des suites de mots suivies de pauses ou d'h  sitations sans pouvoir les

enchaîner (ex. 17), c'est-à-dire sans produire un énoncé fluide et complet, son discours peut produire sur l'auditeur une impression de disfluente. Et cela peut même empêcher ces locuteurs de développer une fluence native<sup>69</sup> (Rebhein, 1987 :104).

(ex. 17) maintenant elle est elle est ne (.) n' est vraiment lio (.) elle ne (.) **je n' sais pas parler**(.) quand \*je quand \*je quand elle va sortir (.) **je ne sais pas parler** (NY, (t1), niv1)  
(ex. 18) elle **a trouvé le chemin pour sortir de la maison et après elle est très très** heureuse **parce qu'elle est très dans la nature une autre fois** (NY, (t2), niv1)  
(ex. 19) des des **années soixante / peut-être / à cause de / il y a / un petit** xxx dans une maison l'endroit où où il garde son cheval (HE (t2) niv2)  
(ex. 20) **et disons qu'au début j'étais un peu curieux de** savoir euh **le pourquoi de** de ce titre, (FA (t1) niv3)  
(ex. 21) le film **commence par** une abeille **on voit** une abeille qui vole **à l'extérieur** donc euh on **du côté des** arbres des fleurs et (.) et **juste après on commence à** voir une maison donc une **maison de campagne** (.) (FA (t2) niv3)  
(ex. 22) mais euh **je pense qu' il est un peu fatigué** alors **il décide de : s' endormir un peu** (.) (VI (t1) niv3)

On remarque (ex. 19) l'enchaînement de plusieurs SP chez HE, dont l'emploi syntaxiquement incorrect de la SP *à cause de* qui aurait dû être suivie d'un GN. Néanmoins, l'apprenant ne fait pas d'erreur sémantique. Il a apparemment activé la SP mémorisée comme une séquence introduisant la cause. Ce sujet n'a pas la possibilité, à son stade et dans les conditions de la production orale, de la modifier ou d'en utiliser une autre plus juste dans le contexte, ici on aurait dû avoir « *parce qu'il y a* ». Les deux débuts de narration de FA (ex. 20 et 21) sont différents mais on voit bien que cet apprenant avancé est à l'aise les deux fois. Il emploie en t2 plus de SP pour l'organisation spatiale - *à l'extérieur, du côté de* - l'organisation chronologique - *commence par, juste après* - et une SP lexicale, le mot composé *maison de campagne*.

Dans cet autre exemple (ex. 22) provenant d'un apprenant de même niveau, on constate que les nombreuses SP bien organisées structurent l'énoncé tout en allongeant le segment de parole. VI répète la SP *un peu* deux fois dans le segment. La répétition des mêmes SP est aussi, on l'a vu, une stratégie pour allonger le segment et donne également du temps au locuteur pour la planification des énoncés. On distingue dans ce segment plusieurs ensembles

<sup>69</sup> « However, speech formulae can also prevent learners from developing nativelike fluency. » (Rebhein, 1987 :104)

prononcés d'une traite *je pense qu'il est un peu fatigué* (3 SP), *il décide de* : (1SP), *s'endormir un peu* (2 SP). VI en t1 utilise une stratégie compensatoire pour remplacer *faire la/une sieste*, l'utilisation de deux SP *s'endormir* et *un peu*. Le premier ensemble est une phrase complète, qui peut être comparée à celles produites par les natifs. Forsberg et Fant (2010 : 59) distingue dans les SP lexicales, les syntagmes (*prendre la route, mortelle randonnée*) et les phrases lexicales (*mettre qqn au courant, je vous remercie de votre compréhension*). La grande différence entre les locuteurs natifs et non-natifs serait la propension des natifs à utiliser un plus grand nombre de ces phrases lexicales, ce que certains locuteurs du groupe avancé comme FA sont aussi capables de produire à leur niveau.

#### 8.2.4 Amélioration de la fluence grâce à une utilisation de type méta-discursif des SP

L'utilisation de type métadiscursif des SP (ex. 23 et 24), qui permet de faire des commentaires, aide les locuteurs à parler plus longuement ; elle améliore ainsi la fluence (Allwood *et al.* 1989).

(ex. 23) *la deuxième partie c'est / un peu / différent parce que...* (FA, (t1), niv3)

(ex. 24) *j' suis pas sûr (.) c'est c'est /c'est drôle....c'est /c'est quand même drôle* (VI, (t1), niv3)

FA utilise des SP discursives (ex. 23) *la deuxième partie, parce que* qui servent à organiser, entre des éléments du discours narratif (*la deuxième partie*) et entre les groupes syntaxiques (*parce que*). La SP lexicale, *c'est différent*, donne le point de vue du narrateur FA sur le contenu du film. On note la succession de pas moins de quatre SP qui constituent un segment entier.

VI (ex. 24) produit une SP discursive *j' suis pas sûr* qui joue deux rôles, maintenir la parole et préparer la suite, en nuancant son interprétation d'une des séquences du film. VI utilise des SP pour commenter *c'est drôle c'est quand même drôle*, avec la répétition de *c'est drôle* qui joue le rôle d'une formule d'insistance. La double SP discursive *c'est c'est* permet de produire la SP lexicale *c'est drôle* d'une traite.

#### 8.2.5 Répétition du dernier mot de la SP

(ex. 25) *c'est l'histoire de / d'un homme que (.)* (VI, (t1), niv3)

(ex. 26) *il s'assure que l'homme il dort parce qu'il le touche avec un petit bout de (.) de / de bois (.)* (VI (t2) niv3)

(ex. 27) *s'est transformé \*dans /\*dans un chien* (VI (t2), niv3)

(ex. 28) *il a envie de / de (.) de / de lui emprunter son cheval* (VI (t2) niv3)

(ex. 29) *et il voit que au lieu de / de son cheval à lui (.)* (VO (t1) niv3)

(ex. 30) *et au lieu de:: / de son truc il met le (.)* (VO (t1) niv3)

Les SP (ex. 25 à 30), produites en bloc, sont parfois suivies de la répétition du dernier mot de la séquence. Quand le locuteur n'a pas la possibilité de traiter la séquence et la suite de l'énoncé à ce stade, ou s'il n'a pas planifié la suite, il marque une pause après la SP qui se traduit par la répétition du dernier mot de la séquence, ce qui lui permet de réfléchir à la bonne formulation (impossibilité en temps réel d'appliquer la règle de l'élision par exemple). On se trouve dans l'étape d'auto-surveillance (*monitoring*) du modèle de Levelt où le locuteur peut planifier la suite grâce au temps libéré que permet la répétition.

Dans le premier exemple (ex. 25) la SP *c'est l'histoire de* est dite en bloc, le traitement s'est fait de manière automatique ; VI n'a pas pu appliquer consciemment la règle de l'élision. Ces formes non élidées sont aussi un moyen de vérifier la présence d'une séquence préfabriquée. Les SP dont le dernier mot est répété confirment que la séquence est bien produite en bloc.

Le dernier mot de la SP peut être répété deux fois (ex. 28), juste après la SP dans le même segment et une autre fois dans le segment suivant. Après avoir produit la SP d'une traite et si le sujet n'arrive pas à poursuivre, il s'appuie sur la répétition du dernier mot de la SP qui peut aussi être suivie d'une pause vide et d'une autre répétition afin de produire le mot juste dans le segment suivant. On constate que les SP sont souvent suivies d'une pause vide ou remplie, d'un allongement ou d'une répétition du dernier mot de la séquence. Comme la séquence est dite en bloc, et qu'elle permet de planifier la suite, le locuteur répète le dernier mot de la SP pour arriver à poursuivre. C'est aussi un bon moyen, efficace, d'identifier une SP.

La dernière syllabe d'une séquence peut aussi être allongée. VO (ex. 30) s'appuie sur la dernière syllabe de la séquence pour gagner du temps afin de planifier la suite. L'allongement est renforcé par la répétition du dernier mot de la séquence. Mais ne trouvant pas le mot juste, elle produit « truc », mot passe partout que l'on trouve aussi dans le discours natif en général. D'autres stratégies peuvent être utilisées en cas de difficulté lexicale, comme *petit jaune* pour *abeille* que OL ne connaît pas.

Mais comme nous l'avons dit plus haut, si les séquences sont produites et suivies de pauses, surtout à un niveau débutant, un natif peut avoir l'impression d'entendre un énoncé haché et disfluent, même s'il pourrait être considéré comme fluent d'un point de vue strictement temporel (Rebhein, 1987 :104).



### 8.2.6 Les autocorrections

Nous avons relevé dans notre corpus un certain nombre d'autocorrections. Les apprenants se corrigent lorsqu'ils se rendent compte d'une erreur, de type lexical ou morpho-syntaxique (voir Théophanous, Perez-Bettan et Hilton, 2014). Cependant, « l'autocorrection n'est pas systématiquement imputable à l'erreur, notamment lorsque le locuteur se corrige pour redire la même chose en utilisant des expressions synonymiques » (*ib.* 203) ou des SP. Nous donnons ici des exemples tirés de notre corpus.

(ex. 31) il raconte → il s'agit d'une histoire (VO, F1)

(ex. 32) et il retourne à la bain → à la salle de bains (ZE, F1)

(ex. 33) je ne sais pas comment faire pour écrire → pour décrire la situation (JA, Bd1)

(ex. 34) elle cherche le la bonne fleur pour qu'elle trouve le euh bon → pour qu'elle puisse fabriquer du miel (VI, F1)

(ex. 35) qui fait du &vél pas du &vé → qui fait du cheval (VO, )

(ex. 36) est euh curieux → est curieuse (ZE, F2)

(ex. 37) une maison belle → une belle maison (ZE, F2)

(ex. 38) il voit rien parce qu'il a mis son chapeau sur sa &t → devant ses yeux (MJ, F1)

Dans les exemples 31 à 34, l'apprenant se corrige et produit une SP à la place d'une unité simple (*il s'agit de, salle de bains, décrire la situation, fabriquer du miel*), il dit la même chose de deux façons différentes, l'expression est plus précise et plus juste. L'apprenant n'a plus besoin de pauser après le verbe et l'expression est dite en une fois (ex. 32), ce qui donne une impression de meilleure fluidité.

Dans les ex 35 à 38, l'apprenant se corrige a) parce qu'il a fait une erreur en produisant une SP incorrecte dans le contexte (ex. 35, *faire du cheval* et *non du vélo*), b) parce qu'il a fait une erreur de genre, (ex. 36 *est curieuse* au lieu de *curieux*) c) une erreur dans l'ordre des mots (ex. 37 *belle maison* et non *\*maison belle*), d) parce qu'il a changé d'idée (ex. 38) : la séquence *mettre son chapeau sur sa t(ête)* est venue en bloc, l'apprenant la modifie en l'adaptant au contexte. Les erreurs peuvent alors être une source d'informations sur la préfabrication comme dans l'exemple 38.

Seul exemple de ce type dans le corpus, l'apprenant MJ interrompt la séquence *mettre son chapeau sur la t(ête)* produite de manière automatique, car elle se rend compte de son erreur,

et qui ne convient pas à la situation illustrée par la vignette correspondante (le chapeau est sur le visage et non sur la tête). Cette suite de mots, arrêtée en cours de route, montre bien comment se fait la récupération de la séquence activée en bloc. De même dans l'exemple 35, VO se reprend, s'autocorrige en modifiant la SP beaucoup plus fréquemment utilisée (*faire du vélo*) par une autre SP, cette fois, mais qui est conforme au dessin (*faire du cheval*). Ce type d'erreur peut être un indicateur précieux de la nature « ballistique » de l'activation d'une expression préfabriquée.

Les reprises de parole peuvent, dans certains cas, être dictées par le souci de précision et/ou de correction mais aussi par le désir de maintenir le flux langagier. En effet, l'autocorrection permet à la fois de rectifier ses paroles en cas d'erreur, de les préciser mais aussi d'améliorer sa production en termes de fluence verbale et d'aisance (Théophanous *et al.* 2014). D'une manière générale, les autocorrections aident l'apprenant à parler plus couramment et, notamment, les séquences préfabriquées car « intervenant dans l'autocorrection, en accompagnement de la version rectifiée et/ou à mi-chemin entre la version initiale et la version rectifiée, [elles] allongent le segment de parole ce qui, à son tour, indique une meilleure fluence » (Théophanous *et al.*, 2014 : 205).

### **8.2.7 La réduction phonologique**

Conformément à d'autres chercheurs (Regan, 1997, 2003), on remarque que la préfabrication semble avoir un effet sur la suppression de la particule négative *ne*, notamment chez les apprenants de niveau avancé qui ont tendance à produire des séquences avec réduction phonologique : *je sais pas, j'sais pas, il y a pas, y a pas, c'est pas*. On trouve également cette suppression dans des expressions comme *il comprend pas, il voit pas que*, qui confirmerait que la séquence est préfabriquée. En effet, « la préfabrication permet une réduction syllabique qui rend l'élocution plus rapide tout en indiquant un traitement holistique de la langue » (Théophanous et Perez-Bettan, 2012 : 261). On trouve également une réduction syllabique dans les expressions avec *petit* surtout à un niveau avancé comme d'ailleurs chez les natifs, *p'tit garçon...* nous développerons ce point dans le chapitre 9.

### **8.2.8 Les marqueurs discursifs**

Dans cette catégorie, on trouve les marqueurs textuels (« organisateurs » selon la terminologie de Raupach, 1984), ex. *c'est, il y a, et comme ça* et les marqueurs autorégulateurs

(« remplisseurs » selon Raupach, 1984), par exemple *je pense, en fait, on peut dire*. Un grand nombre de ces séquences appartiennent spécifiquement à l'oral (Erman et Warren, 2000 : 43). Nous présentons dans cette section certains marqueurs discursifs avec *et* que l'on trouve dans le corpus, ainsi que la catégorie discursive des remplisseurs qui servent à gagner du temps et à maintenir la parole.

*a) Les marqueurs textuels ou organisateurs avec et*

Le tableau 43 présente le nombre d'occurrences des séquences discursives qui font partie des organisateurs avec *et* (+ adverbe, conjonction, mot coordonnant) dont la plus courante est *et après*.

SP	t1	t2	natifs
et après	51	50	1
et donc	9	18	6
et maintenant	2	2	0
et ensuite	1	1	4
et là	2	3	4
et puis	1	4	3
et mais	3	2	0
et quand	2	2	0
et alors	0	1	1

Tableau 43. SP discursives avec *et*, groupe natif et non-natif, t1 et t2

Parmi toutes les séquences avec *et*, la séquence *et après* est la plus fréquente dans le corpus non-natif en t1 et en t2, avec une moyenne de trois occurrences par sujet pour chaque série d'enregistrements. On constate des différences entre les sujets, certains s'appuient davantage sur cette SP alors que d'autres, très peu ou pas du tout. Nous n'avons trouvé qu'une seule occurrence de *et après* dans le corpus natif (groupe contrôle). En revanche, on trouve ce connecteur suremployé chez l'enfant lorsqu'il s'agit de récits qui comportent un lien chronologique (Girardin *et al.* 1988 : 191).

La présence de nombreuses occurrences de *et après* dans des corpus oraux, est aussi attestée en L1 (dans ma consultation du concordancier *Lextutor*, Cobb, 1999) mais dans des emplois différents. On remarque que les natifs emploient surtout à l'oral la forme *après* suivie d'un groupe verbal. A l'écrit, on la trouve surtout dans *(et) après que* et *(et) après avoir + V* (Annexe 11, extrait de *Lextutor*, concordancier en ligne). Nous n'avons aucune indication sur la prononciation de ces séquences dans les corpus en ligne, rien ne dit si la suite est dite d'une traite ou pas. Mais ces corpus semblent indiquer certaines tendances du discours oral.

*Et après* est composé du coordonnant *et* marquant la succession de deux événements et de la préposition *après* suivie en principe, en construction absolue, d'un GN (*après les fêtes de Noël, après lui*). La suite *et après* est redondante car les deux mots qui la composent sont à peu près équivalents sémantiquement, avec cependant une légère différence : *et* marque la réunion de deux événements et *après* la succession. Ce qui en fait un connecteur idéal pour marquer la succession d'actions dans un récit chronologique. Regardons maintenant comment les apprenants utilisent cette séquence. Les locuteurs de notre groupe expérimental ne s'appuient pas tous de la même manière sur cette SP et on note aussi des différences dans la fréquence d'emploi de celle-ci. AK produit 18 occurrences de *et après* (10 fois en t1 et 8 en t2) et FA une seule<sup>70</sup>, MJ, MR et NY deux seulement. Il semble donc que cela n'ait rien à voir avec le niveau puisque ces sujets appartiennent aux trois groupes.

Cela relèverait plus d'une stratégie personnelle, cette SP étant utilisée comme organisateur du discours, ou comme appui, sorte de respiration qui ponctuait la narration en différents épisodes narratifs successifs. Comme les locuteurs débutants ont tendance à énumérer des faits, à les juxtaposer, *et après* joue alors le rôle d'un coordonnant. On peut remarquer que deux des quatre sujets asiatiques, OL (8 occurrences en t1 et 10 en t2) et AK, de niveau débutant, s'appuient aussi largement sur cette séquence ; chez les deux autres sujets, on relève six occurrences en t1 et 2 en t2 (QI) mais, curieusement, aucune chez XI.

a. Quel est le rôle de *et après* dans la production orale ?

*Et après* joue un double rôle en tant que marqueur chronologique (succession de deux événements) et coordonnant. Il semble que les locuteurs non-natifs utilisent cette suite de plusieurs façons. Pour la planification : elle est produite en bloc pour gagner du temps, elle est alors suivie le plus souvent d'une pause. Elle permet ainsi à l'apprenant de réfléchir à la suite de son discours. *et après* est soit suivi d'une pause vide ou d'un groupe d'hésitation, soit suivi d'une pause remplie courte qui permet au locuteur de poursuivre la narration sans faire de pause.

*Et après* peut être suivi d'un groupe verbal ou d'une proposition. La répétition de *et après* dégage du temps pour planifier la suite. La séquence, souvent employée seule, forme alors un segment très court (trois syllabes) qui permet de faire une « pause » afin de planifier la suite. Il semble que cette SP facilite le traitement tout en structurant le récit.

Placé en début de segment, *et après* marque un nouvel épisode du récit (ex.39) ; en fin du segment lorsqu'il est suivi d'une pause vide, il peut être la marque de la planification (ex. 40)

---

<sup>70</sup> e poi en italien

; et en milieu de segment pour marquer la succession d'événements entre deux propositions (ex. 41) :

(ex. 39) **et après** il y a un petit enfant qui vienne avec une (.) (MJ, (t2), niv3).

(ex. 40) « elle a volé (.) volé volé **et après** (.) (MA, (t2), niv2) :

(ex. 41) donc il cherche encore **et après** il voit un un peintre (.) ».

On trouve d'autres séquences avec *et* (*et ensuite*, *et là*, *et donc*) produites peu de fois par les non-natifs contrairement aux natifs qui les emploient plus souvent : *et là* produits 4 fois et 3 fois en t2 mais une seule occurrence de *et ensuite* dans le groupe non-natif.

La séquence *et donc*, composée de *et* et du connecteur argumentatif marquant la conséquence/le résultat, est fréquente dans les corpus aussi bien non-natif que natif. Dans le corpus non-natif, on trouve seulement 9 occurrences de *et donc* en t1 et le double (18) en t2, soit une moyenne de une occurrence par sujet, ce qui correspond à ce qu'on trouve dans le corpus natif (6 occurrences). Les non-natifs après six mois dans le pays de la langue cible font une utilisation plus importante de la séquence *et donc*, en calquant d'une certaine manière leur discours sur celui des natifs, peut-être par imitation (Annexe 12, occurrences de « *et donc* » dans Lextutor).

En L2, Rebhein (1987 : 102) remarque que des apprenants turcs ayant l'allemand comme langue cible s'appuient sur un grand nombre de marqueurs tels que *und* (et) et *und dann* (*et alors*) dont le rôle serait d'ouvrir un nouvel « espace de discours » avec une nouvelle focalisation.<sup>71</sup> Dans notre corpus, ce serait plutôt le marqueur *et après* qui joue ce rôle. On note une seule occurrence de *et alors* et de sa variante *mais alors* et aucune chez les natifs dans le corpus. En revanche, les locuteurs L2 produisent *alors* en début de segment 25 fois en t1 et 16 fois en t2 (dans le corpus natif, *alors* seul est produit 8 fois). Les apprenants s'appuient sur ce mot signalétique en début de segment pour amorcer un nouvel énoncé. Mais ils l'utilisent souvent pour planifier la suite ou lors d'une difficulté, le mot sert le plus souvent à éviter une pause bien que chez les débutants *alors* soit souvent suivi d'une pause sonore (euh) ou d'un allongement,

(VI, (t2), niv 3) *alors* eu:h mais sans que lui il sache (.)

(NO, (t2), niv 2) *alors* i::l (.)

#### b) les marqueurs autorégulateurs ou remplisseurs

<sup>71</sup> « [...] is to open a new discours space in which a new process of focusing is possible. » (Rebhein, 1987 :102)

La catégorie des remplisseurs comprend un certain nombre de séquences sur lesquelles l'apprenant s'appuie lorsque surgit une difficulté le plus souvent d'ordre lexical mais aussi au niveau de la conceptualisation.

SP	T1	T2
peut-être	19	18
je sais pas	18	10
je pense	7	2
comme ça / quelque chose comme ça	7 1	5 -
comment dire / comment dit*	6	13
je crois	4	6
j'sais pas	4	1
je suis pas sûre, j'suis pas sûr	2	-
je sais	-	1
Je ne sais pas	8	-

Tableau 44. SP discursives, catégorie des remplisseurs

Les remplisseurs (*fillers*) contribuent à la réduction du temps de pause parce qu'ils sont souvent utilisés à la place d'une pause vide. Les plus employés sont *peut-être* (19 en t1 et 18 en t2), *je sais pas*, (18 en t1 et 10 en t2); *comment dire* (6 en t1 et 13 en t2). On remarque que le nombre d'occurrences de *je sais pas* a considérablement réduit en t2, moitié moins, ce qui semble indiquer une plus grande confiance des apprenants qui ont moins besoin de s'appuyer sur ce remplisseur. L'utilisation de la SP *comment dire* a été commentée plus haut.

b. *en fait*

Une autre séquence discursive, *en fait*, est moins fréquente chez les locuteurs non-natifs (n=17) que chez les natifs (n=6). on relève chez les LNN 10 occurrences de *en fait* en t1 et 11 en t2, et 11 chez les natifs, soit une moyenne de 1,83 occurrence chez les natifs et seulement 0,65 chez les locuteurs L2, soit trois fois moins en proportion.

La séquence discursive *en fait* est utilisée pour introduire un commentaire qui peut prendre la forme d'une justification ou d'une explication (ex. 42 et 43).

(ex. 42) **en fait** *c'est juste* juste un cadre *c'est pas* (.) (HE, (t2), niv2) :

(ex. 43) dans ce champ-là mais **en fait** elle *se casse la gueule* (VO, (t2), niv3) :

(ex. 44) il trouve cet cheval euh il trouve ce cheval euh (.) est vraiment cool euh **en fait** (.) et il veut changer (.) son cheval... (XI, (t1), niv1)

(ex. 45) et à chaque fleur elle *essaye de*::: (.) de trouver (xx) *peut être* le pollen ou **en fait** (.) *et après* (.) (HE, (t2), niv2)

Son emploi sert à libérer du temps. Souvent associé à une pause vide ou remplie, elle marque une pause pour la planification (ex. 44, XI) ou pour gagner du temps parce que l'apprenant ne sait pas quoi dire et abandonne la construction (ex. 45, HE).

### 8.3 SP et structuration du discours

#### a. Les SP qui organisent le discours

Le langage préfabriqué aide les apprenants à mieux structurer leur discours d'une fois sur l'autre. Une meilleure structuration du discours peut se faire, par exemple, grâce à l'agencement de multiples SP.

FA (niv3)

(ex. 46) l'histoire est **divisée en** deux parties **la première partie** dans laquelle **on voit un seul** protagoniste qui est le chevalier

(ex. 47) **voilà donc** euh **c'est** une série comme **on a vu déjà l'année dernière on a** différentes scènes

FA est l'apprenant le plus avancé du groupe 3. Il produit de nombreuses SP (ex. 46 et 47), qui organisent la narration, la structurent. Les segments de parole sont plus longs que ceux des autres apprenants du groupe (32 et 23 syllabes dans ces deux exemples). Les nombreuses séquences produites (en caractères gras), qui organisent la narration, sont réparties en SP lexicales *divisée en, l'année dernière*, mais surtout en SP discursives : *la première, on voit, voilà donc, on a vu, on a*. FA emploie un grand nombre de séquences discursives aussi bien en t1 qu'en t2, plus que les SP grammaticales, mais aussi de nombreuses SP lexicales

#### b. Les formules d'ouverture

(ex. 48) bon (.) **il y a** un homme (.) qui est avec son cheval euh à la campagne peut-être (.) (VI bd2, niv3)

(ex. 49) ok (.) **il y a** un homme (.) (AK 1, niv1)

(ex. 50) donc euh (.) **y a** sur l'image (.) (MA bd2, niv2)

(ex. 51) ah **il y avait** un monsieur avec euh (.) (AK 2, niv1)

(ex. 50) voilà donc euh **c'est** une série euh comme on a vu déjà l'année dernière on a différentes scènes (.) (FA bd2, niv3)

(ex. 52) bon **c'est** l'histoire où (.) (MA bd1, niv2)

(ex. 53) bon le (.) le **c'est** l'histoire de d'un homme **\*que** (.) (VI bd1, niv3)

Nous signalons ici quelques exemples de début de narration (ex. 48 à 53). Les séquences *c'est* et *il y a / y a* sont employés fréquemment pour débiter la narration (formules en usage dans les récits et les contes). La majorité des narrations des apprenants débutent par l'une ou l'autre de ces SP, avec cependant une plus grande occurrence de la formule *c'est* (ex. 50 à 53). On remarque également que chaque début de narration dans ces exemples commence par un mot signalétique (*bon, ah, voilà, donc, ok*) permettant à l'apprenant de produire la SP, de type présentatif (*c'est, il y a, y a, il y avait*) ; parmi ces mots signalétiques, on peut reconnaître la SP *voilà donc*. Il serait d'ailleurs intéressant d'approfondir l'étude de ces mots et le rôle qu'ils jouent en production orale, mais cela dépasse le cadre de cette recherche.

Pour amorcer le récit, quelques sujets emploient une autre séquence, **j'ai vu** (ex. 54 à 60), qui est une autre façon de débiter un récit, surtout en début d'apprentissage et que nous avons identifiée comme préfabriquée. Cette SP n'apparaît jamais dans le corpus natif, ce qui semble signifier que cette SP est propre à la L2.

(ex. 54) **j'ai vu** le garçon avec euh un cheval (.) (JA, bd1, niv1)

(ex. 55) **j'ai vu un j'ai vu** un homme qui est euh (.) (XI, bd1, niv1)

(ex. 56) **j'ai vu** un monsieur (ZI bd1, niv2)

(ex. 57) **j'ai vu** un film je pense c'était une animation il y a une (.) (ZI, F1, niv2)

(ex. 58) **j'ai vu** une bande dessinée (ZI bd2, niv2)

(ex. 59) tout à l'heure **j'ai vu** un film le nom de film est le prisonnier (QI, F1, niv3)

(ex. 60) euh **j'ai juste vu** (MJ, Bd2, niv3)

Quelques sujets surtout de niveau débutant (ex. 54 et 55) ou intermédiaire (ex. 56 à 58) utilisent cette formule. ZI débute trois des quatre narrations par cette séquence (bd1, film1, bd2). La séquence *j'ai vu* est plutôt utilisée en début d'apprentissage, mais on la trouve aussi au niveau avancé (ex. 59 et 60). Elle est enseignée et utilisée en cours de langue et elle fait partie du répertoire des apprenants à un niveau débutant « *j'ai vu un film, j'ai vu la télévision, etc.* ». Nous la considérons préfabriquée dans la mesure où elle répond aux deux critères obligatoires et au critère de fréquence, même appliquée à un seul locuteur. Utilisée fréquemment, elle est un moyen aisé, pour certains apprenants, de commencer l'histoire et de gagner du temps pour la suite. On trouve aussi cette séquence produite par les enfants en L1.

### c. Les formules de clôture du récit

On trouve des structures avec *c'est* pour clore la narration, signaler que le récit est achevé (ex. 61 à 69). La structure avec *c'est*, souvent seule, peut être combinée à d'autres SP en t2 ou



apparaître parfois seulement en t2 (NO, QI). Les constructions avec *c'est* en fin de récit (*c'est tout, c'est fini*) sont utilisées pour marquer d'une manière redondante la fin de l'histoire.

HE (niv2)

(ex. 61) **c'est ça**

(ex. 62) voilà **ça y est** parce que **c'est ça**

En t1, le segment est composé d'une séquence de deux syllabes *c'est ça* alors qu'en t2, HE produit un segment plus long grâce à trois SP de type discursif *ça y est, parce que, c'est ça*. Le sujet HE parle plus grâce aux trois SP qui allongent le segment de parole.

JA (niv1)

(ex. 63) **c'est ça**

(ex. 64) je crois que **c'est tout**

Nous trouvons le même type de clôture chez JA (ex. 63 et 64). En t1, elle produit une seule séquence *c'est ça* (ex. 63) alors qu'en t2 (ex. 64) le segment est plus long et plus complexe grâce aux deux SP qui se suivent *je crois que* suivi de *c'est tout*.

NO (niv2)

(ex. 65) **j'ai fini**

(ex. 66) **c'est tout**

Une autre formule *j'ai fini* permet à NO de clôturer le récit (ex. 65) en t1 alors qu'en t2, elle emploie *c'est tout* (ex. 66).

QI (niv3)

(ex. 67) voilà (.) merci

(ex. 68) voilà (.) **c'est fini**

QI, en t2, (ex. 67) clôture le récit avec la séquence discursive *c'est fini* alors qu'en t1 (ex. 66), elle n'utilise que le présentatif *voilà*, un autre moyen de clôturer le récit (on note que *voilà* est la soudure de *vois là* qui s'est lexicalisée, chapitre 1). En t2, la séquence *c'est fini* est redondante, elle confirme la fin du récit de manière définitive après le mot signalétique, *voilà*.

SU, niv1

(ex. 68) voilà un jouet d'un petit\_enfant

(ex.69) et c'est pour **c'est comme ça** (.) qu'elle a réussi à sortir de la maison

SU clôt la narration en t2 (ex. 69), en utilisant les deux séquences *c'est* et *c'est comme ça*. On ne trouve qu'une seule occurrence de *c'est comme ça* dans le corpus en fin de narration (ex. 69). SU commence le segment avec *c'est pour* (cette raison) mais n'arrivant pas à trouver la suite ou parce qu'elle se rend compte que ça ne marche pas, elle s'auto-corrige et produit la SP *c'est comme ça*.

Les séquences avec *c'est* sont un moyen de signaler la fin de la narration, tout comme la séquence *j'ai fini*, moins courante (une seule occurrence).

#### 8.4 Quelques formes utilisées comme aide au traitement / SP ou pas ?

Nous devons ici faire remarquer qu'un certain nombre de séquences non prises en compte dans le calcul auraient pu être considérées comme des SP. Ce sont des constructions verbales non finies mais considérées comme des formes automatisées et même préfabriquées par Cordier (2013). Elles sont censées faciliter la production en permettant aux apprenants de commencer une expression qu'ils vont ensuite abandonner pour produire l'expression juste. Nous traitons ces formes seulement comme des faux départs, des hésitations. Ces groupes ont certainement un rôle en production, ils peuvent, à un moment donné, faciliter le traitement car il se passe effectivement quelque chose qui s'apparente à une hésitation. Mais nous ne les considérons pas comme des séquences préfabriquées. : *j'ai / il a / il est / il va /* et nous y ajoutons *un petit* qui fonctionne de la même façon.

AK (t1) niv1

**Il a comment il a (.) il a (.) et il a tant pis tant pis non il a tant pis non il est (.) il est désolé il est navré**

Dans cet exemple, Akira utilise de nombreuses fois la forme *il a*, apprise et mémorisée, normalement suivie d'un participe passé. Mais l'activation ne se fait pas, les connaissances déclaratives n'étant manifestement pas disponibles en MLT. Il ne peut activer que la formule *tant pis* qui marque la déception du cowboy, il emploie également *il est* pour trouver le syntagme complet *il est* + adjectif. Cet exemple illustre parfaitement pourquoi nous avons compté ces formes comme des faux départs et non comme des séquences à part entière. La répétition du syntagme incomplet (*il a*) signale la difficulté d'activation, comme le mot attendu n'arrive pas, AK ne peut qu'activer la SP mémorisée *tant pis*.

**j'ai / il a (.) il a un peu sommeil (AK, t1, niv1)**

Comment **il a** il a été un peu sommeil (AK t1, niv1)

**j'ai** / il a connu il a connu (.) (AK, t1, niv1)

et après **il a** (.) **il a** (.) comment dire détacher le grand cheval (MA, t1, niv2)

donc **il a** / il a volé il a volé (.) (MA, t1, niv2)

et **il a** / il a vu de du (.) (MA, t1, niv2)

il **il a** / il voit (.) (OL, t1, niv1)

**il a** / il a mis une petite signe pour l'abeille (PR, t1, niv3)

et l'abeille est/ **il a** / il regarde... (PR, t2, niv3)

et **il a** / il a vu que cet .. (XI, t1, niv2)

il est / **il a** / il a descendu du cheval (.) (JA, t2, niv1)

il a (.) utilisé cette information et **il a** / il a sorti euh la maison (.) (PR, t2, niv3)

**il a** / il a pas senti quelqu'un (XI, t2, niv1)

**il a** / il a \*trouve que (.) (XI, t1, niv1)

Les deux formes, *j'ai* et *il a*, sont considérées comme des SP par Cordier (2013), dans la mesure où ils permettent dans certains cas l'activation d'une forme verbale complète. La forme à la première personne *j'ai* semble être produite comme un bloc que l'apprenant corrige aussitôt en produisant soit la forme incomplète à la troisième personne, suivie ensuite d'une forme verbale correcte, soit une forme correcte directement.

Un autre groupe fonctionne comme de la même façon, *il est*, et permet aussi de trouver la forme correcte qui se trouve être parfois une SP, comme dans l'exemple *il est il est heureux* ou *il est il \*reste étonné*.

Il est (.) **il est** (.) il a dormi (AK, t1, niv1)

Et comme ça **il est** il \*reste étonné... (FA, t1, niv3)

**il est** (.) il est descendu (MR, t1, niv2)

il est venu ici **il est** / il est venu à (.) (AK, t2, niv1)

**il est** / il est heureux... (AK, t2, niv1)

**il est** il dort (.) (JA, t2, niv1)

il **il est** / il regardait le le &che le cheval

donc quand le le cowboy **il est** / il est réveillé (.) (SU, t2, niv1)

et après le mouchoir **il est** / il laisse (.) (PR, t1, niv3)

**il est** // il était très / il était très fatigué et (.) (XI, t1, niv1)

il y a un garçon **un p'tit** un p'tit garçon (ZI, t1, niv2)

**il va** /il va dormir un peu sur un p'tit (.) banc (ZI, t2, niv2)

D'autres exemples avec *petit* et *il va*, moins fréquents, semblent jouer le même rôle. Les sujets s'appuient sur la répétition d'un élément de la séquence pour la planification et la recherche lexicale, ils peuvent activer ainsi le mot ou la SP qui leur fait défaut.

## 8.5 Les séquences interlangagières ou déviantes

Cette section est consacrée aux erreurs interlangagières que les apprenants font dans leur production. La plupart des erreurs sont des erreurs lexicales et des erreurs morphosyntaxiques. Les erreurs morphosyntaxiques sont plutôt des erreurs sur le genre du déterminant (\*à la côté de) ou de l'adjectif (\*bon odeur, \*petit sieste), sur la morphologie du verbe (\*asseye, se \*refuge, sans \*compris) ou des erreurs de syntaxe portant sur la structure grammaticale, notamment lorsque la préposition qui suit le verbe est erronée, *transformé \*dans* pour *transformé en*, *il arrive pas \*de* pour *il arrive pas à*. Nous présentons un tableau des SP déviantes (Annexe 8) et un tableau des différents types d'erreurs (Annexe 9).

### 8.5.1 Séquences interlangagières lexicales

Les erreurs lexicales sont les plus nombreuses dans notre corpus. Prenons l'exemple du mot composé *grains de café* que normalement les locuteurs devraient avoir acquis et produire au moins en t2. Dans la tâche du film, la mouche va chercher des grains de café qu'elle place en forme de flèche dans la baignoire pour indiquer à l'abeille la sortie. La séquence *grains de café* n'étant pas connue en t1 par certains apprenants est produite correctement en t2 ou en partie autour d'un mot composé avec café : \*pièces de café, \*pois de café, \*morceaux du café, \*haricots du café, etc.

Au niveau débutant, certains apprenants ne peuvent pas produire la séquence *grains de café*. Ils évitent de le faire en t1 alors qu'en t2, ils prennent plus de risque et produisent parfois une forme erronée. C'est le cas de AK qui produit \*pois de café (2) en t2, alors que, en t1, il a parlé d'un signe fait par la mouche pour montrer la sortie. XI (niv1) produit la séquence interlangagière, \*pois de café, une fois en t1 et deux fois en t2. Elle ne peut pas se corriger la deuxième fois. OL (niv1) utilise aussi la même séquence interlangagière. On peut se demander si cette erreur ne serait pas propre aux étudiants d'origine asiatique qui ont l'anglais comme première L2 et qui doivent traduire l'expression *coffee beans*, ce qui donne *pois de*

*café*. On trouve aussi d'autres séquences interlangagières de ce type. NO (sujet anglophone) produit en t1 *\*haricots du café* traduction littérale de l'anglais *coffee beans*. En t2, elle emploie même *\*morceaux du café* (sur le modèle de *morceaux de sucre*) et *\*quiz du café*, dont l'origine semble mal définie. Ce qui est intéressant ici, c'est que le sujet utilise la structure du mot composé (N de N) sur le modèle de *grain de café* et qu'il n'hésite pas à prendre le risque de produire une SP erronée.

On trouve d'autres séquences interlangagières chez ZI (niv2) en t1 *\*pièces de café* et en t2 *\*points de cacao*. Là aussi, l'apprenant n'arrive pas encore à produire la séquence correcte mais seulement la structure qu'il va compléter avec ce qu'il a en mémoire.

MA (niv2) produit une séquence composée de *petit* et *truc* (mot passe-partout) pour contourner la difficulté, *petits trucs du café* en t1 et rien en t2, la production de cet apprenant étant beaucoup plus laconique la deuxième fois.

QI (niv3) produit, en t1 comme en t2, *\*graines de café (3)* en faisant une erreur sur le genre *\*graines* au lieu de *grains*. Mais cette erreur semble s'être fossilisée puisqu'on la retrouve toujours après six mois, ou l'était probablement déjà. On peut aussi signaler deux erreurs de ce type dans le corpus natif (*\*graines de café*, CA ; *\*graines de cacao*, VC). L'erreur peut aussi provenir chez les natifs dans les conditions spécifiques de la production orale (parler en temps limité).

Dans notre corpus, on trouve donc différentes tournures de séquences en remplacement de la SP *grains de café* que certains locuteurs L2 peuvent ne pas connaître ou n'arrivent pas à activer dans la forme correcte. En compréhension, ils la reconnaîtraient sans aucun doute.

Les erreurs interlangagières relevées dans notre corpus sont de plusieurs ordres et semblent se situer le plus souvent au niveau du lexique, ce qui semble plutôt normal. Elles peuvent prendre la forme d'une substitution (*\*elle était un peu peur* pour *elle avait un peu peur*), d'une omission (*\*il pas compris* au lieu de *il n'a pas compris*), d'un amalgame (*\*il y atait* fusion entre *il y avait* et *il était*), d'une confusion (*\*il a un peu de sommeil* pour *il a un peu sommeil*). Des chercheurs, qui se sont intéressés au langage préfabriqué, ont relevé des erreurs qui ressemblent à des confusions lexicales (*lexical blends*), *\*he was breathing down my shoulder* pour *he was breathing down my neck* = « talonner » (Peters, 1983 cité par Théophanous et Perez-Bettan, 2012). Ces erreurs indiquent que les séquences sont stockées dans le lexique mental comme des unités, de manière holistique (voir chapitre 3). Ces formules interlangagières ne doivent pas être simplement considérées comme des erreurs mais « suggèrent un pas vers la compétence en devenir » (Théophanous et Perez-Bettan, 2012 :

260). Elles peuvent aussi être le produit d'une forme que l'apprenant a voulu plus soignée comme dans l'exemple rapporté par Dechert et Lennon (1989 : 158-159) *he has got financially broke*, confusion entre *he is broke* et *he has got into financial difficulties*. L'apprenant commet ici une erreur alors qu'il a voulu améliorer son expression en produisant une tournure plus soignée. Les erreurs peuvent aussi se produire lorsque l'apprenant se sent suffisamment à l'aise pour prendre des risques et n'hésite pas à utiliser des séquences qu'il a peut-être mal perçues. Ces erreurs pourraient être considérées comme un « signe de compétence plutôt qu'absence de celle-ci » (Théophanous et Perez-Bettan, 2012 : 260). . Elles sont parfois dues (ex. 1) à l'influence de la L1 sur la L2 (traduction *to take a nap* qui donne *\*prendre une sieste*, *to be afraid* qui donne *\*être peur*, ou d'une autre L2 (l'anglais) en particulier chez les locuteurs asiatiques (ex. 2). La L2 étant stockée dans le même lexique mental que la L1, il peut aussi se produire des erreurs.

#### 1) erreur de substitution

(ex. 1) PR, (t2) *\*il a pris une sieste* → *il a fait une sieste*,

(ex. )2 XI, (t2) *\*elle était un peu peur* → *elle avait un peu peur*

(ex. 3) JA, (t1) *\*elle a passé pour* → *elle est passée par*

(ex. 4) PR, (t1) *\*en ce moment* → *à ce moment*

(ex. 5) NO, (t2) *\*il est dehors à l'autre côté de la fenêtre* → *de l'autre côté de la fenêtre*

(ex. 6) NO, (F2) *\*il n'arrive pas d'avoir une nouvelle amie* → *à avoir une nouvelle amie*

Ces exemples d'erreurs de substitution portent sur les verbes, exemples 1 et 2, *\*prendre* à la place de *faire*, *\*être peur* pour *avoir peur* ; erreur sur le verbe et la préposition *\*a passé pour* à la place de *est passée par* ; erreurs de préposition, exemples 5 et 6, *\*à l'autre côté* pour *de l'autre côté* ; *\*il n'arrive pas d'avoir* pour *il n'arrive pas à avoir une nouvelle amie*.

#### 2) erreur d'omission

(ex. 7) HE, (t1) *\*il pas compris* → *il n'a pas compris*

(ex. 8) PR, (t1) *\*qu'est ce qui passait* → *qu'est ce qui se passait*

(ex. 9) SU, (t1 et t2) *\*il avait* → *il y avait*

(ex. 10) XI, (t1) *ce moment-là* → *à ce moment-là*

(ex. 11) HE, (t2) *\*il vient comprendre* → *il vient de comprendre*

(ex. 12) MR, (t2) *\*à maison* → *à la maison*

(ex. 13) XI, (t2) *\*il veut reposer un peu* → *il veut se reposer un peu*

(ex. 14) FA, (t1) *\*il sent fatigué* → *il se sent fatigué*

Les SP interlangagières dans les exemples 7 à 14 sont produites avec un élément manquant sans doute parce qu'elles ont été mal perçues alors que la SP est dite d'une traite sans pauses internes. Ce qui veut dire que l'apprenant ne l'a pas créée mot à mot.

### 3) erreur d'amalgame

(ex. 15) AK, (t1 et t2) *\*il y atait* → *il y avait*

(ex. 16) OL, (t1) *\*c'est pas un vraiment fleur* → *c'est pas une vraie fleur*

Les apprenants, dans ces deux exemples (ex. 15 et 16), mélangent deux formes. AK (ex 15) mixe les formes *il y avait* et *il était*, à l'imparfait ce qui donne la forme déviante *\*il y atait* que l'on trouve dans toute sa production, en t1 comme en t2. Dans l'exemple 16, la place de l'adverbe *vraiment* est incorrecte ; OL semble mélanger deux structures, *c'est vraiment pas* où l'adverbe se place avant la négation *pas*, et *c'est pas une vraie fleur* où l'adjectif *vraie* est placé correctement devant le nom.

### 4) confusion

(ex. 17) HE, (t2) *\*avoir un peu de sommeil* → *avoir un peu sommeil*

(ex. 18) XI, (t2) *\*sans fait attention* → *sans faire attention*

(ex.19) JA, (t1) *\*elle y a prendre le chemin* → *elle prend le chemin*

HE (ex. 17) fait une confusion entre *un peu* et *un peu de + N* avec la locution verbale *avoir sommeil*, seul l'adverbe de quantité *un peu* peut être utilisé, alors que *un peu de* se construit avec le verbe *avoir* seul « *avoir un peu d'argent, de tranquillité, de temps...* ». *avoir peu de sommeil* existe aussi.

XI (ex. 18) fait une confusion entre le présent de la séquence *faire attention* (*il fait attention*) et la structure prépositionnelle avec *sans + verbe* à l'infinitif, (*sans faire*). L'apprenant (ex 19), confond deux structures, *il y a* et *prendre le chemin*.

## 8.5.2 Évolution des SP interlangagières

AK, niv1, t1 et t2

(t1) <i>*il a descende en cheval</i> (t2) <i>il descend de cheval</i>
(t1) <i>*il a été un peu sommeil</i> (t2) <i>il a sommeil</i>

La construction verbale, *\*descendre en cheval* produite par AK, niveau débutant, en temps 1, est une séquence déviante, sorte de mixage entre les deux structures *descendre de et à cheval*, en cours d'acquisition (t1). L'apprenant n'est pas encore capable de produire la forme correcte de la SP soit qu'il l'ait mal perçue soit qu'il ne l'ait pas encore complètement mémorisée. Il semble hésiter entre deux SP qu'il réunit de manière erronée. En t2, la séquence verbale *descendre de + cheval* est utilisée correctement. D'autre part, la terminaison du participe passé *\*descende* incorrecte en t1 est produite correctement en t2 *descendu*. Il reste un doute sur le choix de l'auxiliaire mais AK utilise la formule au présent en t2, qui est la plus utilisée.

Dans le deuxième exemple, la formule interlangagière en t1 *\*a été sommeil (il a \*été un peu sommeil)*, avec erreur sur l'auxiliaire, est produite correctement en t2 *a sommeil*. La forme du verbe *avoir* au passé composé *il a eu sommeil* que le sujet AK a voulu employer n'est pas encore acquise à ce stade. AK semble vouloir raconter l'histoire au passé en t1, il connaît certainement *il a sommeil* mais veut sans doute utiliser la formule au passé, il fait donc un amalgame entre *avoir sommeil* et *avoir été* qui ont été utilisés ensemble ; mais en t2, AK emploie correctement la locution verbale *avoir sommeil* cette fois dans une narration au présent.

AK (niv1)

(t1) *\*il y atait le (.)*

(t2) *il y avait un \*beille qui (.)*

Lors de la narration du film, AK emploie la construction déviante *\*il y atait* (mélange de l'auxiliaire *avoir* et de l'imparfait du verbe *être*), soit qu'il ait mal perçu la forme phonologique soit qu'il n'ait pas encore automatisé la forme correcte de l'imparfait *il y avait*. On peut aussi supposer que AK ne connaissant pas l'imparfait de *il y a* se soit basé sur l'imparfait du verbe être, erreur peut-être de surgénéralisation. La séquence est non seulement correcte en t2 mais s'inscrit dans une construction plus complexe (*il y avait...qui*). Nous avons identifié la séquence erronée comme une SP déviante car elle présente toutes les caractéristiques d'une SP, cohérence phonologique et forme idiosyncrasique.

a) MR (niv2)

(t1) *\*il en train de dormir*

(t2) *il est en train de &dor dormir*



En t1, MR ne produit pas la forme complète du syntagme verbal *être en train de*. L'omission du verbe *être* vient d'une confusion avec l'expression *en train de* dans des phrases comme *je l'ai vu en train de voler*, ou d'un problème phonologique. En revanche, MR en t2 produit la SP correctement avec le verbe, *être en train de*.

b) NO, niv2

NO (t1)	NO (t2)
Il es : :t (.)	il <i>est fatigué</i> (.)
Fatigué (.)	alors il <i>s'arrête</i>
Alors il *arrête	

Dans ces deux extraits, Naomi veut produire l'ensemble *être fatigué*. En t1, la pause entre le verbe *être* et l'adjectif, accentuée d'un allongement (durée supérieure à 700ms) de la syllabe, indique que, probablement, le sujet n'arrive pas à activer l'adjectif *fatigué*. Les deux hésitations, l'allongement suivi d'une pause vide, vont lui permettre de trouver le mot juste. Mais la place de la pause entre le verbe et l'adjectif attribut, agrammaticale, accentue l'impression de disfluence. Le troisième segment *alors il arrête* montre là encore une difficulté d'ordre morphologique ; la forme réflexive du verbe, une difficulté pour les apprenants du FLE, n'ayant pas été automatisée. En t2, NO produit les deux séquences correctement, la forme réflexive « *il s'arrête* » et la séquence *être fatigué*, identifiée comme préfabriquée, et produite d'une traite cette fois, c'est-à-dire sans pause interne. De plus, la place de la pause en t2 à une jonction grammaticale, entre les deux syntagmes verbaux et juste avant l'adverbe coordonnant, est cette fois grammaticale.

## 8.6 Complexification de la production

Dans les exemples que nous donnons, nous voulons montrer que les apprenants la deuxième fois produisent des segments plus longs (ex. 20 et 21) et utilisent des stratégies pour garder la parole ou pour éviter la pause

L'apprenant NO peut produire une séquence plus longue en t2 (ex.21). Les deux segments sont réunis en un segment plus long

(ex. 20) alors <b>il y a</b> une insecte (.) qui <b>fait le miel</b> NO (F1) niv2, (ex. 21) alors <b>il y a</b> une insecte <b>qui</b> entre dans une maison (.)NO (F2) niv2,
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Cet exemple montre bien ce qui se passe en t2. NO produit en t1 un énoncé (ex. 20) dans deux segments de parole avec une pause entre les deux propositions (pause grammaticale) ; en t2 (ex.21), l'énoncé est contenu dans un seul segment plus long et sans pause. L'apprenant utilise une séquence plus complexe *il y a...qui* et n'a plus besoin de la pause pour produire la suite.

MJ, niv3 (ex. 22)

donc **on peut** le:: **on peut** le tirer (.)

Dans l'exemple 22, la répétition de la SP *on peut* dans le premier segment permet à MJ de produire un segment plus long. La séquence *on peut* a ici une fonction discursive et semble jouer un rôle dans la planification. Son emploi répété permet à MJ de libérer du temps pour la recherche lexicale.

NY, F2, niv1 (ex. 23)

elle était **un peu un peu**: (.) elle **a peur un peu de** la mouche donc elle a (.)

La répétition de la séquence *un peu* dans le premier segment (ex. 23) permet à NY de produire un segment plus long et permet ainsi la planification. Dans le segment suivant, elle produit un énoncé plus clair, plus long et plus complexe, constitué de deux SP imbriquées *avoir peur de* et *un peu* (11 syllabes, 2 SP et 5 syllabes).

### Comparaisons entre t1 et t2

Exemple 1, FA, un des sujets les plus avancés

FA est l'apprenant du groupe 3 le plus avancé, même si en temps 2 on ne constate pas d'amélioration dans les résultats obtenus ; FA utilise en t1 des expressions calquées directement de sa L1 qu'il va produire correctement en t2.

1) correction d'une SP calquée sur la L1

<p>BD1</p> <p>Il <b>*reste étonné</b> devant le / le grand cheval du cowboy et il change son cheval avec le cheval du cowboy (.) (21 mots)</p> <p>et il est // il <b>*reste un peu étonné de</b> euh <i>parce qu'il reste juste en face du petit</i></p>	<p>BD2</p> <p>et <i>juste après</i> le chevalier il <i>se réveille</i> et il <b>est étonné de</b> voir-euh que <i>à la place du gros cheval</i> il a que ce petit // le <i>petit cheval</i> mais <i>il ne *comprene pas pourquoi</i> <i>il y a</i> (.) (38 mots, 50 syllabes)</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

cheval qui était le jouet <i>avant de</i> (.)	
-----------------------------------------------	--

(25 mots, 31 syllabes)	
------------------------	--

Dans ces exemples tirés des narrations en t1 et t2, on constate en t2 un segment de parole plus long en mots et en syllabes et plus riche en SP. Ces SP ont différentes fonctions dans le discours, expression de la chronologie (*juste après*), d'une action exprimée par le verbe réflexif (*se réveille*), d'un état (*il est étonné de*), l'expression d'un changement / d'un remplacement (*à la place de*).

FA, sujet italoophone, utilise deux fois le syntagme *il \*reste étonné* calqué sur sa L1 « *\*rester étonné* » (*rimanere sorpreso*) et *il \*reste un peu étonné* (*rimanere un po sorpreso*). La pause remplie, qui suit la SP en t1, lui permet de trouver une autre structure plus complexe avec *parce que* pour expliquer la cause de l'étonnement du personnage ; en t2, il abandonne la forme idiosyncrasique pour la forme correcte « *être étonné de* » suivi de son complément « voir », structure grammaticale complexe d'un niveau plus avancé.

La production de cet italoophone (Annexe 3, p. 7-8) comprend de nombreuses séquences de type lexical (à la différence des apprenants des deux autres groupes moins avancés). Fabio emploie cependant des expressions qui sont directement traduites de sa L1 comme « *prend \*part* » (*prendere parte*) pour « prendre parti » en français *\*il sent fatigué* (*si sente stanco*) qu'il traduit par « *il semble fatigué* » en t2.

On note quelques traits communs avec les natifs : la réduction phonologique, suppression du *ne* faisant partie de la particule négative de *ne...pas*. La phrase/séquence « *il ne \*comprende pas pourquoi* » devient « *il comprend pas pourquoi* » correction de la forme verbale et réduction phonologique de la particule négative « ne » comme le font souvent les natifs à l'oral, et qui est, nous le rappelons, un des critères d'identification (critère de cohérence prosodique).

La narration, plus précise, structure les différentes phases de l'histoire et présente moins de répétitions. La production de FA est plus fluide et plus précise, le vocabulaire plus riche et plus spécialisé (*dans un contexte un peu rural*).

Si le nombre de SP n'augmente pas de manière significative en t2, la longueur moyenne des segments de parole, en revanche, augmente : 11,9 vs 14,8 syllabes. Mais cela s'accompagne par une augmentation aussi des corrections (26 vs 44 corrections la 2<sup>ème</sup> fois). Pour 100 mots produits, on note 8,5 corrections en t2 (répétitions et allongements) vs 7,3 en t1. Ces corrections ou reprises sont le résultat des hésitations lexicales dues peut-être aux autocorrections pour trouver la bonne formulation. Mais les hésitations peuvent aussi relever

d'une stratégie qui aide à la planification du discours, en libérant du temps, ce qui semblerait correspondre au profil de cet apprenant.

### Exemple 2, HE, niveau 2

Nous présentons les deux débuts de narration du cartoon par HE en t1 et en t2.

seg	Temps 1	seg	Temps 2
1	<i>c'est c'est</i> un ombre (.)	1	bon eu::h (.)
2	sur un cheval(.	2	il <b><i>il y a</i></b> un un un homme sur un cheval (.)
3	e::t(.	3	eu::h il arrive &s <b><i>ça ressemble *d'</i></b> être un un / (.)
4	il arrête(.	4	un ancien (.)
5	e::t il <b><i>est fatigué</i></b> (.)	5	un <b><i>ça a l'air</i></b> d'être (.)
6	et il essaye(.	6	des des <b><i>années soixante peut-être à cause de il y a</i></b> un
7	pou::r(.		petit xxx dans une maison l'endroit où où il garde son
8	pour rester <b><i>un petit peu</i></b> (.)		cheval (.)
9	pour récupérer l'énergie	7	seulement arrivé à (.)
		8	à aux aux <b><i>années soixante</i></b> ou <b><i>je sais pas quoi</i></b> (.)
		9	bon (.)
		10	après il il il <b><i>a l'air</i></b> d' <b><i>avoir un peu de sommeil</i></b> (.)
		11	il <b><i>s'assoit</i></b> et il dort
	9 segments/28 mots 3 mots par segment (moyenne)		11 segments/87 mots 7,9 mots par segment

On remarque que les segments de parole sont en général plus longs en t2 (segments 2, 3, 5, 6, 8, 10 et 11). Le premier segment, en t2, très court, débute par un mot signalétique, *bon*, souvent utilisé par les locuteurs (21 occurrences dans le corpus), qui aide à la planification. Il fonctionne comme un remplisseur/filler et a la même fonction que la pause remplie qui suit (*eu::h*). Elle permet au sujet, de gagner du temps pour trouver la bonne formulation. La SP *il y a*, produite d'une traite dans le segment 2, est une construction du même type que celle utilisée en t1 (*c'est*) et permet aussi d'introduire un nom. Le segment de parole 6 est plus long et cette augmentation est due principalement à l'utilisation de plusieurs SP (en gras et en italiques). Elles sont de trois types, SP lexicale (*années soixante*), grammaticales (*il y a*, *à cause de*) et discursives (*peut-être*). On ne relève pas moins de 13 SP en t2 alors que HE produit seulement 3 SP en t1. La narration en t2 est marquée par l'emploi de SP lexicales : des phrases impersonnelles *ça ressemble \*de* (pour à), *ça a l'air*, mais aussi des syntagmes nominaux *années soixante* (2 fois), et des syntagmes verbaux, *avoir sommeil*, *il s'assoit*.

Les séquences discursives, *je sais pas quoi*, *peut-être*, allongent le segment tout en permettant au locuteur de maintenir la parole. He produit deux séquences successives de type

grammatical à cause de et il y a, on note ici l'emploi inapproprié de à cause de là où on attendrait parce que. Trois SP successives (avoir l'air, un peu de, avoir sommeil) allongent le segment de parole « il a l'air d'avoir un peu de sommeil ».

Exemple 3, QI, niveau 3

a) exemple a

t1	t2
e::t quand il a vu rouky <b>il est en train de</b> dormir et (.)	il a vu rudy (.) qui <b>est en train de faire une sieste</b> (.)

b) exemple b

t1	t2
à à ce moment-là y a un petit garçon (.) <b>il est malin il est mignon</b> (.) <b>il est un peu méchant</b> (.) il vient avec son petit jouet (.) <b>en fait c'est un:::</b> (.) fausse cheval (.)	<b>à ce moment-là</b> (.) <b>y a un petit garçon</b> avec son <b>petit jouet</b> <b>sous forme de petit cheval</b> (.) arrive (.)

Bien que QI produise moins de segments en t2, l'énoncé est plus complexe en t2. Dans l'exemple (a), en t1, on relève une seule SP *il est en train de* + dormir alors qu'en t2, la SP est suivie d'une deuxième *faire la sieste*, ce qui allonge le segment. Dans l'exemple (b), en t1, les constructions syntaxiques sont simples ; QI utilise trois fois la forme du verbe *être* à la troisième personne suivi d'un adjectif (*malin, mignon, méchant*) pour décrire l'enfant.

En t2, le sujet produit seulement trois segments (6 segments en t1) mais l'énoncé est plus riche au niveau du lexique, il est composé de deux SP *être en train de* et *faire une sieste*, plus précises et surtout plus longues que le verbe *dormir*. La construction est plus complexe grâce à la SP *sous forme de* à la place du *fausse cheval* du premier exemple. On note que QI n'a plus besoin de s'appuyer sur les SP discursives *en fait* et *c'est*.

Cet exemple nous montre que les apprenants ne suivent pas tous le même processus de développement. Certains apprenants, les plus avancés, veulent décrire de manière plus précise et plus complexe la deuxième fois. Pour cette raison, ils s'appuient moins sur les SP discursives et plus sur des SP qui organisent leur discours, qui relient les mots entre eux ou les propositions entre elles.

#### Exemple 4, SU, niveau 1

SU est un apprenant de niveau débutant. Ce sujet a appris le français pendant quatre mois au Brésil et a suivi des cours intensifs pendant un mois avant de venir en France.

SU, F1 : **il \*avait un petit insecte (.) je crois qu'il cherche quelque chose (.) je sais pas quoi je crois que c'est** la nourriture (.) et il était dans la &for dans la forêt après il est rentré dans la maison (.) pour chercher les fleurs pour mettre son (.) **je sais pas quoi** [&=rires] son pollen **je sais pas** (.)  
Su, F2 : le\* **petite histoire raconte l'histoire** d'une **petite abeille** (.) qui cherche le pollen (.)

En t1, SU produit de nombreux segments et utilise de nombreuses SP, *je crois que* (2), *je sais pas quoi* (2), *quelque chose*, *je sais pas*, qui montrent que SU cherche à dire quelque chose qu'elle n'arrive pas à exprimer avec un vocabulaire précis. Elle ne semble pas très sûre d'elle. La répétition de remplisseurs (*fillers*), *je sais pas quoi* x 2 et *je sais pas*, lui permet de gagner du temps afin de poursuivre son récit mais aussi d'allonger le segment. L'ensemble de cet énoncé comprend six segments. En t2, SU parle moins. Cela signifie-t-il qu'elle n'a pas progressé ou, au contraire, qu'elle a amélioré son expression ? En t2, SU n'a plus besoin de s'appuyer sur les remplisseurs et produit un long segment plus précis. La pause intervient à une joncture grammaticale juste avant la relative, *qui cherche le pollen*. En t1, la pause, agrammaticale, se trouve entre le déterminant et le nom ; c'est grâce à la pause suivie d'un *filler* : *son* (.) + *je sais pas quoi* qu'elle arrive à trouver le mot juste : le GN *son pollen*. En t1, comme SU n'arrive pas à activer le mot *pollen*, elle contourne la difficulté pour ne pas arrêter de parler en utilisant plusieurs SP qu'elle connaît (*je crois que*, *quelque chose*, *je sais pas quoi*, *je crois que c'est* la nourriture), toutes ces SP allongent le segment de parole, tout en évitant de faire des pauses.

Il faut donc être très vigilant dans les généralisations concernant le langage préfabriqué. Les apprenants ont envie lors d'une tâche imposée de parler sans faire trop de pauses mais surtout, lorsqu'ils avancent dans l'apprentissage, d'être précis. Ils n'ont plus besoin de s'appuyer sur des SP discursives mais préfèrent améliorer le lexique en employant un vocabulaire adéquat et /ou plus de SP.

À travers ces exemples, on constate que certains apprenants produisent une narration plus courte mais plus précise, qui peut parfois se traduire par une production moins riche en quantité de SP. On constate parallèlement un changement dans la qualité des narrations produites, plus riches au niveau du lexique et plus complexes au niveau de la syntaxe. Peut-être la tâche

narrative influe-t-elle sur le comportement langagier. Ce constat s'inscrit *a contrario* des études faites à ce jour sur le LP. Nous pensons que la nature de la tâche favorise ou pas l'utilisation des SP. Lorsque les apprenants sont libres d'aborder un sujet d'actualité ou un sujet personnel, ils se sentent moins contraints, plus libres, d'utiliser le vocabulaire qu'ils connaissent, ils peuvent éviter les difficultés grâce à des procédures d'évitement et des stratégies de remplacement. Lors d'une tâche narrative imposée, ils doivent trouver des stratégies pour contourner les difficultés lexicales. On a des exemples dans notre corpus : *petit jaune, petite jupe, tant pis* utilisés pour désigner l'abeille, le corps de l'abeille et l'idée, le concept de l'étonnement dû à une mauvaise surprise. Les apprenants peuvent comme SU produire des SP discursives pour ne pas arrêter leur discours ou éviter de pauser. La tâche narrative favorise aussi la production de séquences adaptées et propres au récit que nous avons considérées préfabriquées car produites d'une seule traite et qui introduisent des propositions, *il pense que, il trouve que, il voit que*. Les apprenants peuvent aussi produire un discours plus complexe sans que, pour autant, la quantité des SP augmente.

## **8.7 Comparaison corpus natif et non-natif**

Nous allons à ce stade montrer, à travers quelques exemples, les principales différences entre les SP produites par le groupe contrôle et celles du groupe expérimental.

### **SP propres aux natifs**

Alter ego ; s'assoupit ; à la place de ; il s'agit de ; maya l'abeille ; bouquet de fleurs ; corbeille de fruits ; au final : chapeau rond ; petit poucet ; table du salon ; conduit de canalisation ; **être dans son élément** ; se balade ; se cogne contre, se met à ; **être pris au piège** ; au milieu de ; **se place en observateur**, pendant que, se rend compte ; **fait un échange** ; par la main ; au bout de ; court métrage ; pièce attenante ; se heurte à ; **tombe à la renverse** ; à l'air libre ; au passage ; long voyage ; un p'tit somme ; **file à l'anglaise** ; habitat naturel ; (cheval) à roulettes ; **y a de la vie** ; habitat humain ;

yeux globuleux ; **prend ses précautions**, faire confiance à ; **va droit au but** ; **tourner en rond** ; **se poser des questions** ; par inadvertance ; **est en danger** ; de manière ingénieuse ; en ville ça se limite à ; micro-ondes ; dernière séance ; au final ; rebord de fenêtre ; on se doute ; il se trouve que ; **arrive en ville** ; ça se limite à ; j'avoue que ; la raison du pourquoi ; a du mal à ; incapable de ; trouve moyen de ; comme on le voit bien ; **ça ne marche pas** ; **on va le dire** ; sauf que ; faire un bruit de

Tableau 45. Liste des SP spécifiques au groupe natif



On note, parmi les SP présentées dans le tableau 45, propres aux locuteurs natifs, des références culturelles (*petit poucet, chapeau rond, alter ego*). Une référence au conte de Perrault, *Le Petit Poucet*, une référence aux paroles d'une chanson « ils ont des chapeaux ronds vive la Bretagne, ils ont des chapeaux ronds vive les bretons » ; une référence cinématographique, *maya l'abeille*, et une référence classique héritée du latin et utilisé dans la langue courante, *alter ego*. On peut aussi reconnaître dans les séquences natives des séquences appartenant à un domaine plus spécialisé comme *habitat humain, en ville, dernière séance, la raison du pourquoi*. On trouve peu de séquences de ce type dans le corpus non-natif, excepté quelques séquences comme *système de conduite* ou *coup de vent*. Les locuteurs natifs emploient plus de SP lexicales dont le contenu fait référence à de nombreux domaines (scientifique, littéraire, mais aussi appartenant à la langue familière).

On trouve :

- b) des expressions idiomatiques plus ou moins figées, *filer à l'anglaise, tomber à la renverse, à l'air libre, aller droit au but, tourner en rond, par inadvertance ; la raison du pourquoi ;*
- c) des noms composés, *bouquet de fleurs, corbeille de fruits, table du salon, cheval à roulette.*
- d) des collocations : *poser des questions, long voyage, faire un échange, de manière ingénieuse, être en danger, pièce attenante, court métrage, prendre au piège,*

Mais aussi :

- des expressions discursives : *au final, ça semble, on va le dire, on suppose*
- des syntagmes propositionnels : *ça se limite à ça, ça marche pas*
- des formes impersonnelles : *il se trouve que*
- des structures avec *c'est* qu'on ne trouve pas dans le corpus non-natif : *c'est vrai que, le résultat c'est que.*
- des phrases lexicales en plus grand nombre (Forsberg et Fant, 2010)

*Se placer en observateur*

*Y a de la vie*

*Faire un échange*

*Prendre ses précautions*

*Aller droit au but*

*Se poser des questions*

*Etre en danger*

*Arriver en ville*

*On va le dire*

Ce sont ces phrases lexicales qui font la grande différence entre un corpus natif et un corpus non-natif, en caractères gras dans le tableau 45, celles que l'on trouve exclusivement dans le corpus natif. Certaines sont idiomatiques et figées, *filer à l'anglaise*, *tomber à la renverse*, *être dans son élément*.

En conclusion, on constate que les locuteurs natifs utilisent des séquences que les non-natifs ne sont pas encore capables de produire même si les apprenants du groupe avancé possèdent un répertoire de SP lexicales plus étendu que celui des autres groupes. Parmi les séquences lexicales produites par les apprenants avancés, on trouve des expressions plus ou moins figées, *vis à vis*, *entre guillemets*, *en continu*, *coup de vent*, *système de conduite*, *par hasard*, *avoir des points communs*, *avoir une langue commune*, qui se rapprochent de celles des natifs. Les apprenants débutants peuvent aussi produire des séquences figées qu'ils ont mémorisées, (*tant pis*, AK). Mais on reconnaît un décalage entre la variété et la spécificité des SP lexicales produites par les locuteurs natifs et celles des apprenants. Ce qui semble somme toute tout-à-fait normal dans ce contexte particulier d'apprenants du français, qui n'ont pas eu un enseignement spécifique des SP et qui n'ont pas, loin de là, été exposés à la langue comme les natifs. Seules des productions de locuteurs très avancés pourraient être comparables en partie à celles des natifs (Forsberg-Lundell *et al.*, 2014). Cependant ce qui nous paraît le plus intéressant, c'est que les apprenants semblent avoir compris que les SP, en allongeant le segment, les aident à parler plus et plus longtemps et sans faire trop de pauses. Un grand nombre d'entre elles, notamment les discursives, servent d'appui à la planification et de « bouée de sauvetage » lorsque le locuteur se trouve devant une difficulté. Les SP permettent donc une meilleure fluidité et facilitent ainsi la production orale.

# Chapitre 9 Étude de deux SP, « c'est » et « petit »

---

Dans les modèles basés sur l'usage (chapitre 2), nous avons vu que la fréquence et l'input renforcent la mémorisation. En effet, en L2, les expressions ou les phrases entendues fréquemment (cours de langue, exposition à la langue cible) se transforment, à force de répétition, en formules de routine bien pratiques pour se débrouiller à l'oral. Parmi ces différentes structures enseignées dès les premières leçons, certaines, faciles à mémoriser, permettent aux locuteurs de communiquer très vite des idées ou de transmettre des messages. Selon Bartning et Hammaberg (2007) et Biber, Conrad et Cortes (2004), seule une exposition élevée à certaines structures permettrait de développer et d'acquérir une compétence langagière. C'est le cas de la structure *c'est* sur-employée par les apprenants. Nous avons identifié qu'une autre séquence récurrente qui prend la forme d'un patron formé avec l'adjectif *petit*, et non encore identifiée comme préfabriqué dans les recherches connues à ce jour, semble aussi jouer un rôle important en production orale. En effet, le nombre de syntagmes construit autour de l'adjectif *petit*, aussi bien dans le corpus natif que non-natif, justifie que l'on consacre une section à ses emplois et à ses fonctions.

## 9.1. Rôle de *c'est* dans le discours oral

Nous nous intéressons dans cette section aux différents emplois de *c'est* en français. Les grammaires décrivent *c'est* comme une structure à part, non canonique. A la suite de Bartning et Hammaberg (2007), nous avons identifié la séquence *c'est*, formule aisée à mémoriser qui a plusieurs fonctions dans le discours, comme un préfabriqué. Nous rappellerons tout d'abord les différents emplois de *c'est* tels que les grammaires les décrivent. Ensuite, après avoir examiné la distribution de *c'est* dans le corpus natif, nous la comparerons avec celle des locuteurs non-natifs.

### 1. les emplois de *c'est* comme présentatif

En français, la séquence *c'est* joue un rôle majeur dans la phrase et facilite la production d'énoncés. *C'est* fait partie des présentatifs au même titre que *il y a* et *voici, voilà* et permet en tant que tel la prédication. Il peut être employé par insistance afin de dégager le prédicat *c'est lui !*

Les présentatifs sont des mots ou expressions qui permettent de désigner quelqu'un ou quelque chose en rapport avec une situation : *il y a, il est, c'est, voici, voilà* sont des présentatifs [...] (qui) ont pour fonction d'introduire des noms (ou équivalents) [quelle que soit leur fonction], des infinitifs ou des propositions introduites par que... Les présentatifs sont d'un emploi assez fréquent en langue courante, car ils permettent de former des phrases indépendantes de façon simple. (Arrivé, Gadet et Galmiche, 1986 : 565)

On trouve également la formule présentative *il y a* dans les deux corpus mais en quantité plus réduite.

## 2. Une formule inanalysable

Dans *la Grammaire du français*, Wagner et Pinchon (1962 : 516) proposent « de ne pas traiter *c'est, c'est qui* (ou *que*) comme des syntagmes analysables. On fait mieux, semble-t-il, de les considérer comme de simples particules. *C'est...qui* (ou *que*) est une simple particule encadrante, comparable aux adverbes de négation composés *ne...pas* (ou *point*) entre lesquels s'insère le mot sur lequel porte la négation ». On peut noter que cette définition contient tous les arguments pour considérer *c'est* comme une formule (expression) prête à l'emploi, facile à utiliser à l'oral.

## 3. Une structure de phrase non canonique

Riegel, Pellat et Rioul, dans leur *Grammaire méthodique du français*, (1994 : 453-456 et 330-332) présentent dans le chapitre « les types de phrases », les phrases atypiques du français : « certaines structures de phrases fréquemment employées, en particulier à l'oral, non seulement ne correspondent pas au modèle canonique mais sont difficiles ou impossibles à dériver d'une phrase canonique ». Ils donnent entre autres exemples les phrases à présentatif.

Les présentatifs servent à présenter un groupe nominal ou un constituant équivalent qui fonctionne comme leur complément. L'ensemble présentatif + GN forme une phrase irréductible au modèle canonique. Cette structure est fréquemment employée à l'oral, car elle sert à désigner un référent dans la situation d'énonciation.

Comme par exemple *c'est mon mari*. Là encore, la grammaire ne peut expliquer la formation du groupe et nous renvoie à une catégorie d'expressions qu'il est difficile d'analyser syntaxiquement. Riegel et ses collègues (1994) constatent que *c'est* ne connaît pas de restrictions d'emploi si ce n'est qu'il ne peut s'employer seul. On trouve des formes au pluriel dans la langue plutôt soutenue suivies d'un GN au pluriel. Il varie également en temps et en mode, on trouve des formes à l'imparfait, des formes au pluriel mais la forme au singulier *c'est* est la plus fréquente. Notons qu'aucun des articles de ces grammaires ne parlent d'expression ni de locution, encore moins de préfabriqué.

*C'est* et *il y a* ne fonctionnent pas comme des constructions verbales et perdent de ce fait leur autonomie phrastique.

#### 4. Rôle de *c'est* dans les phrases emphatiques

*C'est* peut être utilisé dans d'autres structures de phrases comme dans les phrases emphatiques par extraction du constituant. *C'est le donneur qui ouvre les enchères. C'est une conversation d'ennemis que nous avons là !* A un niveau plus soutenu, on trouve des phrases avec des formes de détachement. L'extraction met en œuvre le procédé emphatique qui associe une locution identifiante et une relative pour en extraire un constituant et qui permet d'obtenir ainsi une phrase dite *clivée*. Mais on peut relier à cette structure standard une construction homologue, dite *pseudo-clivée* (ou semi-clivée), qui combine en fait l'extraction et le détachement d'un constituant.

##### a) Les phrases clivées

1. *Claire aime le chocolat = c'est Claire qui aime le chocolat*

2. *J'ai commandé le livre de Balzac = c'est le livre de Balzac que j'ai commandé*

##### b) Les pseudo-clivées

3. Ce que j'ai acheté, c'est une péniche

4. Celui qui a gagné, c'est Pierre

La phrase ici est séparée en deux parties.

##### c) dispositif d'extraction

5. C'est ça que je voulais dire

#### 9.1.1 C'est en production native

Nous avons trouvé dans les productions natives un certain nombre d'expressions construites autour de la forme du présentatif *c'est*. La haute fréquence d'emploi de *c'est* (comme celle d'ailleurs de *il y a*) chez les locuteurs natifs a été relevée par des chercheurs comme Blanche-Benveniste (1990), Gadet (1989), et Morel et Danon-Boileau (1998).

Les principales constructions avec **c'est** en production native

A. Un présentatif, une construction grammaticale (SP gram)

a) *c'est* + nom, rôle d'un présentatif (14)

**c'est l'histoire de** (.) (on trouve ici deux SP imbriquées *c'est* et *l'histoire de* sur le modèle de *le début de..., la fin de...*)

**c'est l'histoire d'** une /d'une abeille

voilà **c'est la fin du** film

**c'est la vie** d'une abeille

on ne sait pas si **c'est un ennemi** ou (.)

(persuadée que) **c'est une issue**  
**c'est l'histoire d'** un cowboy un peu solitaire  
**c'est** un p'tit film des minuscules  
**c'est** des peintures  
**c'est le but de** /(de libérer le grand cheval)  
**c'est** des fleurs en plastique

b) *c'est*, construction de mise en relief dite clivée (7)  
**c'est** un p'tit court métrage d'animation **qui**  
**c'est** la première ville **qu'**il voit  
**c'est l'histoire d'**une petite abeille **qui...**  
**c'est** vraiment la vie naturelle **qu'**elle peut avoir à l'extérieur  
**c'est** un cowboy **qui** arrive en ville  
**c' est** une porte **qui** se ferme  
**c' est donc la mouche qu'** on voit à l' intérieur

c) *c'est* + dislocation à gauche (3), procédé de mise en relief du premier élément  
*le problème c'est* comme on le voit bien dans la vidéo  
*le cheval à roulettes c'est...*

d) *c'est*, introduisant une proposition à valeur consécutive  
Le résultat **c'est que** la mouche  
**c'est vrai que** l' abeille *du coup* (dérivé de la forme soutenue *il est vrai que*)

*B. une séquence lexicale à valeur propositionnelle (SP lex)*

**c'est** + groupe adjectival  
**c'est pas facile de** savoir  
**c'est marrant**  
**c'est bien plus drôle**

*C. une séquence discursive (SP disc)*

a) *c'est* + groupe adverbial (ou adverbe)  
donc **c'est vraiment** ( *eah* l'animal **qui** est dans un &habi)

**c'est** pas vraiment un : : (champ digne de ce nom)

b) *c'est* suivi d'une pause vide ou remplie, valeur discursive

**c'est** (.)

**c'est** *euh* là où il habite

c) *c'est* + répétition, valeur discursive de *c'est*

**c'est** des /des vrais tournesols

d) *c'est* + *c'est* (*bien plus drôle que*)

le premier *c'est* est un marqueur discursif, il prépare la suite

e) abandon pour une autre structure

**c' est c' est** (*elle était tellement bloquée dan:s*)

*bon le dialogue on peut pas dire qui y' a un dialogue entre eux hein c' est (.) ils s' regardent...*

Dans ce corpus natif (cartoon et film confondus), on relève 32 occurrences de *c'est*, soit une moyenne de 5,3 occurrences par sujet. La construction *c'est* suivie d'un nom, la plus courante à l'oral, permet de présenter un objet d'une manière aisée sans trop d'effort, d'ouvrir un récit ou de le clore. On trouve 14 occurrences de la construction *c'est* + N, c'est-à-dire une moyenne de 2 par sujet. La construction plus complexe dite « clivée » est employée 7 fois, soit moitié moins que la construction simple *c'est* + GN.

Nous avons classé *c'est* dans les trois catégories selon le rôle que joue cette séquence dans le discours. Elle peut être considérée comme une construction qui introduit un groupe nominal, pronominal ou une proposition. *C'est* est une séquence lexicale lorsque le groupe qu'il introduit a une valeur propositionnelle, *c'est pas facile*. *C'est* est une séquence discursive lorsqu'elle facilite le traitement, en libérant du temps pour la recherche lexicale (ou autre). *C'est* est alors suivi d'une hésitation (pause vide, pause remplie, groupe d'hésitation, répétition) la pause permet au locuteur le temps de trouver le mot juste ou la bonne formulation, *c'est que* (GH), *c'est* (.) *c'est* des / des, *c'est euh*. Dans les exemples en d) *c'est* est un marqueur discursif lorsque le locuteur abandonne cette construction pour une autre formulation ; on trouve deux cas d'abandon : la construction avec *c'est* est abandonnée (*c'est c'est // elle était tellement bloquée dans*) et le locuteur choisit une autre formulation ; et le premier *c'est* de la séquence redoublée (*c'est c'est drôle*). De même, la séquence *c'est* est une séquence discursive lorsque elle est suivie d'un adverbe lui-même suivi d'une hésitation *c'est vraiment* (.)

Les locuteurs natifs malgré leur L1 commune, ont des comportements langagiers différents. Certains sujets comme MG, s'appuient sur la séquence *c'est* beaucoup plus souvent (16 occurrences) que d'autres, JE en produit une seule, TH seulement deux, CA trois. On a déjà évoqué JE au profil différent des autres natifs. Il semble avoir moins besoin de s'appuyer sur cette SP. Il produit cependant une structure en clivée pour ouvrir la narration du film *c'est l'histoire d'une abeille qui*. On trouve une autre structure pour débiter la narration de la bd (*on peut voir*). Son discours est plus construit, peut-être moins spontané que d'autres.

Tous les sujets sans exception utilisent une formule pour l'ouverture du récit ; mais tous ne choisissent pas la SP *c'est*, la plus courante néanmoins pour débiter un récit oral. On trouve dans le corpus natif sept occurrences de *c'est*, formule d'ouverture du récit : des formules complètes qui comprennent deux SP imbriquées comme *c'est l'histoire de* (3), *c'est la vie de* (1), la séquence *c'est* présentatif suivi d'un nom (1), *c'est un petit film*, ou la construction plus complexe dite clivée (2), *c'est un p'tit court métrage d'animation qui s'appelle prisonnière, c'est un cowboy qui arrive en ville*. Cinq des narrations ne débutent pas par *c'est* mais par une autre formule, toutes identifiées comme des SP : « *sur la bd on peut voir un cowboy...* », *il s'agit probablement de, alors... on voit un homme, ça raconte l'histoire*, dans cette histoire *on a* (*on peut voir, il s'agit de, on voit, on a, ça raconte*). Nous ne trouvons aucune formule de clôture avec la SP *c'est*.

### 9.1.2 *C'est* dans les productions non natives

Bartning et Hammaberg (2007), dans un article qui fait date, se sont intéressées à la construction avec *c'est* en français et son équivalent *der ärt* en suédois. *C'est* remplit de nombreuses fonctions chez les locuteurs natifs mais aussi chez les non-natifs (Bartning et Hammaberg 2007 : 2). Placé souvent en début de phrase (énoncé à l'oral) pour amorcer une narration ou pour commencer un énoncé d'une manière aisée, *c'est* joue donc un rôle central dans la structuration du discours et de la phrase.

L'emploi des présentatifs et en particulier *c'est* permet donc au locuteur non-natif de s'exprimer en l'absence souvent d'un vocabulaire précis / juste notamment lorsqu'il y a un manque lexical, en particulier un déficit en verbes. La séquence « *c'est* » facilite le processus langagier : tout en évitant la parole discontinue, elle permet rapidité et fluidité de la parole, et son emploi peut, de ce fait, être une aide à l'encodage. Enseignée dès les premières leçons en L2, les locuteurs s'approprient cette séquence très tôt dans l'apprentissage, par exemple dans les présentations (*c'est un professeur, c'est une fille, c'est mon ami etc.*).



Comme dans le corpus L1, la construction avec *c'est* est présente dans différentes structures dont les deux principales en L2 comme d'ailleurs en L1 sont *c'est + nom* (*c'est un cowboy*, *c'est une série*) et *c'est + adj* (*c'est marrant*, *c'est amusant*) (voir aussi Towell *et al.* 1996 ; Raupach, 1984 ; Bartning et Hammaberg 2007 : 22-23 ; Théophanous et Perez-Bettan, 2012).

*C'est + nom* est le plus souvent utilisé pour introduire un protagoniste dans un récit. Cet emploi est souligné par Véronique (2013 : 54) qui cite l'exemple d'une apprenante de français de niveau avancé. Il remarque que l'introduction d'un protagoniste peut être effectuée à l'aide de *c'est* (et *d'il y a*), alors que d'autres constructions comme la dislocation n'apparaît que dans le corps du récit et avec une fréquence faible.

La structure *c'est + adj*, utilisée le plus souvent pour exprimer un commentaire à un stade peu avancé, peut devenir par la suite plus complexe : *c'est pas facile à expliquer* (FA). Ici FA complète l'expression par un groupe prépositionnel infinitif (*adj + Vinf*). Certains sujets débutants tentent cette construction au risque de faire une erreur comme par exemple JA (niveau 1), *c'est difficile \*pour sortir* pour *c'est difficile de sortir* (*c'est difficile de sortir*, *difficile à expliquer* mais *difficile pour* toi de +V).

De plus, *c'est* étant polyfonctionnel, il joue un rôle significatif dans la planification des énoncés et dans la structuration des phrases : produire des expressions de façon très simple (présentation d'un objet, lieu, personnage), contourner une difficulté lexicale (méconnaissance d'un verbe), débiter un récit, le clôturer. L'apprenant L2 peut ainsi arriver à se débrouiller même si son lexique est limité. Cette séquence associée à un adjectif permet à l'apprenant d'exprimer un sentiment, *c'est drôle*. Avec peu de vocabulaire quelques noms et quelques adjectifs et grâce à la structure *c'est*, l'apprenant L2 peut arriver à dire quelque chose. Il peut en outre s'appuyer sur elle pour planifier la suite de son discours. *C'est* peut avoir aussi une fonction discursive comme dans *c'est c'est c'est drôle la conclusion*, où la formule *c'est c'est* permet de dégager du temps pour trouver la formulation appropriée.

La séquence *c'est* est considérée comme une séquence préfabriquée dont le rôle serait de faciliter le traitement de l'expression (Wray, 1999 ; Wray et Perkins, 2000 ; Forsberg, 2006, Bartning et Hammaberg, 2007 : 3). Au niveau débutant, *c'est* permet de pallier une difficulté lexicale et peut être employée en bloc à la place d'une autre structure que le locuteur ne peut pas activer. JA produit « *un carré que : \* le\* garçon : c'est (.), pour le garçon est dans le carré* (carré désignant la vignette de la BD). Et dans le film « *elle est c'est difficile \*pour sortir de la maison* », JA abandonne la structure *elle est + adj* car l'adjectif *enfermée* ou *prisonnière*, qui

s'imposerait ici, n'a pu être activé. Elle poursuit sa narration grâce à la SP *c'est*, sans toutefois réussir à trouver la formulation exacte qui pourrait être *c'est difficile pour elle de sortir de la maison* ou encore *c'est difficile de sortir de la maison*. *C'est* peut donc remplacer un mot ou une structure non encore acquise ou maîtrisée. Même à un niveau plus avancé, l'apprenant peut s'appuyer sur cette structure, *c'que c'que \*c'est arrivé pour ce qui est arrivé* (VI, niv avancé). *C'est* a sans aucun doute un impact sur le développement de l'interlangue des apprenants. En conclusion, nous dirons que *c'est* est, parmi toutes les stratégies de recherche lexicale, la stratégie favorite des apprenants. Souvent répétée et combinée avec des pauses remplies ou vides aussi bien qu'avec des faux départs, cette formule facilite la planification.

#### Occurrences de *c'est* dans le corpus non-natif

Les apprenants dans notre étude utilisent au moins une fois le syntagme *c'est* quel que soit leur niveau en L2. 209 occurrences de *c'est* ont été relevées dans le corpus non-natif, pour les deux temps, soit une moyenne de 5,8 par apprenant en t2, ce qui est à peine plus élevé que la moyenne des natifs. Nous présentons ici les exemples les plus caractéristiques de notre corpus (Bartning et Hammaberg 2007 : 7-20)

#### Les différentes fonctions de *c'est*

*C'est* peut avoir plusieurs fonctions dans l'organisation du discours

a) un marqueur textuel : il est accompagné d'un élément textuel (structures 1-2-3-4-5)

- *c'est + nom*
- *c'est + adj.*
- *c'est + adv.*
- *c'est + p. passé*
- *c'est + pronom*

Il peut être suivi du relatif *qui/que* à un niveau plus avancé d'apprentissage (rôle d'emphase).

Cette structure dite « clivée » est peu fréquente au niveau débutant :

- *c'est le vent qui a fermé la porte* (PR, F2, niv3)

b) un marqueur autorégulateur (structure 10) : sorte de rampe de lancement, il permet au locuteur de planifier la suite, de gagner du temps afin de produire l'expression juste.

- *c'est c'est (c'est drôle)* (VI, BD1)

*C'est* fonctionne comme un remplisseur, il peut jouer le même rôle que *and then* en anglais (Lennon, 1984). La séquence peut être suivie d'une pause remplie (*eah*) mais parfois elle n'est

pas suffisante, la répétition de la séquence permet alors de réduire les problèmes de planification : *c'est c'est / c'est drôle* dit cette fois en bloc.

*c'est* peut être un moyen de contourner une difficulté syntaxique

Ex : (l'abeille) fait quelques mouvements (.) *mais je ne sais pas \*c'est pourquoi* (= *je ne sais pas pourquoi* (elle les fait) (OL, F1).

*c'est* est parfois utilisé sans suite (structure 14), l'expression est abandonnée, ou suivie d'une pause vide ou remplie (structure 13). On trouve d'ailleurs ce même phénomène chez les natifs. Le locuteur commence une expression avec *c'est* mais il l'abandonne pour des raisons multiples, dont la plus courante pourrait être un problème d'ordre lexical. Mais produire cette structure permet à l'apprenant de réfléchir et de garder la parole encore un moment.

c) une structure propositionnelle

- *c'est marrant*
- *c'est drôle*

*c'est* peut être considéré comme une SP lexicale à structure propositionnelle comme dans les structures 1 et 2. En général, cette séquence clôture la narration en donnant le point de vue du narrateur.

Nous présentons ci-dessous les différentes structures avec *c'est* que l'on trouve dans le corpus non-natif. Nous ne ferons pas ici une étude comparative exhaustive des emplois de *c'est* en temps1 et en temps2. Nous présentons quelques exemples afin d'illustrer notre propos. *C'est* a été identifié comme une SP qui peut être lexicale, grammaticale ou discursive.

#### A. **C'est : une construction grammaticale**

##### **Structure 1.** *c'est + nom*

**c'est** l'histoire

**c'est** un film

**c'est** une salle de bains

**c'est** les fleurs/la fleur

**c'est** la mouche

**c'est** un petit \*cheveu

**c'est** le héros

**c'est** juste un cadre / **c'est** juste un dessin

La construction la plus courante *c'est + nom* introduit le prédicat.

##### **Structure 2.** Construction en clivée

le deuxième film *c'est* l'histoire *qui c'était* faux... (Henrique F1)

*c'est* l'histoire où (.) (Ma BD1)

*c'est* l'histoire de l'abeille *qui* est (.)

*c'est* \*le mouche *qui* vole (OL F2)

*c'est* le vent *qui* a fermé la porte (Pr F2)

Le ligne *c'est* le chose *qui* est dans le... (OL, BD1)

*C'est* l'histoire \**qui* la mouche aide le petit jaune... (OL, F1) (\**qui* pour où)

*C'est* l'histoire d'une petite abeille *qui* euh qui était emprisonnée dans une maison... (PR, F2)

*c'est* le mouche *qui* est très sympa... (QI, F1)

et *c'est* la porte \**que* se ferme... (VI, F2)

*c'est* l'histoire de d'une homme *qui* est avec-euh son cheval... (VI, BD1)

*c'est* l'histoire de d'une abeille *qui* se promène parmi les fleurs... (VI, F2)

*c'est* une mouche *qui* est dedans... (VO, F1)

*c'est* la mouche *qui* est fermée dans une maison... (VO, F2)

*c'est* elle *qui* a indiqué le chemin (VO, F2)

*c'est* la première fois *que* (PR, F2)

*c'est* comme ça *qu'*elle est est sortie de la maison... (QI F2)

L'énoncé avec *c'est* n'est pas toujours dit d'une traite. Une pause peut précéder le relatif (dans ce cas nous n'avons pas compté *qui* dans la SP). La pause après *mouche* permet à Jacqueline de réfléchir à la suite de son énoncé. La phrase clivée met en relief un élément de l'énoncé, qui peut être un nom, un infinitif ou un pronom. Nous n'avons trouvé qu'un seul exemple de pronom mis en relief *c'est elle qui* (VO, F2) mais aucun exemple avec infinitif.

**Structure 3 :** *c'est* + Proposition relative indéfinie

*c'est* ce que la mouche fait (Vinicius F2)

*c'est* ce que je pense

**Structure 4 :** *c'est* + construction pseudo-clivée

Ce que j'ai vu (.) *c'est* une petite histoire d'une abeille (MJ, F1)

On trouve une seule occurrence de cette construction, typique de l'oral.

**Structure 5 :** *c'est* + circonstance

*c'est* dehors

*c'est* en plastique

*c'est* \*en ce moment-là

*c'est* dans un jardin

Cette structure permet d'indiquer une circonstance de l'action (temps, lieu, matière) d'une manière plus simple en l'absence d'un verbe précis.

**Structure 6.** *c'est* + dislocation à gauche

*le ligne c'est* (.) [le chose qui est dans sa main]

*la mouche c'est* pas mon amie (OL F2)

Et *la maison c'est* elle

**Structure 7.** *c'est* + dislocation à droite

*c'est drôle* la conclusion

Dans les structures 6 et 7, la mise en relief se fait par extraction d'un élément détaché soit à gauche soit à droite. Ce procédé de mise en relief par détachement d'un élément est peu utilisé (voir Véronique, 2013).

**B. c'est, séquence lexicale à valeur propositionnelle**

**Structure 8.** *c'est* + adj.

**C'est un peu différent**

**C'est amusant**

**C'est drôle**

**C'est marrant**

La structure *c'est* suivie d'un adjectif permet au locuteur de donner son point de vue, son opinion ou de proposer une interprétation. Nous les considérons comme des séquences référentielles.

**Structure 9.** *c'est* + groupe adjectival prépositionnel

On trouve quelques exemples de *c'est* suivi d'un groupe adjectival. Ces structures déjà plus complexes sont suivies le plus souvent d'un verbe à l'infinitif.

**c'est pas facile à** expliquer

**c'est impossible pour** elle

**c'est difficile \*pour** sortir (de sortir)

**c'est difficile \*pour** expliquer

Dans ces deux derniers exemples, l'apprenant utilise une préposition incorrecte à la place de « *difficile de, à* » ; il ne semble pas en mesure de la produire correctement à son niveau.

**c'est pas bien** pour elle (c'est + groupe adverbial)

Cette structure avec *c'est* permet d'exprimer ou d'atténuer une opinion (*c'est pas bon, c'est mal*).

### C. c'est : une séquence discursive

**Structure 10.** *c'est* + adverbe (introduit un commentaire)

**c'est pas vraiment**

**c'est pas bon** (pour elle)

**c'est pas mal** (pour)

**c'est mal**

**c'est pas trop**

**c'est pas très**

**Structure 11.** *c'est* + participe passé

**c'est fini**, marqueur discursif de clôture

**c'est réussi**, utilisation de type méta-discursif

Ici le participe passé utilisé comme un adjectif marque la conclusion de l'événement ou de la narration.

**Structure 12 ;** *c'est* + pronom indéfini

**c'est ça** (4)

**c'est tout**

*c'est* est suivi du pronom neutre *ça* ou de l'indéfini *tout*. Nous considérons l'ensemble comme une séquence à valeur discursive permettant au locuteur de signaler la fin de sa narration.

**Structure 13 :** *c'est comme ça* (SP)

Cette SP est la réunion de deux SP *c'est* et *comme ça*. Là encore, le locuteur résume ce qui vient d'être dit et/ou apporte une conclusion à sa narration, marqueur discursif lui aussi. Nous avons considéré la SP dans son ensemble *c'est comme ça*.

**structure 14.** *c'est* répété, rampe de lancement

**c'est / c'est** un [homme]

**c'est c'est / c'est** drôle

**c'est / c'est** quand même drôle

il croit que **c'est / c'est** des vrais tournesols

**c'est / c'est** un film

Cette structure répétée avec ou sans pause peut être interprétée comme un moment de planification alors que la séquence n'est pas complètement terminée. La répétition de *c'est* permet au locuteur de gagner un peu de temps pour trouver la bonne formulation, en l'occurrence une autre SP, afin de terminer la séquence.

**Structure 15.** *c'est* + pause vide ou remplie

**c'est** + euh

**c'e:st** (.) **c'est** la &pl la place de:

*c'est* peut être suivi de plusieurs types d'hésitation : la séquence *c'es:t*, comportant une voyelle allongée, est suivie d'une pause vide et d'une répétition (déterminant + fragment de mot) ; ces hésitations permettent au sujet de trouver le mot juste et sont une aide à la planification ; la pause après *c'est* indique que le groupe est dit en bloc et cela malgré la syllabe allongée (chapitre 5, méthodologie).

**c'est c'est** (.) &pro produit un peu triste

La répétition suivie d'une pause et d'un fragment de mot montre le processus de planification à l'oeuvre.

**Structure 16.** *c'est* + abandon

la nature **c'est**...

Le locuteur produit un début d'énoncé avec dislocation à gauche qu'il abandonne sans doute en l'absence du vocabulaire approprié ou parce qu'il veut passer à autre chose.

**Structure 17 :** *c'est* + reformulation

**c'est** de::s (bon) (.) // **c'est** un vrai champ

**c'est** un // ce &mous moustiche (HE, F1)

Les structures 14 à 17 avec *c'est* ne sont pas complètes et marquent une rupture dans la formulation. En effet, tout discours oral comporte des ruptures, des disfluences dues à la nature même de l'oral qui est, par définition, spontané et pas toujours planifié avant l'exécution. A la différence de l'écrit où le scripteur peut revenir en arrière et se corriger pour ne laisser que la version finale expurgée de ses scories, l'oral ne permet pas cela et nous nous trouvons devant

un énoncé privé de ces corrections. A l'oral, les énoncés sont souvent produits avec des ruptures dans leurs structures syntaxiques, qui peuvent être parfois dues à un blocage lexical ou/et un blocage syntaxique. Cela est d'autant plus vrai lorsqu'il s'agit de productions L2. Ces ruptures peuvent prendre l'aspect d'une répétition comme dans *c'est c'est*, d'une pause vide ou remplie, d'un allongement de syllabe qui peut aussi être renforcé par une pause. Bien souvent ces pauses permettent au locuteur de planifier la suite de son discours tout en gardant la parole. Il arrive aussi que le locuteur abandonne après *c'est* pour une autre structure ou pour passer à autre chose. On trouve ces phénomènes aussi bien chez les natifs que chez les non-natifs mais plus souvent, bien évidemment, chez les locuteurs non-natifs.

*C'est* est présent dans les récits narratifs, à l'écrit dans les *incipit* notamment mais aussi et surtout à l'oral. Morel et Danon-Boileau (1998) notent qu'on peut trouver la séquence *c'est* dans le préambule qui annonce la partie rhématique.

*C'est* chez les apprenants L2 est une SP utilisée de manière récurrente. Facile à mémoriser, elle facilite la production en étant une alternative à des structures plus complexes non encore acquises ou difficilement utilisables en production spontanée et en temps réel. Cette séquence enseignée en cours de langue dès les premiers stades de l'apprentissage est sur-employée par les locuteurs non-natifs en t1 et à peine moins en t2. En L2, l'une des thèses de Deulofeu (1986) se trouve ainsi justifiée : « (...) les schémas d'énoncés ont un caractère plus universel que les constructions grammaticales et [...] (ils) vont donc se développer plus rapidement » (Deulofeu, 1986 : 276). On trouve pratiquement les mêmes structures que celles des natifs avec cependant plus de structures 14 et 15 (*c'est* comme « rampe de lancement » ou suivie d'une pause) dans le corpus L2.

Pour conclure, nous dirons que la séquence *c'est* est une structure utilisée aussi bien par les locuteurs natifs que les non-natifs avec cependant, en L1, des structures complexes plus nombreuses. La séquence *c'est* à l'oral est donc une structure polyfonctionnelle (présentatif, mise en relief, marqueur discursif, remplisseur). Elle fait partie des stratégies de l'oral en L2 (structuration du discours et planification des énoncés).



## 9.2 Les séquences construites avec l'adjectif « petit »

La fréquence d'emploi des syntagmes construits avec l'adjectif *petit* par les locuteurs natifs et non-natifs nous a amenée à les identifier comme des séquences préfabriquées ou en cours de préfabrication. Ces séquences polylexicales, comportant une partie fixe et une partie variable, sont prononcées avec un contour mélodique et parfois une réduction phonologique. Elles semblent également jouer un rôle dans le discours. Nous ferons référence à quelques chercheurs qui ont montré que cet adjectif en français avait des caractéristiques sémantiques bien particulières. Devant la fréquence d'emploi de cet adjectif, voici ce que dit de façon humoristique Charles Timoney (2007) dans un ouvrage de la même veine sur le français<sup>72</sup> : « On pourrait sans doute écrire une thèse sur l'usage du mot *petit* en français ! ». Cet adjectif, sans doute l'un des plus courants de la langue française, a une place à part parmi les adjectifs. Employé de manière récurrente à l'oral, associé à des substantifs très variés, il est facile à retenir et ne pose aucun problème de compréhension en L2. Il ferait même partie des soixante mots qu'on retrouve dans toutes les langues naturelles.

Des recherches entreprises depuis les années soixante-dix par Wierzbicka (1972, 1984, 1992, 2006), et aidée par Goddard plus tard (Peeters, 2012a : 1896), ont permis d'élaborer la « métalangue sémantique naturelle » (MSN). Dans cette métalangue, Wierzbicka a classé soixante mots dits « primitifs sémantiques », dont l'adjectif *petit* à côté de *grand* (ce sont les deux seuls adjectifs de la MSN), qu'on trouve dans toutes les langues du monde, pouvant être traduits dans toutes les langues naturelles sans risque de détournement sémantique. Enseignés dès les premiers stades de l'apprentissage, ces mots ne posent aucun problème de sens, en principe, pour les apprenants d'une langue étrangère.

Si l'adjectif *petit* en fait partie, il tient certainement une place à part en français, contrairement aux autres adjectifs de la métalangue. C'est une des raisons, peut-être, qui explique le nombre important de groupes nominaux avec *petit* qu'on trouve dans les corpus natif et non-natif. Dans la métalangue sémantique naturelle, l'adjectif *petit* a le sens de réduction (taille ou quantité). Dans la version française de la métalangue sémantique naturelle, l'adjectif *petit* est très différent des autres primitifs car il semble avoir perdu en quelque sorte son sens originel. En effet, en français, il est utilisé très souvent sans aucune référence à la taille ou à une quelconque réduction. On le trouve, d'ailleurs, souvent associé à des situations où, selon Peeters (2012a :

---

72 « You could probably write a thesis on the use of the word *petit*. » (Timoney, 2007 : 127 cité par Peeters, 2012, en exergue de SON article)

1898), il prend « la forme de scénarios culturels qui expriment avec le même degré de précision que toute autre explicitation formulée en MSN, les normes, les valeurs et les usages que partagent la plupart des ressortissants d'un univers particulier ». Autrement dit, l'emploi de l'adjectif *petit* par les locuteurs français serait fortement lié à leur culture. Kerbrat-Orrechioni (2005 : 262) a constaté que, dans les rapports quotidiens, *petit* a une place à part qui dépasse largement son sens diminutif. Voici un exemple de dialogue - tiré de Kerbrat-Orrechioni (2001) - entre un vendeur (V) et une cliente (C) qui semblent partager les mêmes valeurs et font partie du même univers.

V madame, bonjour

C bonjour, moi j'aurais voulu un *petit* bifteck haché s'il vous plaît

V un gros ?

C moyen. »

En trois répliques, on passe de l'adjectif *petit* à gros et à moyen. Que veut réellement la cliente, un bifteck « *petit* » ? Certainement non, la taille du bifteck n'est pas l'objet de la demande de la dame mais une façon détournée de minimiser la requête. L'adjectif *petit* ici ne peut pas être pris dans son sens dimensionnel alors que dans d'autres exemples, il peut encore avoir ce sens. Dans la question suivante qu'on a toutes entendues (ou qu'on entendra un jour) « *et la p'tite dame qu'est-ce qu'elle veut ?* », il y a peu de chance pour que le locuteur s'adresse ainsi à une femme de haute taille. Cependant on voit bien que *p'tite dame* ici ne s'adresse pas à une femme *petite* en taille, non plus. Ce phénomène semble, d'après les chercheurs, propre au français.

Avant de présenter notre corpus, nous donnons quelques exemples de syntagmes construits avec *petit* qui a perdu son sens primitif : on dit « *comment s'est passé ton petit week-end ?* », même si le week-end est prolongé, « *mon petit chéri* » à son amoureux même si celui-ci est plutôt grand en taille ! des *petits soucis*, alors qu'ils sont énormes, un *petit problème*, pour un véritable casse-tête, *ma p'tite dame*, *mon p'tit cœur*, où *petit* ne renvoie nullement à l'idée dimensionnelle. Lorsqu'on désigne un individu par sa nationalité, invariablement on parle d'un « *petit Belge* », d'une « *petite Anglaise* », d'une « *petite Française* », etc. sans aucune idée de réduction. La fréquence d'emploi de *petit* en français dans des cas où il ne fait pas référence à la taille justifie qu'il ne faut pas prendre le mot systématiquement dans son sens littéral. En effet, le syntagme avec *petit / p'tit* n'est jamais anodin. Si on peut trouver peut-être une connotation dimensionnelle dans certaines expressions, un *p'tit café*, un *p'tit tour*, un *p'tit weekend*, un *petit noir*, un *petit bébé*, elle est le plus souvent redondante et associée à l'expression d'un sentiment : tendresse, affection, bienveillance. L'expression avec *petit* peut

parfois être ironique, voire méprisante, *tu oublies toujours tout, tu as une petite tête* ! D'après Garnier Mathez (2005 : 11), il serait par sa nature un « marqueur privilégié de la subjectivité du locuteur ».

Si le français a tendance à rajouter *petit* à tout va, certains syntagmes avec *petit* se sont aussi lexicalisés, comme *petit déjeuner*, *à la petite semaine* (au jour le jour), *au petit bonheur* (au hasard), *chercher la petite bête* (être tatillon, critique à l'excès), *en petit comité* (en petit nombre, entre intimes) le *petit écran* (la télévision), *petit boulot* (travail précaire, sans qualification), *petite annonce* (bref texte d'offre ou de demande publiée dans la presse), *petite main* (couturière qui débute dans le métier), *petite vérole* (variole) et d'autres en voie de lexicalisation comme *petit ami* ou *petite amie* qui, par ailleurs, n'existent pas dans d'autres langues, en anglais on parle d'un *boyfriend* ou d'une *girlfriend*.

Une des explications à cet emploi de *petit* a sans doute une origine historique dans la formation des diminutifs des langues romanes. D'un point de vue diachronique, le français se distingue des autres langues romanes, comme l'italien et l'espagnol, qui utilisent de nombreux suffixes diminutifs là où le français utilise l'adjectif *petit*. Dans son sens historique, l'adjectif *petit*, joue un rôle de diminutif qui aurait la même valeur que le suffixe diminutif –et/-ette (*garçonnet*, *fillette*), seul à avoir gardé pleinement ce sens dimensionnel. Si l'italien et l'espagnol ont deux adjectifs correspondant à *petit*, *piccolo* et *pequeño*, ils sont employés moins fréquemment que *petit* en français (Hasselrot, 1957). En revanche, l'italien utilise des suffixes qui apportent une subjectivité au sens dimensionnel, *un caffè* deviendra *un caffettino* (un petit café, qui me fait plaisir), *una caramellina* (un petit bonbon), *un vestitino* (une petite robe), *un cagnolino* (un petit chien mignon).

Compte-tenu de l'emploi récurrent de *petit* dans notre corpus, nous avons cherché si des études avaient été entreprises sur ce sujet. Des chercheurs, des philologues comme Hérisson (1954), se sont penchés sur la dérivation diminutive, mais peu de linguistes ont approfondi cet emploi en français mis à part Hasselrot (1957), Togeby (1982) et Kerbrat-Orrechioni (2004a). Togeby (1982 : 52) constate que « *petit* » forme souvent avec le substantif une sorte de mot composé, voire une « collocation ». Certains parlent même d'expressions idiomatiques. Peteers (2012a) évoque les propos (écrits) d'une libanaise, installée en France, qui va jusqu'à dire de façon très humoristique que l'on peut connaître très facilement « les expressions idiomatiques de France », à commencer par « la formule qui consiste à placer “*petit*” devant toute chose : *petite laine*, *petit café*, *petit dîner*, *petit tour*, *petite pause*, *petite mine*, *petit temps*, etc. » (Khouri-Dagher,

2008 : 158-159). Ces formules ne renvoient nullement au sens dimensionnel de l'adjectif mais sa présence traduit / exprime quelque chose de l'ordre de l'affectif.

Dans notre corpus, les nombreuses séquences avec *petit* ne sont pas toujours utilisées dans un sens dimensionnel même si, comme le dit Kerbrat-Orrechioni (2004a), on ne peut évacuer totalement ce sens. Si l'adjectif *petit* est le marqueur d'une certaine subjectivité, il jouerait aussi un rôle en communication dans les actes de langage. D'après Kerbrat-Orrechioni (2005 : 198), l'adjectif *petit* a différentes fonctions, dont celle (très répandue) de l'acte de langage valorisant *Face-Flattering Act* (FFA) ou « actes flatteurs », « en face à face », par opposition aux actes menaçants. On peut ainsi trouver cette fonction dans des syntagmes tels que « *un petit Belge* » ou « *ma petite dame* » qui auraient une fonction valorisante ou flatteuse (voire affectueuse !).

L'adjectif *petit*, dans les actes de langage, peut prendre d'autres valeurs dont la valeur affective, c'est-à-dire qu'il « indique l'estime, la tendresse, l'amour de celui qui l'emploie envers la personne ou l'objet ainsi qualifié, et la gentillesse, la beauté, l'excellence ou la qualité de ces derniers ». (Hérisson, 1954 : 49) ; « *Petit* « notre façon à nous de dire mignon » ajoute une dimension affective au nom qu'il qualifie » ajoute Perrin Debock (2006 : 250)

Il peut aussi avoir une fonction minimisante, sorte d'euphémisme, comme dans cette séquence lexicalisée : « J'ai ma *petite* idée là-dessus » (qui exprime un point de vue personnel auquel on tient beaucoup). En revanche, c'est *ta petite idée* !, aura une valeur dépréciative. Le sens exprimé par *petit* est donc fonction du contexte.

Bartens et Sandström (2006), attribuent huit fonctions à l'adjectif *petit*. Une fonction affective AFF, une fonction diminutive DIM (lorsque *petit* a son sens dimensionnel) et six autres correspondant à différents suffixes de l'espagnol et/ou de l'italien. En français, en l'absence de suffixation (à part les diminutifs -et/-ette qu'on a vus plus haut) c'est l'adjectif *petit* qui en assume les fonctions.

- fonction minimisante **MIN** :

Pourriez-vous me donner un *petit coup de main*, vous n'auriez pas une *petite cigarette* ?

- fonction euphémique **EUPH** :

Elle relève de la politesse négative comme dans la séquence lexicalisée « *petite mine* » ou « vous avez pris un peu de poids ou vous avez un *petit ventre* rebondi naturellement ? » dans les critiques ou remarques « une *petite remarque* » « un *petit problème* » « une *petite objection* à ton projet »

Ces trois dernières étant plus rares,

- fonction humoristique **HUM** : *petit tour de piste*

- fonction ironisante **IRON** : *petite tête* ! ça ne tourne pas rond dans ta *petite tête* !
- fonction péjorative **PEJ** : *petit esprit, petit personnage, petites gens, petite nature, mon petit monsieur* (Hérisson, 1954 :50), *petit glandeur* (fam.) « ouvre tes *petits yeux* et regarde ! » ; « vous avez chacun un vélo au cas où vos *petites jambes* se fatigueraient » ; « Continue à te faire ton *petit ulcère* dans ton coin, ça divertit, mais rien de plus. Allez ciao bozo ! » ; « Arrête de faire ton *petit numéro*, ça ne marche plus ! »

- fonction nourriture **NOUR** : souvent des diminutifs en italien et en espagnol s’ajoutent à des noms d’aliments, de boissons ou de plats et peuvent aussi avoir un sens affectif. En français nous entendons souvent des séquences du type *petit bifteck, petite baguette, petit apéritif, petit dessert, petit café* (« un *p’tit café* ? »).

Ces emplois recouvrent, à notre avis, la fonction AFF ou MIN et non NOUR. Mais nous ne développerons pas plus cette approche, très intéressante, notamment lorsqu’on a un corpus composé de nombreux syntagmes avec l’adjectif *petit*.

Les corpus de productions natives et surtout non natives contiennent un grand nombre d’expressions construites avec l’adjectif *petit*. Tous les syntagmes avec *petit* prononcés avec fluidité, sans hésitation interne, répondent aux critères nécessaires de polylexicalité et cohérence prosodique et à un critère gradué, celui de fréquence. Ils ont donc été identifiés comme des SP. Comme nous ne pouvions pas distinguer cette fonction avec certitude, nous avons fait le choix de compter tous les syntagmes avec *petit* comme des SP. D’autre part, les syntagmes avec *petit* facilitent le traitement en permettant à l’apprenant, du fait de son emploi, de résoudre une difficulté lexicale, d’allonger le segment et de parler avec plus d’aisance. Toutes les conditions étaient requises pour identifier les syntagmes avec *petit* comme des préfabriqués au sens où ils facilitent le traitement langagier.

#### A. Emploi de l’adjectif *petit* dans le corpus natif

Nous présentons sous forme de tableau les emplois de l’adjectif *petit* chez les locuteurs natifs.

Sujet	BD		Film	
AN	<i>un p'tit enfant</i>	1	<i>un p'tit court métrage</i>	1
	<i>un p'tit cheval en bois</i>	2	<i>une p'tite abeille</i>	1
			<i>un p'tit recoin</i>	1
			<i>les p'tits grains de café</i>	1
			<b>une petite flèche</b>	1
CA	<i>un p'tit garçon</i>	1	<i>la p'tite abeille</i>	1
	<i>un p'tit cheval en bois</i>	1	<b>le petit poucet</b>	1
	<i>le p'tit garçon</i>	1		

	<b>le petit garçon</b>	1		
<b>JE</b>	<b>un petit enfant</b>	1	<b>une petite abeille</b>	1
	<b>le petit enfant</b>	1	<b>un petit moment</b>	1
	<b>son petit cheval</b>	1		
<b>MG</b>	<i>un p'tit somme</i>	1	<i>un p'tit film</i>	1
	<i>le p'tit garçon</i>	1	<i>un p'tit film muet</i>	1
	<i>le p'tit garçon</i>	1		
<b>TH</b>	<b>un petit garçon</b>	1		
	<b>le petit garçon</b>	1		
<b>VC</b>	<i>un p'tit garçon</i>	1		
	total BD	16	total film	11
	20 avec plus petit			
	<i>p'tit</i>	8	<i>p'tit</i>	7
	<b>Petit</b>	8	<b>Petit</b>	4

Tableau 46. Occurrences de "petit" chez les locuteurs natifs

On trouve 27 occurrences de *petit* dont 15 occurrences de *petit* prononcé avec réduction phonologique (*p'tit*).

**BD** : 16 occurrences, *petit* (6) et *p'tit* (10)

<i>petit enfant</i> (2)	<i>p'tit enfant</i> (1)
<i>petit garçon</i> (3)	<i>p'tit garçon</i> (5)
<i>petit cheval</i> (1)	<i>p'tit cheval</i> (2)
	<i>p'tit somme</i> (1)

**Film** : 11 occurrences *p'tit* (7) et *petit* (4)

<i>petit poucet</i> (1)	<i>p'tite abeille</i> (2)
<i>petite abeille</i> (1)	<i>p'tits grains de café</i> (1)
<i>petite flèche</i> (1)	<i>p'tit court métrage</i> (1)
<i>petit moment</i> (1)	<i>p'tit recoin</i> (1)
	<i>p'tit film</i> (2)

*Petit* dans le corpus natif est prononcé 12 fois avec deux syllabes et 15 fois avec réduction phonologique /pti/, une seule syllabe. Il est associé à des substantifs qui désignent un objet déjà connoté *petit* par la taille, l'âge ou la durée, *enfant*, *garçon*, *cheval* (jouet à roulettes) dans la bd ; et *abeille*, *grains de café*, *moment*, *court métrage* dans le film. On observe un seul emploi de *petit* avec la seule valeur minimisante, *petit film*, qui a le sens de film de courte durée. Une séquence figée *petit poucet*, avec référence culturelle, est produite par un seul sujet. Les natifs ajoutent *petit* à des noms qui évoquent déjà la petitesse, ce qui est redondant. On peut

reconnaître à l'adjectif *petit* une nouvelle fonction, celle de renforcement (RENF), une de plus peut-être à rajouter à celles déjà mentionnées par Peeters.

#### B. emplois de l'adjectif *petit* dans le corpus non-natif

La bande dessinée, premier support de la production orale, illustre l'histoire d'un enfant qui surprend un cowboy pendant la sieste – et qui décide de lui faire une blague. Il échange son jouet, un cheval en bois, avec le cheval du cowboy. Les apprenants ajoutent l'adjectif *petit* à presque tous les substantifs, excepté *cowboy*, ce qui pourrait donner : « cette *petite bande dessinée* raconte la *petite histoire d'un petit enfant* qui rencontre un cowboy à qui il va faire une *petite blague* en échangeant son *petit jouet* qui est un *petit cheval* en bois avec le cheval du cowboy pendant que celui-ci fait une *petite sieste* ». Tous ces syntagmes avec *petit* ont été produits dans les narrations.

##### a) fonction renforcement

L'adjectif **petit** est redondant avec *enfant, jouet et cheval en bois* (l'idée dimensionnelle est déjà présente dans ces noms qui renvoient à l'idée de petitesse en taille ou en âge). Dans le cas de *petite blague*, *petit* a pour fonction d'atténuer l'expression « *faire une blague* » afin de la rendre moins agressive, d'en atténuer l'effet et de la rendre plutôt bienveillante voire affectueuse (du point de vue du narrateur). Quant au dernier emploi, « il fait une *petite sieste* », là aussi on ne sait pas si *petit* renvoie à la durée, sachant que *sieste* évoque un somme dans la journée d'une durée plutôt courte. En tous cas, son emploi est redondant et renverrait à la fonction de Renforcement.

Ainsi avons-nous des emplois de *petit* qui renvoient à l'idée de taille ou de durée contenue déjà dans les noms mais, à notre avis si la notion de taille est présente, on peut y déceler un emploi affectif ou affectueux de *petit*, comme si le locuteur voulait rendre les faits plus légers, ou en atténuer les effets.

On trouve également d'autres emplois de « petit » hors syntagme nominal comme le syntagme verbal « *devenir petit* » et la forme grammaticale « *plus petit* » qui ne feront pas l'objet de remarques dans cette étude.

##### b) Fonction discursive de l'adjectif *petit*

*petit* est suivi d'une pause dans [*p'tite* (.) *blague*] et [*\*p'tite* (.) *\*jou*]. *petit* peut être suivi de deux pauses, une pause remplie (euh) et une pause vide, permettant ainsi de formuler le syntagme complet [*un p'tit euh* (.) *un p'tit cheval*]; l'adjectif *p'tit* et les deux pauses donnent



un temps pour la planification. Le fait que *p'tit* soit suivi de deux pauses indique le travail de recherche lexicale ; *un p'tit* est produit d'abord aisément comme la partie fixe de l'expression mais sans aucune possibilité d'activation de la partie variable. On pourrait donc lui attribuer la « fonction d'appui » pour la recherche lexicale : *petit* ou *p'tit* permettant au sujet de s'auto-corriger pour arriver au syntagme complet *p'tit cheval* produit en bloc. On reconnaît un processus semblable dans *un p'tit (.) banc*, sans la répétition de *p'tit* cette fois, mais suivi d'une pause vide qui facilite la recherche lexicale mais qui n'a pas permis au sujet de produire la formule complète.

	BD1	BD2
<b>petit garçon</b>	10	6
<b>p'tit garçon</b>	3	3
<b>petit enfant</b>	5	10
<b>p'tit enfant</b>	0	0
<b>petit cheval</b>	14	19
<b>p'tit cheval</b>	3	1
<b>petite sieste</b>		2
<b>p'tite sieste</b>	1	1
<b>petit jouet</b>	3	2
<b>p'tit jouet</b>	1	
<b>petit bout</b>		1
<b>p'tit chien</b>		1
<b>petite blague</b>	1	
<b>p'tit animal</b>	1	
<b>p'tit gamin</b>	1	
<b>petite bande dessinée</b>		1
<b>petit village</b>		1
<b>petit banc</b>	1	
<b>petit* maquette</b>		4

Tableau 47. Distribution des syntagmes avec « petit » et « p'tit » dans le cartoon

AK (niv1) produit quatre fois la séquence « *\*petit maquette* » (de cheval) pour désigner le cheval en bois. Pour ce locuteur, *maquette* fait sans doute référence au jouet en bois. Comme il ne connaît pas ce mot, il utilise une métonymie *maquette* qui renvoie à la fois aux notions de jouet, de matière et de dimension, une maquette étant « une représentation partielle ou complète d'un système ou d'un objet (existant ou en projet) afin d'en tester et valider certains aspects et/ou le comportement (maquette fonctionnelle) » (Wikipédia) ; ici, *petite maquette* désigne un « petit cheval en bois », un jouet.



### Distribution des syntagmes avec *petit* et *p'tit*

Comme chez les natifs, on trouve chez les non-natifs, l'adjectif *petit* prononcé avec réduction phonologique par certains sujets (15 occurrences de la forme réduite « *p'tit* » contre 11 de la forme « *petit* »). Les syntagmes les plus représentés avec la forme non réduite *petit* sont *petit cheval* 14 fois en Bd1 et 19 fois en Bd2 ; *petit garçon* 10 fois en Bd1 et 6 fois en Bd2, en compétition avec *petit enfant* 5 et 10 fois. La forme réduite « *p'tit* » est moins utilisée par les locuteurs non-natifs contrairement aux natifs. Nous avons trouvé dans les corpus écrit et oral du site canadien *lxtutor* plus d'occurrences du syntagme *petit garçon* (Annexe 10).

Contrairement aux natifs qui produisent plus de formes avec réduction phonologique, les apprenants produisent seulement le syntagme *p'tit garçon* avec réduction, trois fois en t1 et t2 mais jamais *p'tit enfant*. Cependant, on peut souligner que *petit garçon* ou *p'tit garçon* sont aussi plus utilisés par les natifs que *p'tit enfant*.

Les syntagmes avec *p'tit* comme *p'tit cheval* (3 fois en BD1 et 1 seule en BD2) sont produits par des sujets à un stade d'apprentissage plus avancé (groupes 2 et 3). En revanche, aucun syntagme avec réduction n'est produit par les sujets du groupe 1. Les non-natifs du groupe avancé se rapprocheraient ainsi du parler natif grâce aux variations que constituent les emplois avec réduction phonologique (Théophanous et Perez-Bettan, 2009).

A noter qu'un des sujets débutants, JA, place de manière incorrecte l'adjectif *petit* après le nom en t1 (ex.1), alors que, en t2 (ex.2), l'adjectif *petit* est placé correctement dans le syntagme. En t1, cet apprenant n'a sans doute pas encore automatisé la séquence.

(ex.1) *et les autres garçons \*petits* (.) et le *garçon \*petit* il part (.) ;

(ex.2) le *petit enfant* il courir pour (.)

Les locuteurs non-natifs produisent d'autres syntagmes avec la forme non réduite *petit* que les natifs. Par exemple, dans la BD on trouve *petit village*, *petit banc*, *petite blague*, *petit chien*, *p'tit gamin*, alors que les locuteurs natifs associent *petit* ou *p'tit* aux personnages (*petit garçon*, *petit enfant*, *petit cheval*), on trouve néanmoins une occurrence de « *p'tit somme* ». Dans le parler natif, on associe presque toujours *petit* à *somme*, quand on évoque une petite sieste et ce syntagme semble être en voie de lexicalisation.

Le tableau 48 présente les différents syntagmes construits avec *petit* que quelques apprenants produisent dans les narrations du film muet *Prisonnière*. [Rappelons qu'une abeille prisonnière dans une maison va retrouver la liberté grâce à une mouche lui indique la sortie avec des grains de café disposés en flèche dans la baignoire en direction de la bonde]. Nous avons relevé toutes

les occurrences de syntagmes avec *petit* dont plusieurs n'ont pas été comptés comme des SP (pause interne). Dix sujets sur dix-sept ont utilisé *petit* au moins une fois dans leurs narrations.

Tableau 48. Emplois de *petit* dans le film

Sujet	F1	F2	
<b>FA</b>		petits grains de café	<b>1</b>
<b>MJ</b>	petite histoire	<b>1</b>	
	des p'tits trucs	<b>1</b>	
	(petit euh mouche)		
<b>MR</b>	p'tit sionair	<b>2</b>	
<b>NY</b>		petits grains de café	<b>1</b>
<b>OL</b>	petit jaune	<b>7</b>	petit tableau <b>1</b>
	petite jupe	<b>1</b>	(petit (.) tableau)
	petit pois	<b>1</b>	
	(p'ti : :t (.) animau)		
	p'tit verre	<b>1</b>	
<b>PR</b>	petite film	<b>1</b>	(petite : :(.)abeille)
	petites mouches	<b>1</b>	petite abeille <b>1</b>
	petites abeilles	<b>1</b>	petits grains de café <b>1</b>
	petits grains	<b>1</b>	
	petite signe	<b>3</b>	
	petite sortie	<b>1</b>	
	petit abeille	<b>1</b>	<b>1</b>
<b>QI</b>	petit histoire	<b>1</b>	petit abeille <b>1</b>
	petit ab/ij/	<b>1</b>	tout petite étiquette <b>1</b>
	petite flèche	<b>1</b>	petit /z/abeille <b>1</b>
<b>SU</b>	petit insecte	<b>1</b>	petite histoire <b>5</b>
			petite abeille <b>1</b>
			petite mouche <b>1</b>
<b>VO</b>			petite abeille <b>1</b>
			p'tit moment <b>1</b>
<b>ZI</b>	petits* pièces de café	<b>1</b>	petit village <b>1</b>
	p'tits pièces de café	<b>1</b>	un petit peu <b>1</b>
	(après avoir dit 4 x p'tits + euh)		p'tits points de caco <b>1</b>
			p'tite & sor <b>1</b>
			p'tite sortie <b>1</b>
	<b>F1</b>	<b>29</b>	<b>F2</b> <b>23</b>

Dans le film, on relève de nombreux emplois de « *petit* » avec *grains de café* et *abeille*. Ces deux noms représentent des « objets » de dimension réduite et ne devraient pas nécessiter d'être précédés de *petit*. Mais les apprenants utilisent certainement *petit* pour marquer leur sympathie envers le personnage ; *petite histoire* renvoie-t-il à une histoire qui serait petite (longueur, durée) ou à l'idée d'une histoire sympathique et drôle ? L'emploi de *petit* devant un nom ne traduirait-il pas plutôt le côté affectif que le locuteur/narrateur semble vouloir exprimer ? Les apprenants semblent avoir compris le rôle de *petit* et se comportent ainsi comme les natifs. L'adjectif *petit* joue, semble-t-il, plusieurs rôles en production orale : il permet souvent d'amorcer la recherche lexicale, de contourner une difficulté lexicale et de parler plus longtemps sans faire trop de pauses. Il permet ainsi d'allonger le segment de parole, et c'est sans doute la raison de son emploi récurrent dans le corpus non-natif. L'emploi de *petit* serait alors une stratégie de l'oral pour gagner du temps pour la planification et pour une meilleure fluence.

Voici quelques exemples de l'utilisation de *petit* comme aide à la recherche lexicale. ZI, groupe intermédiaire, produit quatre fois *p'tits* suivi d'une pause remplie (euh) avant de produire l'expression idiosyncrasique « *\*pièces de café* » (l'activation du nom *pièces* étant sans doute le seul mot disponible en MLT qui évoque une ressemblance de taille avec le grain de café, un trait commun entre les deux noms (nom comptable, concret, etc.), ou encore la ressemblance avec la formation d'un groupe de noms relié avec de, autrement dit d'un mot composé.

OL, apprenant du groupe débutant, utilise le syntagme *petit jaune* à la place du nom *abeille* qu'elle ignore, sans aucun doute. L'emploi de « *petit jaune* » fait partie d'une stratégie lexicale. L'apprenant associe *petit* (taille) à *jaune* (couleur) de l'insecte pour le différencier du deuxième insecte (la mouche), dans un emploi métonymique. De même, lorsque le sujet ne connaît pas le syntagme *grains de café*, il trouve des syntagmes de remplacement, on l'a vu, *\*p'tits points de cacao* ou *\*p'tits pièces de café*. D'autres sujets semblent avoir adopté la même démarche lorsqu'ils utilisent les syntagmes *p'tits trucs*, *petite jupe*, *p'tit sionair* (MR est catalan et *sionair* est probablement un mot de sa L1). L'adjectif *petit* permet aussi d'allonger le segment et de garder la parole quand surgit une difficulté qui peut être lexicale mais pas seulement. Les apprenants peuvent tout simplement ajouter l'adjectif pour gagner du temps, ce qui permet aussi d'allonger le segment de parole.

Les natifs emploient les mêmes types de séquences avec *petit* : *p'tite abeille*, *p'tits grains de café*, mais ils font aussi d'autres associations, on trouve une référence culturelle *petit poucet*, un

groupe construit avec un nom plus spécialisé que film, *p'tit court métrage*. *Court-métrage* certes fait référence à un film court, cinq minutes pour *Prisonnière*, mais associé à *p'tit*, le syntagme renvoie aussi à l'idée d'un film sympathique non dépourvu d'humour et à une attitude favorable du narrateur pour les personnages et l'histoire. On relève, dans les productions L2, des séquences en voie de lexicalisation, *p'tits trucs*, *p'tit verre*, et *p'tit moment* avec réduction phonologique qui est un des traits natifs que les locuteurs L2 tentent de reproduire. C'est le cas de certains locuteurs non-natifs plus avancés. Ce n'est d'ailleurs pas un hasard si ces réductions phonologiques se produisent dans des expressions lexicalisées ou en cours de lexicalisation.

On remarque que tous les locuteurs n'ont pas le même comportement avec cet adjectif. Pourquoi trouve-t-on des emplois récurrents de séquences avec *petit* chez certains locuteurs non-natifs alors que chez d'autres on n'en trouve aucun ? Dans les narrations du film, HE, JA, VI et XI ne produisent aucune séquence avec *petit* contrairement à d'autres sujets. OL et PR produisent pas moins de sept syntagmes construits avec *petit* en F1. On trouve la même différence chez les natifs. VC, locuteur natif, n'emploie aucun syntagme avec *petit* lors de la narration du film. Certains locuteurs non-natifs s'appuieraient davantage sur cet adjectif à des fins stratégiques : soit pour allonger leur temps de parole et produire des segments plus longs, soit parce que ces séquences les aident à parler de manière plus aisée et plus fluide (rôle des SP), soit encore parce qu'elles servent d'appui et d'aide à la planification. C'est moins d'ailleurs une question de niveau qu'une stratégie personnelle. Et, peut-être aussi, le désir de calquer son discours sur celui des natifs. Les apprenants qui emploient *petit* de manière récurrente sont certainement ceux qui veulent paraître « le plus natif » et ceux qui le produisent avec réduction phonologique encore plus (Regan, 1997 ; Théophanous et Perez-Bettan, 2012).

# Conclusion générale

---

Notre étude longitudinale, dont l'objectif était d'examiner l'emploi et l'évolution des séquences préfabriquées par des apprenants du français L2, apporte un nouvel éclairage sur l'importance des SP dans le discours oral. La plupart des études et recherches sur le LP, entreprises à ce jour, ont été faites principalement en anglais langue seconde. En français, on compte deux études d'envergure, celles de Forsberg (2006) et de Cordier (2013). Toutes les études, qui se sont intéressées au LP, ont examiné des corpus de locuteurs de niveau avancé et/ou très avancé (Bartning *et al.* 2009 ; Forsberg-Lundell *et al.* 2014). Contrairement aux études citées, la nôtre s'intéresse au développement du LP chez des apprenants de niveaux différents et notamment chez des apprenants débutants du français.

Dans cette étude, les séquences préfabriquées ont été décrites comme des suites de mots mémorisées et stockées dans le lexique mental au même titre que les unités simples. Ces suites de mots ou expressions toutes faites peuvent prendre différentes formes et avoir diverses fonctions dans le discours. Nous avons montré que le langage préfabriqué était composé en L2 des séquences décrites en linguistique (chapitre 1) mais pas seulement. En L2, les SP peuvent être des séquences propres à un locuteur donné, à un moment donné de l'apprentissage ou de l'acquisition. Nous rappelons que l'identification des SP s'est faite selon des critères nécessaires comme la polylexicalité et la cohérence prosodique, mais aussi d'autres critères comme la complexité syntaxique, l'opacité sémantique, la fréquence ou la systématisme de l'item déviant (chapitre 2). Les séquences identifiées ainsi que les séquences idiosyncrasiques, qui font partie de l'interlangue de l'apprenant, aident le locuteur en production orale. Les SP jouent un rôle au niveau référentiel - elles servent à nommer, à caractériser - mais aussi à organiser le discours. Elles permettent alors de parler avec plus d'aisance, « à la manière des natifs », en temps réel et limité, dans les contraintes de la production orale. Comme elles sont stockées dans le lexique mental comme de longs mots facilement accessibles (chapitre 3), elles contribuent à rendre le discours plus fluide en évitant au locuteur d'avoir à assembler laborieusement des mots, ce qui représenterait une charge cognitive qui l'empêcherait de s'exprimer aisément. En L2, le locuteur doit arriver à effectuer un certain nombre de processus de bas niveaux, et les SP servent à libérer ses capacités attentionnelles pour les processus de haut niveau, comme la planification des énoncés et la gestion du sens.

Nos analyses ont montré que la production orale du groupe expérimental s'est améliorée après six mois dans le pays de la langue cible, c'est-à-dire que les apprenants emploient plus de SP (en quantité) avec une part de séquences lexicales qui augmente dans le temps.

Les apprenants dans l'ensemble parlent plus longtemps et plus vite la deuxième fois et produisent des segments de parole plus longs (chapitres 6 et 7) grâce, entre autres, à une utilisation plus importante des SP. Leur répertoire de séquences s'est également modifié : au début, ils utilisaient des SP discursives en grand nombre mais après six mois leur nombre a diminué au profit des SP lexicales. C'est-à-dire que leur discours est plus riche quantitativement parlant mais aussi plus diversifié. Quand on examine les résultats des sous-groupes, c'est le groupe des locuteurs débutants qui a le plus progressé dans l'emploi des séquences en quantité mais aussi en diversité, ce que les études antérieures ont également constaté : «...le développement lexical est plus notable dans les premiers stades de l'acquisition »<sup>73</sup> (Forsberg-Lundell *et al.* 2014 : 262). Pour les groupes des niveaux intermédiaire et avancé qui possèdent un répertoire plus riche et plus varié, les résultats sont plus nuancés. Ils n'utilisent pas autant de séquences préfabriquées (en proportion) que les débutants. Cette tendance s'affiche pour certains d'entre eux seulement. Quelques apprenants ont progressé dans l'emploi qu'ils font des SP alors que, pour d'autres, nous n'avons pas noté de changement pour diverses raisons comme le manque de motivation ou le désir d'être plus précis et/ou plus créatif lors de la deuxième narration (chapitre 7). L'analyse des productions non-natives a certainement confirmé que l'usage des SP est d'abord une stratégie efficace en production orale : en début d'apprentissage, les locuteurs peuvent, grâce aux SP, se débrouiller avec un minimum de connaissances lexicales ; ils peuvent ainsi développer des stratégies en s'appuyant sur quelques séquences pour gérer les difficultés lexicales (emploi de *c'est* et de *petit*, chapitre 9) et/ou morphosyntaxiques (phrases et expressions mémorisées *qu'est-ce qui s'est passé, qu'est-ce qu'il fait, ça marche pas*, etc.), ce qui leur permet aussi de maintenir une certaine fluidité. A un stade plus avancé, il se peut aussi que l'utilisation des SP les aide ainsi à planifier des énoncés plus créatifs. L'utilisation de ces séquences permet aussi aux locuteurs et ce, à tous les stades, de maintenir une certaine fluidité.

En effet, le lien entre SP et fluence est certainement l'un des constats les plus intéressants de cette étude. Dans notre corpus, l'aisance ou la fluence de la production en L2 est corrélée à un emploi plus ou moins important de préfabriqués, et cela surtout en temps 1. En t2, les tests statistiques ne confirment pas cette tendance malgré l'augmentation du nombre de SP. Nos

---

<sup>73</sup> « [...] lexical development is more notable in the early stages of acquisition. » (Forsberg-Lundell *et al.*, 2014 : 262)

données montrent, néanmoins, que les locuteurs, dans l'ensemble, sont plus à l'aise à l'oral, grâce en grande partie à l'usage qu'ils font des SP. Elles illustrent aussi la variabilité qui caractérise les productions individuelles - les individus n'ont pas tous le même comportement langagier. Nous avons d'ailleurs constaté les mêmes écarts dans le groupe contrôle (chapitre 6), tendance confirmée par d'autres recherches (Forsberg-Lundell *et al.* 2014). Cette recherche s'est focalisée sur le LP produit à six mois d'intervalle, afin de montrer comment il évolue dans le temps. Nous avons choisi la tâche narrative qui nous semblait plus facilement exploitable pour comparer les productions individuelles à deux moments. Néanmoins, si c'était à refaire, nous examinerions sans doute le LP des apprenants dans d'autres types de tâches (discussions, interviews ou autres types de discours), ce qui permettrait d'avoir une vision plus large et plus globale du LP mais aussi d'examiner quelles stratégies langagières les apprenants mettent en œuvre selon le type de tâche.

Les modèles psycholinguistiques de la production orale (chapitre 3) mettent l'accent sur l'importance des traitements automatiques dans la communication. L'effort attentionnel, que les apprenants doivent fournir aux différents stades de l'encodage formel (lexical, morphosyntaxique, phonologique), rend la communication du sens laborieuse (Hilton, 2011b). Selon les approches basées sur l'usage, c'est à travers des milliers d'interactions que les séquences sont mémorisées en L1 : c'est sans doute un apprentissage implicite – qui a lieu sans que l'individu soit conscient de la mise en mémoire de ces ensembles. Il est donc problématique de croire que les locuteurs non-natifs pourraient arriver à mémoriser les SP d'une langue cible sans une exposition prolongée à cette langue – un contact soutenu qui manque dans la plupart des contextes d'apprentissage institutionnels (2 ou 3 heures de cours par semaine, par exemple). Les chercheurs qui s'intéressent au LP concentrent leurs recherches autour de l'identification des SP, de la quantité de SP produites pour montrer qu'un groupe ou un niveau, voire un individu a progressé. Si ces études sont très intéressantes et nécessaires pour la recherche scientifique, elles ne disent rien de la façon dont un locuteur L2 peut les acquérir en milieu institutionnel ni surtout comment les enseigner. En effet, peu de place est accordée à l'enseignement des SP dans la recherche en didactique du FLE. Il nous paraît donc important de réfléchir à la façon dont on pourrait concevoir cet enseignement et aux moyens à mettre en œuvre en classe de langue pour la mémorisation et l'utilisation plus systématique des séquences préfabriquées par les apprenants. Quelques expériences ont montré une amélioration en production orale (fluidité, richesse lexicale) grâce à la mémorisation explicite de séquences préfabriquées dans des contextes d'apprentissage spécifiques (Boers *et al.* 2006 ; Fitzpatrick *et*



Wray, 2006), et il est intéressant de clore cette thèse avec quelques considérations didactiques basées sur cette recherche.

L'approche communicative (ou « actionnelle ») en vigueur en Europe n'a pas encore accordé une méthodologie théorisée pour les apprentissages lexicaux ni pour le langage préfabriqué (Hilton, 2011, communication personnelle). L'école du début (par exemple, « l'approche naturelle » de Krashen) postulait une « acquisition » implicite de ce type de connaissances. Or, nos classes de langue – avec quelques heures par semaine - ne sont pas suffisantes pour l'émergence « implicite » des connaissances préfabriquées en L2. Cependant, les chercheurs en acquisition des langues et en particulier en L2 ne sont pas tous d'accord ; pour certains l'apprentissage doit être implicite (Krashen, 1982) et pour d'autres explicite (Yuan et Ellis, 2003 ; Ellis, 1996, 2005).

En L1, un locuteur natif n'est pas conscient de cette acquisition (qui a lieu grâce à l'utilisation constante de la L1, réceptive *et* productive) ; cet apprentissage est implicite. En classe de langue, les apprenants L2 en France (que ce soit lors de l'éducation secondaire ou même dans des établissements prestigieux comme l'École des Mines) reçoivent environ trois heures de cours de langue par semaine ; l'écoute extensive n'est donc pas possible. Seuls les apprenants qui ont pu voyager ou les plus motivés (qui se mettent en situation d'écoute extensive et qui lisent en L2) – arriveront sans doute à acquérir un nombre élevé de SP, par leur contact plus soutenu avec la langue. Ils pourront de ce fait parler avec une plus grande aisance et précision. Dans leur étude, Derwing *et al.* (2004) constatent que, malgré les 25 heures de cours par semaine qu'ils ont reçues pendant un an, le discours oral des apprenants chinois ne s'est pas amélioré. Ils remarquent également que ce sont les sujets qui ont le plus de contact avec la langue cible qui réussissent le mieux (interactions avec des natifs, écoute quotidienne de la radio et de la télévision, films dans la langue cible, *etc.*). Un apprenant de niveau très avancé, qui aurait été exposé de cette façon pendant des années à la langue, pourrait sans doute être capable d'employer les bonnes séquences dans les bonnes situations à condition de les avoir mémorisées et de maîtriser d'autres compétences (sociolinguistique, pragmatique et discursive) (voir Forsberg-Lundell *et al.*, 2014 : 256 pour les différentes définitions des niveaux avancé (*advanced learner*), très avancé (*near-native speaker*) et niveau natif (*nativelike speaker*)).

Pour combler le manque de contact avec la langue, il faudrait une méthodologie plus ciblée pour l'apprentissage du LP. Mais comment arriver à enseigner le langage préfabriqué, vu la quantité d'informations à acquérir ? Rossiter, Derwing, Manimtim et Thomson (2010 : 583-584) constatent (dans le cadre de l'apprentissage de l'anglais langue seconde dans un contexte



particulier, au Canada) que les ressources en matière d'enseignement ne sont pas suffisantes pour faciliter le développement et le perfectionnement de la compétence orale des apprenants. En effet, les activités que proposent les manuels (des apprenants mais aussi ceux des enseignants), en vue d'améliorer la compétence orale, sont essentiellement des tâches de production libre avec peu d'insistance sur l'emploi des séquences préfabriquées et sur la répétition de celles-ci pour les fixer en mémoire. Ces auteurs proposent donc un certain nombre d'activités pour l'anglais langue seconde qui pourraient être adaptées au FLE en classe de langue, mais aussi des activités de contact « pour aider les apprenants à perfectionner leurs compétences à l'extérieur de leurs cours ». Pour améliorer la fluence, Rossiter *et al.* (2010) recommandent l'apprentissage explicite et en particulier des SP, qui favorisent la fluence pour toutes les raisons que nous avons déjà évoquées (aide à la planification, réduction du nombre de pauses et de leur durée). Les propositions concrètes d'activités (voir Rossiter *et al.* 2010) reposent essentiellement sur la répétition et l'utilisation des SP lexicales mais aussi des marqueurs discursifs qui permettent d'organiser le discours et de dégager du temps pour la planification. Comme les tâches planifiées (Foster et Skehan, 1996) montrent une amélioration dans l'emploi des SP, la méthode de Nation (1989) - la technique 4/3/2 - est préconisée, dans cette étude, pour aider les apprenants à parler d'une manière fluide en temps limité. On laisse quelques minutes à un apprenant pour préparer, sans prise de notes, une petite intervention orale sur un sujet (d'intérêt général ou en rapport avec le cours). Le travail se fait par paire ; l'apprenant parle pendant 4 minutes, pendant ce temps son camarade écoute et doit reproduire le même discours en trois minutes puis on fait la même chose une troisième fois avec un troisième locuteur qui doit reproduire cette fois le discours en deux minutes. Dans cette tâche de production en temps limité, Nation constate non seulement une amélioration de la vitesse de parole des locuteurs mais aussi une diminution du nombre de pauses et d'hésitations. De Jong et Perfetti (2011) ont adapté le test de Nation en l'améliorant. Les sujets, répartis de façon aléatoire en deux groupes, doivent lors d'une première tâche de répétition (1) parler du même sujet de plus en plus vite (4/3/2) – sans confondre vitesse et précision - puis lors d'une deuxième tâche de répétition (2) de trois sujets différents ; la troisième tâche de répétition (3) est réalisée deux semaines plus tard afin de comparer les progrès réalisés par les deux groupes entre (1) et (3) et entre (2) et (3). Ils constatent une meilleure fluidité chez les apprenants mais aussi que les SP les aident dans la réalisation de ces tâches. Toutes ces activités, testées chez des apprenants en anglais L2, seraient parfaitement réalisables dans une classe de langue, en France.

En France, quelques enseignants s'interrogent sur la façon dont on pourrait enseigner la « phraséologie » du français en L2. Pour certains d'entre eux, cet enseignement est un véritable défi et sa mise en place suscite de nombreuses interrogations. Quels types de séquences enseigner ? Doit-on mettre l'accent plutôt sur les expressions idiomatiques, par exemple ? Quelle méthodologie utiliser pour que cet enseignement/apprentissage soit efficace (et pas trop ennuyeux) ?

Si le Conseil de l'Europe fait quelques recommandations pour l'enseignement des langues étrangères et même du langage préfabriqué, il ne dit pas explicitement comment l'enseigner. Un des objectifs définis par le Cadre Européen Commun de Référence pour les Langues (CECRL) « est de développer un répertoire langagier dans lequel toutes les capacités linguistiques trouvent leur place » (CECRL, 2001 : 11) ; c'est-à-dire qu'il préconise le développement de compétences langagières en relation avec le lexique. Pour arriver à une compétence lexicale, il faut d'abord connaître le lexique et comprendre les structures lexicales pour arriver ensuite à les manipuler « Il s'agit de la connaissance et de la capacité à utiliser le vocabulaire d'une langue qui se compose : 1. d'éléments lexicaux et 2. d'éléments grammaticaux et de la capacité à les utiliser » (CECRL, 2001 : 87). Dans le *Cadre Européen Commun de Référence pour les Langues* (CECRL), le Conseil de l'Europe recommande l'enseignement du langage préfabriqué au même titre que le lexique. Les expressions toutes faites et les locutions figées contribuent à la compétence de communication (2001 : 87-88) en lien avec les actes de langage (compétence linguistique, sociolinguistique et pragmatique) mais aussi avec les archaïsmes et les proverbes. Pour le Conseil de l'Europe, « les locutions figées comprennent [...] des collocations figées constituées de mots fréquemment utilisés ensemble, par exemple *faire un discours, faire une faute* » (Conseil de l'Europe, 2001 : 88). L'accent est mis par les auteurs du CECRL sur les collocations, « phénomènes omniprésents dans la langue et qui relèvent d'une compétence linguistique difficile à acquérir » (Mel'çûk et Polguère, 2007 : 21) ; ce qui veut dire qu'il faut tenir compte de la combinatoire (lexique et syntaxe), des associations lexicales et des relations sémantiques (aspect métaphorique de certains éléments). On voit donc aisément qu'on ne peut pas enseigner séparément les éléments d'une collocation sans le risque de perdre le sens de l'association. Il est important de noter que ces recommandations concernent les niveaux B (« intermédiaire ») et C (« confirmé »). Il nous paraît cependant essentiel de faire connaître et apprendre, dès les niveaux A, les collocations, vu l'importance du phénomène dans le langage écrit et parlé et, notamment, la « langue » de la salle de classe (les consignes, la gestion de la classe impliquant un grand nombre d'expressions

préfabriquées). L'exemple suivant « *organiser une compétition, une opération publicitaire* », relevé dans un texte proposé pour l'étude dans une méthode au niveau 1 (Labascoule *et al.* 2004 : 78) montre bien la nécessité de les faire connaître (et de les faire apprendre) (Cavalla, 2009 : 6). Signalons tout de même que le langage préfabriqué n'est pas totalement absent des manuels de FLE. Se conformant en cela aux recommandations du Conseil de l'Europe, certains manuels proposent des rubriques où l'on trouve des expressions, autour d'un verbe par exemple. Dans le manuel *Alors*, niveau B1 (Di Guira et Beacco, 2009 : 119), on relève la consigne suivante « Entraînez-vous à utiliser les formes de *faire* dans des énoncés », suivie d'une liste d'expressions avec *faire* et leur signification. Cette liste pourrait être intéressante mais elle est incomplète car, pour que ces expressions aient un sens pour l'apprenant, il faudrait la compléter (illustrations dans le manuel, exemples tirés des enregistrements ou des lecteurs, *etc.*).

Au-delà de la description des différents types de séquences plus ou moins figées qui composent la langue, on doit constater que peu de place est consacrée dans le *Cadre* aux moyens pratiques d'enseigner le LP. Comment, en effet, enseigner le LP qui comprend, on l'a vu, de nombreuses facettes, des collocations (*rendre visite, beau gosse*) mais aussi des archaïsmes (*au fur et à mesure, par monts et par vaux*) des expressions figées et opaques (*donner sa langue au chat, couper le cheveu en quatre*) comme des expressions libres (*partir en voyage*) en passant par les constructions (*donner qch à qqn*).

L'enseignement des SP doit avoir un double objectif : améliorer la compétence lexicale qui permet de s'exprimer non seulement avec plus de précision - en choisissant les bonnes associations de mots (collocations) - mais aussi avec une plus grande aisance orale (voir Rossiter *et al.* 2010). Connaître les SP d'un point de vue linguistique, c'est bien mais les réutiliser en production orale, en contexte et en tenant compte de la situation de communication, cela doit être aussi un des défis de la didactique du FLE. Notre thèse n'étant pas axée sur la didactique de manière directe, nous faisons, sans vouloir être exhaustive, quelques retours et propositions sur les méthodes connues et possibles pour l'enseignement des SP.

La méthode la plus connue (qui était même révolutionnaire en son temps) en Anglais langue étrangère, est certainement l'approche lexicale (*The Lexical Approach*) de Michael Lewis (1993). Cette « approche lexicale » proposait de faire prendre conscience de manière explicite donc, de la façon dont fonctionne la langue, à partir de supports écrits (toujours en lien, d'ailleurs, avec un support audio-oral). L'idée de cette approche était basée sur l'importance que l'on devait accorder non seulement aux unités-mots mais aussi et surtout aux SP

(collocations, phrases institutionnalisées, patrons et cadres syntaxiques)<sup>74</sup> ; le but de l'approche lexicale étant d'enseigner l'anglais parlé, de la vie de tous les jours, et non le langage conventionnel qu'on trouve dans les manuels et souvent enseigné dans les cours de langue. La méthode prônait l'observation de la langue à partir de corpus écrits et oraux d'où l'on pouvait extraire des collocations, des patrons grammaticaux. Cet enseignement est plus basé sur l'observation des faits de langue (suites de mots, structures grammaticales) que sur l'explication et, on accorde plus d'importance à l'expression orale qu'à la précision grammaticale. Ajoutons à cela que si l'accent est mis sur la dimension prosodique (et non plus sur des syllabes et des structures) - prendre conscience des groupes rythmiques qui scandent la phrase – les apprenants seront plus sensibles aux ensembles de mots. L'approche lexicale a le mérite de ne pas focaliser sur le mot mais sur les suites de mots, les groupes. Attirer l'attention des apprenants sur le « mot » est une bonne base pour comprendre comment les mots, autour de lui, dans sa périphérie, lui donnent du sens, c'est-à-dire qu'il faut s'intéresser au mot dans son contexte. La deuxième phase consiste à proposer des exercices de répétition systématiques afin de mémoriser les expressions repérées dans le discours aussi bien oral qu'écrit pour tendre vers un niveau de maîtrise plus élevé. Fitzpatrick et Wray (2006) ont montré qu'un travail explicite sur des SP peut aider et renforcer la mémorisation de celles-ci. Les SP que leurs sujets avaient apprises et répétées étaient encore en mémoire après trois mois. Rappelons également l'expérience de Boers *et al.* (2006) qui montre que les SP facilitent l'expression orale. Les apprenants, qui avaient reçu un enseignement spécifique des SP, produisaient un discours plus fluide que le groupe contrôle. La lecture oralisée, un peu délaissée ces dernières années, peut également faire prendre conscience des ensembles préfabriqués puisque la prosodie est un élément non négligeable sinon essentiel dans la langue et, en particulier, dans les expressions qui jouent sur le rythme et les sonorités (assonances, allitérations), ce qui facilite aussi leur mémorisation (*bel et bien, contre vents et marée, blanc et noir, mesdames et messieurs, Pierre qui roule n'amasse pas mousse, chat échaudé craint l'eau froide, etc.*). Ellis (1996) avait insisté également sur l'importance de la phonologie dans le « chunking ».

Des chercheurs en didactique du FLE (Cavalla, 2009 ; 2012) proposent un enseignement des SP, centré autour des collocations et des expressions figées. Elles abordent la question selon deux approches, l'approche sémasiologique et l'approche onomasiologique (voir Blank, 2001). Dans l'approche sémasiologique –on part de la forme pour accéder au sens - on rassemble des

---

<sup>74</sup> « Instead of words, we consciously try to think of collocations, and to present these in expressions. Rather than trying to break things into ever smaller pieces, there is a conscious effort to see things in larger, more holistic, ways. » Lewis, 1997 : 204)

expressions ayant en commun des aspects lexicaux ou syntaxiques pour ensuite constater qu'elles n'ont pas toutes la même signification. L'approche onomasiologique – qui part du sens pour aller vers la forme - permet d'aborder non plus les expressions contenant le même mot (par exemple *faim*) mais la notion dans des expressions qui ne contiennent pas le mot (*avoir l'estomac dans les talons, avoir les crocs*) ; il faut aussi, bien évidemment, tenir compte des aspects sociolinguistiques et pragmatiques (registre de langue, situation de communication, etc.). On part donc d'un concept pour arriver aux mots et aux expressions en se posant la question suivante : « Quel lexique me permet de dire ce concept ? » Cette approche permet aussi d'éviter la traduction littérale (Cavalla, 2009 : 8) qui, de toute façon, serait impossible ou insatisfaisante (pour la plupart des expressions idiomatiques, imagées ou opaques) pour plusieurs raisons (syntaxique, lexicale, sémantique et culturelle). Les apprenants se rendent alors vite compte qu'ils doivent apprendre l'expression en bloc sans passer par la traduction. Apprendre une expression ne doit pas être une fin en soi, il faut apprendre aussi à l'utiliser au bon endroit et au bon moment. Ces approches ne peuvent être intéressantes que dans la mesure où la répétition permet la fixation des formes lexicales (morphèmes) et syntaxiques (structures figées). Mais elle ne doit pas être systématique car cela risque d'être ennuyeux si elles ne sont pas étudiées en contexte d'abord, pour être réinvesties rapidement ensuite. La comparaison avec les expressions de la L1 peut être aussi un moyen efficace de mémorisation. En effet, introduire de manière explicite les SP, l'idée ou le concept, s'intéresser à leur forme, les traduire littéralement pour mieux comprendre leur sens métaphorique, voire les détourner, peut conduire l'apprenant à les fixer en MLT. On trouve d'ailleurs de petits ouvrages très drôles sur le sujet (Blum, 1989), avec des exemples d'expressions idiomatiques prises au pied de la lettre (Annexe 16).

Un ouvrage très récent, *Le Français idiomatique*, propose une méthode, la phraséodidactique pour l'enseignement des expressions figées du français. Il faudra certainement un peu de recul pour se rendre compte des effets dans le contexte d'apprentissage du FLE. L'auteur de la méthode explique les motivations de cette méthode, qui ne traite que les expressions figées de la langue. Or, le phénomène de la préfabrication comprend d'autres types de séquences.

Les expressions figées sont généralement reléguées au chapitre des bizarreries de la langue et, de ce fait, bannies de l'enseignement des langues étrangères. Refoulées aux marges des leçons, elles sont minimisées dans les méthodes et rabaissées à une question de vocabulaire qui apporte un pittoresque « censé distraire » plutôt qu'instruire. Or, l'emploi naturel et spontané des expressions figées dans le discours est la preuve indéniable d'une bonne maîtrise de la langue étrangère chez les apprenants. Cependant, l'art d'employer les expressions figées avec aisance et à bon escient en langue étrangère requiert un entraînement progressif et continu qui exige une mise en place précoce dans le processus d'enseignement pour l'acquisition d'une compétence active chez l'apprenant. C'est dans ce

sens qu'il convient de mettre en action la phraséodidactique du français dès le début de l'apprentissage de la langue. (Gonzalez-Rey, 2012 : 12)

Enseigner les séquences préfabriquées est certainement un objectif ambitieux mais aussi passionnant qui peut se réaliser en partie par la mise en place d'exercices de répétition et de mémorisation. En effet, la répétition permet de renforcer les associations et serait même un des principes élémentaires de l'apprentissage selon Ellis (1996 : 110). Favoriser ainsi la mémorisation holistique des expressions non compositionnelles (sans passer par la traduction) et le traitement automatique de blocs langagiers devrait être le défi de la didactique du LP car il n'existe pas, à notre connaissance et à ce jour, de méthode pour enseigner les SP dans toute leur diversité, excepté quelques expériences pédagogiques isolées.

Pour se débrouiller dans une langue étrangère, il est nécessaire d'avoir accès au lexique qui permet de pouvoir communiquer (en situation de danger savoir dire « médecin » ou « pompiers » peut éviter le pire). Il faut donc apprendre d'abord des mots, leur sens - essentiel pour la communication - pour les associer ensuite à d'autres mots dans des structures pour arriver aux séquences préfabriquées. Les séquences préfabriquées étant un indice sérieux du niveau de compétence d'un locuteur, on reconnaîtra sans doute le débutant par une expression maladroitement traduite directement de sa L1, ou par la création de phrases qui ne résonnent pas tout-à-fait comme les natifs. Des professeurs de collège ont bien compris que l'usage des collocations et pas seulement en français langue étrangère est un indicateur « signifiant pour évaluer le degré de maîtrise de langue » chez l'élève et leur enseignement est nécessaire dans la mesure où ce phénomène est source de nombreuses erreurs chez les apprenants. « [...] travailler les collocations est une entrée pour aborder le langage et son étude d'un point de vue synchronique, dans la capacité d'un terme à organiser autour de lui des constellations d'autres termes que l'on proposera à l'élève de préciser » rapporte un professeur de collège (Piémont, Centre Régional de Documentation Pédagogique, article en ligne).

Dans l'approche « collocationnelle », on doit s'assurer que le mot compris est aussi maîtrisé. Par exemple, on peut rechercher les différentes façons de réaliser le mot *gratitude* dans la langue ; on obtient la liste suivante : *manifester, marquer, montrer, témoigner, faire preuve de... envers qqn* ; d'autres mots comme par exemple le mot *confiance* peuvent donner lieu à des associations lexicales pour exprimer une réalité *donner, perdre, avoir* mais aussi à des structures sans verbe cette fois, « *climat de confiance* », « *homme de confiance* ». On travaille à partir d'un mot « noyau » pour arriver aux structures syntaxiques et lexicales du groupe préfabriqué. Le mot « *passer* » peut donner l'occasion de faire surgir un certain nombre

d'expressions avec des degrés plus ou moins hauts de figement. Il peut être réalisé avec - des noms (*temps, frontière, main, Rubicon*) mais aussi des prépositions (constructions) qui modifient le sens du verbe (par, pour), *passer par la fenêtre, passer pour un idiot* – et cela dès les niveaux A. D'autres structures peuvent surgir comme les phrases figées « *il a passé l'arme à gauche* » (qu'on pourrait aussi aborder d'un point de vue onomasiologique autour du concept « mort »), *le temps passe (vite), il s'en passe de belles* ou des phrases plus banales mais utiles comme « *passe-moi le sel* », cette fois complètement transparente mais conventionnelle et préfabriquée. D'autres verbes très courants, comme le verbe *prendre*, par exemple, peuvent faire l'objet d'un travail intéressant à partir a) de sa base<sup>75</sup> (*petit-déjeuner*, par exemple) b) de ses emplois dans des phrases lexicalisées (*prendre la tangente, prendre son avenir en main, prendre le taureau par les cornes*) dans un continuum allant du plus au moins figé. De ces exercices de rencontre avec les mots peut naître une certaine aptitude (et envie) à les utiliser et un travail collectif et comparatif donnerait sans doute aux apprenants les moyens d'améliorer leur aisance à l'oral (traitement automatique des ensembles préfabriqués).

Un grand nombre de formules ritualisées plus centrées sur la fonction discursive du discours, que soit pour organiser le discours (*dans un premier temps, pour conclure, nous dirons que, etc.*) ou dans un but purement discursif, permettent aussi d'améliorer l'expression et d'arriver à une certaine compétence native : *A demain ? Tu sais quoi ? Ça va ? À quoi ça sert ? Ça marche, à tout à l'heure, qu'est-ce que c'est ?* Ces expressions rendent alors l'apprenant plus confiant et améliorent la fluence et la compréhension. Sans systématiser ce travail, les enseignants peuvent attirer l'attention des apprenants, de manière explicite, sur les SP. Mais il est aujourd'hui très difficile, voire impossible en FLE, en France, de pouvoir mener à bien ce type d'enseignement plutôt chronophage car tous ces exercices demandent du temps, ce qui rend la tâche difficile voire infaisable dans les conditions actuelles de l'enseignement en classe de langue.

Si la didactique de la préfabrication en français langue étrangère est un domaine encore peu exploré (même si Fosberg, Towell et Cordier s'y sont intéressés), on assiste en France depuis quelques années à une véritable révolution dans la transmission des savoirs par l'introduction des Technologies d'Information et de Communication dans l'Enseignement scolaire (TICE) selon les directives émanant du Ministère de l'Éducation Nationale. L'objectif des TICE est de mieux prendre en compte les besoins spécifiques des apprenants et de développer de nouvelles

---

<sup>75</sup>(Voir le chapitre 1, section 1.4.1)



modalités d'enseignement. Le recours au multimédia doit favoriser l'autonomie des élèves dans l'accès aux informations et les rendre responsables face aux savoirs à acquérir.

Cependant les TICE peuvent fournir aux enseignants de langue (primaire et secondaire en particulier) des moyens novateurs pour la diffusion des connaissances mais aussi et surtout pour l'exploration de stratégies d'apprentissage favorisant les compétences : information accessible, échange de courriels instantanés bien sûr, mais surtout échange et communication en temps réel ou différé avec d'autres apprenants, d'autres classes et d'autres pays, en interactivité. Le contact avec la langue cible, non plus seulement en classe mais aussi en prolongement à la maison, renforce les connaissances (linguistiques) et les compétences (discursive, pragmatique). Un certain nombre d'études tendent à montrer que les apprenants manifestent une plus grande motivation pour les activités qui font appel aux TICE que pour les cours traditionnels en classe (Grégoire, Bracewell et Laferrière, 1996 ; Lapierre et Gingras, 2001 ; Ouellet, Couture et Gauthier, 2001). En effet, les TICE permettent aux enseignants de diversifier leurs pratiques et d'intéresser leurs élèves à une nouvelle façon d'apprendre et de pratiquer la langue qui se prolonge au-delà des heures de classe. Le recours au multimédia rend l'enseignement plus intéressant, plus ludique, mais surtout permet aux apprenants de pouvoir travailler en autonomie – la compréhension orale, la phonologie, la combinatoire lexicale et syntaxique - grâce à la création d'exercices mis en ligne par l'enseignant et que l'apprenant pourra faire à la maison (liens internet, exercices interactifs, etc.). Appliquées aux langues étrangères, les TICE présentent des avantages, le son et l'image peuvent être exploités dans le cours par le biais d'un site Web ou d'une présentation PowerPoint. D'une manière plus centrée sur le LP, la recherche d'informations sur la « toile » (plutôt dans le secondaire) permet de valider le statut d'un préfabriqué (grâce, par exemple, à la consultation de corpus oraux en ligne tel que le concordancier *Lextutor* de Tom Cobb - si on trouve un minimum d'occurrences de cette forme. La recherche sur *Google* peut être envisagée de la même façon. L'apprenant, grâce à l'outil informatique, peut commencer à explorer la langue et à constater le phénomène de la préfabrication. L'apprentissage n'est plus seulement axé sur les savoirs mais sur une recherche personnelle et une réflexion sur l'apprentissage de ces savoirs ; car, en effet, « Apprendre ne relève pas simplement de l'acquisition des savoirs, mais d'une compréhension des procédures et des processus qui ont permis cette acquisition » (Guir, 2002 : 6). Les activités métacognitives, essentielles dans l'apprentissage des langues étrangères, doivent aider à réfléchir à la manière d'acquérir ces pratiques. Dès lors, comment sensibiliser les élèves au langage préfabriqué grâce à ces nouveaux outils ? L'exploitation de logiciels audio et/ou de



fichiers vidéo peuvent aider l'apprenant à visualiser des situations et entendre des dialogues qu'il pourra ensuite répéter voire recréer au moyen de l'outil informatique sous la forme de scénarios, de dessins par le biais d'images, de fichiers audio, de petites vidéos trouvés sur la toile ou créées à partir de leurs propres outils. C'est-à-dire que l'apprenant apprendra à apprendre mais aussi à construire son savoir. Cette méthodologie peut être appliquée à l'apprentissage du langage préfabriqué. Comprendre et apprendre à utiliser les expressions d'une langue étrangère peut alors parfaitement s'inscrire dans ces exercices d'écoute et de création.

Cette étude s'achève sur une réflexion didactique qui n'est pas, malgré les apparences, éloignée de notre recherche. La recherche en didactique des langues permet de formaliser un certain nombre de résultats d'expériences dans une intention précise, l'application dans les situations de classe. Les résultats d'une étude comme la nôtre devraient faire prendre conscience de l'importance du LP en L2 et susciter une réflexion sérieuse pour la mise en place de pratiques pédagogiques. Pour l'heure, une véritable méthodologie de l'enseignement/apprentissage du langage préfabriqué en FLE reste à mettre en place.

# Bibliographie

---

AISENSTADT, E. (1981) Restricted collocations in English Lexicology and Lexicography, *Review of Applied Linguistics*, 53, pp. 53-62.

ALLWOOD, J. NIVRE, J. et AHLSEN, E. (1989) Speech Management : On the Non-Written Life of Speech. *Gothenburg Papers in Theoretical Linguistics*, 58. Département de Sciences du Langage : Université de Göteborg.

ALTENBERG, B. (1998) On the phraseology of spoken English : The evidence of recurrent word combinations, in A.P. Cowie (ed.) *Phraseology : Theory, analysis, and applications*, Oxford : Clarendon, pp. 101–122.

ANDERSON, J. R. (1983) A spreading activation theory of memory, *Journal of Verbal Learning and Verbal Behavior*, 22, pp. 261-295.

ANDERSON, J. R. (1985) Language comprehension, in J. Anderson (ed.) *Cognitive Psychology and its Implications*. New York : W. H. Freeman, pp. 335-63.

ANDERSON, J. R. (1996) ACT : A simple Theory of Complex Cognition, *American Psychologist*, 51 (4), pp. 355-365.

ARRIVE, M. GADET, F. et GALMICHE, M. (1986) *La grammaire d'aujourd'hui*. Flammarion.

ATILF (Analyse et traitement informatique de la langue française°.  
<http://atilf.atilf.fr/dendien/scripts/tlfiv4/showps.exe?p=combi.htm;java=no;>

BADDELEY, A. D. (1986) *Working memory*. Oxford : Oxford University Press.

BADDELEY, A. D. (2000) The episodic buffer : a new component of working memory? *Trends in Cognitive Science*, 4, pp. 417-423.

BALLY, C. (1905) *Précis de Stylistique française : esquisse d'une méthode fondée sur l'étude du français moderne*. Paris : Klincksieck.

BALLY, C. (1909) *Traité de Stylistique française*. Genève : A. Eggiman.

BARDOVI-HARLIG, K. (2009) Conventional expressions as a pragmalinguistic resource : Recognition and production of conventional expressions in L2 pragmatics, *Language Learning*, 59, pp. 755-795.

BARDOVI-HARLIG, K. (2010) Exploring the pragmatics of interlanguage pragmatics : Definition by design, in A. Trosborg (ed.) *Pragmatics across languages and culture*. Berlin : Mouton de Gruyter, pp. 219-259.

BARDOVI-HARLIG, K. (2012) Formulas, Routines, and Conventional Expressions in Pragmatics Research, *Annual Review of Applied Linguistics* 32, pp. 206-227.

BARDOVI-HARLIG, K. et BASTOS, M-T. (2011) Proficiency, length of stay, and intensity of interaction and the acquisition of conventional expressions in L2 pragmatics, *Intercultural Pragmatics*, 8, pp. 347-384.

BARRETT, M. D. (1982) Distinguishing between prototypes : The early acquisition of the meaning of object names, in S.A. Kuczaj et M. D. Barrett (eds.) *The development of word meaning : Progress in cognitive development research*. New York : Springer-Verlag, pp. 39-67.

Bartens, A. et Sandström, N. (2006) Towards a description of Spanish and Italian diminutives within the NSM framework, in B. Peeters (ed.), *Semantic primes and universal grammar : empirical evidence from the Romance languages*. Amsterdam : John Benjamins, pp. 331-360.

BARTNING, I. et HAMMARBERG, B. (2007) The Functions of a high-frequency collocation in native and learner discourse : The case of French *c'est* and Swedish *det är*. *IRAL*, 45 (1), pp. 1-43.

BARTSCH, S. (2004) Structural and Functional Properties of Collocations in English. *A corpus study of lexical and pragmatic constraints on lexical co-occurrence*. Gunter Narr Verlag Tübingen.

BECKNER, C. BLYTHE, R. BYBEE, J. CHRISTIANSEN, M.H. CROFT, W. ELLIS, N.C. HOLLAND, J. KE, J. LARSEN-FREEMAN, D. et SCHOENEMANN, T. (2009) Language is a complex adaptive system : position paper, *Language Learning*, 59, pp. 1-26.

BENSON, (1985) Collocations and Idioms, in R. Ilson (ed.) *Dictionaries, Lexicography and Language learning*. Oxford : Pergamon, pp. 61-88.

BIBER, D. (2004) Lexical bundles in academic speech and writing, in B. Lewandowska-Tomaszczyk (ed.) *Practical applications in language corpora* (PALC 2003), Hamburg : Peter Lang, pp. 165-78.

BIBER, D. et CONRAD, S. (1999) Lexical bundles in conversation and academic prose, in H. Hasselgard and S. Oksefjell, *Out of Corpora : Studies in Honour of Stig Johansson*. Amsterdam : Rodopi, pp. 181-90.

BIBER, D. CONRAD, S. et CORTES, V. (2004) Lexical bundles in university teaching and textbooks, *Applied Linguistics*, 25, pp. 377-405.

BLANCHE-BENEVENISTE, C. (1990) *Le français parlé. Etudes grammaticales*. Paris : CNRS.

BLANK, A. (2001) Words and concepts in time : towards diachronic cognitive onomasiologie, in R. Eckardt, K. von Heusinger et C. Schwaere (eds.) *Words in time*. CSLI Publications, [www.metaphorik.de/01/blank.htm](http://www.metaphorik.de/01/blank.htm) (01/2001).

BLUM, G. (1989) *Les Idiomatics*. Paris : Editions du Seuil (collection Point-Virgule).

BOERS, F. EYCKMANS, J. KAPPEL, J. STENGERS, H. et DEMECHELEER, M. (2006) Formulaic sequences and perceived oral proficiency : putting a Lexical Approach to the test, *Language Teaching research*, 10 (3), pp. 245-261.

BOLINGER, D. (1975) *Aspects of Language*. New York : Harcourt Brace Jovanovich.

BOLLY, C. (2010) Flou phraséologique, quasi-grammaticalisation et pseudo marqueurs de discours : un no man's land entre syntaxe et discours ? *LINX*, 62-63, pp. 11-38.

BREAL, M. (1897) *Essai de Sémantique* (Science des Significations). Paris : Hachette.

BROWN, R. (1973) *A First Language : The early stages*. London : George Allen and Unwin.

- BROWN, R. et HANLON, C. (1970) Derivational complexity and order of acquisition in child speech, in J.R. Hayes (ed.) *Cognition and the Development of Language*. New York : Wiley, pp. 11-53.
- BRUMFIT, C. (2000) Accuracy and Fluency : The Basic Polarity, in H. Riggenbach (ed.) *Perspectives on Fluency*. Ann Arbor : University of Michigan Press, pp. 61-73.
- BURGER, H. (2010) *Phraseologie. Eine Einführung am Beispiel des Deutschen*, 4. Auflage. Berlin : Schmidt.
- BUTTERWORTH, B. (1980) Evidence from pauses in speech, in B. Butterworth (ed.) *Language Production*. London : London Academic Press, pp. 155-176.
- BYBEE, J. (2002) Phonological evidence for exemplar storage of multiword sequences, *Studies in Second Language Acquisition*, 24 (02), pp. 215-221.
- BYBEE, J. (2006) From usage to grammar : the mind's response to repetition, *Language et Communication*, 82, pp. 711-733.
- BYBEE, J. (2008) Usage-based grammar and second language acquisition, in Robinson, P. and Ellis, N.C. (eds.) *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York : Routledge, pp. 216-236.
- BYBEE, J. (2010) *Language, Usage and Cognition*. Cambridge : Cambridge University Press.
- BYBEE, J. et HOPPER, P. (2001) *Frequency and the Emergence of Linguistic Structure*. Amsterdam : John Benjamins.
- BYBEE, J. et MCCLELLAND, J.L. (2005) Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition, *The Linguistic Review*, 22 (2-4), pp. 381-410.
- BYBEE, J. et SCHEIBMAN, J. (1999) The effect of usage on degrees of constituency : the reduction of don't in English, *Linguistics*, 37, pp. 575-96.
- CALAQUE, E. (2006) *Approche syntagmatique du lexique. Collocations et image de l'organisation lexicale*. [www.edufle.net/Collocations-et-image](http://www.edufle.net/Collocations-et-image).
- CANDEA, M. (2002) Le e d'appui parisien : statut actuel et progression. *XXIVèmes Journées d'Etude sur la Parole*, Nancy, 24-27 juin 2002.
- CAVALLA, C. (2009) *Comment la combinatoire syntaxique peut-elle aider la compétence lexicale ?* Communication présentée au Colloque international « Enonciation et texte au cœur de la grammaire », Université de Toulouse –Jean-Jaurès les 11, 12, 13 mars 2009.
- CHAFE, W. (1968) Idiomaticity as an anomaly in the chomskyan paradigm, *Foundations of Language* 4, pp. 109-126.
- CHAFE, W. (1980) Some reasons for hesitating, in H. Dechert et M. Raupach (eds.) *Temporal variables in speech*. The Hague : Mouton Publishers, pp. 169-180.
- CHAMBERS, F. (1997) What do we mean by Fluency? *System*, 25 (4), pp. 535-544.
- CHOMSKY, N. (1965) *Aspects of the theory of syntax*. Cambridge, MA : The MIT Press.

- CLARCK, R. (1974) Performing without competence. *Journal of Child Language*, 1, pp. 1-10.
- CLARK, E. et KELLY, B. (2006) Constructions in Acquisition, in E. Clarck and B. Kelly (eds.) *Constructions in Acquisition*. CSLI Publications, pp. 1-15.
- CLARKE, S. (2008) Courant globaliste hier et aujourd'hui, in B. Lechevalier, F. Eustache, et F. Viader (eds.) *Traité de Neuropsychologie Clinique*. Bruxelles : De Boeck Supérieur, pp. 31-37.
- CNTRL (Centre National de Ressources Textuelles et Lexicales)  
<http://www.cnrtl.fr/definition/portail>
- COBB, T. (1999) Breadth and depth of vocabulary acquisition with hands-on concordancing. [Electronic version], *Computer Assisted Language Learning*, 12, pp. 345-360. Retrieved November 20, 2004, <http://www.er.uqam.ca/nobel/r21270/cv/Breadth.htm>
- COHEN, N. J. et SQUIRE, L. R. (1980) Preserved learning and retention of pattern analyzing skill in amnesia : Dissociation of knowing how and knowing that, *Science*, 210, pp. 207-210.
- CONFAIS, J. P. (2005) *Les locutions*, Cours de Master. Université Toulouse-Jean-Jaurès.
- CONFAIS, J. P. (2015) Langage préfabriqué et néologismes, in V. Beliakov et S. Mejri (eds.) *Stéréotypie et figement. A l'origine du sens*. Toulouse : Presses Universitaires du Midi.
- CONKLIN, K. et SCHMITT, N. (2008) Formulaic Sequences : Are They Processed More Quickly than Nonformulaic Language by Native and Nonnative Speakers? *Applied Linguistics*, 29 (1), pp.72-89.
- CONSEIL DE L'EUROPE (2001) *Cadre européen de référence pour les langues*. Paris : Didier.
- COOK, V. (2002) (ed.) *Portraits of the L2 User*. Clevedon : Multilingual Matters, 219-249.
- CORDER, P. (1973) *Introducing Applied Linguistics*. Baltimore : Penguin Books.
- CORDIER, C. (2013) *The presence, nature and role of formulaic sequences in English advanced learners of French : a longitudinal study*. Thèse doctorat. Newcastle University : UK.
- COULMAS, F. (1979) On the sociolinguistic relevance of routine formula, *Journal of Pragmatics*, 3, pp. 239-266.
- COULMAS, F. (1981) *Conversational Routine : Explorations in Standardised Communication Situations and Pre-Patterned speech*. The Hague : Mouton.
- COWIE, A. P. (1981) The Treatment of Collocations and Idioms in Learners' Dictionaries, *Applied Linguistics*, 2 (3), pp. 223-235.
- COWIE, A. P. (1988) Stable and creative aspects of vocabulary use, in R. Carter et McCarthy (eds.) *Vocabulary and Language Teaching*. UK : Longman.
- CROFT, W. (2000) *Explaining Language Change : An evolutionary Approach*. Longman Linguistics library, Pearson Education.
- CROFT, W. et CRUSE, A. (2004) *Cognitive Linguistics*. Cambridge : Cambridge University Press.
- CRUSE, D. A. (2004) *Meaning in Language : An Introduction to Semantics and Pragmatics*. Oxford University Press.

- DAHLMAN, I. (2009) *Towards a multi-word unit inventory of spoken discourse*. Thèse doctorat, University of Nottingham : UK.
- DE BOT, K. (1992) A bilingual production model : Levelt's 'speaking' model adapted, *Applied Linguistics*, 13, pp. 1-24.
- DECHERT, H.W. (1980) Pauses and intonation as indicators of verbal planning in second-language speech productions : two examples form a case study, in H.W. Dechert, M. Raupach (eds.) *Temporal Variables in Speech*. The Hague : Mouton Publishers, pp. 271-286.
- DECHERT, H.W. (1983) How a story is done in a second language? in C. Faerch et G. Kasper (eds.) *Strategies in interlanguage communication*. London : Longman.
- DECHERT, R. et LENNON, S. (1989) Collocational blends of advanced second language learners, in S. Weigle (ed.) *Learning Vocabulary for Different Purposes*. Oxford : Pergamon Press, pp. 98-116.
- DECHERT, H. W. MÖHLE, D. et RAUPACH, M. (eds.) (1984) *Second Language Productions*. Tübingen : Narr.
- DEESE, J. (1980) Pauses, prosody, and the demands of production in language, in H.W. Dechert, M. Raupach (eds.) *Temporal Variables in Speech*. The Hague/Paris/New York : Mouton Publishers, pp. 69-84.
- DE JONG, N. et PERFETTI, C. A. (2011) Fluency training in the ESL Classroom : An Experimental Study of Fluency Development and proceduralization, *A Journal of Research in Language Studies*, 61 (2), pp. 533-568.
- DEKEYSER, R. (1991) Foreign language development during a semester abroad, in B. Freed (ed.), *Foreign language acquisition : Research and the classroom*. Lexington, MA : D. C. Heath, pp. 104-19.
- DEKEYSER, R. M. (2001) Automaticity and automatization, in P. Robinson (ed.), *Cognition and second language instruction*. Cambridge, MA : Cambridge University Press, pp. 125-51.
- DELL, G. S. (1986) A Spreading-Activation Theory of Retrieval in Sentence Production, *Psychological Review*, 93 (3), pp. 283-321.
- DESCHAMPS, A. (1980) The syntactical distribution of pauses in English spoken as a second language by French students, in H.W. Dechert, M. Raupach (eds.) *Temporal Variables in Speech*. The Hague, Paris, New York : Mouton Publishers, pp. 255-262.
- DEULOFEU, J. (1986) Sur quelques processus de hiérarchisation de l'information dans des récits d'apprenants marocains en milieu naturel. Pour une conception souple des rapports entre phénomènes de micro- et de macro-thématisation, in A. Giacomi et D. Véronique (eds.) *Acquisition d'une langue étrangère : perspectives de recherche*. Aix-en-Provence : Publications de l'Université de Provence, pp. 263-283.
- DERWING, T. M. ROSSITER, M. J. MUNRO M. J. et THOMSON, R.I. (2004) Second Language Fluency : Judgments on Different Tasks, *Language learning*, 54 (4), pp. 655-679.
- DEWAELE, J-M. (2002) Individual Differences in L2 Fluency : The Effect of Neurobiological Correlates, in V. Cook (ed.) *Portraits of the L2 User*. Clevedon, UK : Multilingual Matters, pp. 221-249.



- DIESSEL, H. et TOMASELLO, M. (2000) The development of relative clauses in spontaneous child speech, *Cognitive Linguistics* 11, pp.131-151.
- DI GUIRA, M. et BEACCO, J.C. (2009) *Alors ?* Méthode de français fondée sur l'approche des compétences (niveau B1). Didier : Paris.
- DÖRNYEI, Z. (2001) *Teaching and Researching Motivation*. Harlow :Longman.
- DÖRNYEI, Z. et SKEHAN, P. (2003) Individual Differences in Second Language Learning, in C. Doughty et M. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford : Blackwell, p. 589-630.
- DUBREIL, E. (2008) Collocations : Définitions et Problématique, *Texte*, 13 (1).  
[http://www.revuetexto.net/docannexe/file/126/dubreil\\_collocations.pdf/](http://www.revuetexto.net/docannexe/file/126/dubreil_collocations.pdf/)
- DULAY, H. BURT, M. et KRASHEN, S. (1982) Special Constructions, *Language Two*, pp. 232-243.
- EDMONDS, A. (2010a) Je suis vraiment désolé” ou comment s'excuser en interlangue, in B. de Buhann-Brun (ed.) *Altérité, interculturalité : perceptions et représentations de l'étranger en Europe et dans l'Arc Atlantique*. Paris, L'Harmattan, pp. 69-82.
- EDMONDS, A. (2010b) “Qu'est-ce qui s'est passé ?” Des questions stéréotypées chez les apprenants du français, *Revue interdisciplinaire « Textes et contextes »*, 5, *Stéréotypes en langue et en discours*.
- EDMONDS, A. (2011) Chaque chose à sa place : les expressions en classe de langue. Intégration de l'altérité dans l'apprentissage des langues, formes didactiques et procédures psycholinguistiques. *Rives, Cahiers de l'Arc Atlantique*, 3, pp 62-75.
- EDMONDS, A. (2013) Une Approche psycholinguistique des phénomènes phraséologiques : le cas des expressions conventionnelles, *Langages*, 189, pp. 121-138.
- EJZENBERG, R. (2000) The Juggling Act of oral Fluency; A psycho-Sociolinguistic Metaphor, in H. Rigganbach (ed.) *Perspectives on Fluency*. The University of Michigan Press : Ann Arbor, pp. 287-313.
- ELLIS, N.C. (1996) Sequencing in SLA. Phonological Memory, Chunking, and Points of Order, *SSLA*, 18, pp. 91-126.
- ELLIS, N. C. (1997) The Epigenesis of Language : Acquisition as a sequence learning problem, in A. Wray & A. Ryan (eds.) *Evolving models of language*, pp. 41-57.
- ELLIS, N. C. (2002) Frequency effects in language processing : A Review with Implications for Theories of Implicit and Explicit Language Acquisition, *SSLA*, 24 (2), pp. 143-188.
- ELLIS, N.C. (2003) Constructions, Chunking, and Connectionism : The Emergence of Second Language Structure, in C. Doughty et M. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford : Blackwell, pp.63-103.
- ELLIS, N.C. SIMPSON-VLACH, R. et MAYNARD, C. (2008) Formulaic Language in Native and Second Language Speakers : Psycholinguistics, Corpus Linguistics, and TESOL. *TESOL Quarterly*, 42, (3), pp. 375-396.
- ELLIS, R. (1994) *The study of second language acquisition*. Oxford : Oxford University Press.
- ELLIS, R. (2005) Principles of instructed language learning. *System*, 33, pp. 209-224.

ELLIS, R. et FANGYUAN, Y. (2004) The Effects of Planning on Fluency, Complexity, and Accuracy in second Language Narrative Writing, *SSLA*, 26, pp. 59-84.

ERMAN, B. et WARREN, B. (2000) The Idiom principle and the open choice principle, *TEXT*, 20 (1), pp. 29-62

FERNANDEZ, J. (1994) *Les particules énonciatives*. Paris : PUF.

FILLMORE, C. J. (1979) On Fluency, in C. J. Fillmore, D. Kempler, et W. Wang (eds.), *Individual Differences in Language Ability and Language Behavior*. New York : Academic Press, pp. 85-101.

FILLMORE, C. KAY, P. et O'CONNOR, M-C. (1988) Regularity and Idiomaticity in grammatical constructions : the case of let alone, *Language* 64, (3), pp. 501-538.

FIRTH, J.R. (1957) *Modes and Meaning, Papers in linguistics (1934-1951)*. Oxford : Oxford University Press.

FITZPATRICK, T. et WRAY, A. (2006) Breaking up is not so hard to do. Individual differences in L2 memorisation, *The Canadian Modern Language Review*, 63 (1), pp. 35-57.

FONAGY, I. (1997) Figement et changement sémantiques, in M. Martins-Baltar (ed.) *La locution : entre langue et usages*, Ophrys : Paris.

FORSBERG, F. (2006) *Le Langage préfabriqué en français parlé L2*. Thèse doctorat. Stockholm University.

FORSBERG, F. et FANT, L. (2010) Idiomatically Speaking : Effects of Task Variation on Formulaic language in Highly Proficient Users of L2 French and Spanish, in D. Wood (ed.) *Perspectives on Formulaic Language : Acquisition and communication*, pp. 47-70.

FORSBERG-LUNDELL, F. BARTNING, I. ENGEL, H. GUDMUNSON, A. HANCOCK, V. et LINDQVIST, C. (2014) Beyond advanced stages in high-level spoken L2 French, *French Language Studies* 24, pp. 255-280.

FOSTER P. (2001) Rules and routines : A consideration of their role in the task-based language production of native and non-native speakers, in M. Bygate, P. Skehan, et M. Swain (eds.) *Researching pedagogic tasks : Second language learning, teaching, and testing*. Harlow : Longman, pp. 75-93.

FOSTER, P. et SKEHAN, P. (1996) The influence of planning on performance in task-based learning, *Studies in Second Language Acquisition*, 18 (3), pp. 299-324.

FRANÇOIS, J. et MANGUIN, J-L. (2006) Dispute théologique, Discussion oiseuse et conversation téléphonique : les Collocations adjectivo-nominales au cœur du débat, *Langue française*, 2, 150, pp. 50-65.

FREED, B. (ed.) (1995) *Second Language Acquisition in a study abroad context*. Amsterdam : John Benjamins.

FREED, B. (2000) Is Fluency, Like Beauty, in the eyes (and ears) of the Beholder ?, in H. Riggenbach (ed.) *Perspectives on Fluency*. Ann Arbor : University of Michigan Press, pp. 243-265.



FREED, B. SEGALOWITZ, N. et DEWEY, D. (2004) Context of learning and second language fluency in French : comparing regular classroom, study abroad, and intensive domestic immersion programs, *Studies in Second Language Acquisition*, 26 (2), pp. 275-301.

GADET, F. (1989) *Le français ordinaire*. Paris : A. Colin.

GALATANU, O. (1984) *Actes de langage et didactique des langues étrangères*, Bucaresti, TUB Presses Universitaires de Bucarest.

GALISSON, R. (1984) *Les expressions imagées*. Paris : Clé International.

GAONAC'H, D. (2006) *Rôle de l'apprentissage dans l'acquisition précoce*. Paris : Hachette.

GARNIER-MATHEZ, I. (2005) L'Épithète et la Connivence : Écriture concertée chez les Évangéliques français, *Travaux d'Humanisme et Renaissance*, 404, CDIV (1941-2011), pp. 1523-1534.

GASS, S. (1988) Integrating research areas : a framework for second language studies, *Applied Linguistics*, 9, pp. 198-217

GASS, S. et SELINKER, L. (2008) *Second Language Acquisition. An Introductory Course*. London : Routledge,

GERMAIN, C. et NETTEN, J. (2002) La précision et l'aisance en FLE/FL2 : définitions, types et implications pédagogiques. Communication prononcée au Colloque intitulé : *La Didactique des langues face aux cultures linguistiques et éducatives*. Paris, Actes du Colloque.

GIBBS, R.W. BOGDANOVICH, J.M. SYKES, J.R. et BAR, D.J. (1997) Metaphor in Idiom Comprehension, *Journal of Memory and Language*, 37, pp. 141-154.

GIRARDIN, C. HUDELOT, C. PRENERON C. et POWDER. M.-C. (1988) *Acquisition du langage chez l'enfant et conditions de communication*. Recueil d'études sous la direction d'E. Sabeau-Jouannet. Peeters/Selaf. Paris.

GOLDBERG, A. (2005) *Construction grammar : a construction approach to argument structure*. Chicago : University of Chicago Press.

GOLDMAN-EISLER, F. (1958) Speech production and the predictability of words in context, *Quart. J. Exp. Psychol.* (10), pp. 96-106.

GOLDMAN-EISLER, F. (1968) *Psycholinguistics : Experiments in Spontaneous Speech*. London : Academic Press.

GOLDSTEIN, K. (1948) *Language and language disturbances*. New York : Grune and Stratton.

GONZALEZ-REY, I. (2002) *La phraséologie du français*. Toulouse : Presses Universitaires du Mirail.

GONZALEZ-REY, I. (2012) La culture de l'autre : l'enseignement des langues à l'Université - Actes. 2010. *La phraséodidactique en action : les expressions figées comme objet d'enseignement*, *La Clé des Langues*. Consulté le 10 septembre 2015, Mis à jour le 1 avril 2010, [file ://localhost/Url / http/ : :cle.ens-lyon.fr :espagnol :la-phraséodidactique-en-action-les-expressions-figees-comme-objet-d-enseignement-92012.kjsp](http://cle.ens-lyon.fr/espagnol/la-phraséodidactique-en-action-les-expressions-figees-comme-objet-d-enseignement-92012.kjsp).

GREGORY, G. (1981) *Linguistic variation in Brazilian Portuguese : Aspects of the phonology, syntax, and language history*. Thèse doctorat, University of Pennsylvania. Ann Arbor : University Microfilm.

GRANGER, S. (1998) Prefabricated patterns in advanced EFL Writing : collocations and formulae, in A.P. Cowie, (ed.) *Phraseology : Theory, Analysis and Applications*. Oxford : Oxford University Press, pp.145-159.

GRANGER, S. et MEUNIER, F. (2008) Introduction : the many faces of phraseology, in S. Granger et F. Meunier (eds.) *Phraseology : An interdisciplinary perspective*, xix-xxviii.

GRANGER, S. et PAQUOT, M. (2008) Disentangling the phraseological web, in S. Granger et F. Meunier (eds.) *Phraseology : An Interdisciplinary Perspective*. Benjamins : Amsterdam and Philadelphia, pp. 27-49.

GRANT, L. et BAUER, L. (2004) Criteria for re-defining idioms : Are we barking up the wrong tree? *Applied Linguistics*, 25 (1), pp. 38-61.

GRÉGOIRE, R. BRACEWELL, R. et LAFERRIÈRE, T. (1996) L'apport des nouvelles technologies de l'information et de la communication (NTIC) à l'apprentissage des élèves du primaire et du secondaire. *Revue documentaire, Rescol*. Université Laval, Université McGill. [<http://www.tact.fse.ulaval.ca/fr/html/apport/apport96.html>]

GROSJEAN, F. (1980) Temporal Variables within and between languages, in H. Dechert and M. Raupach (eds.) *Towards a Cross-Linguistic Assessment of Speech Production*. Frankfurt : P. Lang, pp. 39-53.

GROSJEAN, F. (1992) Another view of bilingualism, in R. Harris (ed.) *Cognitive Processing in Bilinguals*. Amsterdam : North-Holland.

GROSJEAN, F. (1998) Studying bilinguals : Methodological and conceptual issues, *Bilingualism : Language and Cognition*, 1, pp. 131-149.

GROSJEAN F. (2001) The bilingual's language modes, in J. Nicol (ed.) *One Mind, Two Languages : Bilingual Language Processing*. Oxford : Blackwell, pp. 1-22.

GROSJEAN, F. et DESCHAMPS, A. (1972) Analyse des variables temporelles du français spontané, *Phonética*, 26, pp. 129-156.

GROSJEAN, F. et DESCHAMPS, A. (1975) Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation, *Phonetica* 31, pp. 144-184.

GROSS, G. (1996) *Les expressions figées en français*. Paris : Ophrys.

GROSMANN, F. et TUTIN, A. (2003) *Les collocations : analyse et traitement*. Amsterdam : De Werelt.

GUIR, R. (2002) Pratiquer les TICE, in De Boeck Supérieur « *Pédagogies en développement* », Bruxelles. Lien : [http://www.cairn.info/pratiquer\\_les\\_tice--9782206082110.htm](http://www.cairn.info/pratiquer_les_tice--9782206082110.htm).

GÜLICH, E. (2008) Le recours au préformé : une ressource dans l'interaction conversationnelle. Congrès mondial de Linguistique française, in J. Durand B. Habert et B. Laks (eds.) *Discours, pragmatique, interaction*, pp. 869-879.

HAGEGE, C. (1996) *L'enfant aux deux langues*. Paris : Odile Jacob.

- HAKUTA, K. (1974) Prefabricated patterns and the emergency of structure in second language acquisition, *Language learning*, 24, pp. 284-297.
- HALLIDAY, M.A.K. (1961) Categories of the theory of grammar, *Word*, 17 (3), pp. 241-292.
- HASSELROT, B. (1957) *Études sur la formation diminutive dans les langues romanes*. Uppsala : A. B. Lundequistska Bokhandeln.
- HATCH, E. (1972) Some studies in language learning. *UCLA Working papers in Teaching English as a second Language*, 6, pp. 29-36.
- HAUSMANN, F. J. (1989) Le dictionnaire de collocations, in F. J. Hausmann, O. Reichmann, H. Wiegand, et L. Zgusta (eds.) *Dictionnaires. Encyclopédie internationale de lexicographie*, Berlin/New York : De Gruyter, 1010–1019.
- HAUSMANN, F. J. (1997) Tout est idiomatique dans les langues, in M. Martins-Baltar (ed.). *Langages, La Locution entre langues et usages*. Fontenay Saint-Cloud : ENS éditions, pp. 277-290.
- HAUSMANN, F. J. et BLUMENTHAL, P. (2006) Collocations, Corpus, Dictionnaires, *Revue Langue Française*, 150.
- HENDERSON, A. I. GOLDMAN-EISLER, F. et SKARBEEK, A. (1965) Temporal patterns of cognitive activity and breath control in speech, *Language and Speech*, 17, pp. 119-125.
- HÉRISSON, C.D. (1954) L'hypocoristique « petit » dans les titres de journaux, *Le Français moderne*, 1, pp. 49-58.
- HICKEY, T. (1993) Identifying Formulas in First Language Acquisition, *Journal of Child Language*, 20, pp. 27-41.
- HILTON, H. E. (2007) La compétence lexicale en production orale : diversité, accès et niveau. *Actes du colloque DILTEC, Recherches en acquisition et en didactique des langues étrangères et secondes*. Université Paris 3 (6-8 septembre 2006), communication personnelle le 8 septembre.
- HILTON, H. E. (2008a) Connaissances, Procédures et Production orale en L2, *Aile* 27, pp. 63-89.
- HILTON, H. E. (2008b) *Le corpus PAROLE : architecture du corpus et conventions de transcription*. Disponible dans *BilingBank* de *TalkBank*, <http://talkbank.org/>. Pittsburgh PA : Carnegie Mellon University.
- HILTON, H. E. (2009) *Systèmes émergents : acquisition, traitement et didactique des langues*, Habilitation à diriger les recherches.
- HILTON, H. E. (2011a) What is implicit and what is explicit in L2 speech ? Findings from an oral corpus, in C. SANZ et R. P. LEOW (eds.). *Implicit and Explicit Conditions, Processes, and Knowledge in SLA and Bilingualism*. Washington DC : Georgetown University Press, pp. 145-157.
- HILTON, H. E. (2011b) Psycholinguistique de la production orale, aisance et disfluenza en L2, in Danielle ; Goutéraux, Pascale, *Intégration de l'altérité dans l'apprentissage des langues : formes didactique et procédures psycholinguistiques*. L'Harmattan, pp.75-96.
- HOFFMAN, T. (2011) *Preposition Placement in English : A Usage-based Approach*. Cambridge : Cambridge University press.

- HUNSTON, S. et FRANCIS, G. (2000) *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam : John Benjamins.
- JACKENDOFF, R. (1995) The boundaries of the lexicon, in M. Everaert, E. Vander Linden, A. Schenk, et R. Schreuder (eds.) *Idioms : Structural and psychological perspectives*. Hillsdale, NJ : Erlbaum, pp. 133-166.
- JACKSON, J. H. (1958) On affections of speech from disease of the brain. *Brain* (1878/79), 1, pp. 304-330. Reprinted in J. Taylor (ed.) *Selected writings of John Hughlings Jackson* (2). New York : Basic Books.
- JAMES, W. (1890) *The principles of psychology*. New York : Holt.
- JESPERSEN, O. (1924/1976) Living grammar, in *The Philosophy of Grammar*. London.
- JIANG, N. et NEKRASOVA, T. M. (2007) The Processing of Formulaic Sequences by Second Language Speakers, *Modern Language Journal*, 91 (3), pp. 433-445.
- JOHNSTON, T. et FERRARA, L. (2012) Lexicalisation in Signed languages : When is an Idiom not an Idiom ?, *Selected Papers from UK-CLA Meetings*, 1, pp. 229-248.
- JOHNSTON, T. et SCHEMBRI, A. (2010) Variation, lexicalization and grammaticalization in signed languages. *Langage et société*, 131, pp. 19-35.
- KELLER, R. (1979) Gambits : Conversational strategy signals, *Journal of Pragmatics*, 3, pp. 219-237.
- KEMPEN, G. et HOENKAMP, E. (1987) An incremental procedural grammar for sentence formulation, *Cognitive Science*, 11, pp. 201-258.
- KERBRAT-ORECCHIONI, C. (2001) « Je voudrais un p'tit bifteck » : la politesse à la française en site commercial, *Les carnets du CEDISCOR*, 7, pp.105-118.
- KERBRAT-ORECCHIONI, C. (2004a) L'adjectif petit comme procédé d'atténuation en français, *Travaux et documents* [Université Paris-8 Saint-Denis], 24, pp. 153-175. Cité d'après la reprise dans Kerbrat-Orecchioni (2005, p. 262-274).
- KERBRAT-ORECCHIONI, C. (2004b) Négocier dans les petits commerces. *Négociations*, 2, p.7-22.
- KERBRAT-ORECCHIONI, C. (2005) *Le discours en interaction*. Paris : Armand Colin.
- KLINGLER, D. (2011) Le marquage du topique en français (L2) et en japonais (L1) dans des productions bilingues de trois apprenants avancés, in : *Volume d'articles en vue de l'obtention du diplôme d'habilitation à diriger des recherches*, Université de Provence.
- KOPONEN, M. et RIGGENBACH, H. (2000) Overview : Varying Perspectives on Fluency, in H. Riggenbach, (ed.) *Perspectives on Fluency*. Ann Arbor : University of Michigan Press, pp. 5-24.
- KORMOS, J. (2006) *Speech Production and Second Language Acquisition*. Laurence Erlbaum, London.
- KORMOS, J. et DENES, M. (2004) Exploring measures and perceptions of fluency in the speech of second language learners, *System*, 32, pp. 145-164.
- KHOURI-DAGHER, N. (2008) *Hammam et Beaujolais*. L'échelle : Zellige.

- KRASHEN, S. (1982) *Principles and Practice in Second Language Acquisition*. Pergamon, Oxford.
- KRASHEN, S. (1981) *Second Language Acquisition and Second Language Learning*. Oxford : Pergamon.
- KRASHEN, S. D. et SCARCELLA, R. (1978) On Routines and Patterns in Language Acquisition and Performance. *Language Learning*, 28, pp. 283-300.
- KUIPER, K. (1991) Sporting formulae in New Zealand English : two models of male solidarity, in J. Cheshire (ed.) *English around the world : sociolinguistic perspectives*. Cambridge : Cambridge University press, pp. 200-209.
- KUIPER, K. (1996) *Smooth talkers. The linguistic performance of auctioneers and sportcasters*. Mahwah NJ : Lawrence Erlbaum Associates.
- KUIPER, K. et HAGGO, D. C. (1984) Livestock Auctions, Oral Poetry and Ordinary Language. *Language in Society*, 13, pp. 205-234.
- LABASCOULE, J. LAUSE, et C. ROYER, C. (2004) *Rond-Point 1 – Méthode de français basée sur l'apprentissage par les tâches*. Grenoble, Barcelone : PUG.
- LAMIROY, B. (2008) Les expressions figées : à la recherche d'une définition, in P. Blumenthal et S. Mejri, (eds.) *Les séquences figées : entre langue et discours* Stuttgart : Steiner, 85-98.  
<http://www.ling.arts.kuleuven.be/franitalco/papers/Lamiroy2008.pdf>
- LANGACKER, R. (1987) *Foundations of cognitive grammar* (Vol. 1). Stanford, CA : Stanford University Press.
- LANGACKER, R. W. (1988) A usage-based model, in B. Rudzka-Ostyn (ed.) *Topics in Cognitive Linguistics*. Amsterdam / Philadelphia : Benjamins, pp. 127-161.
- LANGACKER, R. W. (2000) A dynamic usage-based model, in M. Barlow et S. Kemmer (ed.), *Usage-based Models of Language*. Stanford : CSLI Publications, pp. 1-63.
- LAPIERRE, J. et GINGRAS, G. (2001) *Perception des professeurs et des étudiants quant à l'utilisation et l'impact des TIC à l'École polytechnique de Montréal*, Département de mathématiques et de génie industriel, École polytechnique de Montréal.
- LEGALLOIS, D. (2012) Pour une sémantique intensionnelle des constructions grammaticales, in M. Van Peteghem, P. Lauwers, E. Tobback, A. Demol et L. De Wilde *Le verbe en verve*. Gent, Academia Press.
- LEGALLOIS, D. et GREA, P. (2006) La grammaire de construction, Autour des Grammaires de Constructions et de Patterns, *Cahiers du Crisco*.
- LEGALLOIS, D. et TUTIN, A. (2013) Présentation : Vers une extension du domaine de la phraséologie, *Langages*, 189 (1), pp. 3-25.
- LENNON, P. (1984) Retelling a story in English as a second language, university of reading, in W. Dechert, D. Möhle et M. Raupach (eds.) *Second Language Productions*, pp. 50-68.
- LENNON, P. (1990) Investigating Fluency in EFL : A Quantitative Approach, *Language Learning*, 40 (3), pp. 387-417.

- LENNON, P. (1994) Self-correction and Error in Advanced Learner Spoken Narrative, in Barteld Guillermo, *The Dynamics of Language Processes. Essays in Honor of Hans W. Dechert*. Gunter Narr Verlag Tübingen, pp. 85-104.
- LENNON, P. (2000) The lexical Element in Spoken Second Language Fluency, in H. Riggenbach, (ed.) *Perspectives on Fluency*. The University of Michigan Press : Ann Arbor, pp. 25-42.
- LEVELT, W. (1989) *Speaking : From intention to articulation*. Cambridge, MA : Mit Press.
- LEVELT, W. (1999) Producing spoken language : a Blueprint of the speaker, in C. Brown et P. Hagoort (eds.) *The neurocognition of language*. Oxford : Oxford University Press, pp. 83-122.
- LEWIS, M. (1993) *The Lexical Approach*. Hove, UK : Language Teaching Publications.
- LEWIS, M. (2003) Towards a Lexical View of Language – a Challenge for Teachers, *Babylonia* 3/05, pp. 7-10.
- LIEVEN, E. BEHRENS, H. SPEARES, J. et TOMASELLO, M. (2003) Early syntactic creativity : a usage-based approach, *Journal of Child Language*, 30 (2), pp. 333-70.
- LIGHTBOWN, P. et SPADA, N. (2006) *How languages are Learned*. Oxford : Oxford University Press.
- LIN, P. M. S. (2010) The Phonology of formulaic sequences : A Review, in Wood (ed.) *Perspectives on formulaic language : Acquisition and Communication*. London : Continuum, pp. 174-193.
- LOGAN, G. D. (1988) Toward an instance theory of automatization, *Psychological Review*, 95(4), pp. 492-527.
- LOGAN, G. D. (1991) Automaticity and memory, in W. E. Hockley et S. Lewandowski (eds.) *Relating Theory and Data*. Hillsdale, NJ : Laurence Erlbaum. pp. 347-67.
- MCLAUGHLIN, B. (1990) Restructuring, *Applied Linguistics*, 11(2), pp. 113-28.
- MACWHINNEY, B. (1978) Review of “Child discourse”. *Language*, 54, pp. 230-233.
- MARCEL, C. V. A. (1820) *Practical Method of Teaching the Living Languages, Applied to French, in which Several Defects of the Old Method are Pointed Out and Remedied*. London : Hurst, Robinson.
- MARTIN, M. (1984) Advanced Vocabulary Teaching : The Problem of Synonyms. *The Modern Language Journal*, 68 (2), pp. 130-137.
- MEDIADICO (dictionnaire de la famille en ligne).  
<http://www.mediadico.com/dictionnaire/expression/sortie>
- MEJRI, S. (2004) Présentation, *Syntaxe et sémantique*, 1, (5), pp. 9-12.
- MEJRI, S. (2011) Figement, Collocation et Combinatoire libre, in J-C Anscombe et S. Mejri (eds.) *Le Figement Linguistique : la parole entravée*. Champion : Paris, pp. 63-77.
- MEL'ÇUK, I. (1993) La Phraséologie et son rôle dans l'enseignement-apprentissage d'une langue étrangère, *Etudes de Linguistique appliquée*, 92, pp. 82-113.

- MEL'ÇUK, I. (1995) Phrasemes in language and phraseology in linguistics, in M. Evereat, E-J. Van der Linden, A. Schenk et R. Schreuder (eds.) *Idioms : Structural and Psychological Perspectives*, 1, pp. 67-232.
- MEL'ÇUK, I. (1997) *Vers une linguistique Sens-Texte. Leçon inaugurale*. Paris : Collège de France, 78 pages. <http://olst.ling.umontreal.ca/pdf/MelcukColldeFr.pdf>
- MEL'ÇUK, I. (1998) Collocations and Lexical Fonctions, in A.P. Cowie (ed.) *Phraseology, Theory, Analysis and Applications*. Oxford : Clarendon Press, pp. 23-53.
- MEL'ÇUK, I. (2011) Phrasèmes dans le dictionnaire, in J-C Anscombe et S. Mejri (eds.) *Le Figement Linguistique : la parole entravée*. Champion. Paris, pp. 41-61.
- MEL'ÇUK, I. et POLGUERE, A. (2007) *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. De Boeck : Bruxelles.
- MILLER, G.A. (1955) The Magical Number Seven, Plus or Minus Two : Some Limits on Our Capacity for Processing Information, *Psychological Review*, 101 (2), pp. 343-35.
- MILLER, N. (2011) The Processing of Malformed Formulaic Language, *Applied Linguistics*, 32 (2), pp. 129-148.
- MÖHLE, D. (1984) A comparison of the second language speech production of different native speakers, in H. W. Dechert, D. Möhle et M. Raupach (eds.) *Second Language Productions*. Narr : Tübingen, pp. 26-49.
- MOREL, M-A. et DANON-BOILEAU, L. (1998) *Grammaire de l'intonation, l'exemple du français*, Paris : Ophrys.
- MYLES, F. (2004) From data to Theory : The Over-Representation of Linguistic Knowledge in SLA, *Transactions of the Philological Society volume*, 102 (2), pp. 139-168.
- MYLES, F. (2007) Complexité, exactitude et fluidité : Le Rôle que jouent les séquences préfabriquées dans l'interlangue des débutants, in S. Van Daele, A. Housen, F. Kuiken, M. Pierrard et I. Vedder (eds.) *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching Transactions of the philological society*, pp. 167-182.
- MYLES, F. HOOPER, J. et MITCHELL, R. (1998) Rote or Rule? Exploring the Role of Formulaic Language in Classroom Foreign Language Learning, *Language Learning*, 48 (3), pp. 323-363.
- MYLES, F. MITCHELL, R. et Hooper, J. (1999) Interrogative chunks in French L2 : A basis for creative construction ? *Studies in Second Language Acquisition*, 21, pp. 49-80.
- NATION, I. S. P. et WEBB, S. (2011) *Researching and Analyzing Vocabulary*. Boston, MA : Heinle.
- NATTINGER, J. et DeCARRICO, J. (1992) *Lexical Phrases and Language Teaching*. Oxford : Oxford University Press.
- NESPOULOUS, J-L. LECOURE, A. R. (1980) Pourquoi l'aphasique peut-il dire : « je ne peux pas le dire » ET PAS « Elle ne peut pas la chanter ? ». De l'intérêt des dissociations verbales dans l'étude du « parler aphasique », in A.R. Lecours et al. (eds.) *Le parler des parlés*.



NESPOULOUS, J-L. CODE, C. VIRBEL, J. et LECOURE, A. R. (1998) Hypotheses on the dissociation between « referential » and « modalizing » verbal behavior in aphasia, *Applied Psycholinguistics*, 19 (2), pp. 311-331.

NEWELL, A. (1990) *Unified Theories of Cognition*. Cambridge, Massachusetts : Harvard University Press.

NOOTEBOOM, S.G. (1980) Speaking and unspeaking : Detection and correction of phonological and lexical errors in spontaneous speech, in V.A. Fromkin (ed.) *Errors in Linguistic Performance : Slips of the Tongue, Ear, Pen and Hand*. New York : Academic Press, pp. 87-96.

OPPENHEIM, N. (2000) The Importance of Recurrent sequences for Nonnative Speaker Fluency and Cognition, in H. Riggenbach (ed.) *Perspectives on Fluency*. Ann Arbor : University of Michigan Press, pp. 220-240.

OUELLET, J. DELISLE, D. COUTURE, et J. GAUTHIER, G. (2000) *Les TIC et la réussite éducative au collégial*, Groupe de recherche et d'intervention en éducation (GRIE), Chicoutimi, Collège de Chicoutimi, [<http://www.cegep-chicoutimi.qc.ca/recherc/grie/rapportparea.pdf>]

PAILLARD, M. (2000) *Lexicologie contrastive Anglais-Français : Formation Des Mots et Construction du sens*. Paris : Ophrys.

PARADIS, M. (2000) Cerebral representation of bilingual concepts, *Bilingualism : Language and Cognition*, 3, pp. 22-24.

PARADIS, M. (2004) *A Neurolinguistic theory of bilingualism*. John Benjamins, Amsterdam.

PARADIS, M. (2005) Neurofunctional components of the bilingual cognitive system. *Seminar on Bi- and Multilingualism and Cognition* (19 mai). MitteleuropaFoundation, Bolzano.

PAWLEY, A. et SYDER, F. (1983) Two Puzzles for Linguistic Theory : Native selection and native fluency, in J. Richards et R. W Schmidt (eds.) *Language and Communication*. London : Longman, pp. 191-225.

PAWLEY, A. et SYDER, H. F. (2000) The One-Clause-at-a-Time hypothesis, in H. Riggenbach (ed.) *Perspectives on Fluency*. Ann Arbor : University of Michigan Press, pp. 163-199

PERRIN, L. (2011) Figement, Enonciation et Lexicalisation Citative, in J-C Anscombe et S. Mejri (eds.) *Le Figement Linguistique : la parole entravée*. Champion. Paris, pp. 81-94.

PERRIN DEBOCK, K. (2006) *On est heureux comme ça ! Ces idées reçues qui plombent la France*. Paris : First.

PECMAN, M. (2005) De la phraséologie à la traductologie proactive : essai de synthèse des fondements théoriques sous-tendant la recherche en phraséologie. Colloque international du 50e anniversaire de la revue META. *Pour une traductologie proactive*. 7-9 avril 2005. Université de Montréal. Canada.

PETERS, A. (1977) Language Learning Strategies : Does the whole Equal the Sum of the Parts, *Langage* 53, pp. 560-573.

PETERS, A. (1983) *The Units of Language Acquisition*. Cambridge : Cambridge University Press.



- PETEERS, B (2010) La métalangue sémantique naturelle : acquis et défis, in J. François (ed.) *Grandes voies et chemins de traverse de la sémantique cognitive* (Mémoires de la Société de linguistique de Paris, N.S. 18). Leuven : Peeters, pp. 75-101.
- PETEERS, B. (2012) Les petites idées d'un petit Belge, ou quand *petit* ne renvoie pas à la taille. 3<sup>ème</sup> Congrès Mondial de Linguistique Française (CMLF). *EDP Sciences*, pp. 1893-1907.
- PETEERS, B (2012) L'interculturel servi à la sauce MSN, ou À quoi sert la métalangue sémantique naturelle ? In Auger, N. Béal, C. et Demougin, F. (eds.), *Interactions et interculturalité : variété des corpus et des approches*. Bern : Peter Lang, pp. 149-180.
- PIEMONT, N. Enseigner le lexique au collège. *Approche didactique de la lexicologie explicative et combinatoire*. [Consulté le 25 septembre 2015]. [http://www.langages.crdp.accreteil.fr/rubriques/pdf/pratiques\\_pedagogiques/enseigner\\_lexique.pdf](http://www.langages.crdp.accreteil.fr/rubriques/pdf/pratiques_pedagogiques/enseigner_lexique.pdf)
- PINKER, S. (1994) *The Language instinct*. New York : Harper Perennial Modern Classics.
- PORQUIER, R. (2001) Mots-phrases, phrasillons, locutions-énoncés : aux frontières de la grammaire et du lexique en français langue étrangère, *Langue Française*, 131, pp. 106-123.
- POULISSE, N. (1999) *Slips of the tongue : Speech errors in first and second language production*. Amsterdam : Benjamins.
- RAUPACH, M. (1980) Temporal variables in first and second language speech production, in H.W. Dechert, M. Raupach (eds.) *Temporal Variables in Speech*. The Hague/Paris/New York : Mouton Publishers, pp. 263-270.
- RAUPACH, M. (1984) Formulae in second language production, in H.W. Dechert, D. et M. Raupach (eds.) *Second language productions*. Tübingen : Narr, pp. 114-137.
- REGAN, V. (1995) The acquisition of sociolinguistic native speaker norms : Effects of a year abroad on second language learners of French, in B. F. Freed (ed.) *Second language acquisition in a study abroad context*. Amsterdam : Benjamins, pp. 245-267.
- REGAN, V. (1997) Les apprenants avancés, la lexicalisation et l'acquisition de la compétence sociolinguistique : Une approche variationniste, *AILE*, 9, pp. 193-214.
- REGAN, V. (2002) L'acquisition de la variation par les apprenants du français langue seconde, *Aile*, 17, pp. 123-142.
- REGAN, V. (2004) The relationship between the group and the individual and the acquisition of native speaker variation patterns : A preliminary study, *International Review of Applied Linguistics*, 42, pp. 335- 34.
- REHBEIN, J. (1987) On fluency in second language speech, in H. Dechert et M. Raupach (eds.) *Psycholinguistic Models of Production*. Norwood, NJ : Ablex, pp. 97-105.
- REY, A. et CHANTEREAU, S. (1993) *Le Robert, Dictionnaire d'expressions et locutions*. Collection les Usuels.
- RIEGEL, M. PELLAT, J-C. et RIOUL, R. (2004) *Grammaire méthodique du français*. Paris : PUF.
- RIGGENBACH, H. (1991) Towards an understanding of fluency : A microanalysis of nonnative speaker conversation, *Discourse Processes*, 14, pp. 423-441.

- ROBINSON, P. et HA, M. A. (1993) Instance theory and second language rule learning under explicit conditions, *SSLA*, 15, pp. 413-438.
- ROSSITER, M. J. (2009) Perceptions of L2 Fluency by Native and Non-native Speakers of English, *The Canadian Modern Review*, 65 (3), pp. 395-412.
- ROSSITER, M.J. DERWING, T.M. MANIMTIM, L.G. et THOMSON, R.I. (2010) Oral Fluency : The Neglected Component in the Communicative Classroom, *The Canadian Modern Review*, 66 (4), pp. 583-606.
- SAJAVAARA, K. (1987) Second language speech production : Factors affecting fluency, in H. W. Dechert et M. Raupach (eds.) *Psycholinguistic models of production*. Norwood, NJ : Ablex, pp. 45-65.
- SAJAVAARA, K. et LEHTONEN, J. (1978) Spoken Language and the concept of fluency, in L. Lautamatti, et P. Lindqvist (eds.) *Focus on spoken language* Jyvaskyla : Language Center for Finnish University, pp. 23-57.
- SAUSSURE (de), F. (1916) *Cours de Linguistique Générale*. Paris. Payot.
- SCHMÄLE, G. (2013) Qu'est-ce qui est préfabriqué dans la langue ? – Réflexions au sujet d'une définition élargie de la préformation langagière, *Langages*, 189, pp. 27-45.
- SCHMIDT, R. (1990) The Role of Consciousness in Second Language Learning, *Applied Linguistics*, 11, (2), pp. 129-158.
- SCHMIDT, R. (1992) Psychological mechanisms underlying second language fluency, *SSLA*, 14, pp. 357-385.
- SCHMITT, N. (2004) (ed.) *Formulaic Sequences Acquisition, processing and use*. Amsterdam : John Benjamins.
- SCHMITT, N. (2010) *Researching Vocabulary. A Vocabulary Research Manual*. Palgrave Mac Millan.
- SCHMITT, N. et CARTER, R. (2004) Formulaic sequences in action : An introduction, in N. Schmitt (ed.) *Formulaic sequences : Acquisition, processing, and use*. Amsterdam : Benjamins, pp. 1-22.
- SCHMITT, N. GRANDAGE, S. et ADOLPHS, S. (2004) Are corpus-derived recurrent clusters psycholinguistically valid ?, in N. Schmidt (ed.) *Formulaic sequences : Acquisition, Processing and Use*. Amsterdam : John Benjamins, pp. 127-151.
- SCHMITT, N. et UNDERWOOD, G. (2004) Exploring the processing of formulaic sequences through a self-paced reading task, in N. Schmitt (ed.) *Formulaic Sequences*. Amsterdam : John Benjamins, pp. 172-189.
- SEGALOWITZ, N. (2000) Automaticity and attentional skill in fluent performance, in H. Riggenbach (ed.) *Perspectives on Fluency*. Ann Arbor : University of Michigan Press, pp. 200-219.
- SEGALOWITZ, N. (2003) Automaticity and second languages, in C. J. Doughty et M. H. Long (eds.) *Handbook of Second Language Acquisition*. Malden, MA : Blackwell Publishing, pp. 382-408.
- SEGALOWITZ, N. (2010) *The Cognitive Bases of Second Language Fluency*. New York : Routledge.

- SEGALOWITZ, N. et FREED, B.F. (2004) Context, contact, and cognition in oral fluency acquisition - Learning Spanish in at home and study abroad contexts, *Studies in Second Language Acquisition*, 26 (2), pp. 173-199.
- SIYANOVA, A. et SCHMITT, N. (2008) L2 Learner Production and Processing of Collocation : A Multi-study Perspective, *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 64, (3), pp. 429-458.
- SKEHAN, P. (1998) *A Cognitive Approach to Language Learning*. Oxford : Oxford University Press.
- SKEHAN, P. et FOSTER, P. (2001) Cognition and tasks, in P. Robinson (ed.) *Cognition and Second Language Instruction*. Cambridge University Press, pp. 183-205.
- SELINKER, L. (1972) Interlanguage, *International Review of Applied Linguistics*, 10, pp. 209-231.
- SINCLAIR, J. (1991) *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.
- SINCLAIR, J. (2003a) *Reading Concordances. An Introduction*. London : Longman
- SINCLAIR, J. (2003b) *Collins Cobuild Dictionnary*.
- SMITH, R. (2009) A Neglected Applied Linguist ? *Language and History*, 52 (2), pp. 171-181.
- SPIELMANN, G. et RADOFSKY, M. L. (2001) Learning Language under Tension : New Directions from a Qualitative Study, *The Modern Language Journal*, 85, (2), pp. 259-278.
- SVENSSON, M. H. (2004) *Critères de figement. L'identification des expressions figées en français Contemporain*. Umeå university : Institutionen för Moderna Språk Publikationer.
- SWEET, H. (1899) *The Practical Study of Languages : A Guide for Teachers and Learners*. London : Dent.
- TESNIERE, L. (1969) Les éléments de syntaxe structurale, *Langue Française*, 1, pp. 36-40.
- THEOPHANOUS, O. (2008) Polylexicalité et préconstruction lexicale : Pour une vision élargie du lexique en FLE, *Travaux de didactique du FLE*, 59, (1), pp. 45-58.
- THEOPHANOUS, O. (2013) Le traitement de la préconstruction lexicale dans un corpus des méthodes de français langue étrangère, in C. Garcia-Deban, C. Masseron et C. Ronveaux (eds.) *Enseigner le lexique*. AIRDF no 5, Collection Recherches en Didactique du français. Presses Universitaires de Namur, pp. 173-192.
- THEOPHANOUS, O. et PEREZ-BETTAN, A. (2009) Séquences préfabriquées et variation dans le discours spontané des apprenants de français langue étrangère. Communication présentée au colloque international : *Hétérogénéité et variation : quels objets sociolinguistiques et didactiques aujourd'hui ?*, Université Paul Valéry Montpellier, 27-29 mai 2009.
- THEOPHANOUS, O. et PEREZ-BETTAN, A. (2011) Segments de parole et fluence verbale en L2 : l'apport des séquences formulaïques. Communication prononcée au *Colloque Dés-organisation de l'oral*. Rennes.
- THEOPHANOUS, O. et PEREZ-BETTAN, A. (2012) Séquences préfabriquées et variation dans le discours spontané des apprenants de français langue étrangère, in M. Dreyfus et J.-M. Prieur (eds.)

*Hétérogénéité et variation. Perspectives sociolinguistiques, didactiques et anthropologiques.* Paris : Michel Houdiard Editeur, pp. 279-294.

THEOPHANOUS, O. PEREZ-BETTAN, A. et HILTON, H. (2014) Paroles reprises, erreur et fluence dans les narrations orales spontanées des apprenants de français langue étrangère. *TRANEL* (Travaux neuchâtelois de linguistique), 60, pp. 195-207.

THEOPHANOUS, O. (2014) Langage préfabriqué et développement de la fluidité verbale chez l'apprenant d'une L2, in C. Martinot et A. Pegaz Paquet (eds.) *Innovations didactiques en français langue étrangère.* Paris : Cellule de Recherche en Linguistique, pp. 117-128.

THEOPHANOUS, O. et PEREZ-BETTAN, A. (à paraître) Apport des séquences préfabriquées dans la fluence verbale en langue seconde/étrangère, in S. Mejri (éd.) Paris : Champion.

TIMONEY, C. (2007) *Pardon my French : unleash your inner Gaul.* London : Penguin.

TOGEBY, K. (1958) Les diminutifs dans les langues romanes du Moyen Age, *Studia neophilologica*, 30, pp. 192-199.

TOMASELLO, M. (2000) First steps towards a usage-based theory of language acquisition. *Cognitive Linguistics*, 11, pp. 61-82.

TOMASELLO, M. (2003) *Constructing a Language : A Usage-Based Theory of Language Acquisition.* Cambridge, MA : Harvard University Press.

TOWELL, R. (2002) Relative degrees of fluency : A comparative case study of advanced learners of French, *Iral*, 40, p. 117-150.

TOWELL, R. (2004) Complexity, Accuracy and Fluency in second language acquisition research, in S. Van Daele, A. Housen, F. Kuiken, M. Pierrard et I. Vedder (eds.) *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching.* Bruxelles : KVAB, pp. 285-300.

TOWELL, R. et DEWAELE, J.-M. (2005) The role of psycholinguistic factors in the development of fluency amongst advanced learners of French, in J.-M. Dewaele (ed.) *Focus on French as a foreign language : Multidisciplinary approaches.* Clevedon : Multilingual Matters, pp. 210– 239.

TOWELL, R. HAWKINS, R. et BAZERGUI, N. (1996) The development of fluency in advanced learners of French, *Applied Linguistics*, 17, pp. 84-119.

TUTIN, A. (2012) Les collocations dans le champ sémantique des émotions : la régularité plutôt que l'idiosyncrasie, in Apresjan et collègues (eds.) *Meanings, texts and other exciting things : a festschrift to commemorate the 80th anniversary of Professor Igor Alexandrovic Mel'čuk.* Moscou : Jazyki slavanskoj kul'tury, pp. 602-612.

TUTIN, A. (2011) *Vers une extension du domaine de la phraséologie,* Journée d'études Conscila du 21 janvier 2011.

TUTIN, A. (2013) Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument, *Langages*, 189, pp. 47-63.

TUTIN, A et GROSSMAN, F. (2002) Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif, *Revue Française de Linguistique Appliquée*, 7, pp. 7-25.

- UNDERWOOD, G. SCHMITT, N. et GALPIN, A. (2004) An eye-movement study into the processing of formulaic sequences, in N. Schmitt (éd.) *Formulaic Sequences : Acquisition, Processing and Use*. Amsterdam : John Benjamins, pp. 153-172.
- VAN LANCKER, D. (1975) Heterogeneity in language and speech. *UCLA, Working Papers in Phonetic*, 29, pp. 1-220.
- VAN LANCKER, D. CANTER, et G. TERBEEK, D. (1981) Disambiguation of ditropic sentences, *Journal of Speech and Hearing Research*, 24, pp. 330-335.
- VAN LANCKER, et D. KEMPLER, D. (1987) Comprehension of familiar phrases by left- but not right-hemisphere damaged patients, *Brain and language*, 32, pp. 180-205.
- VERONIQUE, G. D. (2013) « La dislocation, l'emploi de c'est et il y a et l'organisation topique - focus en français parlé : des variétés d'apprenants aux usages natifs », *TIPA (Travaux interdisciplinaires sur la parole et le langage)*, 29, mis en ligne le 18 décembre 2013, consulté le 27 août 2015. URL : <http://tipa.revues.org/1022>
- VIHMAN, MM. (1982) Formulas in First and Second Language Acquisition, in L.K. Obler, et L. Menn (eds.) *Exceptional Language and Linguistics*. New York : Academic Press, pp. 261-284.
- WAGNER, R. L. et PINCHON, J. (1962) *Grammaire du français classique et moderne*. Hachette.
- WAGNER-GOUGH, J. (1978) Excerpts from comparative studies in second language learning, in E. Hatch (ed.) *Second Language Acquisition, A Book of Readings*. Rowley, MA : Newbury House.
- WEINERT, R. (1995) The Role of Formulaic Language in Second Language Acquisition : A review, *Applied Linguistics*, 16, (2), pp. 180-205.
- WIESE, R. (1984) Language production in foreign and native languages : same or different? In H.W. Dechert, D. Möhle, et M. Raupach (eds.) *Second language productions*. Tübingen : Narr, pp. 11-25.
- WIDDOWSON, H. (1989) Knowledge of language and ability for us, *Applied Linguistics*, 10, pp. 128-137.
- WIERZBICKA, A. (1972) *Semantic Primitives*. Frankfurt a. M. : Athenäum.
- WIERZBICKA, A. (1984) Diminutives and depreciatives : semantic representation for derivational categories, *Quadernidi semantica*, 5, pp. 123-130.
- WIERZBICKA, A. (1992) *Semantics, culture, and cognition : universal human concepts in culture-specific configurations*. Oxford : Oxford University Press.
- WIERZBICKA, A. (2006) Les universaux empiriques du langage : tremplin pour l'étude d'autres universaux humains et outil dans l'exploration de différences transculturelles, *LINX*, 54, pp. 151-179.
- WILLIAM, G. (2005) *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes.
- WONG-FILLMORE, L. (1976) *The second time around : Cognitive and social strategies in second language acquisition*. Thèse doctorat, Stanford University, Stanford.
- WOOD, D. (2006) Uses and Fonctions of Formulaic Séquences in Second Language Speech : An Exploration of the Fondations of Fluency, *The Canadian Modern Language Review*, 63 (1), pp. 13-33.

WOOD, D. (2010) *Formulaic Language and Second Language Speech Fluency. Background, Evidence and Classroom Applications*. Continuum. London/New York.

WRAY, A. (1999) Future Directions in Formulaic Language Research, *Journal of Foreign Languages*, 32 (6); [http://www.cardiff.ac.uk/encap/resources/wray\\_jfl\\_2009.pdf](http://www.cardiff.ac.uk/encap/resources/wray_jfl_2009.pdf)

WRAY, A. (2000) Formulaic sequences in second language teaching : principle and practice. *Applied Linguistics*, 21 (4), pp. 463-489.

WRAY, A. (2002) *Formulaic language and the lexicon*. Cambridge : Cambridge University Press.

WRAY, A. (2008) *Formulaic language : Pushing the Boundaries*. Oxford, University Press.

WRAY, A. (2009) Identifying formulaic language : Persistent challenges and new opportunities, in R. Corrigan, E. A. Moravcsik, H. Ouali et K. M. Wheatley (eds.) *Formulaic Language : Distribution and historical change*. Amsterdam/Philadelphia : John Benjamins, pp. 27-54.

WRAY, A. et PERKINS, M. (2000) The functions of formulaic language : an integrated model. *Language et Communication*, 20-1, pp. 1-28.

YORIO, C. (1989) Idiomaticity as an indicator of second language proficiency, in K. Hyltenstam et L. Obler (eds.) *Bilingualism Across the Lifespan : Aspects of Acquisition, Maturity and Loss*. Cambridge : Cambridge University Press, pp. 55-72.

YUAN, F. et ELLIS, R. (2003) The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 24, pp. 1-27.

ZELLNER, B. (1994) Pauses and the Temporal Structure of Speech, in E. Keller (éd.) *Fundamentals of Speech Synthesis and Speech Recognition*. Chichester : John Wiley.



Université Paris 8 – Vincennes – Saint-Denis  
Ecole doctorale Cognition, Langage, Interaction  
U.F.R. Sciences du langage

## **THESE**

Pour obtenir le grade de Docteur en Sciences du Langage  
Discipline : Linguistique Générale

Présentée et soutenue publiquement  
par

**Annie Perez-Bettan**

Le 09 décembre 2015

# **Apprentissage et utilisation du langage préfabriqué chez des apprenants de français langue étrangère**

---

Volume 2

Sous la direction de

**Heather HILTON**  
**Olga THEOPHANOUS**

Composition du jury

Michel Billières, professeur, Toulouse-Jean-Jaurès  
Amanda Edmonds, Maître de Conférences, Pays de l'Adour-Pau  
Fanny Forsberg Lundell, Associate Professor, Docent-Stockholm  
Heather Hilton, Professeur, Lyon 2 Lumière  
Olga Théophanous, Maître de Conférences, Toulouse Jean-Jaurès  
Marzena Watorek, Professeur, Paris 8-Vincennes-Saint-Denis

## ANNEXES

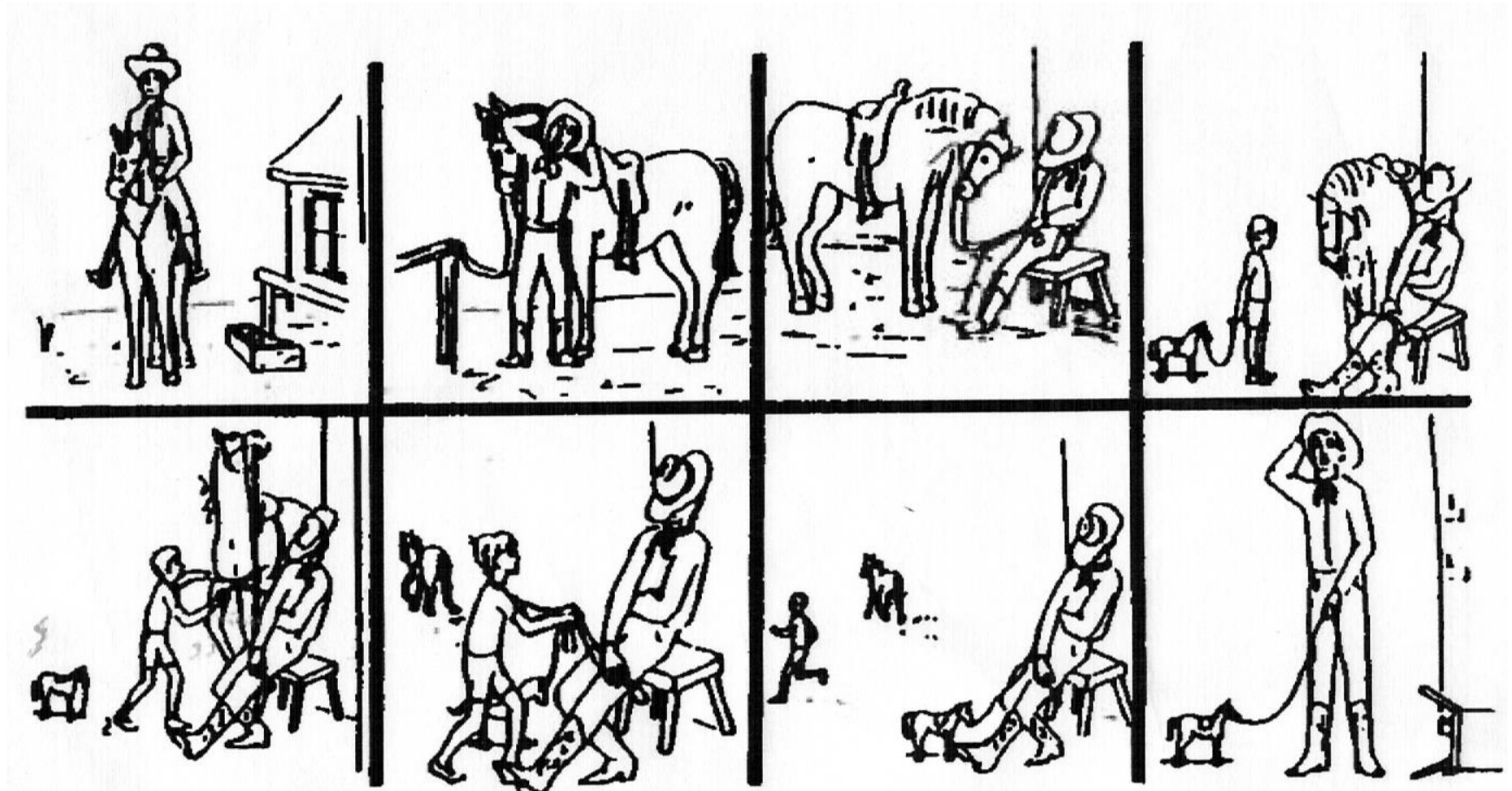
ANNEXES.....	2
ANNEXE 1.....	4
Cartoon.....	4
ANNEXE 2.....	6
Les Minuscules .....	6
ANNEXE 3.....	6
Classement des SP non-natives par sujet .....	7
ANNEXE 4.....	32
Classement des SP natives .....	32
ANNEXE 5.....	41
Narrations non-natives .....	41
ANNEXE 6.....	115
Narrations des natifs.....	115
ANNEXE 7.....	127
Tableau comparatif SP non-natifs et natifs.....	127
ANNEXE 8.....	129
Tableau des SP déviantes .....	129
ANNEXE 9.....	134
Types d'erreurs.....	134
ANNEXE 10 .....	136
Quelques occurrences de « petit » dans <i>lex tutor</i> .....	136
ANNEXE 11 .....	137
Quelques occurrences de “et après” .....	137
ANNEXE 12 .....	138
Occurrences de « et donc ».....	138
ANNEXE 13 .....	139
Présentation des résultats des natifs par tâche, t1 et t2 .....	139



<b>ANNEXE 14</b> .....	<b>141</b>
<b>TextGrid text file</b> .....	<b>141</b>
<b>ANNEXE 15</b> .....	<b>143</b>
<b>Document de travail Excel - Calculs</b> .....	<b>143</b>
<b>ANNEXE 16 <i>Idiomatics</i></b> .....	<b>145</b>

**ANNEXE 1**

**Cartoon**





## Classement des SP non-natives par sujet

AK																	
ID	Bd1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	Bd2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	
AK	petit cheval	BD1	lex	1				SN	petit enfant (2), *petits enfants(2)	BD2	lex	4			2	SN	
AK	se lève	BD1	lex	1				SV	*petit maquette (5)	BD2	lex	5			5	SN	
AK	a *été sommeil	BD1	lex	1			1	SV	descend de	BD2	lex	1				SV	
AK	en* cheval	BD1	lex	1			1	Sadv	a sommeil (2)	BD2	lex	2				SV	
AK	là-bas	BD1	lex	1				Sadv	à cheval	BD2	lex	1				Sprep	
AK	c' est fini	BD1	disc		1				c' est marrant	BD2	lex	1				Sprep	
AK	et après (3)	BD1	dis		3				comment dire (3)comment dit* (2)	BD2	dis		5		2		
AK	et quand	BD1	dis		1				et après (4)	BD2	dis		4				
AK	et pis	BD1	dis		1				peut-être (2)	BD2	dis		2				
AK	et juste	BD1	dis		1				c' est c' est	BD2	dis		1				
AK	tant pis (3)	BD1	dis		3				il a pensé que	BD2	gram			1			
AK	comme ça	BD1	dis		1				un peu	BD2	gram			1			
AK	il est (désolé), il est (né)	BD1	gram			2			je pense que	BD2	gram			1			
AK	un peu	BD1	gram			1			il y avait	BD2	gram			1			
AK	près de (2)	BD1	gram			2											
AK	il y a (2), il y a pas	BD1	gram			3											
AK	y a pas (2)	BD1	gram			2											
F1								F2									
AK	bonne odeur	F1	lex	1				SN	salle de bains (2)	F2	lex	2				SN	
AK	s'appelle	F1	lex	1				SV	*pois de café(2)	F2	lex	2			2	SN	
AK	essaie de	F1	lex	1				SV	sortir de	F2	lex	1				SV	
AK	à la maison (3)	F1	lex	3				SV	abandonné à	F2	lex	1				SV	
AK	c' est amusant	F1	lex	1				Sprop	commencé à	F2	lex	1				SV	
AK	c'est comme ça	F1	dis		1				à la maison	F2	lex	2				Sprep	
AK	d'accord	F1	dis		1				juste en face	F2	lex	1					
AK	c'est	F1	dis		2				se promener	F2	lex	1					
AK	et après (7)	F1	dis		7				peut-être (2)	F2	dis		2				
AK	et donc	F1	dis		1				parce que	F2	dis		1				
AK	je sais pas (4)	F1	dis		4				c' est	F2	dis		1				
AK	je sais	F1	dis		1				et donc	F2	dis		1				
AK	il y a	F1	dis		1				comment dire (6)	F2	dis		6				
AK	peut-être (2)	F1	dis		2				et après (4)	F2	dis		4				
AK	je sais pas pourquoi (2)	F1	dis		2				j'ai oublié	F2	dis		1				
AK	comment faire	F1	dis		1				ah oui	F2	dis		1				
AK	mais après	F1	dis		1				comme ça (2)	F2	dis		2				
AK	je pense (2)	F1	dis		2				au fait (2)	F2	dis		2				
AK	qu'est ce que ?	F1	gram			1			et finalement (2)	F2	dis		2				
AK	qu'est ce que c' est	F1	gram			2			oh mince	F2	dis		1				
AK	il va	F1	gram			1			comment *dit	F2	dis		1		1		
AK	quelque chose (2)	F1	gram			2			par contre	F2	dis		1				
AK	quelqu'un	F1	gram			1			et soudain	F2	dis		1				
AK	il sait pas	F1	gram			1			il y a (+ comment dire)	F2	dis		1				
AK	il peut pas	F1	gram			1			être en train de (2)	F2	gram			2			
AK	au lieu *que	F1	gram			1	1		pas trop	F2	gram			1			
AK	*il y avait	F1	gram			1	1		*tous les, *tout la	F2	gram			3	2		
AK	c' est + N	F1	gram			3			il y avait...qui	F2	gram			1			
									il est (heureux), il était (heureux)	F2	gram			2			
									il a trouvé que	F2	gram			3			
									il y a (5)	F2	gram			5			
									c'est + N	F2	gram			1			
				<b>26</b>	5	11	10	2					<b>30</b>	14	12	4	9
				<b>47</b>	7	26	14	2					<b>57</b>	11	28	18	5
				<b>73</b>	12	37	24	4					<b>87</b>	25	40	22	14
<b>occurrences/types</b>								<b>occurrences/types</b>									
	BD1		17		5	7	5	2		BD2	14		6	4	4	3	
	F1		28		5	13	10	2		F2	33		8	17	8	3	
	t1		45		10	20	15	4		t2	47		14	21	12	6	
		<b>t1 t2</b>	160	37	77	46	18										
		<b>t1 t2</b>	92	24	41	27	10										



BD1	40	22	10	8	3		BD2	49	25	12	12	4
F1	53	24	16	13	1		F2	79	39	26	14	6
t1	93	46	26	21	4		t2	128	62	38	26	10
occurrences/types							occurrences/types					
BD1	31	17	7	7	3		BD2	36	20	9	7	4
F1	45	19	14	12	1		F2	51	22	16	13	4
t1	76	36	21	7	4		t2	87	42	25	20	8
t1 et t2	221	108	64	47	14							
t1 et t2	163	78	46	27	12							

HE																		
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme		BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	
HE	il est fatigué	BD1	lex	1				SV		années soixante (2)	BD2	lex	2				SN	
HE	se réveille	BD1	lex	1				SV		petit xxx, petit garçon, petit cheval	BD2	lex	3				SN	
HE	qu'est ce qui s'est passé (2)	BD1	lex	2				Sprop		a l'air (2)	BD2	lex	2				SV	
HE	il* pas compris (2)	BD1	lex	2			2	Sprop		avoir sommeil	BD2	lex	1				SV	
HE	c'est	BD1	dis		1					s'assoit	BD2	lex	1				SV	
HE	et quand	BD1	dis		1					se réveille	BD2	lex	1				SV	
HE	en fait	BD1	dis		1					arrivé à	BD2	lex	1				SV	
HE	parce que	BD1	dis		1					ça a l'air	BD2	lex	1				Sprop	
HE	c'est ça	BD1	dis		1					ça ressemble *d'(être)	BD2	lex	1			1	Sprop	
HE	un petit peu	BD1	gram			1				je sais pas quoi	BD2	dis		1				
HE	devient + adj (petit)	BD1	gram			1				et après (2)	BD2	dis		2				
HE	c'est (un ombre)	BD1	gram			1				et quand	BD2	dis		1				
										et là (FF)	BD2	dis		1				
										parce que	BD2	dis		1				
										peut-être	BD2	dis		1				
										c'est ça	BD2	dis		1				
										ça y est	BD2	dis		1				
										un peu de* (sommeil)	BD2	gram			1	1		
										un peu	BD2	gram			1			
										il vient* + V	BD2	gram			1	1		
										à cause de (2)	BD2	gram			2			
										il y a (3)	BD2	gram			3			
F1									F2									
HE	porte ouverte	BD2	lex	1				SN		grains de café	F2	lex	1				SN	
HE	a essayé de (2)	F1	lex	2				SV		allers retours	F2	lex	1				SN	
HE	réussi à	F1	lex	1				SV		essayer de (2)	F2	lex	2				SV	
HE	sortir de (la maison)	F1	lex	1				SV		sortir de	F2	lex	1				SV	
HE	là dedans	F1	lex	1				Sadv		se balade sur	F2	lex	1				SV	
HE	là bas	F1	lex	1				Sadv		se cacher	F2	lex	1				SV	
HE	par la fenêtre	F1	lex	1				Sprep		aller aux toilettes	F2	lex	1				SV	
HE	à la maison (2)	F1	lex	2				Sprep		va à *les toilettes	F2	lex	1			1	SV	
HE	(tout) ce qui s'est passé	F1	lex	1				Sprop		ne pas arriver* (à)	F2	lex	1			1	SV	
HE	c'était faux (2)	F1	lex	2				Sprop		en verre	F2	lex	1				Sadv	
HE	plusieurs fois	F1	lex	1						plusieurs fois	F2	lex	1				Sadv	
HE	c'est (2)	F1	dis		2					qu'est ce qui s'est passé	F2	lex	1				Sprop	
HE	c'est c'est	F1	dis		1					c'est + (c'est juste)	F2	dis		1				
HE	parce que	F1	dis		1					peut-être	F2	dis		1				
HE	et après (4)	F1	dis		4					en fait (2)	F2	dis		2				
HE	après ça (2)	F1	dis		2					c'est pas	F2	dis		1				
HE	le deuxième (film)	F1	dis		1					et après (4)	F2	dis		4				
HE	il y a	F1	dis		1					c'est ça	F2	dis		1				
HE	qu'est ce que + X ?	F1	gram			1				c'est juste (2)	F2	dis		2				
HE	en train de	F1	gram			1				un peu	F2	gram			1			
HE	chaque fois que	F1	gram			1				à l'extérieur de	F2	gram			1			
HE	d'autres (fleurs)	F1	gram			1				il y a ...qui	F2	gram			1			
HE	est allée + V	F1	gram			1				il y a (4)	F2	gram			4			
HE	elle croyait que	F1	gram			1				va à + lieu	F2	gram			1			
HE	elle savait pas que	F1	gram			1												
HE	il y a eu (2)	F1	gram			2												
HE	la fin de (le film)	F1	gram			1												
HE	c'était (peinte), c'était (la fin de)	F1	gram			2												
HE	c'est (l'histoire), c'est (un)	F1	gram			2												
	BD1	14		6	5	3	2			BD2	30		13	9	8	3		
	F1	40		14	12	14	0			F2	33		13	12	8	2		
	t1	54		20	17	17	2			t2	63		26	21	16	4		
	occurrences/types									occurrences/types								
	BD1	12		4	5	3	1			BD2	22		9	8	5	3		
	F1	29		11	7	11	0			F2	24		12	7	5	2		
	t1	41		15	12	14	1			t2	46		21	15	10	5		
	t1 et t2	117		46	38	33	6											
	t1 et t2	87		36	27	24	6											





MA																
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme
MA	p'tit cheval (2), p'tit garçon	BD1	lex	3				SN	le p'tit cheval	BD2	lex	1				SN
MA	décider de	BD1	lex	1				SV	s'est réveillé	BD2	lex	1				SV
MA	a attaché à	BD1	lex	1				SV	s'est trouvé avec	BD2	lex	1				SV
MA	s'est réveillé	BD1	lex	1				SV	s'est endormi	BD2	lex	1				SV
MA	s'est endormi	BD1	lex	1				SV	s' est approché	BD2	lex	1				SV
MA	s'est assis sur	BD1	lex	1				SV	je sais pas	BD2	dis		1			
MA	se reposer	BD1	lex	1				SV	et quand	BD2	dis		1			
MA	il est fatigué	BD1	lex	1				SV	et voilà	BD2	dis		1			
MA	en bois	BD1	lex	1				Sadv	venir de + V	BD2	gram			1		
MA	comme j'ai compris (2)	BD1	dis		2				il était (étonné) il était (surpris)	BD2	gram			2		
MA	et après (3)	BD1	dis		3				y a	BD2	gram			1		
MA	je sais pas pourquoi	BD1	dis		1				c'était (le scène)	BD2	gram			1		
MA	comment dire (2)	BD1	dis		2											
MA	et quand	BD1	dis		1											
MA	pour que (2)	BD1	dis		2											
MA	tant que	BD1	dis		1											
MA	c'est comme ça	BD1	dis		1											
MA	c'est ça	BD1	dis		1											
MA	un peu	BD1	gram			1										
MA	il était (surpris)	BD1	gram			1										
MA	au lieu de	BD1	gram			1										
MA	à côté de	BD1	gram			1										
MA	l'histoire de	BD1	gram			1										
MA	c'est l'histoire où	BD1	gram			1										
MA	c'était (le p'tit cheval)	BD1	gram			1										
MA	il a vu que	BD1	gram			1										
F1									F2							
MA	sortir par	F1	lex	1				SV	sens de la sortie	F2	lex	1				SN
MA	chercher la sortie	F1	lex	1				SV	atterri sur	F2	lex	1				sv
MA	est dehors (être dehors)	F1	lex	1				SV	a décidé de	F2	lex	1				sv
MA	quelques instants	F1	lex	1				SV	s'est trouvée ailleurs	F2	lex	1				sv
MA	arrivait pas à	F1	lex	1				SV	est sortie par	F2	lex	1				sv
MA	sortir de (2)	F1	lex	2				SV	s'enfuit	F2	lex	1				sv
MA	commencer à (2)	F1	lex	2				SV	sortir de	F2	lex	1				sv
MA	est libre (être libre)	F1	lex	1				SV	essayer de	F2	lex	1				sv
MA	il arrivait pas à	F1	lex	1				SV	là bas	F2	lex	1				Sadv
MA	à la fin	F1	lex	1				Sadv	par tous les moyens	F2	lex	1				Sprep
MA	un moment après	F1	lex	1				Sadv	sortir par	F2	lex	1				
MA	là bas	F1	lex	1				Sadv	(comment) ça s'appelle	F2	lex	1				
MA	*trucs du café	F1	lex	1			1		c'est	F2	dis		1			
MA	comment ça se passe	F1	disc		1				je sais pas (4)	F2	dis		4			
MA	comment s'appelle	F1	disc		1				et après (2)	F2	dis		2			
MA	c'était	F1	dis		1				parce que	F2	dis		1			
MA	j'sais pas	F1	dis		1				peut-être	F2	dis		1			
MA	il y a (2)	F1	dis		2				j'sais pas	F2	dis		1			
MA	c'est	F1	dis		1				c'était	F2	dis		1			
MA	c'est comme ça	F1	dis		1				d'abord	F2	dis		1			
MA	parce que (2)	F1	dis		2				et voilà	F2	dis		1			
MA	je sais pas (3)	F1	dis		2				elle a compris que	F2	gram			1		
MA	et après (4)	F1	dis		4				elle peut plus (2)	F2	gram			2		
MA	j'ai oublié	F1	dis		1				elle peut	F2	gram			1		
MA	en fait (3)	F1	dis		3				elle pensait que	F2	gram			1		
MA	et donc	F1	dis		1				c'est (2)	F2	gram			2		
MA	pour que	F1	dis		1				c'est l' histoire	F2	gram			1		
MA	le deuxième cadre	F1	dis		1				il peut	F2	gram			1		
MA	et voilà	F1	dis		1											
MA	toutes les, tous les	F1	gram			2										
MA	il peut + V	F1	gram			1										
MA	il puisse (comprendre)	F1	gram			1										
MA	y avait	F1	gram			1										
MA	il pouvait pas +V	F1	gram			1										
MA	au delà de	F1	gram			1										
MA	il comprenait pas	F1	gram			1										
MA	qu'est ce qu'il fait x 2	F1	gram			2										
MA	comprendre que	F1	gram			1										
MA	c'est l'histoire de (2)	F1	gram			2										
MA	c'est (des blés) (2)	F1	gram			1										
	<b>BD1</b>	<b>33</b>		<b>11</b>	<b>14</b>	<b>8</b>	<b>0</b>		<b>BD2</b>	<b>13</b>		<b>5</b>	<b>3</b>	<b>5</b>	<b>0</b>	
	<b>F1</b>	<b>53</b>		<b>15</b>	<b>24</b>	<b>14</b>	<b>1</b>		<b>F2</b>	<b>34</b>		<b>12</b>	<b>13</b>	<b>9</b>	<b>0</b>	
	<b>t1</b>	<b>86</b>		<b>26</b>	<b>38</b>	<b>22</b>	<b>1</b>		<b>t2</b>	<b>47</b>		<b>17</b>	<b>16</b>	<b>14</b>	<b>0</b>	

MA	c'est (des blés) (2)	F1	gram		1														
	<b>BD1</b>	<b>33</b>		11	14	8	0		<b>BD2</b>	<b>13</b>		5	3	5	0				
	<b>F1</b>	<b>53</b>		15	24	14	1		<b>F2</b>	<b>34</b>		12	13	9	0				
	<b>t1</b>	<b>86</b>		26	38	22	1		<b>t2</b>	<b>47</b>		17	16	14	0				
	<b>occurrences/types</b>								<b>occurrences/types</b>										
	BD1	<b>28</b>		11	9	8			BD2	<b>12</b>		5	3	4	0				
	F1	<b>40</b>		13	16	11	1		F2	<b>30</b>		14	9	7	0				
	t1	<b>68</b>		24	25	19	1		t2	<b>42</b>		19	12	11	0				
	<b>t1 et t2</b>	<b>133</b>		43	54	36	1												
	<b>t1 et t2</b>	<b>110</b>		43	37	30	1												



MR																	
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	
MR	petit *cheveu (2)	BD1	lex	2				SN	petit enfant	BD2	lex	1				SN	
MR	se lève	BD1	lex	1				SV	*démonter de	BD2	lex	1			1	SV	
MR	*asseoir à	BD1	lex	1			1	SV	échanger N pour* N	BD2	lex	1			1	SV	
MR	en français	BD1	lex	1				Sadv	*à maison	BD2	lex	1			1	Sprep	
MR	très fatigué (2)	BD1	lex	2				Sadj	et quand (2)	BD2	dis		2				
MR	j'sais pas (s'il y a un mot)	BD1	dis		1				je crois	BD2	dis		1				
MR	et après	BD1	dis		1				c'est tout	BD2	dis		1				
MR	et quand	BD1	dis		1				Être en train de	BD2	gram			1			
MR	c'est tout	BD1	dis		1				la forme de	BD2	gram			1			
MR	très très (fatigué)	BD1	gram			1			il y a (2), il n' y a pas	BD2	gram			3			
MR	en train de	BD1	gram			1			à côté de	BD2	gram			1			
MR	(s'il y a (un mot), il y a + N (3))	BD1	gram			4			c' est	BD2	gram			1			
MR	c'est (un petit cheveu)	BD1	gram			1											
F1									F2								
MR	chemin de sortie	F1	lex	1				SN	grains de café	F2	lex	1				SN	
MR	p'tit sionair (2)	F1	lex	2				SN	salle de bains	F2	lex	1				SN	
MR	indiquer le chemin	F1	lex	1				SV	la porte ouverte	F2	lex	1				SN	
MR	sortir de	F1	lex	1				SV	est prisonnière	F2	lex	1				SV	
MR	est présent	F1	lex	1				SV	sortir à	F2	lex	1				SV	
MR	un jour	F1	lex	1				Sadv	conduit à (au jardin)	F2	lex	1				SV	
MR	en* ce moment	F1	lex	1			1	Sadv	indiquer la sortie	F2	lex	1				SV	
MR	à l'extérieur (2)	F1	lex	2				Sprep	ça s'appelle	F2	lex	1				Sprop	
MR	à la maison	F1	lex	1				Sprep	tout ça	F2	lex	1				Sprop	
MR	parce que	F1	dis		1				dans cette histoire	F2	dis		1				
MR	et finalement	F1	dis		1				c' est 2)	F2	dis		2				
MR	je crois que	F1	dis		1				jusqu'à (2)	F2	dis		2				
MR	l' autre (part*)	F1	gram			1			et après	F2	dis		1				
MR	tous les,toutes les (portes)	F1	gram			4			c'est ça	F2	dis		1				
MR	elle est + place	F1	gram			2			parce que	F2	dis		1				
MR	(voudrait) + V	F1	gram			1			beaucoup de fois	F2	gram			1			
MR	dire que	F1	gram			1			à l'intérieur de	F2	gram			1			
MR	il y a	F1	gram			1			à *la côté de, du	F2	gram			2	2		
									à * l'autre côté	F2	gram			1	1		
									il y a (5)	F2	gram			5			
									c' est (2)	F2	gram			2			
	BD1	16		5	4	7	1		BD2	15		4	4	7	3		
	F1	24		11	3	10	1		F2	29		9	8	12	3		
	t1	40		16	7	17	2		t2	44		13	12	19	6		
occurrences/types									occurrences/types								
	BD1	13		5	4	4	1		BD2	12		4	3	5	3		
	F1	18		9	3	6	1		F2	21		9	6	6	2		
	t1	31		14	7	10	2		t2	33		13	9	11	5		
	t1 et T2	84		29	19	36	8										
	t1 et T2	64		27	16	21	7										

NO																
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme
NO	petit garçon	BD1	lex	2				SN	jeune garçon	BD2	lex	1				SN
NO	se réveillait	BD1	lex	1				SV	il est fatigué	BD2	lex	1				SV
NO	n'est pas là	BD1	lex	1				SV	il s'arrête	BD2	lex	1				SV
NO	*allait au ville	BD1	lex	1			1	SV	il s'est réveillé	BD2	lex	1				SV
NO	*asseye sur	BD1	lex	1			1	SV	il est là	BD2	lex	1				SV
NO	quelques fois	BD1	lex	1				Sadv	c'est	BD2	dis	1				
NO	*chez il chez lui	BD1	lex	2			1	Sprep	je peux *commence	BD2	dis		1		1	
NO	peut-être (4)	BD1	dis			4			peut-être (2)	BD2	dis		2			
NO	c'est	BD1	dis			1			c' est tout	BD2	dis		1			
NO	après ça	BD1	dis			1			pas vraiment	BD2	gram				1	
NO	j'ai fini	BD1	dis			1			il voit que	BD2	gram				1	
NO	la* même	BD1	gram				1	1	à côté de	BD2	gram				1	
NO	il trouve que	BD1	gram				1		y a ...qui	BD2	gram				1	
NO	il y a (2)	BD1	gram				2		il y a	BD2	gram				1	
NO	il y a...qui	BD1	gram				1									
NO	proche de	BD1	gram				1									
NO	c'est,	BD1	gram				1									
NO	ce n'est pas	BD1	gram				1									
F1									F2							
NO	salle de bains (2)	F1	lex	2				SN	petits + N	F2	lex	1				SN
NO	*haricots du café	F1	lex	1			1	SN	salle de bains (2)	F2	lex	2				SN
NO	être dehors (3)	F1	lex	3				SV	*morceaux du café	F2	lex	2			2	SN
NO	sortir de	F1	lex	1				SV	*quiz du café	F2	lex	1			1	SN
NO	*faire le miel	F1	lex	1			1	SV	se pointer dans (2)	F2	lex	2				SV
NO	et quand	F1	dis			1			avoir une amie				1			SV
NO	parce que (2)	F1	dis			2			être son ami	F2	lex	1				SV
NO	peut-être (3)	F1	dis			3			est dedans (être dedans)	F2	lex	1				SV
NO	après ça (3)	F1	dis			3			être dehors (2)	F2	lex	2				SV
NO	le premier(2), *la première (insecte)	F1	dis			3		1	sont là (être là)	F2	lex	1				SV
NO	Une autre, l'autre (4) (insecte), les autres (chambres)	F1	gram				6		essayer de x 2	F2	lex	2				SV
NO	il pense que	F1	gram				1		se cache (2)	F2	lex	2				SV
NO	autour de	F1	gram				1		il *suivre le chemin	F2	lex	1			1	SV
NO	il y a (3)	F1	gram				3		*il arrive pas de	F2	lex	1			1	SV
NO	ce n'est pas	F1	gram				1		ça se passe	F2	lex	1				Sprop
NO	il a vu que	F1	gram				1		dans *le direction, dans la direction	F2	lex	2			1	Sprép
									on ne sait pas où	F2	dis		1			
									je sais pas	F2	dis		1			
									on dira	F2	dis		1			
									et après (5)	F2	dis		5			
									il y a	F2	dis		1			
									parce que	F2	dis		1			
									peut-être	F2	dis		1			
									le premier (insecte)	F2	dis		1			
									à côté de	F2	gram				1	
									*au derrière de (la peintre)	F2	gram				1	1
									*à l'autre côté de	F2	gram				1	1
									il sait pas +V	F2	gram				1	
									un peu(2)	F2	gram				2	
									tous les	F2	gram				1	
									tout le temps que	F2	gram				1	
									l'autre (2) une autre	F2	gram				3	
									les autres	F2	gram				1	
									pas vraiment	F2	gram				1	
									le nom de	F2	gram				1	
									il *voir que	F2	gram				1	1
									on *entendre que	F2	gram				1	1







OL																	
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	
OL	petit cheval	BD1	lex	1				SN	petit cheval (2), petit garçon	BD2	lex	3				SN	
OL	*monter au (à)	BD1	lex	1			1	SV	se réveille	BD2	lex	1				SV	
OL	*en cheval	BD1	lex	1			1	Sadv	il est fatigué	BD2	lex	1				SV	
OL	qu'est ce qui se passe	BD1	lex	1				Sprop	descend de	BD2	lex	1				SV	
OL	c' est	BD1	dis		1				s'assoit	BD2	lex	1				SV	
OL	et après (3)	BD1	dis		3				en bois (2)	BD2	lex	2				Sadv	
OL	il y a	BD1	dis		1				*en poing	BD2	lex	1			1	Sadv	
OL	et maintenant	BD1	dis		1				et après	BD2	dis		1				
OL	et quand	BD1	dis		1				c'est fini	BD2	dis		1				
OL	c'est fini	BD1	dis		1				très vite	BD2	gram				1		
									il peut + V	BD2	gram				1		
									il y a + N (2)	BD2	gram				2		
F1									F2								
OL	petit jaune (7), petite jupe, petit pois, petit ro(nd)	F1	lex	10				SN	salle de bains (5)	F2	lex	5				SN	
OL	salle de bains (2)	F1	lex	2				SN	en verre	F2	lex	3				SN	
OL	p'tit verre	F1	lex	1				SN	petit tableau	F2	lex	1				SN	
OL	s'arrêter (3)	F1	lex	3				SV	sortir de (2)	F2	lex	2				SV	
OL	passer par la fenêtre	F1	lex	2				SV	demander à	F2	lex	1				SV	
OL	entrer à	F1	lex	1				SV	s'écrase	F2	lex	1				SV	
OL	s' appelle	F1	lex	1				SV	s'arrêter	F2	lex	4				SV	
OL	*sort de	F1	lex	1			1	SV	passer par la fenêtre	F2	lex	1				SV	
OL	*pois de café	F1	lex	1			1	SN	se croisent	F2	lex	1				SV	
OL	quelques secondes	F1	lex	1				Sadv	la porte est fermée	F2	lex	1				SV	
OL	ça c' est quoi	F1	lex	1					*récolter du miel	F2	lex	1			1	SV	
OL	je ne sais pas (le mot)	F1	dis		1				*dans ce moment	F2	lex	1			1	SV	
OL	je pense	F1	dis		1				ça fait	F2	lex	1				SV	
OL	c'est (2)	F1	dis		2				qu'est ce qui se passe	F2	lex	1				Sprop	
OL	et après (5)	F1	dis		5				c'est + (c'est réussi)	F2	dis		1				
OL	parce que (3)	F1	dis		3				je peux + V	F2	dis		1				
OL	je pense	F1	dis		1				c'est fini	F2	dis		1				
OL	je ne sais pas +	F1	dis		1				c'est réussi	F2	dis		1				
OL	*c' est pourquoi	F1	dis		1		1		et après (9)	F2	dis		9				
OL	comme ça	F1	dis		1				et maintenant	F2	dis		1				
OL	peut-être	F1	dis		1				je pense	F2	dis		1				
OL	c'est comme ça	F1	dis		1				comment c'est	F2	dis		1				
OL	c' est tout	F1	dis		1				beaucoup de (3)	F2	gram				3		
OL	beaucoup de + N	F1	gram			1			beaucoup de fois (2)	F2	gram				2		
OL	beaucoup de fois (3)	F1	gram			3			*un bouteille de	F2	gram			1	1		
OL	quelque chose (3)	F1	gram			3			il peut + V	F2	gram				1		
OL	une autre + fleur	F1	gram			1			*dehors de	F2	gram				1		
OL	*tous les choses (2)	F1	gram			2	2		c'est + N + qui	F2	gram				1		
OL	tous les (jours)	F1	gram			1			il y a (7)	F2	gram				7		
OL	pas trop	F1	gram			1			à côté de (3)	F2	gram				3		
OL	pas vraiment	F1	gram			1			au dessus de	F2	gram				1		
OL	tous les	F1	gram			2			*à l'autre côté de	F2	gram				1	1	
OL	à côté de (2)	F1	gram			2			c'est	F2	gram				1		
OL	*dehors de	F1	gram			1	1										
OL	*à l'autre côté de	F1	gram			1	1										
OL	il y a rien	F1	gram			1											
OL	c'est + N + qui	F1	gram			1											
OL	il y a (10)	F1	gram			10											
OL	c' est comme	F1	gram			1											
OL	c' est (7)	F1	gram			7											
OL	c'est pas (2)	F1	gram			2											

BD1	12	4	8	0	2		BD2	16	10	2	4	1
F1	84	24	19	41	7		F2	62	24	16	22	4
t1	96	28	27	41	9		t2	78	34	18	26	5
<b>occurrences/types</b>							<b>occurrences/types</b>					
BD1	10	4	6	0	2		BD2	12	7	2	3	1
F1	41	11	12	18	6		F2	33	14	8	11	4
t1	51	15	18	18	8		t2	45	21	10	14	5
t1 et T2	174	62	45	67	14							
t1 et T2	96	36	28	32	13							

PR																	
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	
PR	*petite enfant, *petite cheval (2), petit *enfante, petit cheval, petit *jou	BD1	lex	6			5	SN	petite bande dessinée, petit enfant, petit *jou, *petite + cheval (2)	BD2	lex	5			3	SN	
PR	longue journée	BD1	lex	1				SN	bande dessinée	BD2	lex	1				SN	
PR	(il) *sent très fatigué	BD1	lex	1				SV	journée de travail	BD2	lex	1				SN	
PR	était fatigué (2)	BD1	lex	2				SV	a décidé de	BD2	lex	1				SV	
PR	ne pas avoir de temps	BD1	lex	1				SV	il était fatigué	BD2	lex	1				SV	
PR	se lève	BD1	lex	1				SV	se lève	BD2	lex	1				SV	
PR	faire des choses	BD1	lex	1				SV	il a *pris une sieste	BD2	lex	1			1	SV	
PR	faire une blague	BD1	lex	1				SV	ça c'est (2)	BD2	lex	2					
PR	se mettre	BD1	lex	1				SV	et donc (3)	BD2	dis		3				
PR	pour les enfants	BD1	lex	1				SN	je sais pas	BD2	dis		1				
PR	en français	BD1	lex	1				Sadv	parce que	BD2	dis		1				
PR	transformé *dans	BD1	lex	1			1	Sadj	et après	BD2	dis		1				
PR	c'était bizarre	BD1	lex	1				Sprop	il y a (2)	BD2	dis			2			
PR	*qu'est ce qui passait	BD1	lex	1			1	Sprop	un peu	BD2	gram			1			
PR	comment tu dis ça	BD1	dis		1				il a vu que	BD2	gram			1			
PR	peut-être	BD1	dis		1				il était (étonné)	BD2	gram			1			
PR	je sais pas	BD1	dis		1				c'est	BD2	gram			1			
PR	c'est ça (2)	BD1	dis		2												
PR	parce que (2)	BD1	dis		2												
PR	et après	BD1	dis		1												
PR	je crois	BD1	dis		1												
PR	je sais pas	BD1	dis		1												
PR	un peu (2)	BD1	gram				2										
PR	il a pensé que	BD1	gram				1										
PR	il a vu que	BD1	gram				1										
PR	il y a	BD1	gram				1										
PR	c'est juste (comme un jou)	BD1	gram				1										
PR	c'était une histoire de	BD1	gram				1										
F1														F2			
PR	*petite film, *petite signe (3), petite sortie, *petit abeille, petites mouches, petites abeilles, petits grains	F1	lex	9			5	SN	petite abeille, petits grains	F2	lex	2				SN	
PR	salles de bains	F1	lex	1				SN	grains de café	F2	lex	1				SN	
PR	fenêtre ouverte	F1	lex	1				SN	essayer de	F2	lex	1				SV	
PR	la première fois	F1	lex	1				SN	fermer la porte	F2	lex	1				SV	
PR	sortir de (la maison)	F1	lex	1				SV	faire *une signe	F2	lex	1			1	SV	
PR	essayer de (2)	F1	lex	2				SV	là bas (2)	F2	lex	2				Sadv	
PR	faire un signe (2)	F1	lex	2				SV	tout est faux	F2	lex	1				Sprop	
PR	se regarder	F1	lex	1				SV	en fait (4)	F2	dis		4				
PR	fabriquer du miel	F1	lex	1				SV	parce que (2)	F2	dis		2				
PR	se cachent	F1	lex	1				SV	peut-être (2)	F2	dis		2				
PR	se voient	F1	lex	1				SV	et après (2)	F2	dis		2				
PR	*était peur (2)	F1	lex	2			2	SV	et donc	F2	dis		1				
PR	se ferme	F1	lex	1				SV	et voilà	F2	dis		1				
PR	juste devant	F1	lex	1				Sadv	toutes les (2) (fleurs)	F2	gram			2			
PR	là bas	F1	lex	1				Sadv	un type de	F2	gram			1			
PR	il n'est pas là (ne pas être là)	F1	lex	1				Sprop	il ne peut pas +V	F2	gram			1			
PR	qu'est ce qui *passe	F1	lex	1			1	Sprop	une autre	F2	gram			1			
PR	et donc (2)	F1	dis		2				je crois que +	F2	gram			1			
PR	et après (3)	F1	dis		3				c'est...qui	F2	gram			1			
PR	quelque chose comme ça	F1	dis		1				il n'y a pas	F2	gram			1			
PR	parce que (6)	F1	dis		6				c'était *le histoire de	F2	gram			1	1		
PR	peut-être (5)	F1	dis		5				c'est (une)	F2	gram			1			
PR	je crois	F1	dis		1												
PR	en fait (2)	F1	dis		2												
PR	tu peux +V	F1	dis		1												
PR	c'était + (c'était la fin de )	F1	dis		1												
PR	un peu	F1	gram				1										
PR	un peu de	F1	gram				1										
PR	il va + V	F1	gram				1										
PR	toutes les choses	F1	gram				1										
PR	le même + endroit	F1	gram				1										
PR	une* autre type de	F1	gram				1	1									
PR	il a vu que	F1	gram				1										
PR	elle était (curieuse)	F1	gram				1										
PR	la fin de	F1	gram				1										
PR	il sait que	F1	gram				1										
PR	il est allé	F1	gram				1										
PR	il a pensé que	F1	gram				1										
PR	à côté de	F1	gram				1										
PR	il y a (3)	F1	gram				3										
PR	c'est...que	F1	gram				1										
PR	les choses à (manger)	F1	gram				1										
PR	je crois que +	F1	gram				1										
PR	c'est comme	F1	gram				2										
PR	c'était une histoire de	F1	gram				1										
PR	c'était (2)	F1	gram				1										

PR	à côté de	F1	gram			1													
PR	il y a (3)	F1	gram			3													
PR	c'est...que	F1	gram			1													
PR	les choses à (manger)	F1	gram			1													
PR	je crois que +	F1	gram			1													
PR	c'est comme	F1	gram			2													
PR	c'était une histoire de	F1	gram			1													
PR	c'était (2)	F1	gram			2													
	<b>BD1</b>		<b>37</b>		<b>20</b>	<b>10</b>	<b>7</b>	<b>7</b>		<b>BD2</b>		<b>25</b>		<b>13</b>	<b>6</b>	<b>6</b>	<b>4</b>		
	<b>F1</b>		<b>74</b>		<b>28</b>	<b>22</b>	<b>24</b>	<b>9</b>		<b>F2</b>		<b>31</b>		<b>9</b>	<b>12</b>	<b>10</b>	<b>2</b>		
	<b>t1</b>		<b>111</b>		<b>48</b>	<b>32</b>	<b>31</b>	<b>16</b>		<b>t2</b>		<b>56</b>		<b>27</b>	<b>29</b>	<b>16</b>	<b>6</b>		
	<b>occurrences types</b>									<b>occurrences types</b>									
	<b>BD1</b>		<b>28</b>		<b>14</b>	<b>8</b>	<b>6</b>	<b>3</b>		<b>BD2</b>		<b>17</b>		<b>8</b>	<b>4</b>	<b>5</b>	<b>2</b>		
	<b>F1</b>		<b>46</b>		<b>17</b>	<b>9</b>	<b>20</b>	<b>4</b>		<b>F2</b>		<b>22</b>		<b>7</b>	<b>6</b>	<b>9</b>	<b>2</b>		
	<b>t1</b>		<b>74</b>		<b>31</b>	<b>17</b>	<b>26</b>	<b>7</b>		<b>t2</b>		<b>39</b>		<b>15</b>	<b>10</b>	<b>14</b>	<b>4</b>		
	<b>t1 et t2</b>		<b>167</b>		<b>75</b>	<b>61</b>	<b>47</b>	<b>22</b>											
	<b>t1 et t2</b>		<b>113</b>		<b>46</b>	<b>27</b>	<b>40</b>	<b>11</b>											

Q1																
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type svntaazme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type svntaazme
Q1	petit garçon, petit jouet, *petit sieste	BD1	lex	3			1	SN	petit jouet (3), *petit sieste	BD2	lex	4			1	SN
Q1	longue journée	BD1	lex	1				SN	après midi	BD2	lex	1				SN
Q1	s'en va (s'en aller)	BD1	lex	1				SV	s'en va (s'en aller)	BD2	lex	1				SV
Q1	se promène	BD1	lex	1				SV	se réveille	BD2	lex	1				SV
Q1	a une idée (avoir une idée)	BD1	lex	1				SV	se promène avec	BD2	lex	1				SV
Q1	il *sent fatigué	BD1	lex	1			1	SV	être fatigué	BD2	lex	1				SV
Q1	*assied dessus	BD1	lex	1			1	SV	descend de	BD2	lex	1				SV
Q1	faire une sieste	BD1	lex	1				SV	commence à	BD2	lex	1				SV
Q1	à ce moment là	BD1	lex	1				SV	faire *le sieste, faire une sieste (2)	BD2	lex	3			1	SV
Q1	un jour	BD1	lex	1				Sadv	a des idées	BD2	lex	1				SV
Q1	chez lui	BD1	lex	1				Sadv	faire une bêtise	BD2	lex	1				SV
Q1	très fatigué	BD1	lex	1				Sadj	s'approche de	BD2	lex	1				SV
Q1	il s'appelle rouky	BD1	lex	1				Sprop	dans sa tête	BD2	lex	1				Sprep
Q1	oh mon dieu	BD1	dis		1				en même temps	BD2	lex	1				Sadv
Q1	et après (2)	BD1	dis		2				à ce moment là	BD2	lex	1				Sadv
Q1	en fait	BD1	dis		1				qu'est ce qui se passe	BD2	lex	1				Sprop
Q1	un peu	BD1	gram			1			Un jour	BD2	dis		1			
Q1	être en train de	BD1	gram			1			où il est	BD2	dis		1			
Q1	n'importe où	BD1	gram			1			et puis (2)	BD2	dis		2			
Q1	de.. jusqu'à...	BD1	gram			1			comment on dit	BD2	dis		1			
Q1	il trouve que	BD1	gram			1			je peux + V	BD2	dis		1			
Q1	il est (malin, mignon, mécha	BD1	gram			3			c'est fini	BD2	dis		1			
Q1	(juste) à côté de	BD1	gram			2			tellement *de	BD2	gram			1	1	
Q1	il dit que	BD1	gram			1			être en train de	BD2	gram			1		
Q1	y a (3)	BD1	gram			3			un peu	BD2	gram			1		
Q1	c'est (3)	BD1	gram			3			je vais + V	BD2	gram			1		
									il trouve que	BD2	gram			1		
									il faut que	BD2	gram			1		
									il dit que (3)	BD2	gram			3		
									c'est (le meilleur endroit)	BD2	gram			1		
									y a (2)	BD2	gram			2		
									(à) la fin de	BD2	gram			1		
									sous forme de	BD2	gram			1		
F1								F2								
Q1	table de cuisine	F1	lex	1				SN	la dernière fois	F2	lex	1				SN
Q1	allers retours	F1	lex	1				SN	coup de vent	F2	lex	1				SN
Q1	salle de bains	F1	lex	1				SN	belle maison	F2	lex	1				SN
Q1	eau usée	F1	lex	1				SN	aller retour	F2	lex	1				SN
Q1	petite flèche	F1	lex	1				SN	salle de bains	F2	lex	1				SN
Q1	*petit histoire, *petit abille	F1	lex	2			2	SN	système de conduite	F2	lex	1				SN
Q1	*graines de café (2)	F1	lex	2				SN	fleurs sauvages	F2	lex	1				SN
Q1	porte ouverte	F1	lex	1				SN	*graine de café x 3	F2	lex	3			3	SN
Q1	se promène (3), se promène partout	F1	lex	4				SV	*petit abeille (2)	F2	lex	2			2	SN
Q1	essayer de (2)	F1	lex	2				SV	faire le tour	F2	lex	2				SV
Q1	se cache	F1	lex	1				SV	j'ai pas de langue	F2	lex	1			1	SV
Q1	a envie de	F1	lex	1				SV	*commun	F2	lex	1				SV
Q1	vole vers	F1	lex	1				SV	j'ai pas de points communs	F2	lex	1				SV
Q1	s'arrête (2)	F1	lex	2				SV	entend les bruits	F2	lex	1				SV
Q1	permettre de			1				SV	se cache (se cacher)	F2	lex	1				SV
Q1	arrive plus *de	F1	lex	1			1	SV	commence à	F2	lex	1				SV
Q1	retrouver *son liberté	F1	lex	1			1	SV	permet de	F2	lex	1				SV
Q1	là bas (2)	F1	lex	2				Sadv	évacuer de l'eau	F2	lex	1				SV
Q1	tout à l' heure	F1	lex	1				Sadv	s'en aller (4)	F2	lex	4				SV
Q1	en même temps	F1	lex	1				Sadv	s'approche de	F2	lex	1				SV
Q1	tout de suite	F1	lex	1				Sadv	parler à	F2	lex	1				SV
Q1	par hasard (2)	F1	lex	2			1	Sadv	sortir de	F2	lex	1				SV
Q1	*face en face	F1	lex	1				Sadv	a retrouvé *son liberté	F2	lex	1			1	SV
Q1	super content	F1	lex	1				Sadv	il arrive pas *de	F2	lex	1			1	SV
Q1	plusieurs fois (2)	F1	lex	2				Sprep	super content	F2	lex	1				SV
									très jolies	F2	lex	1				Sadj





VI												
ID	BD1	tâche	SP	SP	SP	type	BD2	tâche	SP	SP	type	
			lex	dis	gr	syntagme			lex	dis	syntagme	
										SP	err	
VI	p'tit garçon (2), p'tit cheval, p'tit animal	BD1	lex	4		SN	petite sieste, petit bout	BD2	lex	2	SN	
VI	est fatigué	BD1	lex	1		SV	p'tit garçon, p'tit chien	BD2	lex	2	SN	
VI	décide de	BD1	lex	1		SV	s'en va	BD2	lex	1	SV	
VI	s'endormir	BD1	lex	1		SV	se réveille	BD2	lex	1	SV	
VI	décide de	BD1	lex	1		SV	remplace N par N	BD2	lex	1	SV	
VI	se réveille	BD1	lex	1		SV	a envie de	BD2	lex	1	SV	
VI	n'est pas là	BD1	lex	1		Sprop	attacher N à N	BD2	lex	1	SV	
VI	*au lieu	BD1	lex	1	1	Sprep	décide de	BD2	lex	1	SV	
VI	c'est drôle (2)	BD1	lex	2		Sprop	faire une sieste	BD2	lex	1	SV	
VI	j' sais pas	BD1	dis	1			se sent fatigué	BD2	lex	1	SV	
VI	j' suis pas sûr	BD1	dis	1			*se transformer dans	BD2	lex	1	1 SV	
VI	d'abord	BD1	dis	1			il est tout seul	BD2	lex	1	SV	
VI	alors que (3)	BD1	dis	3			à la campagne	BD2	lex	1	Sprep	
VI	quand même	BD1	dis	1			il est là (être là)	BD2	lex	1	Sprop	
VI	un peu (3)	BD1	gram		3		(*c'que c'est) *c'est arrivé	BD2	lex	1	1	
VI	il voit que	BD1	gram	1			il comprend pas	BD2	lex	1		
VI	c'est l'histoire de	BD1	gram		1		peut-être (2)	BD2	dis		2	
VI	je pense que	BD1	gram		1		je crois (2)	BD2	dis		2	
VI	il y a	BD1	gram		1		c'est	BD2	dis		1	
VI	c'est (3)	BD1	gram		3		parce que	BD2	dis		1	
VI	ça donne l'impression que	BD1	gram		1		sans que	BD2	dis		1	
							un bout de (bois)	BD2	gram		1	
							il voit pas que	BD2	gram		1	
							il s'assure que	BD2	gram		1	
							au lieu de	BD2	gram		1	
							il a l'impression que	BD2	gram		1	
							il y a (2)	BD2	gram		2	
F1						F2						
VI	fenêtre (est) ouverte	F1	lex	1		SN	grains de café	F2	lex	1	SN	
VI	se ferme	F1	lex	1		SV	salle de bains	F2	lex	1	SN	
VI	se rencontrent	F1	lex	1		SV	la porte entrouverte	F2	lex	1	SN	
VI	sortir de	F1	lex	1		SV	commence à	F2	lex	1	SV	
VI	*fabriquer du miel	F1	lex	1	1	SV	s'envole (s'envoler)	F2	lex	1	SV	
VI	sortir par	F1	lex	1		SV	elle arrive pas à	F2	lex	1	SV	
VI	*un moment donné	F1	lex	1	1	Sadv	se bat contre (3)	F2	lex	3	SV	
VI	et après	F1	dis	1			indique la sortie	F2	lex	1	SV	
VI	je sais pas pourquoi	F1	dis	1			s'écoule	F2	lex	1	SV	
VI	alors que	F1	dis	1			essayer de (2)	F2	lex	2	SV	
VI	comme ça	F1	dis	1			se trouve (devant)	F2	lex	1	SV	
VI	et voilà	F1	dis	1			sortir de	F2	lex	1	SV	
VI	pour que(3)	F1	dis	3			se regardent (se regarder);	F2	lex	1	SV	
VI	parce que	F1	dis	1			décider de (2)	F2	lex	2	SV	
VI	elle sait pas	F1	dis	1			faire un signe (2)	F2	lex	2	SV	
VI	hors de	F1	gram		1		continuent à	F2	lex	1	SV	
VI	c'est l'histoire de	F1	gram		1		entend un bruit	F2	lex	1	SV	
VI	je pense que	F1	gram		1		se ferme (se fermer)	F2	lex	1	SV	
VI	elle va +V	F1	gram		2		n'y arrive pas	F2	lex	1	SV	
VI	pas vraiment	F1	gram		1		se promène	F2	lex	1	SV	
							elle est libre	F2	lex	1	SV	
							à l'extérieur (2)	F2	lex	2	Sprep	
							à la sortie	F2	lex	1	Sprep	
							à un moment donné	F2	lex	1	Sadv	
							*en ce moment là	F2	lex	1	1 Sadv	
							et là	F2	dis	1		
							je sais pas (2)	F2	dis	2		
							comment dire ça	F2	dis	1		
							on connaît pas	F2	dis	1		
							c'est (5)	F2	dis	5		
							peut-être	F2	dis	1		
							et après	F2	dis	1		
							va + voir	F2	gram		1	
							toutes les, toutes *le (vitres)	F2	gram	2	1	
							toutes les deux	F2	gram	1		
							un autre (endroit)	F2	gram	1		
							une plantation de	F2	gram	1		
							une façon de	F2	gram	1		
							elle pense que	F2	gram	1		
							jusqu'à	F2	gram	1		
							en dehors de	F2	gram	1		
							c'est l'histoire de	F2	gram	1		
							c' est (5)	F2	gram	5		
							c'est...*que	F2	gram	1	1	
							c'est ce que ...	F2	gram	1		
							il y a (3)	F2	gram	3		
							je crois que (2)	F2	gram	2		
							il y a ... qui	F2	gram	1		
							de l'autre côté de	F2	gram	1		
											26	
	BD1	31	13	7	11	1	BD2	32	18	7	7	2
	F1	22	6	10	6	2	F2	68	31	12	25	3
	t1	53	19	17	17	3	t2	100	49	19	32	5



							il y a ... qui	F2	gram			1	
							de l'autre côté de	F2	gram			1	
BD1	31	13	7	11	1		BD2	32	18	7	7	2	
F1	22	6	10	6	2		F2	68	31	12	25	3	
t1	53	19	17	17	3		t2	100	49	19	32	5	
<b>occurrences/types</b>							<b>occurrences/types</b>						
BD1	21	9	5	7	1		BD2	28	15	7	6	2	
F1	20	7	8	5	2		F2	49	25	7	17	3	
t1	41	16	13	12	3		t2	77	40	14	23	5	
t1 et T2	153	68	36	49	8								
t1 et T2	118	56	27	35	8								

VO																	
ID	BD1	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	BD2	tâche	SP	SP lex	SP dis	SP gr	err	type syntagme	
VO	p'tite sieste, p'tit gamin, p'tit jouet	BD1	lex	3				SN	se réveille	BD2	lex	1				SV	
VO	petite blague	BD1	lex	1				SN	se demande	BD2	lex	1				SV	
VO	fait du cheval	BD1	lex	1				SV	être transformé en	BD2	lex	1				SV	
VO	descend de	BD1	lex	1				SV	descend de	BD2	lex	1				SV	
VO	a l'air (fatigué)	BD1	lex	1				SV	a l'air fatigué	BD2	lex	1				SV	
VO	s'installe	BD1	lex	1				SV	s'installe	BD2	lex	1				SV	
VO	fait une sieste	BD1	lex	1				SV	*s'approche à	BD2	lex	1			1	SV	
VO	se demande	BD1	lex	1				SV	*en galope	BD2	lex	1			1	Sadv	
VO	s'est transformé en	BD1	lex	1				SV	qu'est ce qu'il s'est passé	BD1	lex	1				Sprop	
VO	se réveille	BD1	lex	1				SV	on voit (2)	BD2	dis		2				
VO	où il est (mon cheval)	BD1	dis		1			Sprop	et donc (2)	BD2	dis		2				
VO	je vois	BD1	dis		1				il voit que	BD1	gram			1			
VO	en fait	BD1	dis		1				une histoire de	BD1	gram			1			
VO	peut-être (2)	BD1	dis		2												
VO	ou quoi	BD1	dis		1												
VO	*sur le premier plan	BD1	dis		1		1										
VO	c'est tout	BD1	dis		1												
VO	il voit que	BD1	gram			1											
VO	au lieu de (2)	BD1	gram			2											
VO	il y a (3)	BD1	gram			3											
F1									F2								
VO	salle de bains	F1	lex	1				SN	dessin animé	F2	lex	1				SN	
VO	*grains du café	F1	lex	1			1	SN	p'tit moment	F2	lex	1				SN	
VO	se retrouve	F1	lex	1				SV	petite abeille	F2	lex	1				SN	
VO	s'appelle	F1	lex	1				SV	salle de bains (2)	F2	lex	2				SN	
VO	n'arrive pas à	F1	lex	1				SV	allers retours	F2	lex	1				SN	
VO	sortir de (2)	F1	lex	2				SV	commence par	F2	lex	1				SV	
VO	essaye de	F1	lex	1				SV	se retrouver (2)	F2	lex	2				SV	
VO	sortir par	F1	lex	1				SV	essaye de (2)	F2	lex	2				SV	
VO	s'installe	F1	lex	1				SV	se casse la gueule	F2	lex	1				SV	
VO	s'en sortir	F1	lex	1				SV	se ferme (se fermer)	F2	lex	1				SV	
VO	s'est libérée	F1	lex	1				SV	sortir par (les fenêtres)	F2	lex	1				SV	
VO	est dedans	F1	lex	1				SV	revient vers (revenir vers)	F2	lex	1				SV	
VO	se casse	F1	lex	1				SV	a indiqué le chemin	F2	lex	1				SV	
VO	se déplace	F1	lex	1				SV	se regardent	F2	lex	1				SV	
VO	trouver un moyen	F1	lex	1				SV	se retrouve	F2	lex	1				SV	
VO	s'approche, *s'approche vers	F1	lex	2			1	SV	avoir l'air (2)	F2	lex	2				SV	
VO	réessaye de	F1	lex	1				SV	s'installe sur	F2	lex	1				SV	
VO	*commence de	F1	lex	1			1	SV	sort de	F2	lex	1				SV	
VO	*a réussi de	F1	lex	1			1	SV	comprendre le problème	F2	lex	1				SV	
VO	on n'a pas peur	F1	lex	1				SV	*on t'inquiète pas	F2	lex	1			1	SV	
VO	rentre dedans	F1		1				SV	au sol	F2	lex	1				Sprep	
VO	*à l'autre côté	F1	lex	1			1	Sprep	à l'extérieur (3)	F2	lex	3				Sprep	
VO	à la campagne	F1	lex	1				Sprep	à la cuisine	F2	lex	1				Sprep	
VO	un par un	F1	lex	1				Sadv	dans les yeux	F2	lex	1				Sprep	
VO	cette fois	F1	lex	1				Sadv	d'après	F2	lex	1				Sprep	
VO	très contente *que (pour de)	F1	lex	1			1	Sadj	de l'autre côté de	F2	lex	1				Sadv	
VO	comment dire	F1	dis		1				sans problème	F2	lex	1				Sadv	
VO	et après (4)	F1	dis		4				plus tard	F2	lex	1				Sadv	
VO	et cetera	F1	dis		1				attirée par (2)	F2	lex	2				Sadj	
VO	et voilà	F1	dis		1				on sait pas encore quoi	F2	dis		1				
VO	c'est (2)	F1	dis		2				c'est impossible	F2	dis		1				
VO	et quand	F1	dis		1				on aperçoit (3)	F2	dis		3				
VO	comment ça	F1	dis		1				d'abord	F2	dis		1				
VO	on aperçoit	F1	dis		1				je rigole	F2	dis		1				
VO	on voit, on voit pas	F1	dis		2				en fait (4)	F2	dis		4				
VO	on observe	F1	dis		1				avant que	F2	dis		1				
VO	y a (une flèche)	F1	gram			1			c'est	F2	dis		1				
VO	c'est ... qui V	F1	gram			1			parce que (2)	F2	dis		2				
VO	il va + V (2)	F1	gram			2			et donc (4)	F2	dis		4				
VO	un peu	F1	gram			1			on voit (3)	F2	dis		3				
VO	un autre épisode	F1	gram			1			je veux dire	F2	dis		1				
VO	une autre N X3	F1	gram			3			peut-être	F2	dis		1				







## **ANNEXE 4**

### **Classement des SP natives**

AN													
ID	BD	tache	SP lex	SP dis	SP gr	type syntagme	Film	tache	SP lex	SP dis	SP gr	type syntagme	
AN	chapeau rond	BD	lex	1		SN	court métrage d'animation	Film	lex	1		SN	
	p'tit enfant, p'tit						p'tit court métrage,p'tite						
AN	cheval (2)	BD	lex	3		SN	abeille, p'tits grains de café,						
AN	s'installe	BD	lex	1		"	un p'tit recoin	Film	lex	4		SN	
AN	s'endort	BD	lex	1		SV	petite flèche	Film	lex	1		SN	
AN	faire une sieste	BD	lex	1		SV	grain de café (2)	Film	lex	2		SN	
AN	s'approche du	BD	lex	1		SV	pièce attenante	Film	lex	1		SN	
AN	s'assoupit	BD	lex	1		SV	salle de bains	Film	lex	1		SN	
AN	s'apercevoir	BD	lex	1		SV	indique chemin de sortie	Film	lex	1		SV	
AN	s'enfuir	BD	lex	1		SV	s'appelle	Film	lex	1		SV	
AN	se réveiller	BD	lex	1		SV	s'ouvre	Film	lex	1		SV	
AN	se rendre compte que	BD	lex	1		SV	s'approche de (2)	Film	lex	2		SV	
AN	fait un échange	BD	lex	1		SV	se rapproche de	Film	lex	1		SV	
AN	accroche N à N	BD	lex	1		SV	essaye de (2)	Film	lex	2		SV	
AN	en bois x 3	BD	lex	3		Sadv	se heurte à	Film	lex	1		SV	
AN	à l'extérieur	BD	lex	1		Sprep	arrive pas à	Film	lex	1		SV	
AN	par la main	BD	lex	1		Sprep	butiner les fleurs	Film	lex	1		SV	
AN	(jouet) pour enfant (2)	BD	lex	2		Sprep	se dirige vers	Film	lex	1		SV	
AN	la première (vignette)	BD	disc	1			se regardent	Film	lex	1		SV	
AN	on voit	BD	disc	1			essaye de	Film	lex	1		SV	
AN	peut être	BD	disc	1			se retourne	Film	lex	1		SV	
AN	et donc	BD	disc	1			continue à	Film	lex	1		SV	
AN	en fait	BD	disc	1			fonce dessus	Film	lex	1		SV	
AN	pendant que	BD	disc	1			s'écrase contre	Film	lex	1		SV	
AN	l'autre	BD	gram		1		tombe à la renverse	Film	lex	1		SV	
AN	y a	BD	gram		2		se remet à	Film	lex	1		SV	
AN	c'est	BD	gram		1		s'élève	Film	lex	1		SV	
AN	au bout de (3)	BD	gram		3		faire des allers retours	Film	lex	1		SV	
	Il sait pas où (est passé						se cacher	Film	lex	1		SV	
AN	son cheval)	BD	gram		1		en plastique	Film	lex	1		Sadv	
AN	il s'aperçoit que	BD	gram		1		à côté	Film	lex	1		Sprep	
AN	on imagine que	BD	gram		1		à l'extérieur (3)	Film	lex	3		Sprep	
							c'est pas facile de	Film	lex	1		S prop	
							ça marche pas	Film	lex	1		Sprop	
							tout ça	Film	disc		1		
							au début (2)	Film	disc		2		
							j' crois	Film	disc		1		
							j'en suis pas sûre	Film	disc		1		
							parc' que	Film	disc		1		
							et donc	Film	disc		1		
							en fait	Film	disc		1		
							et ensuite	Film	disc		1		
							on voit...qui (2)	film	disc		2		
							va + V (3)	film	gram		3		
							toutes les deux	film	gram		1		
							la fin de (film)	film	gram		1		
							c'est (4)	film	gram		4		
							du côté de	film	gram		1		
							y a (4)	film	gram		4		
							j' crois que, je crois que	film	gram		2		
							au dessus de	film	gram		1		
							de l'autre côté de	film	gram		1		
							elle est + lieu	film	gram		1		

bd	38	22	6	10
Film	70	40	11	19
total	108	62	17	29

occurrences/types				
BD	30	17	6	7
film	52	32	9	11
total	82	49	15	3

CA														
ID	BD	tache	SP	SP	SP	type	Film	tache	SP	SP	SP	SP	type	type
			lex	dis	gr	syntagme			lex	lex	dis	gr	syntagme	
CA	cinq minutes	BD	lex	1		SN	maison de campagne	film	lex	1			SN	
CA	p'tit garçon (2), p'tit ch	BD	lex	3		SN	bouquet de fleurs	film	lex	1			SN	
CA	alter ego	BD	lex	1		SN	corbeille de fruits	film	lex	1			SN	
CA	petit garçon	BD	lex	1		SN	chacune de leur côté	film	lex	1			SN	
CA	fait une pause	BD	lex	1		SV	grains de café (3), *graines de café	film	lex	4		1	SN	
CA	s'arrête	BD	lex	1		SV	p'tite abeille	film	lex	1			SN	
CA	s'est enfui	BD	lex	1		SV	au final	film	lex	1			SN	
CA	s'assoupit	BD	lex	1		SV	salle de bains	film	lex	1			SN	
CA	se réveiller	BD	lex	1		SV	petit poucet	film	lex	1			SN	
CA	en bois x 3	BD	lex	3		Sadv	chemin de la liberté	film	lex	1			SN	
CA	et donc	BD	disc		1		table du salon	film	lex	1			SN	
CA	un je ne sais quel	BD	disc		1		conduit de canalisation	film	lex	1			SN	
CA	et là (2)	BD	disc		2		porte ouverte	film	lex	1			SN	
CA	peut être	BD	disc		1		gros yeux	film	lex	2			SN	
CA	en fait	BD	disc		1		se passer	film	lex	1			SV	
CA	nous ne savons pas	BD	disc		1		décide de	film	lex	1			SV	
CA	on peut (2)	BD	disc		2		se met à (4)	film	lex	4			SV	
CA	va + V	BD	gram		1		est dans son élément	film	lex	1			SV	
CA	à la place de	BD	gram		1		se balade	film	lex	1			SV	
CA	il s'agit de	BD	gram		1		butiner les fleurs	film	lex	1			SV	
CA	y a	BD	gram		1		se retrouve	film	lex	2			SV	
							se cogne, se cogne contre	film	lex	2			SV	
							essaie de	film	lex	1			SV	
							s'échapper	film	lex	1			SV	
							y va (y aller)	film	lex	1			SV	
							se place en observateur	film	lex	1			SV	
							retrouver son chemin	film	lex	1			SV	
							s'en sortir	film	lex	1			SV	
							sortir de (la maison)	film	lex	1			SV	
							faire la toupie	film	lex	1			SV	
							pris au piège	film	lex	1			Sadv	
							tout d'un coup (3)	film	lex	3			Sadv	
							à plusieurs reprises	film	lex	1			Sprep	
							de l'extérieur	film	lex	1			Sprep	
							à l'intérieur	film	lex	1			Sprep	
							d'abord	film	disc		1			
							et ensuite	film	disc		1			
							et après	film	disc		1			
							et puis (2)	film	disc		2			
							en fait (3)	film	disc		3			
							on ne sait où	film	disc		1			
							on se demande	film	disc		1			
							on va dire	film	disc		1			
							et donc	film	disc		1			
							on sait pas, on sait pas si (4)	film	disc		4			
							et là	film	disc		1			
							va + V (4)	film	gram			4		
							un peu (3)	film	gram			3		
							très vite (2)	film	gram			2		
							au milieu de	film	gram			1		
							de façon à	film	gram			1		
							grâce à	film	gram			1		
							c'est (3)	film	gram			3		
							c'est l'histoire de	film	gram			1		
							c'était	film	gram			1		
							y a pas (2)	film	gram			2		

BD	27	14	9	4
FILM	82	46	17	19
<b>total</b>	<b>109</b>	<b>60</b>	<b>26</b>	<b>23</b>

occurrences/types				
BD	21	10	7	4
FILM	56	35	11	10
<b>total</b>	<b>77</b>	<b>45</b>	<b>18</b>	<b>14</b>



JE													
ID	BD	tache	SP	SP	SP	type	Film	tache	SP	SP	SP	SP	type
			lex	dis	gr	syntagme			lex	dis	gr	syntagme	
		petit enfant (2) petit											
JE	cheval	BD	lex	3		SN	petite abeille	film	lex	1			SN
JE	s'assoupit	BD	lex	1		SV	grains de café (2)	film	lex	2			SN
JE	se reposer (2)	BD	lex	2		SV	salle de bains (2)	film	lex	2			SN
JE	décide de	BD	lex	1		SV	faire des allers retours	film	lex	1			SV
JE	en bois	BD	lex	1		Sadv	allers retours	film	lex	1			SN
JE	on peut (voir)	BD	disc		1		air satisfait	film	lex	1			SN
JE	pendant que (2)	BD	disc		2		un petit moment	film	lex	1			SN
JE	être en train de	BD	gram		1		se balade	film	lex	1			SV
							se claque	film	lex	1			SV
							essaye de	film	lex	1			SV
							sortir de	film	lex	1			SV
							se frappe, se frappe						
							contre	film	lex	2			SV
							arrive pas à (2), arrive						
							plus à	film	lex	3			SV
							commence à	film	lex	1			SV
							s'arrête	film	lex	1			SV
							a l'air	film	lex	1			SV
							s'observent	film	lex	1			SV
							décide de	film	lex	1			SV
							*se ressemble						
							(étrangement) à	film	lex	1			SV
							continuer à (2)	film	lex	2			SV
							se confond avec	film	lex	1			SV
							entend un bruit	film	lex	1			SV
							continue de	film	lex	1			SV
							se cacher	film	lex	1			SV
							donnent sur	film	lex	1			SV
							faire le tour de (la						
							maison)	film	lex	1			SV
							prendre la direction de	film	lex	1			SV
							à l'air libre	film	lex	2			Sprep
							au passage	film	lex	1			Sprep
							pendant ce temps là			1			Sadv
							il se passe grand chose	film	lex	1			Sprop
							sans que	film	disc		1		
							et là	film	disc		1		
							j' crois	film	disc		1		
							du coup (3)	film	disc		3		
							étant donné que	film	disc		1		
							sans que	film	disc		1		
							un peu	film	gram			1	
							toutes les deux	film	gram			1	
							c'est l'histoire de	film	gram			1	
							y a, y a (plus)	film	gram			2	
							de l'autre côté de	film	gram			1	
							au bout de (2)	film	gram			2	
							en forme de	film	gram			1	
							autour de	film	gram			1	

BD	12	8	3	1
Film	56	38	8	10
total	68	46	11	

**occurrences/types**

FILM	46	31	6	9
BD	8	5	2	1
Total	54	36	8	10

MG													
ID	BD	tache	SP lex	SP dis	SP gr	type syntagme	Film	tache	SP lex	SP dis	SP gr	type syntagme	
MG	long voyage	BD	lex	1		SN	un p'tit film (2)	film	lex	2		SN	
MG	faire un p'tit somme	BD	lex	1		SV	film muet	film	lex	1		SN	
MG	p'tit garçon (2)	BD	lex	2		SN	maya l'abeille	film	lex	1		SN	
MG	chapeau de cowboy	BD	lex	1		SN	habitat naturel	film	lex	1		SN	
MG	sont fatigués, est fatigué	BD	lex	2		SV	porte de sortie	film	lex	1		SN	
MG	décide de (3)	BD	lex	3		SV	habitat humain	film	lex	1		SN	
MG	s'arrêter	BD	lex	1		SV	grains de café	film	lex	2		SN	
MG	faire une pause	BD	lex	1		SV	aller retours	film	lex	1		SN	
MG	s'enfuit	BD	lex	1		SV	yeux globuleux	film	lex	1		SN	
MG	s'assoient	BD	lex	1		SV	se fait plaisir	film	lex	1		SV	
MG	reprend sa liberté	BD	lex	1		SV	ressemble à	film	lex	1		SV	
MG	se réveille	BD	lex	1		SV	se retrouve	film	lex	1		SV	
MG	s'approche	BD	lex	1		SV	essaye de	film	lex	1		SV	
MG	se reposer	BD	lex	1		SV	se rend compte (2)	film	lex	2		SV	
MG	file à l'anglaise	BD	lex	1		SV	se cogne contre	film	lex	1		SV	
MG	(cheval) à roulettes (3)	BD	lex	3		Sprep	trouve pas la sortie	film	lex	1		SV	
MG	accompagné de	BD	lex	1		Sadj	arrive pas à	film	lex	1		SV	
MG	c'est drôle	BD	lex	1		Sprop	trouver la solution	film	lex	1		SV	
	il ne comprend pas trop												
MG	ce qui se passe	BD	lex	1		Sprop	reconnait qqn (la mouche)	film	lex	1		SV	
MG	y a de la vie	BD	lex	1			s'regardent	film	lex	1		SV	
MG	y a de	BD	disc		1		observe le manège	film	lex	1		SV	
MG	il sait pas trop	BD	disc		1		a l'air	film	lex	1		SV	
MG	c'est	BD	disc		1		prend ses précautions	film	lex	1		SV	
MG	la première (ville)	BD	disc		1		se cache	film	lex	1		SV	
MG	comme ça	BD	disc		1		finit par	film	lex	1		SV	
MG	du coup	BD	disc		1		sortir de (2)	film	lex	2		SV	
MG	un peu (3)	BD	gram			3	fait confiance à	film	lex	1		SV	
MG	bien plus	BD	gram			1	décide de	film	lex	1		SV	
MG	c'est l'histoire de	BD	gram			1	retrouve sa liberté	film	lex	1		SV	
MG	c'est (3)	BD	gram			3	ressort par	film	lex	1		SV	
MG	y a..qui (2)	BD	gram			2	va droit au but	film	lex	1		SV	
MG	sans + V + de	BD	gram			1	tourne pas en rond	film	lex	1		SV	
MG	peut + se reposer	BD	gram			1	se poser des questions	film	lex	1		SV	
							habituee à	film	lex	1		Sadj	
							digne de	film	lex	1		Sadj	
							en plastique	film	lex	1		Sadv	
							en tout cas	film	lex	2		Sadv	
							par inadvertance	film	lex	1		Sadv	
							une fois que	film	lex	1		Sadv	
							à facettes	film	lex	1		Sadv	
							(très) à l'aise	film	lex	1		Sadv	
							à l'extérieur	film	lex	1		Sprep	
							le film commence	film	disc		1		
							on a tous connu ça	film	disc		1		
							c'est c'est	film	disc		1		
							on peut	film	disc		1		
							ça semble	film	disc		1		
							et puis	film	disc		1		
							c'est (3)	film	disc		3		
							en tant que	film	disc		1		
							en fait (3)	film	disc		3		
							quand même	film	disc		1		
							p' être	film	disc		1		
							on se doute	film	disc		1		
							on sait pas	film	disc		1		
							du coup (6)	film	disc		6		
							alors que	film	disc		1		
							et voilà	film	disc		1		
							un peu (4)	film	gram			4	
							Être en train de	film	gram			1	
							va + atterrir	film	gram			4	

BD	44	26	6	12
film	104	47	25	32
total	104	47	25	32

occurrences/types				
Film	77	42	16	19
BD	33	20	6	7
total	110	62	22	26

du coup (6)	film	disc	6
alors que	film	disc	1
et voilà	film	disc	1
un peu (4)	film	gram	4
Être en train de	film	gram	1
va + atterrir	film	gram	4
pas trop	film	gram	1
pas vraiment (2)	film	gram	1
Sans + V + trop de	film	gram	1
autour de	film	gram	1
en dehors de	film	gram	1
est ce que (3)	film	gram	3
il se trouve que	film	gram	1
c'est vrai que	film	gram	1
(le résultat) c'est que	film	gram	1
c'est + GN	film	gram	5
on voit...qui	film	gram	1
on peut pas	film	gram	1
dire que	film	gram	1
y a	film	gram	1
j' pense que	film	gram	1
à travers (2)	film		2

TH														
ID	BD	tache	SP	SP	SP	type	Film	tache	SP	SP	SP	type		
			lex	dis	gr	syntagme			lex	dis	gr	syntagme		
TH	petit garçon (2)	BD	lex	2		SN	grains de café	film	lex	1		SN		
TH	s'assoit	BD	lex	1		SV	salle de bains	film	lex	1		SN		
TH	s'endort	BD	lex	1		SV	fleur à butiner			1		SV		
TH	attache à	BD	lex	1		SV	est en danger	film	lex	1		SV		
TH	s'enfuir	BD	lex	1		SV	cherche à	film	lex	1		SV		
TH	a l'air (2)	BD	lex	2		SV	s'enfuir	film	lex	1		SV		
TH	avoir de l'imagination	BD	lex	1		SV	s'ensuit	film	lex	1		SV		
TH	s'amène avec	BD	lex	1		SV	indiquer le chemin	film	lex	1		SV		
TH	a du mal à	BD	lex	1		SV	entendre un bruit	film	lex	1		SV		
TH	laisser N à N	BD	lex	1		SV	se trouve	film	lex	1		SV		
TH	(Arrive) en ville	BD	lex	1		SV	trouve moyen de	film	lex	1		SV		
TH	en retour	BD	lex	1		Sprep	par inadvertance	film	lex	1		Sprep		
TH	à la place	BD	lex	1		Sprep	à la recherche de	film	lex	1		Sprep		
TH	ça se limite à ça	BD	lex	1		Sprop	à l'extérieur	film	lex	1		Sprep		
							de manière							
TH	parce que	BD	disc		1		ingénieuse	film	lex	1		Sadv		
TH	j'ai cru comprendre	BD	disc		1		incapable de	film	lex	1		Sadj		
TH	beaucoup de	BD	gram		1		on a	film	disc		1			
TH	il pourra pas + V	BD	gram		1		on le voit bien	film	disc		1			
							apres avoir +							
TH	c'est ... qui	BD	gram		1		ppassé	film	gram		1			
TH	j'avoue que (2)	BD	gram		2		il comprend que	film	gram		1			
							(Le problème) c'est	film	gram		1			

BD	23	16	2	5
film	20	15	2	3
total	43	31	4	8

#### occurrences types

film	21	16	2	3
bd	20	14	2	4
total	41	30	4	7

VC															
ID	BD	tache	SP	SP	SP	type	Film	tache	SP	SP	SP	SP	type		
			lex	dis	gr	syntagme			lex	dis	gr	gr	syntagme		
VC	un p'tit (garçon)	bd	lex	1		SN	le rebord de la fenêtre (2)	film	lex	2			SN		
VC	commence à (2)	bd	lex	2		SV	*graines de cacao	film	lex	2			SN		
VC	s'asseoir sur	bd	lex	1		SV	fait allers retours	film	lex	1			SV		
VC	décide de	bd	lex	1		SV	micro ondes	film	lex	1			SN		
VC	arrive en ville	bd	lex	1		SV	salle de bains	film	lex	1			SN		
VC	tout simplement	bd	lex	1		Sadv	se retrouve à +Vinf	film	lex	1			SV		
VC	la première	bd	disc		1		commence à	film	lex	1			SV		
VC	la troisième	bd	disc		1		entend un bruit	film	lex	1			SV		
VC	la cinquième	bd	disc		1		se ferme	film	lex	1			SV		
VC	on le voit	bd	disc		1		se retrouve + p.passé	film	lex	1			SV		
VC	pour pas que	bd	disc		1		essaie de (2)	film	lex	2			SV		
VC	y a..qui	bd	gram			1	se font des signes	film	lex	1			SV		
VC	ça raconte l'histoire c	bd	gram			1	continue à	film	lex	1			SV		
VC	y a	bd	gram			1	se diriger vers	film	lex	1			SV		
VC	on (le) voit...qui	bd	gram			1	se cacher	film	lex	1			SV		
							décide de (2)	film	lex	2			SV		
							penser à	film	lex	1			SV		
							passer à	film	lex	1			SV		
							se poser sur	film	lex	1			SV		
							plusieurs fois	film	lex	1			Sadv		
							au final (2)	film	lex	2			Sadv		
							au bout d'un moment	film	lex	1			Sprep		
							à un moment	film	lex	1			Sprep		
							de l'autre côté	film	lex	1			Sprep		
							à l'extérieur	film	lex	1			Sprep		
							ça ne marche pas	film	lex	2			Sprop		
							en fait (2)	film	disc		2				
							au début	film	disc		1				
							et ensuite	film	disc		1				
							en premier	film	disc		1				
							on va le dire	film	disc		1				
							on sait pas encore où	film	disc		1				
							comme ça	film	disc		1				
							et donc (3)	film	disc		3				
							on voit	film	disc		1				
							et là	film	disc		1				
							on suppose	film	disc		1				
							en train de	film	gram			1			
							une autre (pronom)	film	gram			1			
							les autres (fleurs)	film	gram			1			
							la même (fenêtre)	film	gram			1			
							sans trop	film	gram			1			
							non plus	film	gram			1			
							de N en N	film	gram			1			
							un peu	film	gram			1			
							avant de	film	gram			1			
							y a (une mouche) qui	film	gram			1			
							sauf que	film	gram			1			
							c'est (la mouche) qu',								
							c'est (une porte) qui	film	gram			2			
							c'est l'histoire de	film	gram			1			
							à l'intérieur de	film	gram			1			
							je sais plus (lequel)	film	gram			1			

bd	16	7	5	4
film	62	32	14	16
total	78	39	19	20

occurrences/types				
film	52	26	11	15
bd	13	6	5	4
total	65	32	16	19



## ANNEXE 5

## Narrations non-natives

AK (niv 1)				
Seg	BD1	BD2	Film1	Film2
1	ok (GH))	ah il y avait un un monsieur avec euh	e et d' accord c' est c' est un film c' est hum c' est amusant mais	ah il y' avait un: beille* qui
2	il y a un _ homme	cheval euh avec euh &che cheval	comment	qui était en train de cherche:: (GH))
3	homme avec euh cheval euh	cheval	il il y atait* le	le &f le miel peut être
4	et après	et euh:	un /insecte un insecte	par euh par fleur par des fleurs
5	j' ai il a:	et il descend de cheval e::t il garait euh cheval euh hein::	euh et qui cherche: de:: qui cherche des fleurs	et eu::h (GH)
6	il a descendre* descendre:*	avec le: comment dire	euh avec avec bon* /odeur	après il
7	il a descende: *en cheval	cadre	bonne odeur	et e::t il est rentré au
8	e::t	et après eu:h	et::	à la maison de: humain (GH)
9	il a:	il a sommeil	et après	parce que dans la dans la maison il y a des quelques fleurs
10	omment il a: il a *été: un peu sommeil	et il a sommeil et il i	e::t il	et eu::h (GH)[&=bche]
11	et après	et enfant il a:: il a *dor il a dormi	il eu:::h il entrait il *a rentré le: à:: la maison de quelque chose quelqu'un	et soudain: la fenêtre il a fermée
12	et il est	et (GH))	et euh	donc euh il (x) il euh il ne
13	il es:t	ce dont ce don::t le petit enfant il (GH))	mais le cherch euh le contenu cherche le:: fleur	il il n'a pas pu euh euh sortir de: de la maison

14	il a dormi	il est venu ici il est il est venu à	avec euh bon* odeur	et donc euh il a trouvé que
15	et quand il a dormi eu:h (GH))	à à lui	et après	il a trouvé que il ne
16	il y a *autre garçon	et avec euh	e::t je sais pas pourquoi mais	il ne peut pas il ne il ne peut pas euh sortir
17	venir euh	avec euh *petit maquette de cheval *petit maquette	à la maison fermé la fenêtre	donc il cherche le:: le:: le sortir*
18	venir (GH))	*petit maquette de cheval	e:t	[&=bche] sortie (GH)
19	près de	et::	i:: il peut pas trouver le: sortie	par contre:: de tous les fenêtres:: tout la porte:
20	près de::	de des enfants il (GH) &=RG	de à la maison	tous les f protes il est
21	un _ homme	il a vu euh il a vu euh monsieur*	mais eu::h comment	toutes les &f toutes les frottes protes ils_étaient fermées
22	e::t	e:t à et à cheval:	et donc euh il cherche: comment comment commen::t sortie euh à la maison:	donc euh il ne il *n' a pu pas sortir
23	la garçon le la garçon a le garçon euh comment	e :t	et après (GH	et::
24	ah quitte quitté quitté euh quitté le cheval et &ap	je je pense que i*	il trouvait euh bon fleur	et il a trouvé il a trouvé un monsieur (GH)
25	et après ( (GH)	de petits enfants il il euh i i il a pensé que:	mai::s	de dors dors de::
26	la g le garçon eu:h ( (GH)	comment dire	mais je sais pas	*dors
27	comment connecté le connecté le	il veut il veut euh	comment	eteu:h
28	le cheval plus plus petit	un cheval réel	et après e:uh de	et::: il
29	e:t	ré réelle*	il y a autre:	il euh rencontrait euh en juste en face: (GH)
30	e::t &ap	et après eu::h	autre insecte (GH)	comment dire
31	e:::t et pis (GH))	hum le petit enfant il veut échanger le	c' est peut être monsieur mons monsieur*	mais



32	et le* homme	petit *petit maquette de cheval et cheval euh un cheval &re réel	mais je sais pas mons peut être monsieur	mais entre entre les deux il
33	euh &l la le homme:	et après euh comment dit (GH)	non (GH)) euh	il y a deux (GH)
34	se lève	*comment dit le &che un un cheval réel il sort	monsieur mais	fenêtres
35	et j' ai il a connu il a connu euh	e::t *petits enfants ils	des:::: et après:s (GH))	glaces
36	y a pas le il y a pas y a pas le::	il est sort aussi	le monsieur entrain dans la maison je sais pas pourquoi mais	donceu:::h
37	courage grand cheval et juste:	et après euh soudain eu:h comment dire	il eu:h il va faire il ah il a fai::t quelque	il a abandonné à ce monsieur
38	le petit cheval et	le monsieur*	quelque chose	et après il va &sor il va &s il va: &com il a commencé à: trouver
39	là bas et	il peut être cowboy	mai::s la insecte	trouver à *sortie
40	et il a:	cowboy et (GH)	je sais pas	à sortir (GH)
41	comment	il a recommencé juste:	ah ah il il sait pa:s comment faire comment il fait il a fait	et après (GH)
42	il a: comment il a:: (GH))	il avec euh *petit maquette de cheval e:t	e::t	après euh après j 'ai oublié
43	il a:	c' est c' est	mais après:s	et après euh ce monsieur ce monsieur ce qu' ce que ce qu' il a il vu il rentrait il est rentré à la maison
44	et il a tant pis tant pis non il a tant pis non il est	&pro produit un peu triste peut être (GH)	e::uh de insecte	donc eu::h de: le *beille:
45	il es:t il est désolé il est navré	oui [&=rires]	la un un insecte: il il cherche: dans la dans la maison	il a trouvé que il y a le sortir
46	oui comme ça (GH)	c' est marrant	e::t	il y a le sortie
47	oui		il trouvait le::	donc euh il est &ch

48	c' est fini non		comment	il est en train de cherche
49			le::	euh le sortie
50			le:: (GH)	mai::s
51			je sais pas	ah oui
52			qu' est-ce que qu' est-ce que c' est	donc eu:h
53			euh qu' est-ce que c' est	ileu::h
54			qui s' appelle:	comment dit il s' est eu:h suivi le suivi le: le: le monsieur
55			il essaie de de de	et après il a trouvé le:
56			c' es:t	ah il a
57			la pl la place de:	il a ren il est rentré eu::h (GH)
58			d' une personne qui:: lassait qui lassait qui la	une salle de bains
59			qui nettoyage &=rire	eteu:h
60			e::t	il y a de: comment dire
61			et il trouvait le saigne* signe	figure
62			signe symbole si signe	comme ça avec de avec de *pois de café
63			qui::	des pois de café comme ça:
64			qui qui m' indiquait le:: qui	mais mais au fait
65			s' indiquait le: sortie	donc eu::h euh
66			et après il est trouvé le sortie dans le: (GH)	c' es::t une salle de bains il y a comment dire il a &tro &tro &tro trouvé
67			dans l' eau l' eau l' *os l' ossiu	comment dire
68			dans l' ossieu je sais dans dans chemin le [&=rires] et après il trouvait le sortie e:t de il il est au lieu que*	trou:vé

69			de trouver euh	le::
70			que il a trouvé le:: sortie e:t (GH)	comment dire c' est& trou *trofé *troufé
71			comment	e:::t il est rentré eu::h ce trophée
72			c' est comme ça (GH)	et: finalement il est
73			et	il est & sor
74			donc euh il il a il a	il est sorti sur euh le:: comment dire
75			il a pu pas il est il a pu pas il n' a pu pa::s trouver le bon fleur je pense	trophée trouvé
76			dans le film dans *la film (GH)	trophée
77			et mais c' est amusant film je pense	trophée eu::h
78				et finalement il *a sorti à la maison
79				et:::
80				et::
81				il est il est heureux heureux peut être il était heureux [&=rires] oh mince *c'est pas trop promené au fait

FA (niv 3)				
Seg	BD1	BD2	Film1	Film2
1	l' histoire es:t divisée en deux parties la première parti:e dans laquelle on voit seulement un seul protagoniste qui est le chevalier	voilà donc euh c' est une série euh comme on a déjà vu l'année dernière on a différentes scènes	là quand je viens de voir le le film [ [&=bche] ]	le film commence par une abeille on voit une abeille qui vole: à l' extérieur donc euh on: du coté des arbres des fleurs et
2	un cowboy qui: qui descend de son cheval il laisse le cheval euh juste à côté il sent fatigué donc euh il s' assoit sur une: petit *banc	et dans une: *contexte un peu: rural donc eu::h on voit une:	e::t disons qu' au début j' étais un peu: curieux de savoir euh le pourquoi de de ce titre	et juste après on commence à voir une maison donc une maison de campagne

3	il commence à: à se reposer à dormir	un chevalier	prisonnière	et:: très jolie maison et la l' abeille aussi regarde cette maison e:t curieuse commence à rentrer à rentrer dedans
4	la deuxième parti:e c' est un peu différent parce que: [&=bche] elle est elle est basée sur la présence de deux deux protagonistes	un cowboy en anglais je sais	parce que bon on voyait le l' abeille qui::	et donc eu::h on voit le vol de l' abeille qui nous montre aussi le::: l' intérieur de la maison
5	parce qu' il va s' ajouter une* petit enfant	qui: qui: semble fatigué et il arrive: à coté d' une maison on peut imaginer que ça soit sa maison	qui bougeait qui qui volait libre	avec des meubles ca classiques et cetera [&=bche]
6	qui: qui arrive dans l' histoire qui *prend part d' un: partie dans l' histoire avec son petit cheval	e::t	et donc euh on a liberté	et:: justement l' abeille es:t surprise de voir eu::h des des fleurs sur une sur une table
7	c' est comme *une petit jouet	il: descend de:: son cheval et: il va l' attacher à: à la maison par une corde	et de l' autre côté prisonnière pourquoi ça	elle commence à se rapprocher à* &sent à chercher le parfum de ces fleurs
8	e:t comme ça il est reste étonné devant le grand cheval du cowboy et il change son cheval avec le cheval du cowboy	donc euh il &s il est fatigué il s' assoit sur une chaise	et après cette abeille qui rentre dans dans la maison e:t qui reste fermée dans la maison et commence un peu à découvrir euh *tous les choses mais après	et commen et: juste après elle
9	et repart avec euh le le vrai cheval le grand cheval	il commence à: dormir il *s' endorme	et: l'abeille a a peur de ça parce que elle n' arrive pas à sortir e:t se retrouve enfin comme: vis à vis entre guillemets avec euh	elle vole à l' intérieur de la maison
10	comme ça le le chevalier quand quand se lève quand se réveille	et juste après un petit garçon avec un petit cheval euh	une une insecte je pense une mouche qui se retrouve à l' extérieur donc euh les deux sont séparées que par u:ne /K/ par une vitre	elle était rentrée par une fenêtre qu' elle était que c' était qui était un peu ouverte
11	e:t il est il reste un peu étonné de euh parce qu' il reste juste en face du petit cheval qui était le jouet avant de: (GH)	qui qu'il tient par une corde aussi qui le *suive	e::t d' une fenêtre [&=bche]	mais après pa::r le vent ferme la fenêtre et et l' abeille se retrouve: enfermée dans la dans la maison

12	du &pet	il voit le chevalier il est un peu un peu peut être un peu impressionné par un: par le: le: &gr le gros cheval qui: qui est là	et um [&=bche]	et donc euh justement elle commence à avoir peur elle commence à bouger partout à voir s'il y a une sortie et cetera
13	du petit enfant	donc euh il détachE le gros cheval et il attache le petit cheval donc euh	ils n'arrivent pas à communiquer enfin voilà sont d'espèces différentes	et elle se retrouve devant une fenêtre et: juste devant elle il y a une une mouche qui est à l' extérieur justement
14	donc eu::h il y a une* changement de d' animau e:t [&=bche]	il échange son cheval avec* le cheval du chevalier	et la: la mouche eu::h enfin:	et qui et les deux se regardent*
15	et c' est c' est drôle la conclusion et	et le chevalier et on les voi:t s'éloigner: à distance donc euh	elle rentre dans la maison on sait pas comment	et:::
16	c' est tout	et juste après le chevalier il euh se réveille et il est étonné de voir euh que à la place du gros cheval il a: que ce petit le petit cheval mais il ne comprenne pa:s pourquoi il y a	et commence à: à prendre des: des grains de café	et juste après la:: l' abeille commence à: peut être à avoir faim donc euh elle se rapproche *aux fleurs qui étaient sur la table mais les fleurs euh
17		peut être on peut imaginer qu'il comprend pas pourquoi il y a	qu' il qu' il retrouve: sur une sur une meuble	y sont y sont pas des vraies fleurs y sont des fleurs en plastique on les voit par euh
18		y a eu cet échange	et comment on voit seulement la: la mouche qui prend les les grains de café qui les:	la présence du bruit juste::
19		et voilà e:t donc euh il a obligé: à::	et qui vole avec	juste à côté
20		à rester à côté de de ce petit cheval de	et sans savoir la direction[&=bche]	et:: et donc eu:::h elle se retourne l' abeille elle voit une:: une peinture de:: cadre avec euh (GH)
21		qui n'est pas *le sienne	et une fois qu' elle a terminé e::uh l' abeille commence à: être curieuse aussi donc elle bouge elle voit elle va dans la salle de bains elle voit la	[&=bche] des tournesols
22			dans la baignoire la présence de ces grains qui qui montrent aussi une espèce une espèce de de flèche	et donc elle commence à se rapprocher comme si si elle voyait de vraies fleurs
23			donc euh qui von:t flécher le chemin pour retrouver la la liberté	mai:s justement les fleurs aussi ne sont pas des vraies fleurs parce que c' est qu' une peinture

24			donc l' abeille:	et donc après la: la mouche que la l' abeille avait avait vue euh justement après euh avant
25			passé à travers eu:::h ce chemin qui: qui mène: à l'extérieur à travers un robinet et donc euh elle retrouve: enfin la liberté	e:t elle rentre dans la maison
26			et en fonction du du film euh caractérisé pa:r des couleurs euh très: très jolies e::t surtout la présence des fleurs beaucoup de fleurs soit à l'intérieur soit à l'extérieur	et:: l' abeille a peur donc elle se *refuge: à côté du:: du tableau
27			e::t on a pas de son c' est	et:: et la et la mouche commence à bouger en continu et commence à prendre des des petits grains de café et les déplacer
28			pas de communication que le le son de [&=bche] les bruits de la:	et la l' abeille ne bouge pas elle reste justemen:t elle se *refuge elle se *réfuge sur un sur le tableau
29			que l' abeille fait et que la mouche fon:t	et quand elle entend plus rien euh e::lle commence à bouger et et elle rentre dans dans la salle de bains
30			en volant	elle voit une baignoire avec euh justement sur le fond de la baignoire on a une: une flèche qui a été réalisée a été formée par ces grains de café
31				donc eu::h là on comprend que la que la mouche avait for avait créé ça pour un [&=bche]
32				montrer la sortie à à l' abeille
33				donc l' abeille prend la sortie par le tuyau de la: de la baignoire
34				e:t on la voit sortir à l' extérieur par un: un robinet

35				et donc euh voilà la:: la prison la prisonnière qui:: qui est libre et qui est libéré par un:: par la mouche
----	--	--	--	--------------------------------------------------------------------------------------------------------------

HE (niv 2)				
Seg	BD1	BD2	Film1	Film2
1	c' est c' est un ombre	bon eu::h	bon c' es:t	oui il y a une abeille qui se balade sur (GH)
2	sur un cheval	il il y a un: un: un homme sur un cheval	le deuxième film c' est l' histoire	le jardin (GH)
3	e::t	eu::h il arrive ç ça ressemble *d' être un: un:	d' une abeille (GH)	après elle regarde et
4	il arrête	un ancien::	elle étai:t dans le jardin (GH)	elle voit une maison elle entre dans la maison
5	e::t il est fatigué	un ça a l' air d' être::	en train de voler de chercher:	e:t à chaque fleur elle essaye de:: (GH)
6	et il essaye	des des année:s soixante peut-être à cause de il y a un petit xxx dans une maison: l'endroit où où il garde son cheval ( (GH)	une une fleur	de trouver [xxx] peut être le pollen ou en fait
7	pou::r	seulement arrivé à:	e::t elle a cherché sur le jardin et elle a vu une maison avec la porte ouverte	e::t après (GH)
8	pour rester un petit peu	à: aux aux années soixante ou je sais pas quoi	euh donc elle a* entré	elle regarde un cadre
9	pour récupérer l' énergie	bon	à la maison [&=bche]	avec les fleurs dessinées
10	e::t	après il il il a l' air d' avoir un peu de sommeil (GH)	e::t dans la maison il y a eu:: un::	elle va sur la fleur mais
11	un enfant	il s' assoit et il dort	d' autres d' autres fleurs aussi (abandon)	en fait c 'est c' est juste un cadre c' est pas
12	vient avec un petit (GH)	et après arrive un petit garçon avec un::	e::t	c' est juste un dessin

13	cheval (GH)	un petit cheval	après elle a cherché des fleurs elle elle a essayé de de sortir	après il y a un: des fleurs mais sont pas des vraies fleurs sont des f:aux
14	e::t	un un jouet (GH)	de la maison	fleurs (GH)
15	il il change	donc	par la fenêtre eu::h	il y a même le le prix marqué sur la fleur
16	le le chev le cheval	le garçon: (GH)	mai:s chaque fois qu' elle a essayé de sortir elle frappait (GH)	après elle essaie de sortir de la maison
17	de de l' homme	change le: le cheval	à une vitre de la fenêtre parce qu' elle* croyait qu'	mais elle arrive pas il y a:: des fenêtres
18	pour	du monsieu:r avec l son peti:t euh cheval	il était	en verre elle elle voit pas le verre donc elle se
19	un un le cheval	e::t i il laisse le cheval euh libre	bien là-bas	frappe contre le verre
20	le le petit che cheval	et aprè::s le monsieur se réveille	e:t elle a frappé: sur plusieurs des fenêtres	e:t après il y a un mouche dehors la maison (GH)
21	et et quand l' homme*	et quand il regarde la la situation avant il a eu son grand cheval et là il y a un petit il vient* comprendre donc euh	e::t et après il y a eu:	la mouche entre dans la maison
22	euh se réveille	il a l'air un peu:::	un:::	et l' abeille s:e cache un peu
23	e:t il pas compris qu' est ce qui s' est passé	per perdu	une un	la mouche va à la cuisine cherche des grains de café
24	parce que:	à cause de la situation	un &m &mous moustiche eu::h	va *à les toilettes tu fais ses la allers retours plusieurs fois
25	le le che le cheval a	voilà ça y est parce que c' est ça	dehors (GH)	et après:: et après le::: (GH)
26	a:: (GH)		donc	l' abeille va aux toilettes pour regarder qu' est ce qui s' est passé et elle elle voit sur dans la baignoire
27	a a a		c' es::t	une flèche et elle suivie la flèche
28	* être remplacé (GH)		c' est un ce &mous moustiche	et après elle sort dans le robinet à l' extérieur de la maison
29	le en fait le cheval devient petit		a a regardé tout ce qui s' est passé là dedans	et la mouche reste dans la maison



30	et il pas compris qu' est ce qui s' est passé		et:::	voilà c' est ça [&=rires]
31			e::t après ça la l' abeille	
32	bon (GH)		a vu une	
33	c' est ça		une fleur	
34			qui c' était faux	
35			elle savait pas que c' était faux	
36			et	
37			e::t et après elle a vu une cadre il y a c' était peinte une (GH)	
38			une fleur aussi elle est elle a:: cherché sur le cadre	
39			et: après lé lé moustiche a entré sur la maison	
40			e:t a cherché du:	
41			café	
42			e::t e:::t	
43			le le mousstiche a passé plusieurs fois	
44			devant l' abeille	
45			et::: et après l' abeille	
46			est allée chercher *qu' est ce qu' elle faisait	
47			et elle a trouvé un une baignoire baignoire	
48			une baignoire (GH)	
49			où ét où était une une direction	
50			pointée	
51			donc elle a entré*	
52			et elle a réussi à sortir de la maison	

53			par la robinet	
54			donc (GH)	
55			après ça elle a volé	
56			tout haut à la maison	
57			et:: la mouche elle a* resté dans la maison	
58			e:t c' est c' est c' était la fin de le film du film	

JA (niv 3)				
Seg	BD1	BD2	Film1	Film2
1	j' ai vu le garçon avec eu:h un cheval	bon le dessin::	les les fourmis &mon montent	le le film montre *la histoire
2	et::	il montre la histoire	et::	d' une d' *un abeille abeille
3	il y a les:	eu:h du monsieur avec les::	une:	qui: entre dans une dans une maison
4	les &k les carrés	son cheval	abeille	et après::
5	un carré que: le garçon: c' est	le monsieur il arrive: sur &mont monter sur le cheval	qui est:	et:: [&=bche]
6	sur les les les les *cheval	et::	qui entre et danse dans la maison	elle est c' est difficile pour sortir de la maison
7	les autres il y il est à côté du cheval	et après	et::	elle a fait une promenade *sur la
8	et les autre:s	il il est il a il a descendu du cheval	qui *fait les promenades et après elle a fait elle *a entré dans dans la maison	*sur le salon:: *sur la cuisine
9	carrés il y il est	il est assis sur un banc	et elle a passé *pour	su::r [&=bche]
10	en face du cheval	un tabouret	*pour plusieurs places *pour *pour la cuisine *pour les:	et dans toute la maison
11	e:t les autres	je crois parce qu' il est fatigué	les salles *pour les::	et après elle a cherché un endroit pour sortir
12	carrés	il est il dort	et après elle *a resté	elle: se tape: (GH)

13	il y a deux garçon:s	il a dor:t*	à la cuisine	avec la: la fenêtre
14	avec un *cheva cheval	et::	et::	et:::
15	&chev	il arrive après arrive un petit_enfant	il vient les autres les autres animaux	et elle a continué à chercher continué à chercher un un endroit pour sortir
16	cheval	qu' il suit avec lui un petit cheval	qui qui entraîent à la maison aussi	et: elle a regarde les fleurs
17	et après qu' il y:	il change	e::t	elle: sent les parfums (GH)
18	et les autres situations	ton: petit cheval	les: les abeilles elle a elle a part	de fleurs
19	et mais je ne sais pas comment faire pour écrire pour décrire la situation	il prend les *pieds il laisse son petit cheval avec le monsieur	et:::	il est
20	e::t	il prend le grand cheval	elle part	de fleurs
21	hum je pense que c' est ça	du monsieur	&s elle a::	et::: après elle regarde
22	les:::	e::t	n' est pas	où
23	une *un situation que le garçon	quand il fait ça le grand cheval il part	pour sortir elle a pas sortir elle a	dans le salon elle regarde les choses là
24	et c' est un garçon grand	vite	pour partir de la maison elle a::	elle arrête sur la chaise
25	et les autres garçons petits	e::t	c' est pas c' est pas bon pour elle a elle a pas:	et après elle regarde la fenêtre
26	et: le garçon petit il part	le petit enfant il *courir pou:r	cherché les places pour sortir	il y a::
27	et le garçon grand il reste il a e:t il es:t resté	pou:r prendre le &ch le le cheval le grand cheval	elle a	une mouche
28	e::t	e:t au en ce moment	elle a &f elle a fait beaucoup de choses et après elle a elle regarde les fleurs elle a pensé que c'est les fleurs de la	qui vient
29	et le *&chev cheval	le monsieu:r	de la place et a	qui: a regarde aussi
30	aussi	il se réveille	et::	et:::
31	*parti *parte aussi	il voit qu' il es:t avec le petit cheval	a choisi	après la la: l' abeille là sort
32	comme le garçon petit	l' enfant	comment dit*	elle a regarde qu' il y a des tournesols
33	c' est ça	je crois que c' est tout	elle a choisi	et elle a pense je crois qu' elle: pense que c' étai:t

34			à côté de la maison mais c' est pas la la fleur de la place à coté de la maison c' est la fleur de de une image	la nature c' est elle a dit ah je je vais sortir c' est bien
35			et s' il l'a s' il l' a mis	et elle elle a elle commence à marcher à marcher
36			et elle a fait pour	à:: voler très vite très vite
37			quand elle a va regarder les &fleu les fleurs elle a fait /bo/	et: quand elle arrive là bas
38			et y a qu' à et	elle tombe sur
39			et c' est pas mal pour et c' est mal pour la l' abeille	un:: carré
40			et après	qui es::t
41			les les autres animaux qui entraient *pour la cuisine	qu' il est sur la le mur
42			et::	et après: elle a:
43			va aider les les abeilles pour sortir de la maison	elle a &p elle a eu peur
44			elle a elle y	de:: pas
45			les autres animaux il faire les	pas sortir de pas:
46			les chemins	pouvoir sortir
47			où	et:::
48			et &ap elle y a prendre le chemin	en ce moment il y a la mouche qui entre dans la maison
49			pour l' abeille pour la &p elle a	et: l' abeille elle: l' abeille elle:* s' y cache
50			peut sortir de la maison	parce que aussi elle a peur de:
51			et y a et elle va elle va sortir de la maison	de pas::
52			pour les toilettes pour les salles de bains	c' est pas bien pou:r pour elle

53			pour les *baignoires je ne sais je ne sais pas comment on dit	elle: se cache
54			et elle a elle a retourné	et: la mouche elle: va
55			*pour la place du coté de la maison	à la cuisine elle: cherche des::
56			*pour les	des oignons
57			*pour les rues *pour les pour les fleur:s	des haricots
58			pour les places à côté de la maison	pardon des haricots
59			c' est ça	et elle: va
60				elle: cherche des haricots elle: cherche un autre un autre
61				et après elle *sorte de la maison
62				et quand elle sort l' abeille elle va à la salle de bains
63				et::
64				et là bas il y a: (GH)
65				une chose marquée
66				une set
67				marquée pour dire
68				comment fait pour sortir de la maison
69				et l' abeille elle a elle a réussi à sortir (GH)
70				elle a réussi à sortir et c' était la: la mouche
71				qui a a aidé
72				bon comme ça elle a sorti de la maison

MA (niv 2)				
Seg	BD1	BD2	Film1	Film2
1	bon: c' est l' histoire où:		hum donc c' étai:t	euh c' est un
2	comme j' ai compris c' est: l'histoire de d' un homme	donc eu::h	c' est l' histoire de	dessin::
3	qui est un: d' un	y a x sur l'image::	j' sais pas comment s' appe:lle cet insecte: l' abeille	bon [&=rires]
4	cowboy	dont je::	c' est c' est comme ça ok c' est l' histoire de l' abeille qui est:	c' est l' histoire: d' une abeille
5	cowboy c' est comme ça euh qui euh	que: je viens de regarder	qui es::t entrée dans la maison	qui est::
6	qui est eu:h	c' étai::t le scène avec euh	dans la maison:: e::t parce qu' il	qui es:t je sais pas qui es:t
7	qui est avec son (GH)	avec un:	il a vu là bas de	qui est entrée dans la maison
8	son cheval	un cowboy	des fleurs dont il peut respirer il peu:t	donc eu::h et après
9	est venu et il es::t fatigué comme j' ai compris il est fatigué il est euh:	avec euh son cheval	oui il peut:	parce que:: elle a vu elle a vu là bas des
10	e::t il est comment dire e::t	e::t un garçon	prendre:	des:: images de fleurs ou elle pensait que je sais pas elle pensait que peut être que c' est
11	il s' est assis sur euh la banquette pour se reposer un peu il a	e:t ce cowboy s' est endormi	et pour soi des fleurs je sais pas comment ça se passe &=rire donc voilà e:t il est entré il a commencé à:: ( (GH))	des vraies fleurs
12	attaché son: cheval	e::t une ga un: garçon avec un::: je sais pas son joujou	glaner ou:: je sais pas	donc après elle est entrée et:: elle a compris que voilà elle peut plus s sortir:
13	à: côté de lui	joujou le cheval aussi le p'tit cheval il s' est approché	de:: la poussière des fleurs xx ( (GH))	ces fleurs qu' elle a vue:s ne sont pas vraies
14	pour que la:	de ce cowboy e::t il a::	donc euh	donc euh elle a essayé par tous les moyens de sortir de la maison: elle a volé

15	pour qu' elle coure coure pas	il a: échangé	un moment après	volé volé et après
16	[&=rires] voilà et après il s' est endormi	il a changé le cheval	il es::t	elle a compris qu' elle peut plus donc euh c' est: elle a::
17	est endormi* eu:h	le joujou avec eu:h	il est il y a il est compris que:: il arrivait pas à sortir de cette maison parce que	atterri: sur la fenêtre et elle a vu la mouche la mouche
18	et euh (GH)	le cheva::l (GH)	dans tous les	une mouche qui::
19	tant qu' il a il dormait	le vrai cheval donc eu::h après	toutes les fenêtr::s étaient	qui a:
20	venu (GH)	euh le cowboy: dort toujours	y avait des fenêtres et l' avait ouvert dont	je sais pas qui a décidé de lui (GH)
21	venu euh	donc eu::h le garçon est parti avec le cheval normal	il es::t i::l	de la: de l' aider
22	p'tit garçon avec euh son cheval	et quan:d le cowboy s' est réveillé: e:t voilà et	il pouvait pas sortir	à sortir euh
23	son cheval eu:::h son	il s' est trouvé avec un: joujou (&rires) e::t voilà	donc il a: il a volé il a volé	donc la mouche:
24	joujou	et il était étonné	il a cherché: la sortie	la mouche est:: entrée dans la maison
25	son cheval en bois	il était surpris [&=rires]	mais /t/ à la fin il euh	donc (GH)
26	c' était le p'tit cheval euh		( &= rires) il n' arrive: pas &=rire	voilà après la mouche a compris &rires où où il peut &s
27	donc eu:h		&=rire à trouver la sortie e::t il se: il se:	sortir par où
28	e:t après il a regardé cet homme et il a décidé de:		atterrit pour eu::h ( (GH))	donc eu:h elle a dessiné la flèche (GH)
29	de changer euh [&=rires] de cheveu* [&=rires] (GH)		pour regarder ( (GH))	dans le sens de la sortie c' étai:t dans la: dans le robinet
30	donc euh je sais pas pourquoi [&=rires]		hum au delà de la fenêtre	j' sais pas comment ça s' appelle le trou dans le robinet donc
31	pour rigoler euh		et il a vu la mouche qui euh se:	que par des: par la: canalisation elle peut sortir donc euh
32	et après il a		et qui:: qui s' est atterrie* donc euh	elle est sortie elle a::
33	il a		aussi ils sont regardés	elle est partie

34	comment dire détacher le cheval le grand cheval euh le cheval réel		pendant quelques quelques instants et après:s	et:: l' abeille
35	et il l' a attaché à son: son: euh (GH)		euh après j' ai j' ai oublié déjà il est sort euh il est entré aussi dans la maison	hum l' abeille *est d' abord pensé
36	joujou		et il a il a vu de:: du::	après elle a vu la flèche
37	son joujou eu:h il l' a attaché et et il est parti avec un un cheval		du café des::	donc euh
38	normal		des comment c' est des blés	et par des:
39	et quand cet homme s' est réveillé		euh donc	par des: je sais pas
40	il a::		il a vu et il a commencé à prendre ses trucs du café et en fait il lui aussi il comprenait pas qu' est ce qu' il fait	par euh les
41	il a vu que au lieu de: de		e::t en fait i:l	canalisations elle est sortie par le robinet est sortie et voilà
42	son [&=rires] de son beau cheval euh		et lui aussi il il comprenait pas qu' est ce qu' il fait	s' est trouvée ailleurs &=rire
43	euh ce [&=rires]		et après il est compris* hum quand il est	et s' enfuit
44	[&=rires] ce p'tit euh		quand il y a quand	
45	euh ce p'tit cheval quoi		quand il a volé dans la douche dans la chambre	
46	c' est ça		dans la pièce de avec le douche il a vu::	
47	et il était surpris [&=rires]		qu' il a euh	
48			qu' il eu::h qu' il qu' il	
49			lui a dessiné la flèche	
50			dans la: robinet euh (GH)	
51			pas dans le robinet euh xx	
52			pas le bain	
53			et donc eu::h pour qu' il puisse comprendre que: il peut	



54			il peut en fait sortir par la cana canalisation après: il a compris aussi il est	
55			il est descendu dans la::	
56			je sais pas dans le trou	
57			[&=rires] dans l' eau [&=rires] il est sorti et après il a marché marché marché	
58			et après eu::h le deuxième cadre	
59			le robinet qui qui est dehors	
60			de cette maison	
61			et il sort de cet robinet	
62			et voilà elle es::t elle est libre	

<b>MJ (niv 3)</b>				
<b>Seg</b>	<b>BD1</b>	<b>BD2</b>	<b>Film1</b>	<b>Film2</b>
1	eu:h un homme est arrivé sur un chev euh cheval	euh j' ai juste vu euh	ce que j' ai vu euh	eu::h on voit une: abeille
2	e::t il était fatigué	des dessin:s d' un: cowboy qui vient	c' est une petite histoire d' une abeille j' suis je suis pas sure que ça c' est le mot pour	dans un jardin qui cherche: des fleurs
3	donc il s' est endormi à côté d' *une mur sur un *banc	en chevau	il a cherché de::	mais il peut pas trouver de *bons fleurs dans le jardin
4	et un petit garçon est venu avec un petit cheval	cheval	quelque chose: pour manger	donc il euh il voit la maison
5	et il a:: euh la laché le le grand cheval	et eu:h il est un peu fatigué donc il euh place son: son cheval à côté il eu::h (GH)	et il a vu une: maison	et il entre dans la maison

6	et remplacé avec le petit cheval donc quand l' homme es:t	après eu:h	il *a entré	en recherche pour des euh
7	eu:h s' est reveillé	il eu:h il s' endorme	mais eu::h quand il voulait sortir	des fleurs mais il peut pas trouver il cherche il cherche il peut pas trouver
8	il a vu que il a vu juste un le petit cheval	à côté	tout était fermé donc les: les:	alors il entend *une bruit il veut sortir
9	donc il était étonné	et eu::h	fenêtres: étaient fermées il pouvait pas partir par le euh porte (GH)	mais i peut pas euh sor
10	c' est ça je crois	il voit rien parce qu' il a mis son chapeau sur sa t euh devant ses yeux donc i il voit	porte (GH)	euh sortir parce que tout est fermé
11	ça y est [&=rires]	plus rien et il dort	donc eu::h	et euh il veut essayer de: de sortir de la fenêtre mais aussi
12		et après il y a un petit enfant qui vienne* avec une: (GH)	il a cherché dans la maison pour euh	fer il est elle est aussi fermée
13		un petit &che cheval en plastique je crois ou en bois	euh de nourriture il a trouvé des euh fleurs mais	et eu:h
14		parce qu' il a des eu:h des roues	qui	ensuite il trouve un:
15		donc on peut le:: on peut le: tirer	qui sont en plastique	des fleurs mais c' est en plastique
16		et eu::h	il pouvait pas:	e::t
17		on peut le:: (GH)	euh manger ça	donc il cherche encore et après il voit un un peintre
18		déplacer eu:h c' est pas un vrai cheval donc on on il marche pas tout seul	et ensuite il a vu euh une peint	qui eu:h (GH)
19		et euh:	ture	une peinture
20		ce que l' enfant &f	et::	qui eu:h (GH)
21		fait euh eu:h il donne le il prend le le	il a:	ah il croit que: que c' est des fleurs des vraies fleurs donc i est
22		le cheval le vrai	essayé de manger ça aussi mais	il veut euh
23		et: il	euh sans succès	il veut euh

24		remplace le cheval avec le petit euh en bois ou plastique	et il a vu une: il y avait une:	y aller mai:s c' est un peintre* donc euh
25		et eu::h quand l' homme se réveille	un petit eu:h mouche	un peinture donc i peut pas euh (GH)
26		il voit que il n' y a plus de: (GH)	qui euh qui l' a vu euh qui a vu le l' abeille euh	faire quelque chose avec des fleurs
27		plus son: vrai cheval mais un petit en plastique	à l' intérieur	euh ensuite il voit un une mouche
28		et euh il est choqué	e:t il est entré aussi à la maison	à l' extérieur
29		par ça	mais l' abeille ne comprend pas pourquoi comment	et un regard un peu triste
30			il est entré	pour eu:h pour ce qu' il veut d' aide pour sortir de la maison (GH)
31			et pour montrer euh la direction euh	euh le mouche entre dans dans la maison il eu:h
32			euh des euh	il prend des euh grains de café (GH)
33			des sorties	il mette le dan::s dans le: le bain dans le euh salle de bains
34			pour l' abeille	et eu::h (GH)
35			il a indiqué eu:h	il il indique (GH)
36			et le mouche euh a indiqué la route	où il y a *le sortir
37			avec eu:h	le sortie pour l' abeille
38			des p'tits trucs pour faire du café je sais pas le mot	et l' abeille trouve le:
39			et dans et le sortie était dans: la salle de bains (GH)	eu:h le sortie et:
40			et euh par un	il est libéré
41			hum d' une: je sais pas le mot d' une lavabo c' est comme un	par la mouche
42			dic qui y' a dehors donc il a livré par les: les	

43			et euh (GH)	
44			l' installation de l' eau	
45			de:: du salle de bains	
46			c' est ça	
47			c' est tout	

MR (niv 2)				
Seg	BD1	BD2	Film1	Film2
1	il y a un cavalier je j' sais pas s' il y a un mot en français pour dire *jinete	il y a: un	il y a: un p'tit sionair	c' es::t
2	e:::t qu' il est très fatigué très très fatigué	un cavalier qu' arrive à:	qui voula:t	dans cette histoire il y a un mouche il y a un abeille
3	e:::t il laissait son::: euh *chevau	à *maison	es::t/est présent la maisone	et il y a un maison avec jardin
4	eu:h amarré amarré	et il *démontait de du *cavale du cheval	e:::t un jour *lé *lé porte est ouvert	e:::t
5	e:::t	e:::t il a laissé le cheval	et et le p'tit sionair est entré dans sa maison	et:: l'abeille
6	il asseyait *à *une banc banc	le cheval à côté de la maison:*	e:::t elle cherchait tous les tous les tous les fleurs	elle vole pour le jardin
7	ok	avec la &cor*	e:::t	e:::t
8	e:::t	e:::t	mais quan::d i:l	il trouve la la porte ouverte de la maison et entrer dans la maison
9	il: en train de dormir	il est en train de dor de dormir	voulait sortir de la de la maison	i:l (GH)
10	il dormir	il dort et quand il dort	e:::t (GH)	volait
11	et après il y a un enfant avec un jouet un jouet que c' est un petit *chevau	un petit enfant qui pren:ne son cheval	un prisonnière	vole
12	e:::t	échanger son cheval pour un cheval de jou un jouet que c' est un cheval aussi	lui euh ne ne peut pas	dans la maison à l'intérieur de la maison

13	et a profité que: le: le chevalier le *jinet est est très fatigué et il dorme*	e::t et quan:d	parce que::	pour les différents chambres et::
14	pour euh changer le: le cheveu réel	le cavalier est réveillé	toutes les portes et les fenêtres sont fermées	cherchait les différentes fleurs et tout ça
15	pour le: cheveu de jouet	e:t il regardait le: le cheval	et en ce moment eu:::h un mouche	jusqu' à:: *qu' il veule sortir (GH)
16	e::t après le l' enfant partit avec le: le: cheveu réel	et il n' y a pas le cheval il y a *solament le jouet	et:::	vo volait il veut sortir
17	et quan::d le *jinete et le chevalier le chevalier	avec la forme d' un cheval	bou	de la maison
18	es::t	c' est tout non	e::t voit	et::: il essayait pour* la porte essayait pour* la fenêtre il essayait
19	se lève	ouais je crois	e:::lle es::t	et beaucoup de fois
20	e::t trouve que::		l' est prisonnière dans la maison	mais il ne trouve pas
21	son cheveu a *désapparu il y a seulement un petit *cheveu de jouet		pour un fenêtre	un: lieu pour sortir dans (GH)
22	c' est tout non		il est: et le* mouche est:	pour sortir à au jardin
23			à l' autre:	et: jusqu'à:
24			à l' autre part de la fenêtre à: à l'extérieur	en fenêtre:
25			et lé mouche entrait dans sa maison	en fenêtre il es:t
26			et indiquait le chemin de sortie pour *lé::	à l' autre: à la coté de
27			*le prisonnière	dans les dans la maison
28			et lé prisonnière prend les:	et à l'autre côté à la côté du jardin il y a une mouche
29			lé: chemin de sortie	la mouche regardait la:
30			et: finalement	l' abeille
31			elle est à l' extérieur	e:t la mouche voudrait entrer::: veut entrer et l' abeille veut sortir

32			mais eh: elle est à l' extérieur mais lé* mouche	et après
33			désirait rester à la maison	la mouche finalement elle entrait dans la maison
34			je crois que l' hi eu:h l' histoire	et:: l' abeille
35			voudrait dire que lé mouche (GH) euh ( )	et regardait la mouche et la et la mouche
36			préférerait la maison et lé: prisonnière préférerait le gard	cherchait
37				les grains de café pour indiquer la sortie
38				e:t la sortie c' est dans les:
39				dans les: salle de bains
40				et dans la baignoire
41				e::t l' abeille prend les:
42				le:s les forêts qu' il y a à la: à la baignoire
43				et ces forets conduit directement à: au jardin
44				et c' est ça que ça pour ça s' appelle l' a l'abeille prisonnière
45				parce que l' abeille est prisonnière dans la maison et
46				et promène la la la
47				l' animau l' animal que que
48				que vaut que veut
49				et::
50				et et être dans la maison c' est la c' est la mouche pas l' abeille
51				l' abeille préfère le jardin

**NO (niv 2)**

<b>Seg</b>	<b>BD1</b>	<b>BD2</b>	<b>Film1</b>	<b>Film2</b>
1	alors	je peux *commence	alors il y a une insecte	je sais pas encore le nom de cet insecte
2	euh il y a un homme	ok	qui fait le miel	alors [&=rires] (GH)
3	et il: hum &	alors (GH)	et euh [&=bche]	on dira une insecte
4	il a une* animal	y a un homme qui:	il	alors il y a une insecte qui entre dans une maison
5	qui il um	est venu eu:h avec son: cheval	il est dehors il a	et::
6	il allait peut être au* ville sur le* animal	et::	et il	i::l
7	alors hum (GH)	il est fatigué	voit une fleur dans une maison [&=bche]	i:l entre dans le cuis:i:ne dans le salon: il sent les fleurs il sent les fruits
8	i::l	alors il s' arrête	alors il entre le maison (GH)	les fruits qui sont là
9	trouve que il	et:::	pour voir les fleurs	et après il [&=bche]
10	il es::t	il asseye	peut être pour prendre le nectar	essaye de: il essaye de partir de la maison mais tous les fen
11	fatigué	lui sur u::ne chaise pour dormir	e:t	fenêtres sont fermées
12	alors il arrête	alors il dorme et i:l	il euh	alors i:l [&=bche]
13	pour rester pour quelques minutes peut être	laisse le cheval à côté de lui	il entre dans les autres chambres dans la maison	il reste un peu euh dans le cuisine à côté de la fenêtre
14	alors il euh	e::t hum	et quand il veut	et après
15	il s:	il arrive u::ne	sort de la maison il ne peut pas parce que il ne peut	il voit une autre insecte qui est dehors à: l' autre côté de la fenêtre
16	ss' asseyait	jeune garçon	&trouv il ne peut &trou	et il regarde les autres:
17	asseye* sur un ah chaise	qui:: voit que il y a un homme avec son cheval	il ne:	et après (GH)
18	qui es::t	et le garçon il a une cheval pour jouer c' est pas vraiment une ch cheval	trouve pas les (GH)	l' insecte qui est dedans i:l il parte il revient

19	proche d' une maison ou::	alors i::l	une fenêtre qui est ouverte	encore au salon:
20	un magasin peut être	pour rigoler peut être i::l (GH)	alors [&=bche]	et il y a une image des fleurs dans le mu
21	et il dort pour quelques fois	prend le cheval	il entre dans les chambres et	dans le mu:r
22	et il y a un petit garçon qui	le vrai cheval	quand il est dans la cuisine	et: il a encore essayé pour entrer dans l' image mais
23	le voir* et le garçon:	et il laisse ton:: *jou	il voit une autre insecte qui est dehors (GH)	ss il arrive pas parce que c' est pas vraiment
24	a u:ne* animal la* même animal	pour l' homme	mais l' autre insecte ne peut pas	des fleurs
25	mais ce n' est pas *une vraie animal c' est c' est (GH)	et il sorte le cheval il court	lui aider	alors après le:
26	un: jou pour le garçon	peut être dans le désert	alors i:l (GH)	l'insecte qui était dehors il entre dans la maison
27	alors	e::t après le:	entre dans la chambre de	et::
28	quand le: le homme	l' homme il s' est réveillé et il voit que son cheval est parti et	avait	il voit que l' autre insecte il sait pas comment partir
29	*dorme	il est là avec le	où il	alors i::l
30	le petit garçon:	le petite	voit	prend les petits *morceaux du: café
31	*prend l' animau vraiment de le homme	*jou	u:ne	[ [&=rires]] les &hari
32	et il remplace	c' est tout	autour d' une fleur et peut être il pense que ça c' est	les &zari haricots &rires
33	sa animau avec son jou		une fleur vraiment (GH)	*quiz du café
34	alors il		mais ce n' est pas une vraie fleur	e::t
35	part avec le animal vraiment		et [&=bche]	il eu::h il sorte avec les morceaux
36	e:t le: après ça l' homme		après ça l' autre insecte entre dans le maison	e:t après
37	se réveillait		et il prend* les	on ne sait pas où il mais
38	et il trouve que son animal s' es:t euh (GH)		les_haricots du café	on entendre* que le
39	n' est pas là		e::t	eau



40	alors il y a le		il entre dans le salle de bains [&=bche]	tombe dans le salle de bains
41	jouet le jou: e:t		et après ça l' autre insecte part	alors le:
42	peut être il n' est pas		et la première insecte il &ex (GH)	tout le temps que ça se passe le: premier insecte il regarde mais i:l
43	ah il ne		il entre dans le salle de bains	il se cache un peu (GH)
44	peut pas retourner		et il a vu que	après dan (GH)
45	chez *il		il y a	il se cache
46	chez lui		une euh les directions	&ap
47	j' ai fini		qu' il a dit	*au derrière de la peintre
48			tu entres: (GH)	de l' image des fleurs
49			sous le:	alors [&=bche]
50			le douche	après il &é il écoute l' eau
51			alors il entre:	et il entre dans le:
52			là et:	salle de bains
53			après ça	et: il voir que
54			il est dehors parce que il y a	dans la baignoire il y a tous les morceaux du café
55			des u:ne:	et se pointent dans le direction (GH)
56			rue	se pointent dans la direction
57			qu' il (GH)	dans le
58			pas	tu:be
59			sort le maison de ça	alors il descende
60			alors l' autre insecte il	et::
61			restai:t	il suivre: le chemin
62			peut être pour le premier insecte	et après il sorte
63			mais le premier insecte ne le	de:: le: robinet qui est dehors
64			voit pas alors il sortit	et:: il partit

65				e:t le
66				l' autre insecte il regarde lui peut être
67				il a voulu
68				être son ami mais
69				mais
70				il arrive pas *d' avoir une nouvelle amie (&rires)

NY (niv 1)				
Seg	BD1	BD2	Film1	Film2
1	euh il y a un un homme*	il y a un homme et un s et avec s: son: cheval	euh il y a	il y a une abeille
2	e::t e:t il est un peu fatigué	il fait un randonnée avec	une abeille elle cherche une nourriture	elle: vive dans dans la nature
3	donc il *dormit	cet ani	et:	mais un jour
4	et un:: [&=cl]	avec l' animau	et euh elle	elle a:
5	e:::t	après des:: &ap après les les randonnées:	ell:e	elle entrait dans une maison
6	un un un garçon: eu:h (GH)	il est il était un peu fatigué donc	elle cherche	dans cette maison: elle a trouvé beaucoup de sens différents différents sens comme de les fleu::rs
7	fait quelque chose *biar bizarre avec le homme	il arrivait à dormi*	pour tous les les	des sens de la nourriture de la cuis:ne
8	e:t maintenant il est très::s	mais	les places [&=bche] et elle a [&=bche] (GH) [&=bche](RG: )	elle a:
9	très::s	il a: sorti de l' animau il a	et elle a (GH) [&=bche] RG	elle a cherché partout:t
10	très::s je ne sais pas (GH)	après il a essayé pour dormi*	elle euh eu:::h va	par quelque chose différent
11	et je ne il ne s il n' est	donc il a dormi beaucou:p	en une maison	&ap

12	il il va fai:re	parce qu' il était très fatigué:	et elle cherche	mai::s elle a trou elle a pas elle a rien trouvé
13	il	mais	e::t tout tout elle elle seulement	après:
14	va faire quelque chose pou:r [&=rires] (GH)	un moment &=s	pour chercher	elle es::t elle a
15	pou::r		e::t pour les pour le nez	elle a vu les fl
16	rencontrer &ren &ren &ren rencontrer	il a eu:h arrivé un &ga un un garçon petit garçon un garçon	e::t (GH)	fleurs dans les murs
17	son _ animal	avec	mais maintenant elle a::	mais ce
18	[&=rires] c 'est tout	un::	elle veut sortir de la maison mais elle ne ssait elle ne s	mais: c' était c' était pas une vraie fleur
19		avec un cheval	elle	c' es::t
20		de jouet	ne: savoir pas	c' est une &fl une &for une
21		il il est il regardait le le &chev le cheval	le chemin	une &f
22		il il pensait pour &ch	mai::s	une *form &f:: forme ou fleur
23		pou:r changer	elle rencontre une amie	parce que quand elle essayait de sentir
24		pou:r	e::t	les fleurs elle arrive pas de
25		pour prendre les les cheval pour laisser les les	que elle est et l' amie ce n' est pas	de::
26		le::s	e:t dans la maison c' est elle elle est	elle arrive elle elle arrive pour pour s elle elle
27		le: cheval de jouet	elle es::t [&=rires]	elle a vu que c' était pas une vraie fleur
28		après le homme il	elle elle est pas dans la maison	donc elle voit par les par les
29		il se révei:lle il regarde qu' il le cheval n' est pas là pas plus là	e::t	par les s
30		il il est en train pou::r	&s euh l' amie va	par les yeux aussi
31		un peu fâché un peu:	va:: (GH)	par les sens
32		sans sans *compris qu' est ce qui se passe	va aider	après elle était elle était un peu: elle voudrait sortir de la maison et elle elle arrivait pas

33		sans comprendre ce qu' est ce qui se passe	l' abeille	donc
34		c' est tout	et qu' est ce qu' elle va faire elle va elle va: elle prenait*	elle a vu dehors
35			nourriture e::t grains de café parce que	u:ne une mouche
36			les grains de café la abeille elle pose pour	une une mouche
37			pou:r (GH)	après lé* mouche
38			sentir	entrait dans la maison
39			les grains de café	il il *le mouche
40			et elle indi l' indique le chemin	elle était un peu un peu::
41			avec les grains de café	elle a peur un peu de la mouche donc elle a
42			et la l' abeille	elle s' est cassée
43			elle:	de de le mouche mais le mouche voudrait aider la le
44			elle rencontre le le chemin* le chemin pour sortir de la la maison	l' abeille pour sortir de la maison
45			e::t avec cette grain de de:	*qu' est qu' il *qu' est qu' il a fait il a:
46			de:: café	il essaye comme comme le l' abeille elle voit par les yeux et par les
47			et et maintenant elle est	et aussi par les sens
48			elle es::t	il essayait de &p prendre les: les:: les grains de café
49			elle es:::t	pour montrer le chemin pour sortir de la maison
50			il il n' a::	il a il a indiqué
51			n' a vraiment il vraiment	tout (GH)
52			euh e:t elle est	dans dans une baignoire

53			je ne sais pas	ou est ce que que l' abeille pou pouvait trouver le::
54			elle es::t	*le sortie de la maison
55			elle elle est très:s	pour arriver:
56			et je ne sais pas parler	à: à nature
57			maintenant elle est elle est ne	donc après ça il fait le l' indication avec les:
58			n' est vraiment lio	les peti:ts les petits grains de café
59			elle ne [&=rires]	et l' abeille a &s a &s suivre le sens de le grain de café:
60			je n' sais pas parler: (GH)	et après elle a
61			s quand jé quand *je quand elle va sortir	elle a trouvé le chemin pour sortir de la maison et après elle est très très heureuse parce qu' elle est très dans la nature une autre fois
62			je ne sais pas parler	

OL (niv 1)				
Seg	BD1	BD2	Film1	Film2
1	un homme	il y a: un cowboy	un	c' est dan:s un jardin
2	il euh monte	un cheval	le::	l'abei:lle
3	un: cheval	et il eu:h vien:t	vidéo commence	vole:
4	e:t après il descendre	e:t	on:	tige
5	e:t il euh met le ligne	et i il euh il est fatigué	voi:t	pendant les fleurs
6	le ligne c' est	et il peut dormir il descend de	un	e::t
7	le chose qui est dan:s le hum (GH)	zeerrrr	peti:t	il eu:h
8	on	et i il euh s' assoit et il dort	un peti:t	collecte: les

9	quand on monte au en cheval*	e::t euh quand tant pendant il dort il y a un: garçon	animau	*pollens
10	on prend le ligne	il pren:d	mais c' est pas trop animaou c' est comme un mouche mais en plu:s	*pollens
11	sur la main	il eu:h tient	plus grand	*pollens
12	e:t il prend les lignes sur le main	un petit cheval euh *en poing	et il est jaune	les *prolens
13	et il dort	en bois	il mange: le fleur	et après (GH)
14	et après il y a un garçon	et euh	je ne sais pas le mot de ce de cette: cette: animaou	à coté de le jardin il y a une maison
15	il pr il euh	le petit garçon il trou:ve	je pense: je vais	et *le porte est *ouvert
16	il a hum (GH)	le grand cheva:l	s'appe:lle	l' abeille:
17	plus peti:t cheval	es::t joli:	le petit jaune (GH)	il eu:h
18	et il euh vu* il voit	e:t il euh change: euh les deux rênes	eh eh le petit jaune (GH)	entre: au dessus de la:
19	le gran:d cheval	e::t le grand cheva:l	il a vu:	porte
20	et il change le le ligne	il euh	les fleu:rs euh	et (GH)
21	et euh quan:d	&cou:: *cou:rir oui	et il euh mange et il euh mettre *petite jupe sur son jambe	il trouve les fleurs
22	le_homme	il court	et après il y a une maison	e::t
23	dort	il court très vite et après eu:h le garçon il sort aussi	et la porte:	dans un:
24	e:t	e:t	est *ouver:t	non:
25	&ap	le: cowboy:	e:t	sur un petit
26	et maintenant	il euh se réveille	il il a il euh voit	tableau
27	et après	et i::l trou::ve	il y a	c' est le tableau de salon
28	le	il tien:t le rêne de un	une &f fleur euh &s dan:s le: verre	et il y a un boutei:lle de fleurs
29	garçon prend	petit cheval	le fleur est très jolie	et l' abeille: il vole et il écoute
30	le	en bois	et il euh	il collecte des
31	grand cheval	c' est fini	ren entre	*prolins

32	et le_homme le euh l' homme		la maison	pollens
33	dort		e:t	et après
34	et le ligne de le petit cheval		il euh	euh il entre le cuisine *le salle de cuisine il a vu des poivres
35	est sur son main		écoute:	des: basilics
36	et quand il euh		la fleur	&baz
37	se: rél réveille		hum mais il trouve il y a rien	*parsile
38	il euh		c' est pa:s un vraiment fleur	e::t
39	il euh il a il euh voit		et	beaucoup de choses
40	le cheval est plus plus est		il euh	et il écoute il ne trouve pas les fleurs
41	plus petit		commence: fli	e::t
42	et il pense qu' est ce qui se passe		partou:t la chambre: la cuisine	im
43	c' est fini		et il écoute: le fleur euh	et après dans ce moment
44			et le poivre le sel tous* les choses	il écoute: le:: bruit
45			et après	le bruit est le son
46			il euh	de:
47			entendu	la porte est fermée
48			il a entendu	e:t
49			un son: qui	il trouve il peut sortir
50			comment	e:t
51			la &fl	il eu:h il a vu le les fenêtres
52			porte est fermée	et les et la porte mais il y a un verre:
53			et il	pom
54			il doit sortir euh de la maison	dans le la fenêtre et les portes
55			mais	il euh a essayé

56			il ne sait pas comment	passer par le fenêtre
57			il a vu:	mais il euh s'é-
58			euh la fenêtre	crase
59			et il *flee	oui
60			directement sur	il s' écrase:
61			par euh	dans *le fenêtre:
62			la fenêtre	e:t il trouve il ne peut pas sortir
63			il pense il peut passer par la fenêtre	et il entre:
64			mais il a:	la salle:
65			p'tit* verre sur euh	de bains oui
66			la fenêtre	il euh entre la salle de bains
67			et [&=rires]	et il euh s' arrête:
68			il euh comment	à côté de la fenêtre:
69			il *flee et	e:t il eu:h
70			pan [&=rires]	il pense &co comment il peut sortir euh
71			et il euh fait	et maintenant
72			beaucoup de fois	euh là
73			e:t	il a
74			il euh trouve	une: mouche:
75			il ne peut pas sortir	qui vient
76			sur la fenêtre parce que il a quelque chose	et la mouche s' arrête à l' autre côté de la fenêtre
77			et il euh s' arrête	e::t
78			à côté de la fa fenêtre	je pense: l' abeille:
79			e:t	demande à lui comment je peux sortir



80			après quelques secondes	et après la mouche:
81			il y a	il vole
82			une mouche ou un mouche une mouche	et: l' abeille:
83			il y a une mouche:	ne sait pas comment sortir et l' abeille continuait
84			il s' arrête:	trouver: le::s
85			*à l' autre côté de la fenêtre	fenêtre non
86			et je pense: (GH)	les fleurs
87			le petit jaune	et il euh trouve:
88			parle quelque chose:	pendant (GH)
89			avec euh le mouche	sur un: petit tableau
90			parce que:	il y a
91			elle:	*un bouteille de
92			veut	fleurs artistiques
93			veut	et il écoute: il écoute
94			parce que elle:	beaucoup de fois mais il ne *vut pas les:
95			fait quelques mouvements	*prolens
96			mai:s je ne sais pas c' est pourquoi	e::t
97			et après	après il euh trouve un tague
98			elle:	dans le:
99			elle::	dans les &fleu fleurs
100			trouve une autre fleur	il eu:h
101			et elle elle écoute	il a *écrire
102			mais c' est c' est pas vraiment fleur encore	ça fai:t
103			et elle *vut il y a un prix	un euro cinquante [&=rires]

104			sur la fenêtre c' est	et ces: fleurs artistiques
105			un t euro	et après il a vu un: peintre:
106			un_euro cinquante [&=rires]	un peinture
107			e:t	cette dans cette peinture il y a beaucoup de:
108			après elle vut un:	tournesols (GH)
109			ça c' est quoi	il y a beaucoup de tournesols et il euh
110			*print (GH)	euh il a vu les tournesols
111			*un photo	il pense il peu:t
112			comme photo mais c' est pas photo	*récolte du miel
113			où écrit quelque chose sur euh	et il euh
114			la	vole directement sur euh le &f le: peinture
115			et c'es:t	ils euh se
116			il y a beaucoup de fleurs	croisent
117			les fleurs sont jaunes	e::t
118			e:t	ils euh
119			le fleur euh (GH)	se croisent et après
120			le le &fleu la face	il il é:coute:
121			de la fleur	le le son: de: mou:che
122			face de la	c' est *le mouche qui: vole:
123			soleil tous les jours c' est comme ça	le son de vol
124			e:t	et il euh
125			il vut la fleur	s' arrête l' abeille s' arrête au dessus
126			et il euh	de la:
127			flee directement sur la fleur et pan (GH)	de le peinture

128			e:t	et il attend
129			et et après	et après le mouche entre:
130			euh il écoute la mouche:	de:: salle salle de la salle de bains
131			qui est qui entre la:	et il entre:
132			chambre	le la salle de cuisine
133			et il euh	et:
134			il euh se:	le mouche prend un grain de café
135			il euh trouve	e::t
136			une place	il entre: le salle de bains
137			qui est très	encore
138			privée	et il fait beaucoup de f fois il prend un:: grain de café de la salle de cuisine
139			e:t il euh	à la salle de: bains
140			la mouche ne sait pas où elle	et après
141			où	il eu:h
142			il est	*dispart
143			il arrête	il *dispart
144			où il arrête	et l' abeille ne sait pas qu' est ce qui se passe
145			e:t	et l' abeille il vole:
146			il vut la mouche:	et il entre le salle de bains
147			entre la cuisine	e:t
148			et prendre le pois* de café	il trouve il y a un: baignoi:re:
149			le petit pois qui est noir	e:t à côté de le
150			le pois	tuyau:
151			de café	de la baignoire

152			e:t il e	il a *trouve *un flèche de::
153			sort	qui
154			de la: cuisine	qui es:t construit non
155			et il entre à la	qui construit non
156			salle de bains	qui est construi:t
157			et euh il fait beaucoup de fois comme ça	par euh
158			et après	les:
159			il euh ne	de grains de café
160			après beaucoup de fois	e:t il euh entre le
161			il euh ne revient pas	tuyau
162			et euh	et il marche
163			le petit jaune	et après: il sort *de le tuyau:
164			il euh	dehors de la maison
165			il pense: il peut	e:t l' abeille:
166			sortir	es:t
167			su euh	comment c' est
168			dans si il fait	c 'est réussi
169			juste:	de: sortir de la maison
170			ment comme:	c' est fini
171			la mouche:	
172			et il euh sortit la cuisine et il euh entre:	
173			la:	
174			bain	
175			la salle de bains	
176			et il a vu	
177			sur eu:h (GH)	

178			sur euh la:
179			&boi
180			la baignoire
181			sur la baignoire
182			il y a un petit *ro
183			hum euh le &to
184			peut sort de la baignoire
185			et à côté de la roule
186			euh il y a un
187			*haro
188			héros
189			un &ro
190			c' est le héros (GH)
191			qu' il
192			euh la c' est le héros
193			que la mouche:
194			il met
195			la mouche:
196			mettre tous les
197			tous les
198			petits
199			noires pois
200			comme: un *aro
201			et *la héros c' est direct à la:
202			*roule
203			sur euh

204			la
205			baignoire
206			et euh
207			la le petit jaune écoute:
208			et voit
209			et il euh
210			il e marche
211			et il euh passe par la *ro
212			et après (GH)
213			dehors la maison
214			il y a un: peti:t
215			il y a un::
216			pi
217			pi
218			oui
219			peut être c' est *pi
220			il y a un *pipe
221			et il euh marche:
222			il marche:
223			il sort euh
224			la pipe
225			il a vu il est
226			dehors il est dehors de la maison
227			e::t
228			la mouche:
229			s' arrête:

230			à côté de la fenêtre:	
231			et la mouche	
232			voir euh le petit jaune:	
233			sort	
234			c' est le histoire qui	
235			la mouche:	
236			aide:	
237			le petit jaune[&=rires]	
238			sort sort c' est tout	

PR (niv 3)				
Seg	BD1	BD2	Film1	Film2
1	ok donc euh c' était une histoire de: de: une cowboy je sais pas comment tu dis ça en français	ok et donc eu::h ça c' est une petite bande dessinée de: une: cowboy	donc euh	ok donc eu:h
2	il *a retourné après peut être une longue journée: euh: de &f	américaine je sais pas euh qui es::t	cette:: (GH)	c' était *le histoire d' une petite::
3	je sais pas il est il est dans la:	qui est rentré: eu:::h après une jour journée de travail (GH)	petite film c' était une histoire d' une abeille	abeille
4	paysage: il fait des choses	il es::t	qui: cherche::: [&=bche]	qui eu:::h qui était emprisonnée dans une maison (GH)
5	il a retourné il était fatigué donc il n' a pas de temps pour euh	il était fatigué donc eu:h	les choses: dans la fleur	en fait euh je crois que c' est une:
6	mm se mettre son euh cheval euh bien il il	il *a pris une sieste	mai:::s (GH) euh	enfante d' une petite abeille
7	tire: le: (GH)	et il a pris son eu:h	quand il a cherché il a trouvé sur le le fleur qui n' est pas (GH) euh	il es::t

8	le:: euh (GH)	son eu:h cheval eu:h	euh des choses pour lui	il es::t hu:m [&=bche] essayé de: de:
9	le rêne	par une corde euh	et donc euh il cherche il cherche et après il a:	de chercher les fleurs
10	c' est ça le rêne de le cheval	et il a il il dort	il es::t entré dans une maison	mais en fait dans la *maison tout est faux
11	et il es:t il est d il es:t	et après eu::h dedans il y a un petit enfant	et là ba::s	donc euh
12	dormi	qui a passé avec une petite:	il y a des fleurs mais c' est les faux fleurs	il es::t (GH)
13	parce que il était fatigué	une petite cheval qui: qui n' était pas vrai c' est une petit jou	donc euh mais il a vu que c' est comme une fleur donc il a essayé de: (GH)	il es:t
14	et pendant ça il y a une	et donc eu:h il es::t	de:: chercher euh	il es:t hu::m regardé toutes les fleurs
15	une: petite enfant	il est f: il a décidé de faire une petite::	des choses: dans le dans le fleur	il a:
16	qui l' a vu	blague	e:t il n' est pas là et après les portes: (GH)	& s
17	e:t il laisse il a pensé qu' il veut faire une blague:	et eu:h	d' entrée il est f il es:t fermé	soufflé il a
18	donc eu::h [&=bche]	et il a (GH)	donc il ne peut pas sortir	il a &sen il sent
19	il es::t euh (GH)	il es::t	il a essayé il ne comprend pas qu' est ce qui passe: parce que peut être c' est la première fois qu' il était dans dans une:: maison	toutes les fleurs mais en fait il n' y a pas une odeur parce que: ils sont faux ils sont faux
20	il es:t eu::h (GH)	il a &li libéré le le cheval qui qui le cowboy	je crois que peut être c' était une	après eu:h les personnes:
21	ouais il	a eu::h	jeune: abeille parce que elle est	qui habitent là bas il est parti ou peut être c' est le vent qui a fermé la porte
22	il est &decon déconnecté le cheval euh de son rêne euh	a pris	elle était curieuse mais aussi elle *était un peu peur	donc euh il ne peut pas eu:h
23	donc le cheval es:t il est parti	et il a mis son *petite: cheval eu:h dans sa main	parce que après il y a une: (GH)	euh sortir
24	et il a une petite cheval euh	et donc quand le cowboy se lève:	une autre type de::: animal euh qui &s	et eu:h



25	pour lui pour le petit *enfante c' est juste comme un *jou le petit cheval	il a vu que son vrai cheval est est déjà parti	une mouche mouchoir ou mouche quelque chose comme ça	et après il a vu: une:
26	et donc il a *mette le rêne de: le vrai cheval avec le petit *jou	e:t il était un peu étonné parce qu' il y a une petite:	qui est qui est venu il est il est resté euh juste devant la fenêtre mais à côté *d' eau	une autre type de:
27	cheval	*jou d' enfante euh qui était dans sa main	et l' abeille il es::t dans la maison	insecte euh (GH)
28	et euh et après le *l'enfante est parti	et ça c' est l'histoire	et donc ils se regarde:nt cette mouche:: mouchoir	une
29	et le mec euh il se lève		et eu::h	mouchoi:r ou mouche une mouche
30	est réveillé		et eu::h il il veut sortir mais il ne peut pas il ne il ne sait pas	et eu::h
31	et il a vu que:::		donc après il euh juste regarde:: *le chambre	et là bas eu::h le mouche en fait eu:h
32	euh &b il il est un peu eu::h c' était un peu bizarre parce qu' il ne sait pas *qu' est ce qui passait		et les petites: mou::ches: (GH)	peut être il fréquente euh *le maison parce que il connaît *le sortie
33	et son vrai cheval est transformé *dans une		il connaît en fait bien parce que peut être il a passé souvent euh dans cette: [&=bche] cette maison pour trouver les choses à manger euh	euh de le maison
34	petite cheval pour les enfants		et donc eu:h il a pensé que il va_aider les petites abeilles	et donc il a aidé euh l' abeille pour eu:h
35	c' est ça je crois		et donc il est venu	pour euh partir
36			et il a pris les petits grains de::: (GH)	et eu::h
37			café	e:::t il a mis eu:h
38			e:::t il a: il a faire une petite signe	les petits grains grains de café (GH)
39			pour pour le sortie parce qu' en fait quand toutes les choses est sort est fermé (GH)	et: il a *faire une signe:
40			le: mouchoir il il *save que	pour euh pour le sortie
41			euh il y a une: petite sortie euh cachée	et après le abeille il eu:h il a eu:h

42			dans le:	il a (GH)
43			dan::s les	utilisé cette information et il a il a sorti euh la maison
44			les salles de bains le:s (GH)	e:t voilà
45			l 'endroit où le:s le eau sorte: il il sait que tu peux aussi tu peu:::x sortir de la maison avec ça	
46			donc euh il a il a mis une petite signe pour l' abeille	
47			et l' abeille est il a il regarde quand euh le mouchoir il il est prendre le café	
48			et pour faire son petit signe	
49			mais peut être il *était peur donc il il se cache ils ils se voient	
50			et après le mouchoir il est il laisse	
51			sortir peut être il *a resté dans la maison parce qu' avant il est dans la maison	
52			e:t le:: petit abeille quand il n' a pas vu euh::h le mouch:::oir pou:r un peu de	
53			temps	
54			il es::t	
55			il est allé dans le même *endroit et il a vu le petite signe	
56			pour euh sortir la *maison	
57			donc il a essayé de sortir aussi eu:h	
58			il *a sorti donc il était très:s	
59			contente je crois et le mouchoir il reste dans la maison	
60			et il a vu euh qu' il *a sorti bien donc euh c' était c' était la fin de l' histoire	

QI (niv 3)

Seg	BD1	BD2	Film1	Film2
1	y a un	un jour euh	euh tout à l' heure j' ai vu un film le nom de* film es::t le prisonnier	un jour euh (GH)
2	cowboy	cowboy rudy i:l se promène: avec euh son cheval euh	il raconte euh une	une
3	il s' appelle rouky	e::t	une tout* petit* histoire:	une peti:t zabeille euh une *petit abeille (GH)
4	et un jour euh	hop à la fin de l' après-midi il arrive dans un petit village il dit que:	euh sur un:: un petit abille* qui rentre:	qui se promène dans le jardin
5	il euh pren:d son cheval de: chez: lui: jusqu' à l' autre village	ah je suis *tellement de fatigué il faut que: je *fais une::	par hasard dans un* maison et	et puis il a vu:: une belle maison il dit que:
6	e::t après une longue journée de:::	*petit sieste (GH)	comment il il sortir à l'aide d' *un mouche (GH)	ah il faut que je rentre
7	voyage et il *sent très fatigué donc eu:h quand il a vu:	donc euh	tout d' abord euh	hum donc euh
8	à côté: de (GH)	il *descendre: il descen:d de:: sa: son:: cheval il a vu un: (GH)	cet *abille il se promène: dans un dans un jardin e:t	et il rentre:: dan:::s cette maison
9	d' une maison y a une: chaise il dit que ah c' est le momen::t de pause donc euh	un *banc sur contre un mur il me &f il dit que ah c' est le meilleur	il cherche: le::s	par la porte
10	il *descendre de::	endroi:t pour euh faire une sieste	les fleurs pour prendre: le::s	il rentre dans le salon
11	son cheval et il &as *assied dessus (GH)	donc euh (GH)	comment on dit ça euh pollen	il a trouvé que (GH)
12	il fait une *petit sieste:	il *a assis dessu:s e:t il baisse:::	et par hasard euh il a vu (GH)	y a des:::
13	e::t le tige de son cheval c' est toujours eu:h (GH)	son chapeau il commence à faire *le sieste	y' a une: maison	fleurs sauvages violets
14	dans sa main	à ce moment-là	dans *ce maison la porte est ouverte donc euh il rentre dans sa maison	dan:s un vase dans une vase

15	e::t	y a un garçon avec son: petit jouet sous forme de:: petit cheval	il se promène: dans des chambre:s des cuisines des salons et cetera	oh il est super content
16	à: à ce moment-là y a un petit garçon	arrive	et	il euh vole ver:s les fleurs
17	il est malin il est mignon	[&=bche] et il a	euh i i::	il sen:t
18	il est un peu méchant	il a vu Rudy	il quand il est en train de regarder le	les parfums ah ça sent super bon
19	e::t il vient avec son: petit jouet	qui est en train de faire une sieste	euh les fleurs dans le salon il entend tuQUE:	et puis il fai::t
20	en fait c' est un:::	il a des idées dans son:	euh y' a u:ne porte qui est fermée	il fait le tour &aut (GH)
21	*fausse cheval	dans sa tête il il dit que a:h je peux rigoler un peu je: je vais faire une bêtise	donc i il essaie de sortir tout de suite mais malheureusement la porte es:t	alentour des fleurs
22	e::t quand il a vu: rouky il est en train de: dormir e:t	donc euh	est fermée e::t (GH)	mais
23	le:: cheval il est juste à côté de lui	i:l s' approche de Rudy	il arrive plus de* sortir donc euh	à ce moment-là y y' a un coup de vent
24	il a une idée	il détache: le cheval de: Rudy euh	il est un peu mm:: *inquiété il il se promène partout il il vole partout pour voir: si y' a l' autre piste qui permet de sortir	la porte ferme
25	il enlève le tige de vrai: cheval euh	et puis il il branche: le fil de eu:h comment on dit le fil avec euh	et mais malheureusement y' en a pas	là bas y eu:h l' abeille:
26	e::t	son: jouet	et	est coincée dans euh dans la maison
27	il euh les fixe sur euh son:	et puis	à ce moment-là (GH)	donc eu:h il commence:: cherche:
28	son *fausse cheval	il laisse: le cheval de rudy eu:h partir	il a vu une fleur fenêtre comme: la vitre de fenêtre est transparent donc il arrête là bas il regarde dehors	il commence à chercher les sorties
29	e:t et après il s' en va et eu:h	et en même temps	et	il rentre: dans le dans la cuisine
30	le:: cheval de rouky i:l il se promène: n' importe où e:t	il s' en va e::t (GH)	en même temps y' a::	e::t (GH)
31	et après quand rouky il se &ré	après une heure	un mouche::	il a vu la fenêtre
32	réveille il trouve que	quand Rudy i il se réveille il trouve que où il est mon: mon cheval il	qui est à *extérieur qui s' arrête:	mais il sait pas que y a des vitres sur la fenêtre

		est parti et pourquoi ici y' a un jouet		
33	o my &g euh oh mon dieu où il est mon cheval	qu' est ce qui se passe	en face de:	il pense que (GH)
34	voilà	voilà	cette *abi:lle	selon euh via cette: fenêtre:
35	merci	c' est fini	les deux euh regardent *face en face	il peu:t sortir de cette maison
36			et après le *l'abille qui recule qui essaye de	donc euh
37			retrouver les les autres chemins pour sortir ( (GH) )	il vole vers le::
38			au début il a:	la fenêtre
39			il est attiré par euh	mais malheureusement y a::
40			les fleurES (GH)	des: vitres
41			mai:s euh malheureusement c' est pas les vraies &fl fleurEs c' est des fleurES	il fait le::
42			comment on dit ça artificielles (GH)	collision entre:
43			il sent rien et y' a pas de:: (GH)	bon y a une: &colli collision entre:: l' abeille et: la fenêtre
44			des odeurEs	euh il il a essayé plusieurs fois mai:s il n'a pas réussi
45			donc il abandonne	e::t
46			et	à ce moment-là à l' extérieur euh de la fenêtre:
47			il a vu aussi (GH)	y a mouch:e
48			&su euh contre *la mur y' a un tableau	qu'* il arrive:
49			qui::	donc euh juste a avan:t le::
50			qui a un dessin de champ de tournesols	l' abeille (GH)
51			il est super content il pense que il a trouvé le endroit pour sortirE donc euh	il il regarde:: (GH)
52			il vole vers le: tableau mai::s euh	à double sens

53			c' est la même chose il il touche le tableau mais sans sortirE	la mouche il regarde l' abeille l' abeille elle regarde la mouche
54			et (GH)	le problème c' est
55			quand il a vu le: tournesol il il a envie de prendre de:s (GH)	l' abeille trouve que la mouche c' est pa:::s mon amie et::: il est sale il n' est pas joli:
56			pollens de tournesol	e::t j' ai pas de point commun: j' ai pas de langue commun* avec lui donc euh
57			il eu:h	euh je veux pas parler: &a:: avec
58			il se promène plusieurs fois sur tun sur un fleur e de tournesol	euh je le veux je ne veux pas parle:
59			mai::s	je ne veux pas parler à::
60			sans avoir euh des résultats	à &s: à ce mouche donc il s' en va
61			et à ce moment-là il entend les bruits	il fai::t
62			et après il se cache: (GH)	le tour eu:h dans le sa:lle dans le dans le salle et dans le:: la cuisine aussi
63			dans &s euh à: derrière: de:: *de la cadre	il a trouvé des: j fleurs très jolies
64			il a vu	il s' approche de: cette fleur
65			la mouche:	mais
66			le mouche il rentre:	il arrive pas de* trouver le le parfum de ce mur euh fleur [&=bche]
67			et après	qu' est ce qui se passe il regarde à gauche il regarde à droite
68			euh le mouche i:: il rentre et il a vu	et finalement il il a trouvé une tout petite étiquette
69			sur e:::uh	il marque
70			une table de cuisine: (GH)	un virgule cinquante euro
71			y' a des *graines de café donc il prend des *graines de café et puis il *sortir par une: porte	oh: c' est des fleurs en plastique

72			et	bon c' est pas des vraies je m' en vais (GH)
73			il fait *aller retour plusieurs fois et finalement il *sortir	et après l' abeille es:t i::: qui rentre dans le salon il a vu
74			donc à ce moment-là	que: y a *des tournesones & tourne tournesols dan:s mm::: contre un mur
75			*l' abille: i::: *il veut sortirE	il est super content il dit que ah j' ai plein j' ai trouvé plein de nourriture
76			il suivit le le:::	j' arrive [xxx] il vole vers le:: le tournesol
77			le:::	mais malheureusement
78			le::: trajet de:::	c' est un: table
79			mouche:	donc euh c' est pas des vrai:s *tournesones
80			il arrive dan:::s *unsalle de bains	comme la dernière fois:
81			et	y a une collision entre les abeilles et
82			il il euh [&=bche]	les tournesols [&=bche]
83			il s' arrête: dans un baignoire	finalement (GH)
84			et	l' abeille i::: il trouve que c' est pas des vrais tournesols
85			il a vu	et: en même temps
86			juste: (GH)	il entend les bruits de:::
87			comment on dit ça dans le::: (GH)	mouche
88			la sortie de la baignoire	et après
89			là bas un trou ça permet de: (GH)	il dit que comment i rentre
90			éliminer de l' eau usée et cetera	je vais voir
91			et	donc euh l' abeille i:::
92			à sa	il se cache:
93			&s	derrière: de::: la:::

94			sortie de baignoire il a vu une petite *flash flèche:	de de la table
95			qui est construit par euh des *graines de café	non de la tableau voilà
96			en fait c' est le: c' est le mouche qui est très sympa qui lui guide	e::t
97			pour lui dire que c' est tu partis par euh	il regarde
98			c euh ce chemin tu peux sortir donc euh	bon (GH)
99			*l' abille:	le mouche il rentre dans la cuisine
100			il rentre dans le trou et finalement il euh	il pren:d
101			il retrouve il retrouve le:: la nature par un robinet de:: l' eau	un *graine de::: un grain de::
102			et et après il retrouvait son liberté voilà	non une graine de café
103				et après il s' en va
104				il revient
105				il *reprendre une autre *graine de café
106				il s' en va
107				aprè:::s
108				les: allers retours plusieurs fois comme ça
109				euh la mouche:
110				est sortie
111				il sort
112				donc euh
113				l' abeille il le suit
114				e:::t
115				l' abeille il arrive dans le salle de bains euh (GH)
116				dans le lavabo il a vu



117				un:: un flèche
118				qui es:::t [xxx]
119				conçu par eu:::h des graines de café
120				a:h c' est la mouche il a fait ça pour euh lui dire que
121				ici c' est la sortie bon
122				l' abeille il *prendre: (GH)
123				comment on dit ça le::: (GH)
124				l' &ab euh (GH)
125				l' abeille il rentre dans le trou (GH)
126				qui permet de évacuer de l' eau
127				voilà et qui sort euh
128				il rentre dans le système de conduite
129				et finalement il sort euh de:::
130				de:: d' un robinet
131				à extérieur de: la maison finalement euh (GH)
132				l' abeille il i: il a retrouvé *son liberté

SU (niv 1)				
Seg	BD1	BD2	Film1	Film2
1	il *avait un cowboy	il *avait un cowboy	il *avait un petit insecte	*le petite histoire raconte l' histoire d' une petite abeille
2	cowboy	bon	je crois qu' il cherche quelque chose	qui cherche le pollen
3	e:t il est arrivé de son: de son boulot	il est arrivé je crois d' une: [&=bche]	je sais pas quoi je crois que c' est la nourriture	pa:r

4	e::t il était très fatigué [&=rires]	d' une promenade	e::t il était dans la &for dans la forêt après il est rentré dans la maison	dans son &ja un gros jardin euh (GH)
5	donc il avait sommei:l	il était fatigué	pour chercher les fleurs pou:r mettre son:	pendan::t s::a:: sa journée
6	e::t	il a envie de dormir	je sais pas quoi [&=rires] son pollen je sais pas	elle a::
7	il: dort	donc euh	e::t il commençait à à chercher dans la maison une fleur	elle commence à chercher le pollen
8	sur un* banc	il a *dormait su:r un *banc (GH)	il *seulement rencontrer les fleurs at at artificielles	dans une: maison
9	e::t	avec son cheval	et après quan::d	elle a elle rentre elle est rentrée elle est entrée dans une maison
10	un un petit garçon est arrivé	il y avait un &ga un petit garçon qui arrivait	il veut il étai::t dans la maison fermée	e::t elle commence à chercher partout dans cette maison le:: le pollen dans les fleurs
11	et je crois que le petit garçon *soyette un chevau	avec son jouet	il ne peut pas sortir	dans les fleurs &s dans les *sasso dans les salons sur une table basse
12	un chevau::	e::t	e::t (GH)	elle cherche aussi le:: [&=bche]
13	un chevau:	que c' était un un petit cheval aussi	hors de la maison il a rencontré	le pollen dans la cuisine
14	pas le	et je crois qu' il a: envie de d' avoi:r	autre insecte	sur la table où *il avait les fruits
15	pas comme le le::	un grand un: cheval plus grand	qu' il *vut qui *vut qu' il étai:t dans la maison	donc eu:h elle cherche partout et après elle s'est rendue compte
16	le chevau que le: petit garçon avait	il a:: il a profité que le cowboy il dormai:t il a changé le: le petit cheval	et l' autre &inse::	que: la porte que:: elle est entrée
17	que c' était un un un chevau pour les enfants	pou::r	avec le chemin de de café	la porte elle est fermée
18	e::t le petit garçon	le grand cheval du cowboy	il essayait un &che un chemin pour sortir de la maison	et après elle::
19	il::: [&=bche]	il a pris le cheval du cowboy il a laissé son petit cheval	e::t	arrive pas à sortir de la maison
20	il &lai il laissait	donc quan::d le::* le cowboy il est il est réveillé	il a:: il souvient	et mais quand même elle continue à chercher

21	son son petit chevau avec le cowboy	il a vu que son: cheval	le chemin	le pollen sur les fleurs
22	il prend il prit les::	était pas là	et après	et aussi elle s' est rendue compte que les fleurs sont pas les fleurs naturelles
23	il a pris le: le chevau de de cowboy	il avait sur sa main	il est sorti de la maison	sont les fleurs artificielles
24	et le &cow cowboy quand il est réveillé	un petit cheval	il était libre	elle cherche aussi le pollen sur une toile
25	i:l il *vut	voilà un jouet d' un petit enfant		que c' était sur le mur
26	que son:::			et après (GH)
27	son chevau était pas là			il *avait une mouche
28	c' est tout je crois			que:::
29				il a vu sur la fenêtre
30				et la mouche: elle a:
31				*percevu
32				perçu
33				elle a perçu
34				que la: petite abeille était *emprisonnée dans la maison
35				qu' elle pouvait pas *sorti
36				donc eu:h la petite mouche
37				el:le aide la:: la petite abeille à sortir de la maison
38				elle cherche elle a:
39				elle est entrée dans la maison aussi
40				par un &chem un chemin:
41				un chemin
42				et elle: *faire *la: indication de *cet chemin

43				pour la petite abeille avec de::
44				de *semence de café
45				e::t après la petite abeille elle::: sui:t cet chemin
46				pa:r
47				pa:r la baignoire
48				e::t c' est comme ça qu' elle est sortie de la maison par un tuyau
49				e::t le chemin (GH)
50				par le tuyau elle est sortie dan:s &s
51				par un robinet
52				dans un jardin
53				et c 'est pour c' est comme ça qu' elle:
54				qu' elle a réussi à sortir de la maison

VI (niv3 )				
Seg	BD1	BD2	Film1	Film2
1	bon le:	bon	bon eu:h c' est l' histoire de d' une abeille	eu:h là c' est c' est l' histoire de:: d' une abeille
2	le: c' est l' histoire de d' un homme que:	il y a un homme	*qu' elle est dans un jardin	qui se promène parmi les fleurs
3	qui est avec euh son cheval	qui est avec son:: cheval eu::h à la campagne peut être (GH)	euh devant une maison elle cherche parmi les fleurs	euh à un moment donné elle: se trouve devant une maison
4	mais euh je pense qu' il est un peu fatigué alors il décide de: s'	là euh il se sent fatigué	elle cherche le la bonne fleur pou:r qu' elle trouve le (GH)	avec la porte entrouverte

	endormir un peu			
5	il décide de dormir eu:h (GH)	et comme il est tout seul il décide de: de faire une petite sieste	bon pour qu' elle puisse fabriquer du miel	elle décide d' y entrer eu:h (GH)
6	alors il euh laisse euh il attache son cheval	là le il accroche eu:h il attache son cheval *à le	et bon alors qu' elle: qu' elle cherche pendant les fleurs elle	e::t elle commence::
7	e::t i::	à une maison je crois	el:le elle entre d dans la maison parce que la fenêtre est ouverte	à chercher les fleurs qui qu' il y a dedans
8	bon euh voilà il se:::	et il dort (GH)	et après le::	mai::s après
9	il dort un peu	c' es::t bon à	bon mm je sais pas pourquoi la fen la fenêtre se se ferme	elle ouvre elle elle entend un bruit
10	mais alors qu' il dort eu:h	puis eu:h il y a un::	et euh bon alors elles se:: elles se rencontrent &renferm enfermées dans la maison	et c 'est la porte *que:: se ferme
11	il vien:t euh un un p'tit garçon	un:: un garçon qui arrive	elle sait pas vraiment comment elle va sortir de: (GH)	peut être par le vent je sais pas
12	avec un: p'tit cheval un p'tit animal j' sais pas	un p'tit garçon	comment qu' elle comment elle va sortir	mai:s après elle:: elle se bat contre:: elle se bat contre
13	c' est un chien euh je	et il eu:h voit le::	et bon (GH)	toutes le vitres de des fenêtres pour essayer de sortir
14	j' suis pas sûr mais	le:: le: le:	alors &k (GH)	mais elle n' y arrive pas
15	euh alors que le le: le l' homme dort euh le le p'tit garçon le il il euh change: euh il il prend son le son cheval euh	l' homme qui dorme qui dort	un moment donné elle: elle regarde:: par la fenêtre une mouche	e:t après il y a::: un moment quand elle est devant la fenê une fenêtre (GH)
16	il laisse le le chien	i::: il veu::t	elle trouve une mouche que::	elle voit une mouche
17	alors euh attaché euh où était le cheval d' abord	il a envie de de:	je pense qu' elle connaît le le chemin pour sortir	qui est de l' autre côté de la fenêtre de à l' extérieur (GH)
18	alors quan:d quand l' homme se réveille il voit que:: le son cheval n' est pas là e:::t	de de: lui emprunter son cheval	de la maison	bon (GH)
19	*au lieu il y a un *cheva un un chien euh (GH)	alors eu:h mais sans que lui il sache	alors	toutes les deux elles se regardent mai::s elles décident de de elles continuent à

				essayer de trouver euh
20	bon c' est c' est c' est drôle: euh ça ça lui donne l' impression que alors qu' il dormait euh	mai:s alors eu:h	euh la mouche: fait une indication avec de des grains (GH)	un: une façon de sortir de la maison
21	son cheval s' es:t	il eu::h	comme une flèche pour que l' abeille puisse trouver le le la sortie (GH)	mai:s c' est *en ce moment-là
22	bon a changé euh bon	il s' assure que l' homme il dort parce qu' il le touche: avec un petit bout de:	comme ça le le l'abeille sort par par *le baignoire	qu' elle voit un: une peinture de de tournesols d' une plantation de tournesols
23	c 'est c' est quand même drôle	de:: de bois	euh on et	et* elle eu:h je crois qu' elle:
24		je crois	et voilà elle est elle est hors de de la maison	qu' elle pense que c' est c' est une ouverture c' est c' est la sortie
25		et lui il le il le:: il voit pas que que le garçon il est là		et:: elle s' envole: elle elle vole elle elle
26		alors ce qu' il fait il détache le:: le cheval		juste eu::h contre cette peinture mai:s toujours elle arr elle arrive pas à sortir elle se bat contre cette (GH)
27		i:::		cette:
28		il le remplace par un p'tit chien		peinture
29		il s' en va avec le cheval (GH)		e::t après il y a la mouche qui entre dans la maison
30		donc après le::		par une façon bon::: dans ce moment &ap jusqu' à ce moment-là on connaît pas
31		quand le::: l' homme il se réveille		et elle commence à chercher des grains de café
32		au lieu d' un cheval euh il voit un chien attaché eu:h à lui		et à:: les les emporter su euh dans un autre endroit de la maison
33		il comprend pas c' que c' que c' que c' est arrivé		après quan::d
34		il euh: peut être il a l' impression que le: le cheval eu:h		elle fait un signe je crois qu' elle fait un signe:
35		s' est transformé *dans dans un chien		euh sonore

		[&=rires]		
36				eu:h la: et l' abeille elle elle::
37				elle va: elle va voir ce que ce c'est ce que la mouche fait
38				e:::t* elle &ar quand elle arrive dans le le::: la salle de bains
39				elle voit sur le::: sur le:::
40				sur la baignoire le: u:ne une flèche
41				qui indique la sortie que c' est eu:::h c' est le trou eu::h par le:
42				l' eau s' écoule: je sais pas comment dire ça
43				et alors euh elle entre dan:s dans ce trou
44				et
45				et à la sortie c' est c' est vraiment le robinet c' est un robinet qu' il y a en dehors de la maison
46				euh à l' extérieur
47				et là elle est libre

VO (niv 3)				
Seg	BD1	BD2	Film1	Film2
1	donc je vois un dessin (GH)	donc euh dans ce BD on voit une histoire d' un cowboy qui:	un autre épisode des minuscules qui s' appelle prisonnière	donc dan::s ce:: [&=ahem]
2	qui raconte peut être un une petite blague	arrive: dans une village	il raconte il s' agit d' une histoire d' un	dessin animé [&=bche]
3	e:::t il y a &=cl	il descend de:	d' une abeille [&=bche]	on aperçoit une histoire d' amitié de deux insectes
4	*sur le premier plan il y a:	son:: cheval il a l' air fatigué donc il s'	on aperçoit	qui commence par le fait qu' une:

		installe		abeille
5	un:: un homme qui:	sur un banc	sur les champs qui qui(GH)	[&=ahem] rentre dans une maison dans la campagne
6	qui fai:t du vé à	il tien:t les fils euh de	qui: (GH)	attirée par euh
7	pas du vé qui fait du cheval	de: son: cheval euh dans une dans sa main	comment dire	les odeurs eu::h
8	e::t après il euh &=cl	ensuite on voit un enfant	qui cherche: de::s	des fleurs d' abord elle entre:
9	descend de son cheval	avec euh un jouet du bois aussi de cheval	des des fleurs [&=bche]	dans la salle: eu::h
10	et il a l' air fatigué	qui s' approche *à ce cowboy	voilà	dans le salon
11	il euh s' installe: sur le banc	et donc euh pour euh	et après on voit qu' elle s' approche vers une maison	attirée par les odeurs de fleurs
12	et il fait une p'tite sieste &=cl	pour rigoler apparemment il eu:h	à la campagne	après elle passe: à la cuisine
13	et le p'tit gamin arrive	il eu::h libère le cheval d' un cowboy	elle: elle entre dedans [&=bche]	où: il y a des fruits eu:h sur la table [&=bche]
14	avec eu:h	qui part en galope et::	e::t elle trouve des	ensuite: (GH)
15	le:: jouet	donc il change les fils euh ah de il met	fleurs elle	on entend euh le bruit
16	de cheval	le fil euh de son jouet	trouve des des fruits [&=bche]	de la porte qui se ferme
17	e::t &=cl	[&=bche] dans une main de cowboy	mai::s	donc euh on s' imagine que
18	en faite* il prend les	et donc euh quand le cowboy se réveille	après	que une petite abeille
19	les fils de ce cheval (GH)	il voit que son cheval a::: disparu et il voit	elle n' arrive pas à sortir de la maison	est enfermée dans une maison
20	de cet homme	que il tient le fil euh d' un jouet	elle essaye de:: de::	elle essaye de sortir par les fenêtres mais toutes les fenêtres sont fermées
21	et il la libère	donc il se demande qu' est ce qui s' est passé comment	sortir par de sortir	[&=bche] donc eu::h c' est impossible pour elle
22	et au lieu de:: son truc il met le::	le: son cheval vivant *pourrait être transformé en jouet	par des fenêtres	e::t elle a l' air désespéré d' après le son: qu' elle qu' elle fait
23	le fil de son jouet &=cl		dans le *selon salon	elle a l' air désespérée [&=bche]
24	donc [&=rires] (GH)		après à la::	ensuite:



25	quan:d ce cowboy il se réveille (GH)		dans une *kitchen	elle s'installe: sur un bord d' une fenêtre
26	euh il voit que au lieu de: de son cheval à lui		cuisine et cetera mais elle n' arrive pas toutes les fenêtres sont fermées [&=bche]	et on voit qu' il y a une mouche
27	il y a un p'tit jouet		e::t après il s' installe	un: eu::h de l' autre côté de la fenêtre donc à l' extérieur
28	voilà		devant:	qui: vient voir eu::h une abei:lle e::t ils se regardent
29	donc il se demande peut être		une fenêtre	dans les yeux pour un p'tit moment
30	et où il est mon cheval il &s il s' est transformé en jouet ou quoi		il voit *à l' autre côté	ensuite apparemment une mouche comprend le problème:
31	c' est tout [&=rires]		dehors	d' une abeille [&=bche]
32			une	de cette &a
33			mouche [&=bche]	abeille donc eu:h (GH)
34			qui:: qui qui le regarde	quelque:
35			qui le regarde	temps plus tard on voit que
36			et voilà e:t cette mouche elle voit que	cette mouche eu::h s' est retrouvée
37			pauvre:	également dans cette maison [&=bche]
38			abeille elle n' arrive pas [&=bche]	e::t
39			à partir	donc eu::h
40			e::t voilà et après une abeille	en fait
41			elle es:t	oui
42			elle réessaye elle réessaye de	avant que que cette mouche se &re
43			de de trouver un moyen pour partir	se retrouve
44			elle voit [&=bche]	dans cette maison on voit que
45			*une tableau	une abeille: voit
46			avec des tournesols	aperçoit [&=bche]

47			il croit que c' es:t c' est des vrais tournesols	un tableau
48			et quand elle s' approche	avec les tournesols
49			*au tableau [&=bche]	elle est avec les champs de tournesols
50			elle::	elle croit que c' est de::s bon
51			elle se casse un peu	c' est un vrai champ donc elle &es
52			e::t après elle tombe	essaye de: partir
53			elle réessaye donc euh c' est	dans ce champ là mais en fait elle se casse la gueule
54			c' est un piège comment ça elle	je rigole parce que c' est un tableau
55			jamais euh vu que tournesols [&=bche]	c 'est pas
56			il sont pa:s	un un vrai champ et donc ensuite il y a une mouche qui rentre
57			hum des vrais	et on voit elle est elle est entrée depuis (GH)
58			ils sont *fausses	depuis la salle de bains [&=bche]
59			donc et après elle voit que:	la mouche
60			une mouche	elle voit on voit que::
61			elle a réussi	elle::: elle vole dans une cuisine
62			aussi d' entrer dans la maison [&=bche]	e::t elle: prend le:s grains
63			e:::t	du café et en fait elle fait
64			une abeille elle: [&=bche]	plusieurs allers retours (GH)
65			se déplace au en haut de du tableau pour	et:::
66			observer c' qu' il va faire	ensuite on entend [&=bche]
67			une mouche	le
68			une mouche on voit que	brui::t de:::s
69			elle:: mm [&=bche]	je veux dire des horloges

70			elle trouve des grains *du café	qui signifient peut être:
71			et elle commence *de les (GH)	la fin de::: de quelque chose on sait pas encore quoi
72			transporter	et donc dès qu' on entend ce: son
73			dans une autre chambre	cette abeille elle:: elle entre dans la salle de bains [&=bche]
74			on voit on observe (GH)	et ensuite on aperçoit
75			une porte [&=bche]	*une bain
76			euh dans une autre chambre	e::t au sol
77			donc euh e:t il prend des grains un par un	de ce bain
78			les	on voit une flèche
79			amène dans une autre chambre	qui est *fait
80			on voit pas	du::: des grains
81			pourquoi	du café
82			et après	et donc elle &f cette flèche elle::: [&=bche]
83			une abeille	elle montre: le:: trou
84			elle suit	le trou de: canalisation eu:::h dans cette* bain
85			cette mouche	et donc eu::h [&=bche]
86			et elle se retrouve dans la salle de bains	cette abeille:: elle: elle entre dans ce trou
87			e:::t [&=bche]	et on voit que:
88			elle se: déplace vers le:	en fait elle sort de cette maison par robinet
89			vers une douche	donc elle se retrouve
90			et elle voit	à l' extérieur
91			que	l' extérieur donc euh

92			au sol	cette abeille est sauvée (GH)
93			au sol de cette douche	elle:: revient vers la maison
94			y a une flèche	et maintenant on aperçoit que c' est
95			qui montre une direction	la mouche qui est enfermée
96			direction	dans une maison
97			direction	e:t là l' abeille elle est à l' extérieur
98			dans un robinet [&=bche]	mais euh
99			et voilà e::t	on *t' inquiète pas [&=bche]
100			après il suit	pour cette mouche pas mouche
101			cette flèche donc elle rentre dedans	oui
102			e::t grâce à ça	parce que c' est elle
103			elle: elle a réussi de sortir	qui a indiqué le chemin donc on se dit que elle pourrait sortir sans problème
104			de cette maison [&=bche]	
105			voilà et elle est très contente que	
106			elle s' est libérée	
107			elle revient	
108			vers la maison	
109			cette fois on voit que	
110			c' est une mouche qui est dedans	
111			voilà mais	
112			cette mouche elle sait comment sortir de cette maison donc	
113			on n' a pas peur pour cette mouche	
114			va	
115			elle va s' en sortir [&=rires]	

XI (niv 1)

Seg	BD1	BD2	Film1	Film2
1	euh j' ai vu un j' ai vu un homme qui est euh	une	eu::h dans eu::h dans un jardin	une abeille: (GH)
2	sur un cheval euh	un jour	i euh::	euh il voit le:
3	il il veu:t il veu:t se	un homme:: &k	i i il y a une: abeille:	dans dans *une: jardin
4	se:: &repo reposer	un homme eu::h	qui euh qui::	de la &m de la &mais euh d'une maison
5	sur eu:h	*assois sur un cheval	volait	et eu:h la porte de la maison euh (GH)
6	sur un:: un (xxx)	il est passé par eu:h	il vole qui vole	n' est pas n' est pas bien un bien fermée
7	j' sais pas (GH)	il passait par euh il passait par une maison	e::t	du coup eu:h cette *beille:
8	euh il est il était très il était très fatigué et eu:h	il trouve qu' il est	elle a elle a vu (GH)	il est euh elle est rentrée
9	i::l &s::: se repose ::	qu' il est	dans dans une dans un cette non dans un jardin	elle est eu:h
10	après il *est dormi (GH)	il trouve	dans dans son jardin il y a une maison et eu:h	elle est
11	*ce moment-là un un un garçon	qu' il est fatigué	et eu::h	rentre:
12	qui est passé	e:t il veut *reposer un peu (GH)	la porte de de la maison est ouverte	dans cette maison
13	là bas	du coup euh il euh	donc eu::h l'abeille	et eu:h
14	et euh il a: il a vu que: cet euh ce ce ce ce nomme est en train de dormir	il attache: il attache:	entre: entre: entre:: la maison	elle eu:::h elle eu:h elle vole:
15	eu:h il veut voler sa* cheval euh son cheval	son cheval	par eu::h	elle euh
16	mais eu:h le le la renne de de cheval euh	euh à côté de::	par (GH)	elle vole:: partout
17	es::t dans la main de::	à côté de cette maison (GH)	entre la maison eu::h (GH)	dans la maison
18	de de son_homme (GH)	et euh il repose:	je eu::h peut être: l' abeille:	et euh elle *sentit des::

19	et	il *assoit sur une chaise:	il euh il sent le il sent le le le le le:	de::s des pollens
20	eu:h ce ce ce petit garçon es::t	juste &av juste euh devan:t devant cette maison	le le pollen pollen	et eu:h
21	* est changé le:::	mm parce qu' il était vraiment fatigué	*du fleur	et eu:h soudain
22	la rêne	du coup il dort eu::h pendan::t euh il dort	des fleurs	& s euh la porte de cette maison
23	[xxx] (GH)	sans *fait attention	euh il euh il vole:	a été fermée
24	eu:h ce: ce garçon a changé:	à: son cheval (GH)	dans::	et (GH)
25	a changé *la rêne de: de son cheval	*en ce moment (GH)	il vole il vole: partout dans euh dans cette maison	cette *beille:
26	par eu:h celle de:::	une &ga	et soudain la porte::	il a entendu euh
27	celle de lui	un garçon (GH)	*il ferme *tout seul	il il sait que: la porte a été fermée
28	en fait *cet garçon a:::	à un garçon qui:	e::t cette: abeille:	du coup elle étai:t euh elle était euh elle est eu:h stressée
29	a	qui est passé	et si et cherche::	et euh elle veu:t
30	un cheval	par euh cette maison	cherche: la &sort	elle veut trouver la sortie
31	en jouet	i eu:::h	cherche::	elle: elle vole partout et elle s::e elle se::
32	donc euh il il lui a changé avec eu:h ce: *ce homme (GH)	il voit euh son cheval euh (GH)	la sortie	*heurter
33	pendan::t ce homme es:t en train de dormir	il trouve *cet cheval euh il trouve ce cheval euh	il veut il veut sortir de la maison euh (GH)	elle se heurte: euh (GH)
34	euh après quelques minutes eu:h ce euh ce ce nomme:* es:t	est vraiment cool euh en fait (GH)	i il euh il vole partout euh	contre: la fenêtre
35	se réveille	et il veut changer (GH)	mais il ne trouve pas le la s&sort	elle cherche elle a elle a essayé eu:h
36	et i:l	son cheval &s il veut changer son cheval avec eu::h	il ne trouve pas la sortie	de chercher la sortie mais euh
37	il il a *trouve eu::h	oh il veut changer	et il eu::h	elle trou eu:h elle ne trouve pas (GH)
38	il euh il a il a *trouve que:::	non il veu:t	il eu::h	donc euh elle s' arrêtait (GH)
39	il a *trouve que /s/	voler	parce que:	elle s'arrêtait eu:h

40	son cheval *est disparu	ce cheval avec eu::h	euh il il eu::h	juste à à côté de de la fenêtre:
41	et e::t dans sa main	son cheval en bois	*comment dit cassé non	et euh elle a vu (GH)
42	eu::h dans sa main il y a un:	e::t	veut i	elle a elle a vu eu::h (GH)
43	euh *une rêne *une rêne de: cheval en jouet	du coup cet:te euh du coup ce garçon il a changé euh la laisse	pour euh chercher euh la sortie euh il a essayé (GH)	*en ce moment
44		il euh	i il a essayé eu:h	elle a vu eu:h une mouche
45		&en *enlevé	il a il a &essa il a essayé trouver (GH)	qui eu:h qui s' arrête à l' extérieuR de la fenêtre
46		la laisse: dans la main de:: cet	dans dans dans dans la salle de cuisine et dans la salle de cuisine: il y a des fenêtres partout euh il euh	et euh (GH)
47		ce homme	et dans la salle de cuisine: il y a des fenêtres partout euh il eu:h	elle regarde cette mouche (GH)
48		avec eu:h la laisse de::	j' sais pas euh [&=rires]	pendant quelques secondes (GH)
49		de la cheval	et euh	il semble que cette mouche: (GH)
50		en bois	parce que il ne trouve pas et euh	cette mouche:: elle parle avec eu::h
51		parce que: ce monsieu:r	i il ne trouve pas la sortie euh il regarde il euh il tourne sur euh les les fleurs euh	avec euh cette *beille
52		il était vraiment fatigué	e:t *ce moment-là il a vu euh: une tableau de	mais euh cette *beille:
53		il a il a pas senti quelqu' un	non	elle ne comprend pas ce ce que::: la mouche: &k qu' elle a qu'elle a *dise (GH)
54		qui es::t	non	donc (GH)
55		qui a changé (GH)	pas encore	donc euh cette *beille
56		cet &ga euh *cet garçon (GH)	non	a::: reessayé de: trouver eu:h
57		euh il a enfin cet* garçon il a &r	euh *ce moment-là il euh	euh reessayé de chercher
58		* réussite::: *change: de: changer	euh &j je pense que il est fatigué et il eu::h il se repose: sur eu:h	la &sor eu:h la sortie (GH)

59		de de voler euh *cet cheval	euh sur eu::h	elle::
60		et eu::h	à côté de de la fenêtre	elle:: s:e vole sur
61		après eu:h après quelques minutes (GH)	et::	autour euh d' une:: euh des fleurs artificielles
62		ce monsieur	il a vu euh un mouche	et euh elle a vu euh euh une table
63		il euh il se:: réveille	euh une mouche:	un &d tournesols
64		e:t i::	qui est à l'extérieur euh de la fenêtre: euh (GH)	elle croit que c' est un:: c' est c' est des (GH)
65		il a il a *trouve que::	ils sont	une::: c' était un un tournesol euh vrai
66		l::	euh je pense que ils son::t (GH)	et elle::
67		il a *trouve le: le cheval	non ils se regarde:nt	elle:: s' arrêta:it sur euh c:::e tableau
68		la laisse d' un cheval non il a trouve euh &s::: son cheval a été volé	ils se regarde:nt (GH)	et euh en ce moment
69		et eu:h	eu:::h l' abeille: reste: là ba:s	cette mouche euh la mouche euh (GH)
70		le cheval la laisse non	quelques minutes après il eu::h	rentre: dan:s dans cette maison
71		la laisse	euh: il euh il euh revole	non
72		le cheval est devenu un:: &ch	*e:t ce moment-là i il a: il a vu euh	cette euh la mouche::
73		un cheval en bois	il a vu une tableau de tournesols (GH)	la mouche est venue
74		c' est bon merci	il eu::h	euh dans cette maison
75			il vole:	et euh (GH)
76			sur le s s sur e ce tableau	euh cette euh l' abeille:
77			e:::t	l' abeille: elle trouve: pourquoi il est pourquoi &s euh la mouche:
78			il tourne: il tourne::	est rentrée dans cette maison
79			après euh après euh &j (GH)	et du (GH)
80			après	cette *beille euh cette: non cette abeille: (GH)
81			quelques minutes:	*était un peu: eu:h peur de la mouche



82			cette: cette mouche:	donc eu::h il se cache elle se cache:: sur eu::h sur le &tabl euh sur eu::h le côté de::
83			i:l	de::
84			il es::t	du:: tableau tournesols
85			il est *apparaître::	et elle eu::h &s
86			il il vole:	cette \$beille regarde:: qu' est ce qu' il
87			dans cette maison et il euh transporte:	le qu' est ce
88			il euh	qu' est ce que: la mouche il fait eu::h la mouche
89			il *transport le le le *pois de: de café	il vole il eu::h il déplace des:: *pois de café eu::h (GH)
90			il eu::h	euh la mouche il déplace des pois de café euh par
91			il *transpor euh le pois de café et	un par un (GH)
92			et l' abeille (GH)	*en face en face euh la mouche il ne revient plus euh (GH)
93			et aborde: cette mouche: donc il euh	cette *beille:
94			il eu::h	elle veut savoir euh qu' est ce que qu' est ce que la mouche qu' il fait
95			cache il se cache: (GH)	la mouche: fait
96			il se cache sur euh sur euh	euh fait fait fait fait quoi [&=rires] (GH)
97			le: s s sur le sur euh tableau sur euh le tableau	et euh
98			et il regarde: qu' est ce qu' il a fait euh cette mouche	donc eu::h
99			et eu::h cette mouche: (GH)	cette abeille:
100			partie et retournée plusieurs fois	*il es::t (GH)
101			il a transpor euh quelques il a transport quelques pois de &caf &ca euh café (GH)	elle es::t

102			l' abeille: se (GH)	elle es::t
103			l' abeille se euh trouve que c' est bizarre:	elle *est vole
104			donc eu::h	dans la salle de:: dans la: dans la *salle du bain
105			il attende quelques minutes et cette mouche: ne re ne revient plus	et elle a euh elle a trouvé
106			et euh l' abeille veut voir qu' est ce qui se passe eu::h (GH)	une: une flèche: sortie
107			il euh	en mm en u
108			il euh il euh il euh sorti euh la porte que:: la mouche	euh une flèche en sortie
109			par par laquelle la mouche: est sortie	qui::: qui *son:t
110			e:t e:t	qui son:::t eu::h qui son:::t dessinée par eu::h par la mouche (GH)
111			et l' abeille: (GH)	par la mouche
112			l' abeille: &s	en utilisant le la *pois en utilisant les *pois de café (GH)
113			trouve: trouve les pois de particule dans un bain	au début (GH)
114			dans un bain	l' abeille:
115			euh dans un	elle était un peu::: toute:
116			bain	est ce que on peu::t
117			dans le toilette dans un bain	elle aidait à
118			e:t euh il voit une: flèche:	elle aidai:t
119			pour	elle s' arrête: su:r la flèche
120			ça veut dire: tu peux euh (GH)	et eu::h
121			une flèche sortie	enfin:
122			comme ça et euh	mais finalement (GH)

123			il il euh existait euh quelques minutes m	l' abeille:
124			par contre: il euh il euh il euh rentre:	et: rentre le trou de:: bassin
125				e::t finalement
126			il se	euh l' abeille a réussi de sortir euh
127			non	cette maison
128			il rentre: dans le canal de:: du du bain	
129			du bain	
130			e:t enfin il il *a sorti de la maison	
131			voilà	

**ZI (niv 2)**

Seg	BD1	BD2	Film1	Film2
1	hu:m j' ai vu un monsieur	hum j' ai vu une bande dessinée dans la	euh j' ai vu un film je pense c' était une animation il y a une	um j' ai regardé le film que j' avais vu
2	sur une &ch sur un cheval (GH)	dans le dessiné il y a une euh	un animal	des mois d' avant et euh
3	il a:	monsieur un monsieur sur un cheval	dedans et euh cet animal:	cette fois cette fois j' ai remarqué que:
4	il s' arrête eu:h	il a:	comment dire il vole c' est petit euh il vole et il a	c' est eu:h
5	um chez une chez une ferme e:t	je pense qu' il est fatigué car il eu:h	entre une:	une abeille
6	il va hum faire une:	il laisse son euh *chevau cheval	maison une maison belle et	se trouve dans un petit village et euh
7	*une sommeil faire il va dormir	e:t après il va il va dormir un peu sur un p'tit euh	une belle maison et euh	il euh voit un jardin et euh
8	il va dormir et après il y a	*banc	il va euh voler tout euh	et eu:h
9	vient	mais eu:h après qu' il euh	et euh	une très belle maison et il entre le la maison
10	un::	il *a endormi il vient un un [xxx] un p'tit garçon	il voit	il euh vole un petit peu dans la maison il regarde: dans la cuisine et dans le

				salon
11	il y a un garçon un p'tit un p'tit garçon et il a un une p'tit cheval	et le p'tit garçon il a un jouet	il veut	et après il veut
12	chez toi	un jouet et euh le jouet c' est un p'tit euh	aller dehors	il entend quelque chose et il se dit oh je doi:s
13	chez soi et eu:h il euh [&=bche]	un p'tit cheval euh de	mais il ne peut pas et euh il:	je dois sortir d' ici
14	il laisse le cheval le grand cheval	de marrant je pense	cherche:	et il veut sortir mais euh
15	cheval aller (GH)	et il *vienne avec ça et eu:h	dehors et il a *trouve:	il voit les euh vitres de:
16	et là comme le grand cheval est parti	il regarde le cheval et eu:h	les euh	il voit les vitres et il pense ah c' est dehors je peux sortir mais bou:m il se trouve contre les vitres la vitre
17	il a: *laisse le p'tit cheval là et il euh	il euh	vitres et il euh veut	il &re il euh essaye quelquefois les les autres vitres mais euh
18	fait	le cheval était euh	aller	& tou:: c' est euh toujours eu:h
19	&re il euh réveille le monsieur	lié avec une corde	ah dehors donc euh	clos
20	et le monsieu:r est choqué par son p'tit euh	et mais euh	il va tout euh (GH)	c' est toujours fermé
21	cheval	le garçon eu:h	il va	et euh après il euh *se réalise que:: il est euh
22		tire:	il va contre les vitres chaque fois qu' il veut aller	fermé il eu:h ne peut pas sortir
23		la corde	mais euh après il euh	et il se trouve triste et il va
24		des mains de:	voit *une	il va:
25		de monsieur et après il euh	tableau	*s' asseyer euh il va se
26		il euh prend ses euh son jouet et il euh	un tableau euh avec les euh	il va *s' asseyer* sur euh
27		il	un tableau avec les:	devant une vitre
28		il donne:	fleurs	et qu' est ce qu' il il voit dehors une: un insecte volant et euh l' insecte volant ne vient chez euh lui
29		le::	il &av	il euh les regarde il euh regarde

30		la corde: de son euh	pense wouah les fleurs et euh il va juste	il regarde l' abeille:
31		jouet dans les mains de monsieur	voler très vite et après il euh	et euh
32		et il va se et il euh part	ioups contre les euh &tabl contre le tableau	l' abeille elle regarde lui ils se ils se regardent *l'une à l'autre et euh
33		après que: après une	il va	et après euh quelques instants euh l'abeille:
34		quelques minutes le: monsieur euh	il eu:h tombe	veut sortir donc il va voler et il cherche:
35		se relève et qu' est ce qu' il voit il voit que:	et après il eu:h	des places ou il &ch il cherche une: une sortie
36		son cheval est euh a eu	il avait aussi vu une autre: euh animal dehors	mais ils ne peuvent pas trouver et il euh voit une:
37		est devenu un jouet	et euh cet animal est &ve	quelques fleurs et eu:h il eu:h
38			est venu ici est venu euh dedans	veut les sentir il va les euh
39			et euh l' autre animal le premier animal va euh se	il va les sentir mais eu:h c' est pas très c'est ce sont des fleurs euh (GH)
40			va se	ce ne sont pas des fleurs *frais
41			va se cacher	et &a après qu' il *se réalise pas il euh se retourne et il voit un grand tableau
42			va se:	sur le tableau il y a des euh fleurs d' soleil
43			cacher oui je pense	des tournesols
44			et après le: deuxième animal	il pense ouah des tournesols je peux euh trouver du miel des euh
45			entre et il euh [&=cl]	des ingrédients de miel là
46			il prend des euh p'tits euh	donc il va voler euh
47			p'tits euh [&=rises]	à t à toute vitesse et boum
48			comment dire	il rencontre les tournesols ils ne sont pas vrais
49			il prend de::s (GH)	après que euh
50			de café il prend de les p'tits euh	après un instant eu:h

51			p'tits euh	il entend euh le zoumzoum de l' insecte volant
52			il prend des p'tits pièces de café de café et euh	volant et eu:h il va se cacher sur le tableau
53			il eu:h	il eu:h
54			*prende il vole et après il	il reste *muette et euh il regarde euh ce qu' il faut ce qu' il fait euh l' insecte
55			il vient il prend il euh parte et euh	volant
56			après quand il est parti	et l' insecte il euh va euh
57			le premier animal va euh suivre	à la cuisine il eu:h
58			suit euh &ss	il prend des euh p'tits euh
59			va suivre les euh	des p'tits points de *caco
60			va va regarder euh ce qu' il euh la deuxième anim le deuxième animal a fait	et euh il euh
61			il trouve: le *chiam la chambre de &b	il vole::
62			la salle de bains	à la
63			et euh il va regarder et dans *le: couette dans la couette	salle de bains mais euh
64			il *vut	l' abeille elle ne sait pas
65			il voit le:	et euh il vole il euh retourne à la cuisine et il retourne à la *bain salle de bains quelquefois et après un instant
66			il voit un	il ne retourne pas (GH)
67			il voit un signe (GH)	il ne revient pas de: salle de bains (GH)
68			un signe de: petits pièces de café	et euh il reste là
69			il euh	donc l' abeille:
70			va le suivre et euh après il euh	est eu:h curieux

71			et	est curieuse: ce qu' il a fait euh
72			i::l	le:: l' insecte là il se demande où l' insecte est
73			il vient il est dehors	donc il euh va voler à à la salle de bains et après
74			car c' était euh	dans la salle de bains il euh trouve dans le: la couette
75			il y *atait une:	une::
76			il y avait *une: crane	*une signe:
77			qui sortait [&=rires]	hu:m et il ne comprend pas
78			[&=rires] et il est libre:	l' abeille elle ne comprend pas elle regarde
79				il eu:h
80				il &s il marche un peu sur euh la couette et après il voit *une gran:de
81				point
82				un grande non pas un grande mais un p'tit eu:h
83				sur une p'tite &sor une p'tite sortie
84				et il va sortir de là et euh il se trouve dans le elle se trouve dans les les tuyaux de: la salle de bains
85				et après quelques instants elle se trouve dehors car euh la salle de bains avait un tuyau euh
86				qui se menait vers euh l' extérieur dans un
87				avec un crâne
88				et euh *l'as
89				l' abeille a &s est sortie
90				et il peut voler





## ANNEXE 6

### Narrations des natifs

AN		
seg	BD	Film
1	alors dans la première vignette on voit un homme qui est peut être un cowboy qui a un grand chapeau rond	alors c' est un p'tit court métrage d' animation qui s' appelle prisonnière (GH)
2	qui:: arrive sur un cheval euh donc un cavalier qui	donc au début le film s' ouvre avec aussi une p'tite abeille qui vole à l' extérieur
3	qui ensuite euh accroche son cheval à une maison	e:::t elle entre dans une maison on la voit qui erre dans plusieurs pièces elle va s' approcher de:
4	où y a un abreuvoir donc on imagine que c' est euh là où il habite: et où:	d' une lampe j' crois au début ensuite de d' un:
5	et le cheval habite aussi (GH)	d' un compotier avec des fruits hum ensuite elle essaye de de ressortir et elle se heurte aux carreaux de la fenêtre et elle arrive pas à trouver le chemin vers l' extérieur
6	et il s' installe sur un banc à l' extérieur et il s' endort	e:::t elle aperçoit de l' autre côté de la vitre une mouche donc qui est à l' extérieur qui arrive et qui le
7	e:::t pendant qu'y qu' il fait une sieste y a un p'tit enfant qui arrive en tenant euh	et ensuite eu::h (GH)
8	par euh le bout d' une ficelle par la main euh au bout d' une ficelle un p'tit cheval en bois un jouet pour enfant	elle se dirige vers un: vase où y a des fleurs alors j' crois que c' est des fleurs en plastique enfin j' en suis pas sûre parc' qu' avec l' animation c' est pas facile de savoir
9	euh l' enfant s' approche du: du cavalier qui s' assoupit il s' aperçoit que le	mais y a une: donc elle essaye de butiner les fleurs mais apparemment ça marche pas y a une étiquette sur une des fleurs
10	que: l' autre ne répond pas et qu' il dort et il fait un échange en fait il	elle se retourne elle continue à voleter et elle voit un tableau avec des énormes tournesols
11	il met dans la main de::	donc elle fonce dessus et euh (GH)
12	de l' homme qui dort du cowboy un	en croyant donc que c' est des des vrais tournesols et elle euh
13	la ficelle au bout de laquelle est accroché euh le p'tit cheval en bois et il délivre le le vrai cheval	elle s' écrase contre le papier elle tombe à la renverse elle: se remet à voleter elle s' élève elle se rapproche du tableau
14	qui s' enfuit et l' enfant euh part aussi voilà et donc euh quand le:	et euh là elle voit qu' la que la mouche entre dans la maison donc elle va se se cacher dans un p'tit recoin au dessus du tableau

15	le:: cavalier se réveille et il se rend compte qu' il a au bout de la main un p'tit cheval en bois un jouet pour enfant	la mouche: fait des allers retours entre: la cuisine et une pièce sur le côté qu' on n' a pas encore vue
16	et il sait pas où est passé son son cheval	à plusieurs reprises euh contre: les fenêtres contre le les carreaux (GH)
17		et donc euh le l' abeille observe tout ça attentivement
18		et puis tout d' un coup euh (GH)
19		et elle s' approche de la baignoire et en fait avec les p'tits grains de café
20		la mouche a fait une petite flèche indiquant eu:h
21		et eu::h on sait pas si c' est un ennemi ou:
22		la la bonde: le trou d' évacuation de la baignoire donc qui indique un chemin de sortie donc eu:h l' abeille:
23		entre descend dans dans les tuyaux et euh ressort à l' extérieur (GH)
24		voilà et ensuite je crois que dans la dernière séquence on
25		on voit la mouche qui est du côté d' une vitre donc apparemment elle est restée dans la maison
26		voilà c' est la fin du film

CA		
seg	BD	Film
1	alorseu:h	alors c' est un p'tit court métrage d' animation qui s' appelle prisonnière (GH)
2	il s' agit probablement d' un cowboy &=rire	une abeille avec de très gros yeux
3	qui:: ou d' un lucky luke ou d' un: je ne sais quel cowboy qui:	qui: qui vit dans la campagne qui est dans la campagne et qui::
4	fait une pause e:t mm qui s' arrête euh avec son cheval e::t qui s' assoupit	qui: tout d' un coup euh passe devant une maison de campagne e::t
5	et le cheval habite aussi (GH)	décide de:: rentrer à l' intérieur enfin d' abord elle est autour et elle est
6	et là y a un p'tit garçon eu:h	dans le jardin: dans la campagne elle butine les fleurs de l' extérieur
7	qui arrive euh	et ensuite eu::h (GH)
8	qui est peut être son alter ego nous ne savons pas &=rire	tout d' un coup elle arrive au milieu du salon et là elle se met à butiner euh les fleurs du sal d' un bouquet de fleurs sur la table du salon

9	avec euh un p'tit cheval en bois un: jouet	et euh
10	et en fait le petit garçon substitue le:: le vrai cheval euh	et très viteeuh::h() elle se elle est dans son élément elle se balade elle arrive dans la cuisine là elle se met à
11	et remplace le:	elle est attirée par la les fruits la corbeille de fruits
12	le cheval &p du cowboy par son jouet	et après
13	euh on peut d on peut imaginer que ce soit un cheval en bois un jouet en bois	elle: continue sa vie mais la pauvre: se retrouve très vite prise au piège
14	et donc le:: le cowboy se:: quand le cowboy va se réveiller il ne trouvera plus son cheval qui s' est enfui	elle est bloquée dans la maison elle ne peut plus sortir y a pas de fenêtre ouverte y a pas de porte ouverte
15	et le p'tit garçon est parti et là il ne il n' aura plus qu' un:	e:t donc euh cette pauvre abeille se se cogne:
16	qu' un vieux jouet à la place de son cheval &=rire	à plusieurs reprises euh contre: les fenêtres contre le les carreaux (GH)
17		elle es:t bon visiblement un peu paniquée
18		et puis tout d' un coup euh (GH)
19		une mouche arrive dans la maison aussi
20		e:t une mouche avec des gros yeux rouges [&=rises]
21		et eu:::h on sait pas si c' est un ennemi ou:
22		ou une amie ou enfin on sait pas
23		c' qui va se passer avec cette mouche
24		et en fait euh
25		cette mouche se met à agir euh bizarrement euh enfin elles continuent encore un peu leur vie chacune de leur côté
26		le l' abeille essaie de:
27		de s' échapper encore elle trouve un cadre: (GH)
28		représentant des tournesols un cadre: de salon et donc persuadée que c' est une issue elle y va mais elle se cogne contre les tourne les tournesols
29		elle tournoie elle fait la toupie &=rire
30		et en fait euh cette euh
31		la mouche continue sa vie aussi dans le dans la maison
32		et la mouche se met à transporter des grains de café on ne sait _ où

33		et euh donc on on se demande bien et la p'tite abeille se place en observateur
34		et elle observe la mouche e::t elle observe:
35		son on va dire son amie la mouche &=rire
36		et puis au final euh &=bche (GH)
37		l' abeille se:: se retrouve dans la salle de bains de la maison
38		dans la baignoire et là trouve
39		les grains de café de sa copine la mouche &=rire et
40		elle a placé euh comme: finalement un peu comme le:
41		le petit poucet euh pour retrouver son chemin &=rire elle a placé les *graines de café euh de façon à former une flèche
42		qui indique: le:
43		le conduit de canalisation de la de la baignoire
44		et donc là on sait pas c' que l' abeille va faire si l' abeille va: va partir dans dans le trou de la baignoire ou pas
45		e:t on sait pas si elle va s' en sortir mais finalement elle suit ce chemin
46		et en fait c' était le chemin de la liberté e:t
47		et de: et de la campagne donc elle retrouve: voilà elle retrouve sa vie dans la campagne et elle arrive
48		à sortir de la maison grâce à la mouche

JE		
seg	BD	Film
1	sur la bd on peut voir un cowboy qui s' assoupit (GH)	alors eu:h c' est l' histoire d' une petite abeille qui &f::: euh qui se balade autour de plusieurs fleurs qui: sent les différente: euh les différentes fleurs
2	qui se repose un petit enfant qui arrive pendant que pendant que celui ci est	puis qui rentre dans un: &=bche dans une maison
3	est en train de se reposer le petit enfant a un petit euh cheval en bois et il voit le	elle commence à faire le tour de la maison::
4	cheval du:	à sentir les:

5	du grand cowboy et il décide i décide de	les différentes fleurs de la maison: dans les différentes piè:ces e::t eu::h elle arrive dans la cuisine
6	d' intervertir les les deux chevaux et d' en	elle continue à: à:: &=bche
7	et de donner son son petit cheval [&=rires] au cowboy voilà	à sentir les différentes fleurs
8		et elle entend la porte qui::
9		qui se claque
10		du coup el::le
11		elle essaye de sortir de la maison
12		elle se frappe elle se:: se frappe contre les fenêtres
13		elle arrive pas à sorti:r
14		elle::
15		ouais elle: elle elle arrive donc donc contre différentes fenêtres
16		&=bche elle se
17		elle:
18		elle arrive plus à sortir (GH)
19		puis elle retourne dans la cuisi::ne elle arrive eu:h
20		devant une fenêtre elle s' arrête elle a l' air un peu désespérée
21		étant donné qu' elle arrive pas à sortir et contre la fenêtre
22		de l' autre côté de la fenêtre y a une mouche qui arrive
23		elles s' observent toutes les deux
24		sans qu' il se passe grand chose
25		eteu::h l' abeille cont décide de de repartir au bout d' un moment
26		dans un petit moment pou:r (GH)
27		continuer à sentir d'autres fleurs
28		elle:eu::h sent d' autres fleurs et là elle
29		tombe devant un:
30		un tableau de de tournesols
31		qui se ressemble étrangement à ceux de van gogh au passage
32		e::t (GH)
33		arrivée devant ces tournesols elle confond les tournesols avec de véritables fleurs du coup elle rentre dans les
34		dans le tableau

35		plusieurs fois et euh elle entend une:: elle entend un bruit
36		&=bche du coup elle se
37		cache sur le:::
38		haut du cadre du tableau
39		et (GH)
40		qui voit la la mouche
41		quieu::h débarque
42		et la mouche prend des::: rentre dans un::
43		dans la pièce et passe va directement dans la cuisine
44		et prend différents grains de de café
45		on la voit faire des allers retours
46		et l' abeille continue de se:
47		de se cacher (GH)
48		pui:s euh au bout de plusieurs allers retours y a plus de bruit l' abeille (GH)
49		vaeu::h
50		va dans la direction qu' avait pris la mouche et elle arrive dans la
51		dans la salle de bains j' crois
52		elle arrive dans la salle de bains et elle voit les:: euh
53		les grains de café qui sonteu::h disposés en forme de flèche
54		el::le et les les flèches donnent sur euh sur la tuyauterie
55		et elle: euh
56		elle comment s' appelle
57		elle: rentre dans la tuyauterie et elle arrive enfin à l' air libre
58		et arrivée à l' air libre elle: peut retourner vagabonder au:x différentes fleurs
59		et pendant ce temps-là la mouche eu::h l' observe d' un air satisfait

MG		
seg	BD	Film
1	alors euh c' est l' histoire d' un d' un cowboy un peu solitaire qui:	oui alors c' est un p'tit film des minuscules (GH)
2	euh donc accompagné de son cheval après un long voyage ils sont un peu fatigués et c' est la première ville qu' ils voient depuis	un p'tit film muet euh qui raconte euh qui se fait plaisir à raconter enfin qui montre en tout cas la

3	des kilomètres et des kilomètres dans le désert américain (GH)	la vie supposée de:
4	donc il décide de s' arrêter de faire une pause le cheval est fatigué le cheval aussi	de:: d' insectes
5	donc ils s' assoient euh sur un banc	et notamment une: abeille et une mouche
6	et euh attache: son cheval au banc comme ça son cheval euh peut se reposer	alors euh le film commence autour d' une abeille on peut l' appeler maya maya l' abeille
7	et euh lui décide de faire un p'tit somme avec euh	ça semble elle ressemble un peu à maya l' abeille
8	avec euh son chapeau de cowboy ramené sur euh ramené sur son visage	euh qui qui butine donc c' est vraiment la vie naturelle qu' elle peut avoir à l' extérieur sans se poser trop de questions
9	mais le le village n' est pas n' est pas désert euh y a des enfants qui jouent y' a de y' a d' la vie	et puis euh elle rentre par inadvertance voilà dans une maison ce qui arrive souvent on a tous connu ça euh
10	et euh un des enfants est intrigué par euh par cette personne endormie s' approche	en tant qu' humain voilà des abeilles qui rentrent dans les maisons
11	avec en traînant son son jouet son cheval à: à roulettes (GH)	et qui se retrouvent du coup vraimen:t eu::h
12	et euh une idée lui vient lui vient c' est que (GH)	enfin vivre en dehors de leur habitat naturel complètement désarmées un peu::
13	le cheval à roulettes c' est c' est bien plus drôle qu' un:	un peu perdue donc eu::h on la voit qui:
14	qu' un vrai qu' un vrai cheval	une fois qu' elle est rentrée elle cherche voilà elle butine les fleurs mais les fleurs sont en plastique ou euh
15	euh il décide donc de de substituer le:	ou c' est des peintures
16	le: cheval du cowboy par son jouet	et euh quand elle essaye de sortir puisqu' elle se rend bien compte que c' est pas vraimen:t
17	alors euh en libérant le la lanière qui attache le cheval le cheval forcemen:t	c' est pas vraiment un:: un champ euh digne de ce nom
18	reprend sa liberté euh et et &s et s' enfuit	elle se ben elle se cogne contre les contre les vitres euh elle trouve pas la sortie
19	et euh et le p'tit garçon décide de donner euh son son jouet de	donc c' est vraiment euh l' animal qui est dans un habi euh un: environnement qui ne lui correspond plus et du coup euh

20	de cheval à roulettes au::	enfin elle n' arrive pas à trouver la solution quoi c' est c' est elle était tellement bloquée dan:s
21	au au cowboy qui lui est profondément endormi sur le banc	elle a elle elle trouve pas de porte de sortie en fait alors qu' elle y est rentrée quand même (GH)
22	e:t et du coup le p'tit garçon part euh	e::t &=bche à travers une à travers la la fenêtre elle: elle rencontre une mouche
23	part à: eu:h	on reconnaît à ses gros yeux globuleux rouges là à facettes (GH)
24	part un file un peu à l' anglaise mais euh	bon le dialogue on peut pas dire qui y' a un dialogue entre eux hein c' est
25	part sans sans faire de bruit	ils s' regardent
26	e::t quelques heures après le le:	mais la mouche eu:::h
27	quelques minutes après	j' pense qu' elle est b
28	le cowboy se réveille et euh il ne comprend pas trop ce qui se passe	b' être plus habituée au: au: à l' habitat humain enfin elle a l' air en tout cas de connaître son chemin et pour rentrer et pour en ressortir
29	et il sait pas trop si i rêve ou: si c' est vraiment la réalité	et c' est vrai que l' abeille du coup euh prend ses précautions elle se cache et elle observe le manège de la mouche
30	son cheval a été transformé en cheval à roulettes	qui rentre: qui sort du salon:
31		euh en prenant des graines de: en prenant des grains de café
32		donc elle se rend bien compte que la mouche: paraît vraiment très à l' aise elle va droit au but elle tourne pas en rond euh
33		elle: elle elle sait où elle va
34		e:::t
35		et du coup elle est intriguée
36		et elle finit par sortir de sa cachette pour alle:r
37		pour aller voir euh un peu ce qu' a manigancé la mouche
38		et eu:::h et le résultat c' est que la mouche en fait par ses aller retours était en train de constituer une flèche
39		euh de grains euh une flèche constituée euh par euh plusieurs grains de café mais qui font le motif d' une flèche



40		et qui indiquait euh le::
41		le trou de sortie de la baignoire
42		e::t mm &=bche
43		et du coup elle elle fait confiance à la mouche e:t décide de rentrer dans le tube: de:
44		dans le tube de la baignoire
45		alors on se doute enfin: on on sait pas trop quoi penser est ce qu' elle va atterrir dans les égouts est ce qu' elle va:
46		est ce que c' est un traquenard mais en fait non
47		il se trouve qu' elle ressort par le:
48		par le robinet d' extérieur euh d' arrosage et voilà du jardin
49		et du coup elle retrouve sa liberté euh
50		dehors

TH		
seg	BD	Film
1	donc eu:h à ce que j' ai cru comprendre c' est un cowboy qui arrive en ville	dans cette histoire on a: une abeille euh qui recherche à:: une une fleur à butiner et qui rentre: visiblement par par inadvertance
2	et euh une ville qu' il visiblement ne connaît pas donc euh il s' assoit	dan:s une maison
3	euh il s' endort	euh toujours à la recherche: de de d' une plante à butiner euh il passe suce il passe suce successivement de fruit à::
4	et un petit garçon euh s' amène avec euh un cheval euh plus petit (GH)	à diverses fleurs (GH)
5	et eu:h il l' attache à sa main et il libère le le plus grand cheval et euh étrangement il le laisse s' enfuir et et il le prend pas en retour	et eu::h
6	donc eu::h j' ai	&=bche après avoir entendu un bruit dans la maison
7	je j' avoue que j' ai du mal à im à savoir pourquoi euh quel est le but de de de libérer le plus grand cheval et de lui laisser un cheval plus petit	euh il comprend qu' il est en danger et cherche à: s'enfuir euh le problème c' est comme on le voit bien dans la vidéo

8	à la place euh parce qu' il pourra pas le monter	euh après s' être successivement eu::h
9	alors euh ça se limite à ça (GH)	cogné contre les vitres::: de de la maison
10	j' avoue que j' ai pas beaucoup d' imagination sur la raison du pourquoi euh le petit garçon a laissé un cheval plus petit au au au cowboy	il est incapable de sortir et eu::h s' ensuit ensuite euh une sorte de confrontation avec une mouche qui elle se trouve à l' extérieur
11	qu' a l' air qu' a l' air assez dépité (GH)	et qui euh trouve moyen de rentrer dans la maison
12	donc euh je non je n' ai pas de signification particulière à donner	et euh qui en se servant euh de manière assez ingénieuse
13		euh humoristique oui
14		aussi de grains de café pour ensuite dans la salle de bains (GH)
15		tracer une flèche afin eu:h d' aider eu:::h le l' abeille
16		à lui indiquer le chemin de la sortie qui est euh par la les canalisations

VC		
seg	BD	Film
1	donc y a huit vignettes euh ça raconte l' histoire donc d' un: cowboy qui arrive en ville donc dans la première vignette: on le voit tout simplement arriver	donc l' histoire c' est l' histoire d' une: d' une abeille et d' une mouche
2	dans la:: deuxième et la troisième on le voit donc qui commence à attacher son cheval donc euh	donc en fait l' abeille au début elle butine tranquillement dans le jardin elle passe de fleur en fleur
3	pour pas qu' il parte	et au final elle se retrouve à aller donc dans une maison
4	et donc après il commence à s' asseoir sur un banc et à dormir	dù là pareil elle commence à butiner euh les fleurs qu' elle trouve:
5	donc dans la cinquième vignette ensuite y a un p'tit garçon qui arrive: avec euh un: jouet	dans différentes euh pièces
6	un cheval [&=rires]	et ensuite on entend un bruit donc euh c' est une porte qui se ferme et donc l' abeille se retrouve coincée dans la maison
7	et euh qui regarde le cowboy avec son vrai cheval	elle essaie donc de sortir donc par les différentes fenêtres
8	et qui décide donc de euh remplacer euh le cheval du vrai enfin le cheval du cowboy avec euh son son jouet	elle essaie plusieurs fois par le par la même fenêtre
9		avant de passer à: à une autre mais ça ne marche pas
10		et: à un moment donc y a une mouche qui vient se poser sur le rebord de la fenêtre de l' autre côté d' elle (GH)

11		qui se font un peu des signes mais [&=rires]
12		sans trop non plus
13		et eu:h alors je sais plus lequel des deux part en premier
14		mais au final euh donc l' abeille continue à:
15		à vagabonder on va le dire (GH)
16		dans la maison:
17		et décide de se diriger vers un tableau avec des tournesols
18		donc évidemment elle essaie d' aller euh les butiner comme avec les autres fleurs
19		sauf que ça ne marche pas &=bche et là eu:h la mouche rentre dans la maison
20		donc l' abeille elle part se cacher derrière le tableau
21		et euh l' a
22		la mouche donc passe devant elle et va dans
23		la cuisine chercher donc des *graines de cacao
24		qu' elle part ramener on sait pas encore où
25		donc elle fait plusieurs aller retours comme ça
26		et euh au bout d' un moment la mouche ne revient plus et donc l' abeille décide de partir après euh
27		fait un bruit de
28		moi il me faisait penser à un bruit de micro ondes mais bon (GH)
29		et donc elle va dans la salle de bains où elle voit donc euh une flèche avec les graines de cacao
30		que donc on suppose la mouche lui a laissées
31		euh elle décide donc euh d' y aller euh de suivre: la flèche de passer dans les tuyaux de la baignoire
32		pour ressortir en fait par un robinet
33		et euh donc la dernière: séance séquence (GH)
34		c' est donc la mouche qu' on voit à l' intérieur de la maison à la sur le rebord de la fenêtre en train de regarder à l' extérieur



## ANNEXE 7

Tableau comparatif SP non-natifs et natifs

NON NATIFS		NATIFS	
collocations verbales	collocations nominales	collocations verbales	collocations nominales
aller aux toilettes	autre fois	aller droit au but	air satisfait
aller dehors	à toute vitesse	avoir du mal à	aller-retour
aller en ville	aller-retour	avoir l'air	alter ego
avoir des points communs	années soixante	butiner fleurs	bouquet de fleurs
avoir faim	après-midi	entendre un bruit	chapeau rond
avoir l'air	bande dessinée	être dans son élément	chemin de sortie
avoir peur	belle maison	être en danger	cheval à roulettes
avoir sommeil	bonne odeur	faire confiance à	cinq minutes
comprendre le problème	chemin de sortie	faire le tour de la maison	conduit de canalisation
entendre un bruit	choses à manger	faire sieste	corbeille de fruits
être dehors	coup de vent	faire un bruit de	court métrage
fabriquer du miel	dessin animé	faire une pause	de manière ingénieuse
faire + promenade	eau usée	filer à l'anglaise	film muet
faire attention	face à face (en)	indiquer le chemin	grains de café
faire du café	fleurs *frais	observer le manège	habitat humain
faire du cheval	fleurs artificielles	prendre la direction de	habitat naturel
faire du miel	fleurs d'soleil	prendre ses précautions	l'autre côté
faire du vélo	fleurs naturelles	pris au piège	la dernière séance
faire la sieste	fleurs sauvages	reprendre sa liberté	la raison du pourquoi
faire le tour	grains de café	retrouver sa liberté	long voyage
faire une blague	histoire d'amitié	retrouver son chemin	maison de campagne

indiquer route/chemin	insecte volant	se faire des signes	maya l'abeille
montrer la sortie	journée de travail	se placer en observateur	micro-ondes
passer par la fenêtre	la dernière fois	se poser des questions	p'tit somme
prendre la sortie	langue commune	se rendre compte	petit moment
raconter + histoire	longue journée	tomber à la renverse	petit poucet
raconter une blague	maison de campagne	tourner en rond	pièce attenante
rester *muette	petit moment	trouver la solution	porte de sortie
retrouver la liberté	petite blague	trouver la sortie	salle de bains
se casser la gueule	porte entrouverte	trouver moyen de	table du salon
se rendre compte	porte ouverte		yeux globuleux
sentir + parfums	salle de bains		
vouloir de l'aide	système de conduite		
	table basse		
	table de salon		

## ANNEXE 8

### Tableau des SP déviantes

	BD1	BD2	F1	F2
AK	a été* sommeil (a sommeil)	comment dit* (2) (comment dire)	il y atait*	pois* de café (2)
	descendre en* cheval (de cheval)	petits* enfants (2) (erreur phonologique)	au lieu que*	comment dit* (comment dire)
		*petit maquette (5)		tous* les, tout* la
FA	*sent fatigué (se sent fatigué)	s'endorme* (s'endort)	*tous les choses (toutes les choses)	se rapproche *à (2)
	*prend part (prend parti)	il ne comprenne* pas pourquoi (ne comprend pas)		se *réfuge (sur) (3) (se réfugie)
	*rester étonné (être étonné)	obligé *à (de)		tait rentrée par *une fenêtre (la fenêtre)
		dans *une contexte (un contexte)		
HE	*il pas compris	ça ressemble *d'être (à)	porte ouvert* (ouverte)	va à *les toilettes (aux toilettes)
		<b>*un peu de (sommeil) (un peu sommeil)</b>	qu'est ce qu'elle faisait (ce qu'elle faisait)	reste *dans la maison (à la maison)
		il *vient + V (il vient de)		elle *arrive pas (elle y arrive pas)
JA	C'est	*a descendu du (est descendu)	*a passé *pour (est passé par)	*sorte de
		*en ce moment (à ce moment)	elle *y a	*sorte de la maison
			*Fait les promenades	*comment fait

			Comment *dit	difficile *pour
MA			*trucs du café	
MJ		*en *chevau (à cheval)		entendre *une bruit
				veut *d' aide
MR	*asseoir à (s'asseoir sur)	*démonter de (descendre de)	*en ce moment	*à la côté de, du
		à* maison		*à l'autre côté
		échanger N *pour N		
NO	allait *au ville	je peux commence*	la* première (insecte)	*morceaux du café (2) *quiz du café
	asseye* sur (s'assoit sur)		Haricots du café	il *suivre le chemin
	*la même		Faire *le miel	il arrive pas *de
	chez *il (chez lui)			dans *le direction
				*au derrière de (la peintre)
				*à l'autre côté de il *voir que on *entendre que
NY		fait *un randonnée		arrive pas de* (à) elle *voudrait
		sans *compris *sorte de		*qu'est qu'il a fait (qu'est ce qu'il...)



OL	monter *au (à)	*en poing	*sort de	*dehors de (en dehors de)
	*en cheval		*dehors de	*à l'autre côté de (de l'autre côté)
			à* l'autre côté de	*dans ce moment (à ce moment)
			*Pois de café c'est pourquoi *tous les choses	<u>un*</u> bouteille de <u>récolter du miel</u>
PR	transformé *dans	petit *jou	*petite film, *petite signe (3), *petit abeille	faire *une signe
	*petite enfant, *petite cheval (2), petit *enfante	petite*+ cheval (2)	*était peur (2)	C'était *le histoire
	petit *jou	il a pris* une sieste	qu'est ce qui *passe	
	qu'est ce qui passait*		*une autre type de	
	Transformé *dans			
QI	*sent fatigué (se sent)	petit* sieste (petite)	*petit histoire, petit* abille*	det + tout* petite + N (étiquette)
	*assied dessus (assis)	faire le* sieste (la sieste)	tout* petit* histoire (toute petite)	graine* de café x 3
	fait un petit* sieste (une petite)	tellement de*+ fatigué (tellement fatigué)	arrive plus de* (à)	petit* abeille (2)
			retrouver *son liberté (sa liberté)	langue *commun
			*graine de café (3)	a retrouvé *son liberté
				il arrive pas *de (à)
				à *extérieur de (à l'extérieur de)
SU	il *avait (il y avait)	il *avait (il y avait)	il *avait (il y avait)	il *avait (il y avait)
VI	*au lieu (au lieu de)	se transformer *dans	*fabriquer du miel	*en ce moment-là (à ce moment-là)
		*c'que c'que c'est arrivé (ce qui est arrivé)	*un moment donné (à un moment donné)	toutes *le (toutes les)
				c'est... *que (qui)

VO	sur le premier plan (au premier plan)	s'approche *à (de), *en galope (au galop)	grains *du café (de café) s'approche *vers (de)	on *t'inquiète pas (s'inquiète) grains *du café
			commence *de	
			réussi *de	
			*à l'autre côté	
			très contente *que (de)	
XI	*en jouet (en bois) *ce moment-là (à ce moment-là)	sans *fait attention (faire) *assois sur (s'assoit sur) *en ce moment (à ce moment-là)	*pois de café (grains de café) *a sorti de (est sorti de) *ce moment-là (à ce moment-là)	*pois de café (grains de café) a réussi *de (à) *en face à face (face à face)
			*comment dit (dire)	salle *du bain
			*qu'est-ce qu'il (ce qu'il)	
ZI	chez *toi (chez lui)		*pièces de café (grains) *la deuxième (le) *il y avait (il y avait)	fleurs d'soleil (tournesols) *point de cacao (grains de café) fleurs *frais (fraîches)
				Reste muette (muet)
				*il se réalise pas (il ne réalise pas)
				*s'asseyer x2 (s'asseoir)
				*se menait vers (menait vers)



## ANNEXE 9

### Types d'erreurs

types d'erreurs					
déterminant	erreur genre ou personne	morpho verbe	Phonologie	préposition	lexique
à *la côté de, du	indiquait le* sortie	se *réfuge (sur) (3)	il * avait (il y avait)	se rapproche *à (2)	*démonter de
*à maison	bon* odeur	il pas compris	petits_enfants	*en ce moment	pois* de café(2)
à extérieur de	tous* les, tout* la	a* descendu du	*qu'est qu'il a fait	*asseoir à	*morceaux du café
veut d' aide	La* même	asseye* sur	*comment fait	allait au* ville	*sent fatigué
	entendre *une bruit	sorte* de	sans fait* attention	va à *les toilettes	*fermée dans la maison
	fait *un randonnée	a *commence à		reste *dans la maison	en *poing
	*petite enfant, *petite cheval (2), petit *enfante	*sorte de la maison		il arrive pas *de	points de caco
	*petit sieste	*sorte avec (les morceaux)		à* l'autre côté de	grains du café
	faire *le sieste	il *suivre le chemin		monter au* (à)	en jouet
	*la première (insecte)	je peux *commence		en* cheval	en galope
	*un bouteille de	*était peur (2)		transformé *dans	fabriquer du miel
	D + *tout petite + N (étiquette)	sans *compris		*dehors de	fleurs d'soleil
	faire *une signe	on t'inquiète pas		*dans ce moment	pieces de café
	graine* de café x 3	assoit sur		ça ressemble *d'être	haricots du café
	petit* abeille (2)	il y atait		un peu de* (sommeil)	trucs du café
	langue commun*	il se réalise pas		il vient* + V	en jouet
	a retrouvé *son liberté	s'asseyer		par)	
	*une autre type de	se menait vers		difficile pour*	
	*elle y a	a été* sommeil		tellement de*+ fatigué	

	petit *jou	*asseoir à		*au lieu	
	dans *une contexte	*qu'est ce qui passait		un moment donné	
	fruits frais			s'approche *vers	
	reste muette			s'approche à	
	*la deuxième (le)			Commence de	
	chez *toi			Réussi de	
				A l'autre côté	
				Très contente que (de)	
				sur le premier plan	
				ce moment là	
				en face en face	

## ANNEXE 10

### Quelques occurrences de « petit » dans *lexutor*

002. , quoi! MG492 – T(u) as ton Brevet ! [rire] Brave **PETIT** ! LW493 – Ou sinon, i(l) fauuut, sinon, bref  
012. n, il essayait de me faire comprendre en français **PETIT**-nègre **alors** e, e, moi e, moi Vincent, ee, m  
013. as mangé l'ail et qu'après tu dois voir ton ou ta **PETIT** **euh ami(euh)** (et) b(i)en, bon bonjour l'hal  
014. ueule. TJ128 – Attends, déjà e je veux dire mon e **PETIT** **ami** e dépendra, (et) b(i)en franchement, [ri  
015. un magazine eee de rock et dedans y'avait un tout **PETIT** **article**. J'adore lire les petits article cac  
016. 'est bien ! [rire] LB89 - Et i(l) y avait un tout **PETIT** **article** qui dis(ait), qui parlait de Miyavi  
021. ussi ! eeee, c'était un couple, qu'avait pris un **PETIT** **bébé** lion. Eeee comme animal de compagnie. AB  
022. s] (II) faut regarder les messages, que e que mon **PETIT** **bonhomme** (m') a laissés. Qu'est-ce qu'il rac  
023. our faire un truc quoi ! GP273 – Bon, on mange un **PETIT** **bout**, enfin tu prends l'air. AF276 – Mais ou  
026. i toujours rêvé d'avoir une cabane. Quand j'étais **PETIT** **c'était** mon rêve. J'en ai eu une mais  
028. ue-le [bruits] # ouais c'est-à-dire que tu as ton **PETIT** **carré**, il le remplit, quand il a fini de le  
029. le jeu}, on se fait un carré déjà un **PETIT** **carré** on prend la photo [rire]  
030. aux et lui expliquer comme ça ## de lui faire, le **PETIT** **carré** orange, le losange vert et le gros car  
031. lèvres] SB33 – Bon en gros c'est, c'est, c'est un **PETIT** **coin** paumé pas très loin ensuite j'ai vécu d  
032. oi, j'étais à l'internat et on avait déjà not(r)e **PETIT** **confort**. C'est-à-dire LB266 - mm MR268 - qu'  
033. choppé un, j'avais réussi à dédédé, à décrocher un **PETIT** **contrat** et tout, j'arrive, première réunion  
034. le mec en fait te sais e il est comme ça avec son **PETIT** **costard** il ne bouge pas il est touj(.) [brui  
035. finalement, je l'ai ai(.) je lui ai donné un, un **PETIT** **coup** de main et il a pu ee, avec quelques di  
036. par exemple les nouvelles générations comme {mon **PETIT** **cousin**}, lui qui est vraiment né avec cette  
037. là quoi ! En parlant de ça je commence à avoir un **PETIT** **creux**. [rises] MW416 – Tu as faim ? Tu veux  
038. fin bref, on prend un petit déjeuner, on prend un **PETIT** **déjeuner**, je me prends genre trois, trois ta  
039. u hier à courir et tout (en)fin bref, on prend un **PETIT** **déjeuner**, on prend un petit déjeuner, je me  
040. six heures et demie, mais je me lève après! (x)le **PETIT** **déjeuner** et un peu sur l'ordi, genre regarde  
042. deux heures et quart. Je vais peut-être faire un **PETIT** **détour** à la cafet quand même hein. BB321 – D  
043. ir. On va pas dire, ouaiis, ça fait quand même un **PETIT** **diplôme** PN131 – Ouais, c'est comme ee le, le  
046. ue dans l'avion pour aller en Guyane, chacun a un **PETIT** **écran** # dans l(e) siège # dans l(e) siège de  
047. 41 – Voilà ! RL 254 – (II) fallait qu'il mette un **PETIT** **écriteau**, e, volez-moi ma voiture, quoi, car  
049. c(e) (n)'est pas sérieux, c'est de la fiction! Le **PETIT** **enfant** qui rentre dans le film, tout ça ! No  
053. s mais elle est bien faite surtout avec le **PETIT** **flipper** et tout là hein NC145 – Tu sais je s  
054. rence moi, et e, a alors il m'a parlé dans un, un **PETIT** **français**, e parce qu'il pensait # enfin, com  
055. nd il a commencé à me par(.), à me parler avec un **PETIT** **français** pour essayer de me faire comprendre  
056. se porte ! [rire] SB82 – D'accord ! ça c'est mon **PETIT** **frère**, c'est pareil ! # {Mon frère} c'est pa  
057. nt, ouais, c'est ça, hein, je t'ai dit (pour) mon **PETIT** **frère**, mes parents ils ont tout essayé. Alor  
058. rires] SB40 – Ouais ! [rire] (Et) b(i)en j'ai mon **PETIT** **frère**, moi, c'est ma sœur qui lui avait pris  
059. urtout les jeux sur e, sur l'ordi e. Moi j'ai mon **PETIT** **frère**, son ordinateur il s'en sert que pour  
060. eux violents, quoi SB274 – Oui ! (et) b(i)en {mon **PETIT** **frère**} c'est Dofus, après j(e) (ne) sais pas  
061. l'énervait. SB272 – Ah oui d'accord [rises] {Mon **PETIT** **frère**} ça l'énervait, mais ça va pas l'empêc  
062. e # non {moi} c'est e, les jeux en réseau, ah mon **PETIT** **frère** c'est pareil, quoi. Mais, il pourrait  
063. , d'accord JG84 - mm. AI84 – {Moi} j'ai dit à mon **PETIT** **frère** qui est venu à la fac parce qu'il est  
064. avait, un grand [rire], un grand, un moyen et un **PETIT** **genre**, comme une petite blague, RL548 - Oua  
065. si ee, « Hé, Gervais, plein de calcium. [rire] Et **PETIT** **Gervais**, regarde, tu vois l'espèce de halo b  
066. ? ». TJ269 « Mais c'est normaale, j'ai mangé un **PETIT** **Gervais**. » BB269 Ge (x) Le mec avec d 068 <j  
067. J'ai mangé un **PETIT** **Gervais** >> [rises] TJ270 . il (ne) faut pas que tu oublies de lui donner ce **PETIT** **gr**  
**ain** de folie là. MW310 – mm, ouais, ouais

## ANNEXE 11

### Quelques occurrences de “et après”

241 user selected lines from 448 hits (normalized to **624 per million** for comparison)

Click **keyword** for Larger Context

084. – Aaah. EL447 – Et, eee, elle était gratuite. Et APRÈS, au début, { moi } j(e) pensais qu(e) tu payai
085. – Forcément, un gâteau au chocolat... TJ348 Et APRÈS ee BB348 ça (ne) se fait pas de laisser un
086. {lui}, i(l) sort une arme, i(l) tue sa femme, et APRÈS, eeee, alors je (ne) sais plus si i(l) s(e)
087. 5000€, hop là. Qu'est-ce que ça fait ? CC88 Et APRÈS, nan, il faut trouver un titre quand même, ?
088. a qu'un bus, il doit passer vers 6 h et demie et APRÈS (il) y a. CA392 – Mais l'année dernière tu v
089. amis, ils nous ont fait découvrir leur Japon, et APRÈS, { les autres parties }, c'était, on se débrou
090. -Ben ce qu'on peut faire c'est aller chez toi et APRÈS tu remontes avec moi chercher le sandwich #
091. – C'est les Etats-unis. CA241 – Les Etats-Unis et APRÈS Canada, je crois que c'est la prem
092. ça pour première partie [bruits de crayon] ## et APRÈS eee deuxième partie [bruits de crayon] #bo
093. case peuvent aller dans n'importe quel bouton et APRÈS c'est lui si il se trompe qui, qui
094. comment tu le prononces ? # [bruit de chaise] et APRÈS on appuyant sur espace A386 mm B387 -  
T'as
095. - D'accord ! AB414 - Il l'a décroché tu vois ? Et APRÈS, il s'est mis à parler de cette personne [ri
096. d'ailleurs. KDC12 Oui. Donc voilà et NA12 Et APRÈS tu fais quoi ? KDC13 Après je vais rentrer
097. du coup je l'ai mangé. BB350 – [rire] TJ351 et APRÈS ee par contre après quand on a fait la cuisi
098. du plus petit au plus grand et après on voit, et APRÈS on positionnera après nos, nos cases #
100. entier quand tu les vois passer tous ensemble et APRÈS en fait tu la réentends jamais en entier jus
101. et droit des entreprises # {ça} c'est commun. Et APRÈS (il) y a genre e y a genre une langue et e #
102. être c'est moi qui l'a pendant toute ee DW54- Et APRÈS, bien sure s'ils (ne) parlent paas français
103. eu je me suis dit il est mineur c'est bizarre et APRÈS j'ai su qu'il l'avais adopté et tout mmm et
104. eux leur plus grande, tout ça c'est matériel. Et APRÈS leur aide ee, à côté, c'est qu'ils forment,
105. fait troisième interface, quatrième interface et APRÈS e cinquième interface, résultats EM225 –
106. font un clip. FB726 – Ah ! D'accord ! EL727 – Et APRÈS, i(ls) te disent « Bientôt sur Arte » (en)fi
107. i(ls) restent à Mac Do(nald), du coup MB538 – et APRÈS, i(l)s arrivent, CT539 Ouais, mais être ma
108. j'ai deux pulls et j'ai un haut en dessous. # Et APRÈS j'ai mon manteau. ## Ben, j(e) suis frileuse
109. jusqu'à la surface. DW12 mm, d'accord. AR1 Et APRÈS on e DW13 Donc alors vous vous examinez le
110. la rigueur ce serait un exercice préliminaire et APRÈS il y aurait un exercice intéressant ce serait
111. le # mm, y avait l'interface avec ee le pendu et APRÈS en haut y avait marqué palmarès, il a cliqué

## ANNEXE 12

### Occurrences de « et donc »

36 user selected lines from 499 hits (normalized to **695 per million** for comparison)

225. aussi en quelques sortes. MW223 – mm. CD223 – **Et DONC** c'est pour ça que nous on travaille, cet été  
226. avaient des problèmes après je pense. CA343 – **Et DONC**, ils ont eu une 3e année. NV344 – Ils ont pas  
227. bon elle s'est suicidée pour lui, par amour # **et DONC** e à partir de là il mène une vie de débauché  
228. cas, enfin, si les dires sont fondées ou non. **Et DONC** Dorian Gray il explique que oui et il lui exp  
229. cette époque, cette génération de jeux vidéo. **Et DONC** quand on sera plus âgé, ce sera normal aussi  
230. e, justement ce peintre était amoureux de lui **et DONC** e, c'est lui, il l'a trop il l'a trop idolâtr  
231. enfin elle est par terre, en larmes et tout # **et DONC** euh, il rentre chez lui et en fait il re(.) i  
232. fait à Mac Donald je faisais 23h par semaine, **et DONC** du coup (il) y a plein de e (il) y a plein d'  
233. japonais en fait. MW74 – Oui d'accord. CD75 – **Et DONC** il dit « ouais, il (ne) faut pas parler en fr  
234. kanjis ont été un peu laissées à l'abandon ## **et DONC**... TJ259 – BB259 – Oui,  
235. la peau. CD294 – Ah, quelle horreur ! MW295 – **Et DONC** en fait cette jeune e cette jeune policière l  
236. mais c'était vieux SB2 – Ouais, logique NC3 – **et DONC** y avait certains, certains grands (x) de mes gr  
237. sais pas limite comme une parole divine, quoi **et DONC** il fait un il fait un, enfin il émet le souhait  
238. vois, j'avais jamais pris l'avion toute seule **et DONC** c'est ça c'est pour moi c'était écrit sur mon  
239. , et ça recommence à zéro mais le licence xxx. **Et DONC** voilà, {le truc ce qui me sauve}, c'est que j  
240. , oui la 1ère semaine de juillet, normalement, **et DONC** on pourrait faire e d'une pierre deux coups e  
241. 'ai pas fini, donc dès que je rentre chez moi. **Et DONC** ? GP116 – Parce qu'il y a eu il y a eu à la f  
242. amoufle le meurtre et donc e [bruits de table] **et DONC** oui après e, après il continue sa vie mais to  
243. crois. Voilà, on n'est pas en France ! [rires] **Et DONC**, et quand je suis venue ici, je vais en fac q  
244. is voté oui. GP94 – Il y avait Favenec et tout, **et DONC** on vote et, je ne sais pas si tu te rappelles  
245. it et tout, il entre et il camoufle le meurtre **et DONC** e [bruits de table] et donc oui après e, après  
246. ller au Duty-Free acheter des clopes pas cher, **et DONC** je croise ce gars qui était dans le même truc  
247. llons qui sont pas de filles, qui sont mixtes, **et DONC** disons qu'il a des grosses fêtes, des trucs c  
248. nesse enfin toute la vieillesse # **et DONC** ouais au final, euh enfin au fur et à mesure  
249. ont convenu en gros de, de se marier et tout # **et DONC** ils vont la voir et en fait elle # elle impre  
250. ouze, quinze personnes. TJ356 – Ouais. BB356 – **Et DONC** t(u) avais un gros truc au mili(.) t(u) avais  
251. pas de recherche particulière ou quoi, tu vois **et DONC** ouais, je sais après, je me suis arrêté e, je  
252. r de la nuit et euh et le rictus ne part pas # **et DONC** il commence à comprendre que c'est ça, et le  
253. sa beauté avant Lord Henry justement **et DONC** e d'un accès de colère, il va et il lui plant  
254. société et que tout le monde rit de toi, enfin **et DONC** il se demande il lui demande si c'est le cas  
255. t] AF171 – Ah non mais il n'y a pas de soucis. **Et DONC**, euh, je (ne) sais plus quoi on disait, tac,  
256. ttres. MW136 –ouais, ouais, c'est sûr. CD137 – **et DONC** c'est plus agréable de lire un truc sur l'ord  
257. u'il # pourra tout avoir dans le monde, quoi # **et DONC** il lui en fait part et tout et ee vu qu'il es  
258. ux] il veut passer e il passe chez Dorian Gray **et DONC** e ils commencent à avoir ils ont une discussi  
259. xiste, j(e) suis désolée [rires] 37 :00 B328 - **Et DONC** ee, eh ee Zeus n'est pas content # parce qu'i  
260. y aller B417 - Oui A417 - Donc ee voilà B418 - **Et DONC** t'es allée les récupérer le lendemain matin ?



## ANNEXE 13

### Présentation des résultats des natifs par tâche, t1 et t2

#### 1. Présentation des résultats temps 1

Sujet	BD1	BD1	Film1	Film1	Total t1	Total t1	Diversité des SP
	occurrences	/100 mots	occurrences	/100 mots	occurrences	/100 mots	Types/100 mots
AK	26	14,05	47	12,60	73	13,08	8,06
FA	39	18,75	53	14,89	92	16,82	13,48
HE	14	12,50	40	13,11	54	12,81	9,83
JA	19	13,01	39	10,86	58	11,49	7,72
MA	33	14,16	53	12,50	86	13,33	10,35
MJ	13	16,05	24	9,88	37	12,96	10,19
MR	18	12,68	24	13,48	42	13,08	9,69
NO	24	13,19	33	12,60	57	12,89	7,66
NY	10	14,71	21	8,24	31	9,60	6,19
OL	12	8,39	84	11,43	96	9,91	5,81
PR	37	15,81	74	14,12	111	14,97	9,76
QI	36	17,31	88	14,40	124	15,86	10,13
SU	16	12,90	18	12,59	34	12,75	10,11
VI	31	16,94	23	12,37	54	14,65	11,11
VO	26	16,88	65	14,44	91	15,66	11,75
XI	17	8,10	39	6,67	56	7,38	5,28
ZI	13	12,50	34	10,46	47	11,48	7,93

Sujets	Niveau de français	Fréquence des SP	Diversité des SP
ID	1=débutant 2=inter 3=avancé	/100 mots	/100 mots
AK	1	13,08	8,06
FA	3	16,82	13,48
HE	2	12,81	9,83
JA	1	11,49	7,72
MA	2	13,33	10,35
MJ	3	12,96	10,19
MR	2	13,08	9,69
NO	2	12,89	7,66
NY	1	9,60	6,19
OL	1	9,91	5,81
PR	3	14,97	9,76
QI	3	15,86	10,13
SU	1	12,75	10,11
VI	3	14,65	11,11
VO	3	15,66	11,75
XI	1	7,38	5,28
ZI	2	11,48	7,93

## 2. Présentation des résultats temps 2

Sujet	BD2	BD2	Film2	Film2	Total t2	Total t2	Diversité des SP t2
	occurrences	/100 mots	occurrences	/100 mots	occurrences	/100 mots	/100 mots
AK	30	14,78	57	14,25	87	14,51	7,79
FA	49	19,29	79	15,31	128	17,30	11,30
HE	30	17,34	33	13,92	63	15,63	11,21
JA	15	9,93	51	13,39	66	11,66	8,46
MA	13	10,92	34	12,55	47	11,74	10,77
MJ	31	17,51	27	11,25	58	14,38	10,79
MR	15	13,64	29	9,67	44	11,65	8,05
NO	15	10,87	54	15,52	69	13,19	11,11
NY	21	12,57	44	11,73	65	12,15	8,30
OL	16	13,79	62	10,39	78	12,09	6,59
PR	25	14,71	31	14,55	56	14,63	10,18
QI	42	18,18	94	13,35	136	15,77	9,84
SU	16	11,51	37	11,49	53	11,50	6,94
VI	32	14,29	68	16,96	100	15,62	12,32
VO	15	10,14	84	16,18	99	13,16	10,19
XI	28	9,76	46	8,23	74	8,99	5,67
ZI	19	10,61	71	11,53	90	11,07	7,30

Sujets	Niveau de français	Fréquence des SP	Diversité des SP
ID	1=débutant 2=inter 3=avancé	/100 mots	/100 mots
AK	1	14,51	7,79
FA	3	17,30	11,30
HE	2	15,63	11,21
JA	1	11,66	8,46
MA	2	11,74	10,77
MJ	3	14,38	10,79
MR	2	11,65	8,05
NO	2	13,19	11,11
NY	1	12,15	8,30
OL	1	12,09	6,59
PR	3	14,63	10,18
QI	3	15,77	9,84
SU	1	11,50	6,94
VI	3	15,62	12,32
VO	3	13,16	10,19
XI	1	8,99	5,67
ZI	2	11,07	7,30

## ANNEXE 14

### TextGrid text file

"Praat chronological TextGrid text file" de FA, Bd1, groupe 3

0 77.10185941043085 ! Time domain.  
1 ! Number of tiers.  
"IntervalTier" "TRANSCRIPTS" 0 77.10185941043085  
1 0 3.819688530745877  
"-3.819"  
1 3.819688530745877 10.39924232164389  
"l' histoire es:t divisée en deux parties la première parti:e dans laquelle on voit seulement un seul protagoniste qui est le chevalier"  
1 10.39924232164389 11.092762462109548  
"693"  
1 11.092762462109548 19.317179258394482  
"un cowboy qui: qui descend de son cheval il laisse le cheval euh juste à côté il sent fatigué donc euh il s' assoit sur une: petit banc\*(bank)"  
1 19.317179258394482 19.773062186066515  
"455"  
1 19.773062186066515 22.072898411892684  
"il commence à.: à se reposer à dormir"  
1 22.072898411892684 22.586937121662412  
"514"  
1 22.586937121662412 29.31543766796654  
"la deuxième parti:e c' est un peu différent parce que: &=bche elle est elle est basée sur la présence de deux deux protagonistes"  
1 29.31543766796654 29.687737762910054  
"372"  
1 29.687737762910054 31.827366403963648  
"parce qu' il va s' ajouter une\* petit enfant"  
1 31.827366403963648 32.492592969548  
"665"  
1 32.492592969548 37.96812416078842  
"qui: qui arrive dans l' histoire qui prend part d' un: partie dans l' histoire avec son petit cheval"  
1 37.96812416078842 38.397119701662426  
"428"  
1 38.397119701662426 39.57182059716062  
"c' est comme une petit jouet"  
1 39.57182059716062 40.265394885710215  
"693"  
1 40.265394885710215 48.28892203741587  
"e:t comme ça il est reste étonné devant le grand cheval du cowboy et il change son cheval avec le cheval du cowboy"  
1 48.28892203741587 48.70067271348602  
"411"  
1 48.70067271348602 51.41567601240827  
"et repart avec euh le le vrai cheval le grand cheval"

1 51.41567601240827 51.85979033886373  
"444"  
1 51.85979033886373 55.55397936314977  
"comme ça le\* le\* chevalier quand quand se lève quand se réveille"  
1 55.55397936314977 56.14361200082681  
"589"  
1 56.14361200082681 63.42690388912955  
"e:t il est il reste un peu étonné de euh parce qu' il reste juste en face du petit cheval qui était le jouet  
avant de:"  
1 63.42690388912955 64.04110455337647  
"614"  
1 64.04110455337647 64.52331252286832  
"eu:h (482)"  
1 64.52331252286832 64.73389415846556  
"du &pet"  
1 64.73389415846556 64.99658305794347  
"262"  
1 64.99658305794347 65.79331431390507  
"du petit enfant"  
1 65.79331431390507 66.17317380163932  
"379"  
1 66.17317380163932 69.8652825916925  
"donc eu::h il y a une\* changement de d' animau e:t"  
1 69.8652825916925 70.18277708890322  
"317 &=bche"  
1 70.18277708890322 72.07184060312552  
"et c' est c' est drôle la conclusion et"  
1 72.07184060312552 73.30512679275188  
"1.233"  
1 73.30512679275188 73.76247005659113  
"c' est tout"  
1 73.76247005659113 77.10185941043085  
"-3.339"

## ANNEXE 15

### Document de travail Excel - Calculs

FA							
BD1	durée	nb mots	nb syl	nb SP	nb syll	SP	SP
	pauses						
1	l' histoire es:t divisée en deux parties la première parti:e dans laquelle on voit seulement un seul protagoniste /qui est le chevalier	693	22	36	3	12	divisée en - la première partie -on voit...qui
2	un cowboy qui: qui descend de son cheval il laisse le cheval+euh juste à côté il sent fatigué donc+euh il s' assoit sur une: petit banc*.455	455	26	36	5	18	qui descend de - juste à côté - *sent fatigué- s'assoit sur - petit banc
3	il commence à: à se reposer à dormir .514	514	8	12	2	7	commence à - se reposer
4	la deuxième parti:e c' est un peu différent parce que: &=bche elle est elle est basée sur la présence de deux deux protagonistes .372	372	22	33	6	20	la deuxième partie - c'est un peu différent - parce que - elle est basée sur - la présence de
5	parce qu' il va s' ajouter *une petit enfant	665	9	13	4	11	parce que - il va - s'ajouter - petit enfant
6	qui: qui arrive dans l' histoire qui *prend part d' un: partie dans l' histoire avec son petit cheval	428	19	24	2	6	*prend part - petit cheval
7	c' est comme *une petit jouet	693	6	7	2	6	c'est comme - petit jouet
8	et: comme ça il est reste étonné devant le grand cheval du cowboy et il change son cheval avec le cheval du cowboy	411	23	32	2	7	comme ça - * reste étonné
9	et repart avec euh le le vrai cheval le grand cheval	444	10	15			
10	comme ça le le chevalier quand quand se lève quand se réveille	589	12	15	3	7	comme ça - se lève - se réveille
11	et:il est il reste un peu étonné de euh parce qu' il reste juste en face du petit cheval qui était le jouet avant de: [GH (.) 614 +euh (482)]	1096	25	32	6	18	un peu- étonné de - parce qu' - en face du- petit cheval - avant de
12	du &pet	262	1	2			
13	du petit enfant	379	3	5	1	4	petit enfant
14	donc+euh: il y a *une changement de d' animau et: &=bche	317	10	15	1	3	il y a
15	et c' est c' est drôle la conclusion et	1233	9	10	2	3	c'est - c'est drôle
16	c' est tout		3	2	1	2	c'est tout
	<b>totaux</b>	<b>8551</b>	<b>208</b>	<b>289</b>	<b>40</b>	<b>124</b>	0,43

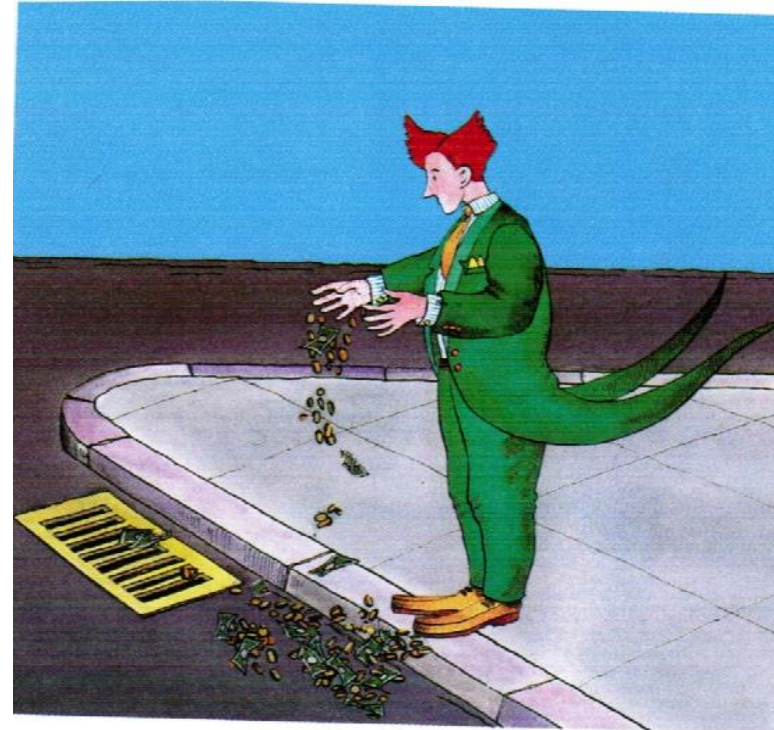
CALCULS	
segments	16
segments contenant des SP	14
rapport entre segments	0,88
nb de SP par segments	2,5
nb de SP par segment contenant des sp	2,9
nb de syllabes/syllabes total	0,43
nb de syllabes de sp ds segments avec SP	8,9
longueur moy des SP en syllabes	3,1
nb de syllabes de SP par segment	7,8

ANNEXE 16 *Idiomatics*



Jeter l'argent par les fenêtres

*To throw money out of the window*



To throw money down the drain

*Jeter l'argent dans l'égout*