

Les séquences préfabriquées à travers les stades de développement en français L2

Bartning, Inge et Forsberg, Fanny, Université de Stockholm

Acquisition des langues

Il sera question, dans ce qui suit, d'une mise en rapport de deux domaines d'étude acquisitionnels, à savoir l'hypothèse de stades de développement dans une langue seconde d'une part, et les séquences préfabriquées d'autre part. Dans les recherches acquisitionnelles, nombre d'études essaient de dégager des stades de développement ou des itinéraires acquisitionnels, notamment dans le domaine de la morphosyntaxe. Parallèlement, on s'intéresse aussi de plus en plus à des séquences qui échappent aux règles de syntaxe et de lexique, et qui sont acquises en tant qu'unités. Se peut-il que les séquences préfabriquées (SP), tout comme la morphologie verbale, entre autres, suivent un itinéraire acquisitionnel généralisable ? Des données du français parlé des apprenants à différents niveaux d'acquisition et des locuteurs natifs nous permettront de considérer de plus près cette question. Seront étudiés des paramètres comme la quantité de SP dans la production langagière et la distribution et l'apparition de différentes catégories de SP. Ainsi, les résultats montreront que la quantité de SP augmente à mesure que l'apprenant avance, et que certaines catégories de SP sont des traits précoces alors que d'autres sont des traits tardifs. Il est donc mis en évidence que les SP sont d'une grande importance tout au long de l'acquisition et dans l'usage des locuteurs natifs.

1. Introduction

Dans les recherches acquisitionnistes on s'est depuis longtemps intéressé au développement des deuxièmes langues et au fait que ces dernières semblent suivre un certain ordre d'acquisition. Ces études concernent surtout d'autres langues que le français (voir Clahsen et al. 1983, Klein & Perdue 1997, Pienemann 1998). Récemment ont été proposés des stades de développement du français L2 acquis par les adultes suédophones, décrits comme des profils grammaticaux d'apprenants à différents niveaux. Ces recherches se fondent sur des résultats empiriques des travaux menés au sein de deux projets différents sur l'acquisition du français oral des apprenants suédophones (Universités de Stockholm et de Lund). À partir d'*itinéraires acquisitionnels* on a identifié six *stades*, qui vont des débuts de l'acquisition jusqu'à la production d'apprenants quasi-natifs. Ces itinéraires, et les stades qu'on a tenté d'en déduire, reflètent l'acquisition de phénomènes morphosyntaxiques du français par l'apprenant suédophone, dans des situations orales spontanées où il/elle doit avoir recours à ses connaissances « automatisées ». Le corpus InterFra comporte des productions d'apprenants (semi-)guidés et le corpus de Lund surtout des productions d'apprenants non-guidés. Or, il est fort intéressant de constater que les stades et les itinéraires acquisitionnels (susmentionnés) semblent valoir pour ces deux types d'apprenants (pour plus de détails, voir Bartning & Schlyter 2004a). Pour approfondir ses réflexions sur la notion de stades et de continuum en acquisition, voir Sharwood-Smith & Truscott 2005, Granfeldt & Schlyter 2005.

Dans la présente étude, de nouveaux résultats (Forsberg 2005) concernant l'acquisition des séquences préfabriquées (SP) seront examinés à la lumière des stades d'acquisition proposés. Les séquences préfabriquées (*formulaic sequences*, Wray 2002) se sont vu attribuer un statut de plus en plus important dans la théorie linguistique et des chercheurs tels que Schmitt & Carter (2004) soulignent le besoin urgent d'études traitant de leur fonctionnement et de leur rôle dans l'acquisition d'une langue seconde. Certains (comme Myles et al. 1998) estiment que les SP constituent avant tout un élément caractéristique des premiers stades, alors que d'autres soutiennent que la maîtrise des SP constituent le dernier pas vers la compétence native (cf. par ex. Yorio 1989 et Wray 2002), ce qui met en évidence le caractère très hétérogène de la notion de séquence préfabriquée. Nous avons élaboré une catégorisation des SP incluant à la fois des SP caractéristiques des productions initiales et des expressions typiques présentes surtout aux stades avancés et dans la production des locuteurs natifs: cette catégorisation nous permettra de discuter si l'emploi des SP permet de caractériser le développement d'une deuxième langue.

2. Les deux corpus

Les données proviennent de deux grands corpus constitués de productions orales en français de locuteurs adultes suédophones: le corpus InterFra (de Stockholm), sous la responsabilité d'Inge Bartning, et celui de Lund, sous celle de Suzanne Schlyter. Les deux corpus sont présentés plus en détail dans les tableaux 1 (InterFra) et 2 (Lund) ci-dessous.

Tableau 1. Le corpus *InterFra*, Stockholm

	Long/ Trans	Nombre d'appr	Années d'étude	Âge	Période suivie	Occas. d'enreg	Tâches: Int 15-20 min, BD, Vid	Type d'ac- quisition
LNN								
Gr 1 Débutants	a) long b) trans	a) 6-10 b) 18	0	19-30	2 semest.	5-10 1	+	guidée
Gr 2 Lycéens	a) long b) trans	a) 13 b) 20	3,5	16-18	2 ans	2 1	+	guidée
Gr 3 Étudiants	a) long b) trans	a) 6-8 b) 18	(4,5) – 6	19-26	4-5 semest.	10 1	+	semi-guidée
Gr 4 Futurs profs	long	6	7-8	23-34	2 semest. +stage	4	+	semi-guidée (sauf BD)
Gr 5 Doctorants	trans	5	8-9	23-26	-	1	+	semi-guidée
LN								
Gr 6	-	20	-	23-26	-	-	+	

Légende: LN=locuteur natif, LNN=locuteur non natif ; Long=groupe longitudinal ; Trans= groupe transversal ; Étudiants=étudiants universitaires, du 1er au 4e semestre, Futurs profs= étudiants futurs professeurs; Int=Interview, BD=récits de bandes dessinées, Vid=récits de films vidéo muets.

(Pour plus de détails sur le corpus InterFra, voir Bartning 2002)

Tableau 2. Le corpus *Lund*

	Long/ Trans	Nombre d'appr	Années d'étude de français	Âge	Période suivie	Occas. d'enreg	Tâches : Int 30-40 min, récits, traduction	Type d'ac- quisition
LNN								
Adultes séjournant en France	long	8	quelques semaines ou 0	20-50	2 à 15 mois	2 à 5	+	non-guidée
Étudiants d'université	long	7	6 ans de lycée	19-22	5 à 12 mois	4 à 9	+	guidée
Débutants (Corpus Thomas)	trans/ long	5	25 - 50 heures de cours	35-70	6 mois	1 à 2	interv	guidée

(Pour une description détaillée des parties de ce corpus, voir Granfeldt, 2005)

3. Bilan des stades – les phénomènes prototypiques de chaque stade

Ci-dessous seront présentés les traits les plus caractéristiques de chaque stade. Ces traits ont été répertoriés dans 80 interviews des deux corpus selon leur apparition chronologique, ce qui a permis de créer une échelle implicationnelle détaillant les traits morphosyntaxiques en

groupements de stades (voir Bartning & Schlyter 2004a, p.294; cette échelle se trouve aussi, sous forme abrégée, en Appendice ci-dessous; pour la méthodologie, voir aussi Bartning & Schlyter, 2004b en version étendue sur le web). La procédure méthodologique est celle du « bottom-up », c'est-à-dire que nous sommes parties des données empiriques de chaque interview ; nous les avons répertoriées selon des itinéraires acquisitionnels, dans lesquels une forme A précède une forme B qui précède une forme C du même phénomène grammatical, par ex. l'accord sujet-verbe, la négation etc. Ensuite, ces différents phénomènes, grâce à leur présence simultanée dans chaque production, ont été rassemblés en faisceaux de traits qui forment ainsi les stades.

Ces stades ont aussi été utilisés et testés sur d'autres corpus de français L2, notamment ceux de Ågren 2005 et de Granfeldt 2005 (CEFLE, Flloc), avec un résultat prometteur quant au potentiel généralisable des stades proposés.

Stade 1 – le stade initial

À ce premier stade, l'apprenant se sert d'une structuration en grande partie nominale (*je // des amis* pour *j'ai rencontré des amis*) mais aussi de moyens grammaticaux. Il/elle emploie aussi bien des « formes finies » que des formes non-finies des verbes, ce qui montre que cette opposition n'est pas maîtrisée et que l'apprenant ne fait pas la distinction entre les personnes des formes verbales. On trouve, pourtant, à ce stade l'article défini et indéfini ainsi que le pronom sujet *je*, bien que très souvent accentués et non élidés (Granfeldt et Schlyter, 2005). L'apprenant utilise aussi la négation Nég X (*non grand-lit*) et des expressions non-analysées telles que *je(ne)sais pas* et *je m'appelle*. L'apprenant utilise quelques rares formes du passé composé mais très peu de formes du passé sont marquées. Il y a aussi un début d'emploi des connecteurs *et*, *mais* et *puis*. Ce stade peut être comparé au *Niveau Prébasique* ainsi qu'à la transition vers la *Variété basique* (Klein & Perdue 1997, ESF).

Stade 2 - le stade post-initial

Ce stade est caractérisé par l'apparition de quelques phénomènes grammaticaux, mais encore très variables. La subordination simple introduite par *quand*, *parce que*, *qui* et *que* apparaît. Il y a également apparition de la négation préverbale *ne* sans *pas*, à côté de la négation postverbale. L'apprenant commence à utiliser des formes verbales modales (suivis d'infinitifs) et le futur périphrastique. Son emploi du PC est devenu plus productif. Ainsi on voit que certains morphèmes grammaticaux apparaissent tôt. Quelques apprenants guidés utilisent parfois l'imparfait dans les formes *était* et *avait*. Les formes verbales non-finies dans

un contexte fini (*je donnE, je faire*) sont encore fréquentes mais le nombre de « formes finies » augmente. L'accord sujet-verbe est marqué par l'opposition entre la 1ère et la 2ème personne au singulier pour les verbes non-thématiques *être, avoir* ainsi que par l'alternance de *nous V-ons* et de **nous donnE*, sans désinence. Les pronoms objet sont généralement postposés. L'apprenant à ce stade a aussi abondamment recours à *c'est* en tant que constructeur d'énoncés et de structure passe-partout (Bartning, 2005). Une comparaison avec le projet ESF suggère que ce niveau corresponde en partie à la *Variété basique* : la plupart des énoncés contiennent un verbe, mais la morphologie de ces verbes est loin d'être établie.

Stade 3 - le stade intermédiaire

Au stade intermédiaire l'apprenant semble avoir développé une variété plus systématique et régulière, bien que très simple encore. Cette interlangue est caractérisée par des surextensions et des régularisations de sorte qu'elle n'est pas toujours conforme aux normes de la langue cible. L'apprenant utilise cependant la négation (*ne Verbe fini pas*) comme dans la langue cible. De plus, il/elle emploie le passé, marquant la plupart des références au passé, ainsi que le futur (périphrastique; le futur simple, dans des cas isolés) pour les références à l'avenir. Les formes verbales non-finies dans des contextes finis existent encore mais se raréfient. Les formes *nous V-ons* sont marquées dans la plupart des cas pour les verbes thématiques et non-thématiques. L'opposition entre la 3e personne du singulier et du pluriel pour *ont, sont, etc.* commence à s'établir relativement bien (dans la moitié des cas environ) mais l'apprenant à ce stade continue à employer en alternance *ils donnE, ils *prendre, ils *prend* et parfois *ils prennent*. Les pronoms objet sont placés devant le verbe thématique pour les temps composés et simples - souvent incorrectement après l'auxiliaire *est/a*. La subordination se complexifie aussi avec l'apparition des subordonnées causales, temporelles, relatives, complétives et interrogatives. Il y a même quelques premiers emplois du subjonctif.

Stades 4-6 - les stades avancés

Ces stades sont représentés par un éventail plus large de structures d'énoncés et sont caractérisés par la grammaticalisation de la morphologie flexionnelle, qui devient fonctionnelle mais avec des zones « fragiles » de développement. La raison pour laquelle nous divisons la variété avancée en trois niveaux tient au fait qu'au fur et à mesure que l'interlangue se développe, cette évolution crée une richesse d'expressions permettant un choix dans le répertoire.

Stade 4 - le stade avancé inférieur

Au stade avancé inférieur, des structures spécifiques du français, plus complexes et variées, apparaissent : le pronom clitique avant l'auxiliaire, le conditionnel, le plus-que-parfait et le subjonctif.

Des formes verbales plus complexes commencent donc à apparaître - formes qui sont liées à des structures syntaxiques multipropositionnelles (comme le subjonctif), ou impliquant une relation discursive (relations temporelles et conditionnelles); ces contextes ne sont cependant pas encore systématiquement marqués par les formes pertinentes. Les négations deviennent complexes, contiennent *rien, jamais, personne* et elles sont, pour la plupart d'entre elles, bien placées et correctes. À ce niveau, les formes verbales non-finies (*ils *parlE*) au lieu d'une forme finie n'existent pratiquement plus. Les formes *ont* et *sont* sont plus fréquentes que *ils *a* et *ils *est*, et le pluriel est souvent marqué pour les verbes thématiques, du type *ils prennent* (à côté de *ils *prend*). La clitisation de l'article est acquise ainsi que celle du pronom sujet. Le genre du nom pose cependant toujours des problèmes à l'apprenant. Il y a diversification des connecteurs ainsi que suremploi significatif par rapport aux natifs des marqueurs polyfonctionnels *mais* et *parce que*.

Stade 5 - le stade avancé moyen

L'un des traits les plus caractéristiques du stade avancé moyen est la poursuite du développement de la morphologie flexionnelle : les formes verbales *ils sont, ont, etc.* sont employées correctement de même que, dans la plupart des cas, le plus-que-parfait, le conditionnel et le futur simple. C'est à ce niveau que le subjonctif devient plus productif. L'accord en genre du déterminant et surtout de l'adjectif pose toujours des problèmes. Pourtant, même à ce niveau, il reste quelques résidus de la forme au pluriel du présent des verbes thématiques irréguliers *ils *sort*. La négation est utilisée en tant que sujet. On trouve aussi des relatives introduites par *dont* et le gérondif.

Stade 6 - le stade avancé supérieur

Ce stade avancé supérieur est caractérisé par une morphologie flexionnelle stabilisée, même dans les énoncés multipropositionnels. On y observe un emploi natif des connecteurs *enfin* et *donc*, un haut degré d'empaquetage, d'ellipse et d'intégration des propositions caractérisant une capacité à gérer plusieurs niveaux informationnels au sein du même énoncé. Il y a aussi un emploi presque natif des relatives macro-syntaxiques. Le fait que le développement soit

très similaire chez les différents types d'apprenants, guidés et non-guidés, indique qu'il ne s'agit ni d'un développement aléatoire ni d'un développement qui est très lié aux facteurs externes.

Dans la section suivante nous montrerons les premiers résultats de l'établissement d'un lien entre l'acquisition et de l'usage des SP et des stades développementaux qui viennent d'être présentés (cette fois basée exclusivement sur un sous-corpus d'InterFra). Ces résultats seront précédés par un aperçu théorique et une catégorisation préliminaire des SP. Les résultats montreront ainsi l'émergence et la fréquence des différentes catégories de SP répertoriées à travers les stades. À notre connaissance, ce type d'étude n'a pas de précédent.

5. Les séquences préfabriquées – repères théoriques, catégorisation et méthode

Ci-dessous suit un survol général des séquences préfabriquées et un compte rendu de la catégorisation employée. On traitera ensuite de notre méthode et des données examinées dans cette étude (cf. Forsberg 2005, Forsberg et al. à paraître).

5.1. Repères théoriques

Au cours des dernières années, on s'est de plus en plus intéressé aux séquences préfabriquées dans la production langagière, à savoir des suites de plusieurs mots qui sont stockées en tant qu'unités holistiques et qui ne sont pas générées par la grammaire ou le lexique au moment de l'énonciation (cf. Wray 2002, Forsberg, 2005). Les SP en français peuvent être aussi bien des collocations comme « regarder de haut » que des marqueurs discursifs comme « c'est vrai que ». Wray (2002) postule que les SP constituent une *stratégie de production*. Selon le modèle de Wray, les SP facilitent à la fois la production orale et la compréhension de la part de l'interlocuteur, d'une part, parce qu'il est plus facile d'extraire un bloc tout fait pour le locuteur que de créer un nouvel énoncé et d'autre part, parce que ce bloc sera conventionnalisé et donc facilement reconnu par son interlocuteur.

Erman & Warren (2000, p.31) ont également proposé une définition :

*A prefab is a combination of at least two words
favoured by native speakers in preference
to an alternative combination which could have been
equivalent had there been no conventionalisation.*

Cette dernière définition s'applique uniquement aux séquences du répertoire natif, étant donné que le noyau de la définition est « préférée par le locuteur natif ». Comme nous travaillons principalement sur un corpus de locuteurs non-natifs, nous ne pouvons pas nous servir entièrement de cette définition, il a donc été nécessaire de la modifier (Forsberg 2005, p. 35) :

*Une séquence préfabriquée dans l'interlangue
est une séquence d'au moins deux mots qui est ou semble être stockée
comme une unité et qui n'est pas générée ou analysée par la grammaire
ou le lexique. Elle peut être une combinaison préférée par
les locuteurs natifs, mais elle peut aussi être une combinaison
qui n'est utilisée en tant qu'unité non-segmentée que par l'apprenant d'une L2.*

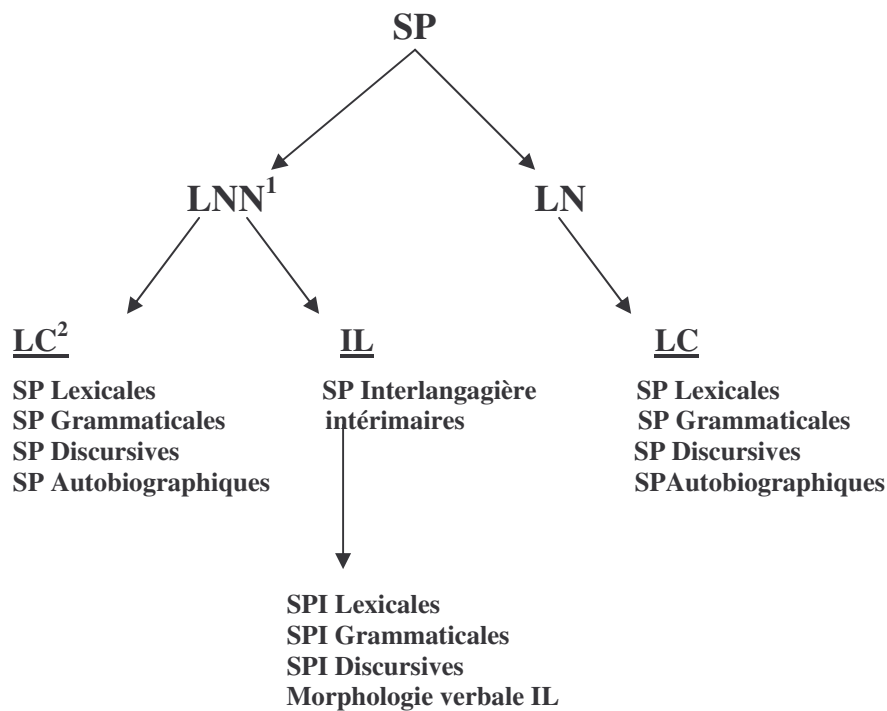
L'intérêt des acquisitionnistes pour ce phénomène porte sur deux domaines (Weinert 1995, p.184). Le premier concerne le langage préfabriqué utilisé par les locuteurs natifs et sa présence dans la production des apprenants. On considère souvent la maîtrise des séquences préfabriquées comme la « clef » de l'idiomaticité, le dernier pas vers la compétence native. La différence d'emploi des séquences préfabriquées entre apprenants très avancés et locuteurs natifs permet de distinguer ces deux groupes. Le deuxième domaine d'étude est purement concentré sur l'apprenant et son interlangue. Si la langue cible a des suites préfabriquées, l'interlangue en a également et ce ne sont pas toujours les mêmes. Ces séquences sont la plupart du temps préfabriquées pendant une certaine période pour ensuite être segmentées, c'est-à-dire que l'apprenant « ouvre » l'unité et commence à employer les éléments inclus de façon indépendante . Un exemple de notre corpus est fourni par la suite « c'est tout passé bien », produite à plusieurs reprises par Robert, un apprenant débutant.

- (1) I : qu'est-ce que tu as pensé ?
E : **c'est tout passé bien** . (I :mm) c'est eh # eh je / j'ai eh eh eh eu beaucoup de faire.
(Robert, INT4)

On peut supposer qu'il s'agit d'une inversion de l'ordre des mots dans l'expression « tout s'est bien passé », et que cette forme reste non-analysée pendant une certaine période dans l'interlangue de cet apprenant.

La figure 1 ci-dessous rend compte de la distribution des séquences préfabriquées en français parlé L2 ainsi qu'en L1.

Figure 1. Les SP en français parlé L2 et L1



Dans la figure, on peut observer l'existence de différentes catégories de SP comme les SP Lexicales et les SP Discursives. On voit qu'il y a des recoupements importants entre les catégories présentes dans la production des LN et celles de la production des LNN. Ces derniers ont toutefois une catégorie qui leur est propre, à savoir les SPI (SP Interlangagières intérimaires). Passons maintenant à une description et discussion de ces catégories, fondamentales dans l'analyse que nous mènerons par la suite.

5.2 Identification: les SP Lexicales, Grammaticales et Discursives, les SP Interlangagières et les SP Autobiographiques

Étant donné qu'une étude de ce type n'a pas encore été effectuée en français L2, ni en français L1, nous avons développé notre méthode d'identification et catégorisation à tâtons, et en ayant recours à plusieurs études anglo-saxonnes. Comme il ressort de la figure 1 et de la section 5.1, l'étude des séquences préfabriquées chez les locuteurs non-natifs est un domaine vaste. Pour le chercheur, il s'agit, d'une part, de détecter les SP propres à la langue cible, vers laquelle l'apprenant évolue, et, d'autre part, de décrire et d'identifier les séquences purement interlangagières. Nous touchons ainsi à deux extrêmes de l'acquisition d'une langue : l'un ayant trait à la préférence native, à savoir l'idiomaticité, l'autre ayant une origine

¹ LNN = locuteur non-natif et LN = locuteur natif

² LC = langue cible et IL = interlangue

idiosyncrasique, et, par conséquent, non-idiomatique. Nous étudions donc à la fois le développement des unités qui rendent le discours plus idiomatique et celles qui le rendent moins idiomatique, le dénominateur commun étant la combinatoire des mots. En étudiant ces unités à la lumière des stades de développement, nous espérons découvrir la manière dont l'usage de ces unités évolue, tout en mettant l'accent sur la typologie des SP et sur la quantité de SP dans la production.

Les trois premières catégories sont inspirées des catégories de Erman & Warren (2000). Nous les avons adaptées à la langue française.

SP Lexicales

La première catégorie possède les caractéristiques principales suivantes:

- Les SP réfèrent à un phénomène extralinguistique
- Ce sont des mots de contenu / classes ouvertes

Nous proposons, à l'instar de Erman & Warren une sous-catégorisation syntaxique :

- SN: *coup de foudre, centre ville*
- SAdv: *à tâtons, à pied*
- SV : *faire la fête, faire un retour sur soi-même*
- SP : *au bout de, à côté*
- Réflexes conversationnels : *c'est pas évident, c'est pas mal*

Pour l'identification des SP, Erman & Warren se servent du critère de la commutation restreinte (*restricted exchangeability*). Ce critère signifie qu'au moins une entité de la SP ne peut pas être remplacée par un synonyme sans changer le sens ou la fonction de la combinaison et faire perdre l'idiomaticité de la séquence. Erman & Warren se sont principalement servis de l'intuition comme mesure lorsqu'il fallait choisir entre deux options. Comme l'intuition s'est avérée problématique comme outil d'identification (cf. Wray 2002), nous avons opté pour une méthode d'identification sur corpus. Jusqu'ici, faute de grands corpus oraux informatisés facilement accessibles en français, nous nous sommes servis du moteur de recherche www.google.fr. Nous procédons comme suit :

- 1) Intuition. Dans notre recherche des séquences préfabriquées, nous détectons à l'aide de notre intuition, suivant le critère de la commutation restreinte, les séquences qui nous semblent être préfabriquées. Ensuite nous passons à la recherche sur Google.
- 2) Recherches sur www.google.fr, option *Pages francophones*

a) Recherche de la SP potentielle, par exemple *Poser une question* qui obtient 682 000 occurrences.

b) Recherche d'une séquence modifiée, où l'un des mots a été changé contre un synonyme, dans ce cas *Demander une question* 2160 occurrences, *Planter une question* 5 occurrences.

Au vu d'écart aussi importants, nous considérons qu'il y a une préférence native dans la combinatoire du mot *question*. Dans la plupart des cas, les écarts sont du même ordre et le choix est ainsi fait assez facilement.

En voici des exemples de SP Lexicales en contexte. Cette fois-ci il s'agit de l'apprenant universitaire Christina. On observe trois SP différentes : « essayer de », « faire du sport » et « de temps en temps », la dernière étant sans doute la plus idiomatique. On notera que les trois séquences sont préfabriquées pour différentes raisons : dans le cas de « essayer de », c'est la préposition qui ne tolère pas de changement ; dans le deuxième cas, « faire du sport », il s'agit d'une construction à verbe support (cf. Gross 1996), où l'on ne peut pas remplacer *faire* par un autre verbe équivalent (qui est d'ailleurs difficile à trouver) et dans le dernier cas, « de temps en temps », nous avons affaire à une séquence entièrement figée : on ne peut changer ni le nom, ni la préposition, ni la duplication du nom. Ces divers exemples de de préfabrication figurent ainsi dans cet extrait.

- (2) I : mhm . / et quand tu n'étudies pas :s / quand tu n(e) travailles pas à *Systemet / que fais tu ?
E : (SOUPIR) e:h / je vois mes amis . **j'essaie de faire un peu d'sport**. je nage **de temps en temps** .
et je cours aussi **de temps en temps**. (Christina, apprenant universitaire)

SP Grammaticales

Les caractéristiques principales de cette catégorie sont les suivantes :

- Les SP n'ont pas de référence extralinguistique
- Elles déterminent et quantifient au sein des structures syntaxiques
- Mots fonctionnels / classes fermées

Elles sont sous-divisées selon le modèle ci-dessous:

- Déterminants (pronoms indéfinis): *une sorte de, le/la/les même(s)*
- Quantificateurs (pronoms indéfinis): *un petit peu, la plupart*
- Marqueurs aspectuels: *être en train de, venir de*
- Intensificateurs: *pas du tout, beaucoup plus*

Nous tenons à souligner que les phénomènes visés ici n'englobent pas le concept de grammaire en général ; il s'agit tout simplement d'une petite partie de la grammaire qui contient aussi des formes préfabriquées. Voici un exemple du marqueur aspectuel « venir de », tiré du sous-corpus de lycéens :

- (3) I : est-ce que tu as le temps de : sortir un peu
d'aller au cinéma de :...
E : oui (I :mm) **je viens de** : / de voir de regarder des films .
I : mm que- quel est le dernier film que tu as vu ?
E : st / (RIRE) maintenant nous+ / sommes regarder le film
Camille Claudel dans le / leçon . (Oskar, Lycéen, INT2)

SP Discursives

Ses caractéristiques principales sont les suivantes :

- Elles n'ont pas de contenu extralinguistique
- Elles se placent à l'intérieur et à l'extérieur des structures syntaxiques
- Mots fonctionnels
- Ce sont souvent des « tics », employés prototypiquement comme des remplisseurs ou des îlots de sécurité
- Certaines n'appartiennent qu'à la langue parlée

Elles se subdivisent de la façon suivante :

- a. Marqueurs textuels
 - *Parce que*
 - *Par contre*
 - *C'est*
- b. Marqueurs autorégulateurs
 - *Je crois (que)*
 - *Je veux dire*
 - *C'est-à-dire (que)*
- c. Marqueurs interactifs
 - *Bien sûr*
 - *D'accord*
 - *Tout à fait*

L'exemple (4) ci-dessous contient un grand nombre de SP Discursives, à savoir « c'est-à-dire que », « c'est » (fonction textuelle est autorégulatrice) et, enfin, « quand même ». Les nombreuses répétitions de « c'est » montrent que c'est une séquence qui a une influence importante sur la planification de la production orale et que c'est un vrai îlot de sécurité.

- (4) E : **c'est-à-dire que c'est c'est c'est**
c'est eh la meilleure c'est d'avoir / tout lu.
mais / mais ça marche **quand même**. (Lena, apprenant universitaire)

Il convient cependant de noter qu'il existe des chevauchements à plusieurs niveaux. Le premier concerne la division entre SP Grammaticales et Discursives, où il n'est pas toujours facile de trancher entre fonction grammaticale et fonction discursive. Le deuxième concerne les sous-catégories des SP Discursives. Il est bien connu qu'un grand nombre de marqueurs discursifs sont polyfonctionnels : nous nous référons donc à la fonction la plus fréquente d'une séquence donnée pour la ranger dans une des sous-catégories.

Ces trois premières catégories sont développées à partir de la langue cible et suivent de près la catégorisation proposée pour l'anglais par Erman & Warren (2000). La quatrième et la cinquième catégorie ont, quant à elles, leur origine dans les données interlangagières.

SP Interlangagières

Cette catégorie est notamment inspirée des travaux de Myles et al. (1998), mais aussi des critères proposés par Peters (1983). Il se trouve que, au début de l'acquisition d'une langue, que ce soit en L1 ou en L2, la notion de séquence préfabriquée implique souvent des formes qui semblent non-analysées, mais qui divergent syntaxiquement ou sémantiquement de la langue cible. Nous en trouvons, logiquement, notamment chez les apprenants lycéens et débutants de notre corpus. Les caractéristiques principales en sont les suivantes :

- Ce sont des versions déviantes syntaxiquement ou sémantiquement des SP Langue cible (*c'est tout passé bien*, ayant comme origine *tout s'est bien passé*)
- Ce sont des constructions idiosyncrasiques, n'ayant pas son origine dans une SP Langue cible (*une l'église*). Dans l'exemple donné, il est visible que la forme libre « l'église » n'a pas été segmentée par l'apprenant. Il se peut donc que des formes, qui ne sont pas figées pour le locuteur francophones, soient figées pour le locuteur non-natif.
- Elles sont souvent utilisées par un seul apprenant

Leur mode d'identification est bien entendu tout autre que celles de la langue cible, étant donné qu'il ne s'agit pas de préférences natives, pour lesquelles il est possible d'avoir recours à un corpus de référence native, mais de particularités de l'interlangue. Nous les considérons comme des séquences potentiellement préfabriquées, s'il y a au moins deux occurrences de la

même suite, dans la production d'un même apprenant. La catégorie présente ensuite une sous-division, contenant entre autres des versions déviantes des catégories dans la langue cible :

- SPI Lexicales
- SPI Grammaticales
- SPI Discursives
- Morphologie IL

Ci-dessous figure un exemple d'une SPI Lexicales. Cette fois-ci l'apprenant débutant Robert ne semble pas avoir acquis la structure interne de la séquence « ça va », qu'il traite comme un verbe simple.

- 5) I : comment ça va depuis euh le mois de : déc
E : euh :
I : depuis la dernière fois ?
E : euh co ss co comment **je ça va** ou ?
I : oui comment ça va ?
E : ah **je ça va** bien . (Robert, apprenant débutant)

SP Autobiographiques

Cette catégorie se fonde sur les travaux de Nattinger & DeCarrico (1992), qui présentent un inventaire des séquences préfabriquées, dont la sous-catégorie *Sujets nécessaires, Autobiographie*, leurs catégories étant construites à partir de critères fonctionnels. Les séquences incluses dans cette sous-catégorie sont :

- *Je m'appelle, J'habite, J'ai X ans, Je travaille à*

Ces séquences sont saillantes dans les tout premiers enregistrements des apprenants débutants, dans la mesure où elles se distinguent du reste de la production par leur conformité aux normes de la langue cible, alors que le reste de la production est peu développé. En voici un exemple illustratif où l'apprenant lycéen Clara semble maîtriser la séquence auto-présentatrice « avoir X ans », mais elle n'arrive pas lorsqu'il s'agit d'une autre personne, en l'occurrence sa sœur.

(6) E : euh*** / . et **j'ai seize ans** (I :oui) / . et + je (Clara, Apprenant lycéen, INT1)

(7) E : euh ma soeur est 23 ans (I :mm) / . (Clara, Apprenant lycéen, INT1)

Afin d'examiner leur statut préfabriqué, nous avons effectué une étude des formes verbales dans la première interview avec les apprenants débutants, dans laquelle nous avons répertoriées toutes les formes finies et toutes les formes non-finies. Nous avons pu conclure

qu'une grande partie des formes finies, produites conformément à la norme de la langue cible, appartiennent à cette catégorie. Mais cela justifie-t-il de les considérer comme préfabriquées à des niveaux plus avancés, où la morphologie verbale est plus développée ? Nous adopterons ici l'hypothèse de Wray sur la « non-segmentation » (*needs-only-analysis*), selon laquelle toute unité une fois acquise de façon holistique ne sera pas segmentée, puisqu'il est plus commode et efficace de la garder en bloc et de l'extraire lorsque une même occasion de l'employer se présente.

5.3 Méthode et données

Nous avons eu recours aux interviews réalisées avec 13 apprenants (25 interviews). Ces interviews ont été classées, selon les critères morphosyntaxiques proposés par Bartning & Schlyter (2004a), à cinq stades de développement différents. Ont également été étudiées cinq interviews réalisées avec des LN du français, soit un total de 30 interviews. La raison pour ne pas inclure le sixième stade présenté dans Bartning & Schlyter (2004a) est que, nous n'avons pas, à ce jour, étudié à fond l'emploi des SP des apprenants appartenant à ce stade.

Dans toutes ces interviews, nous avons repéré toutes les séquences préfabriquées appartenant aux catégories susmentionnées. Nous avons calculé la distribution des différentes catégories ainsi que le pourcentage de SP dans la production orale de chaque apprenant. Les différentes catégories et le taux de langage préfabriqué dans la production sont donc les paramètres que nous utilisons pour décrire l'usage des SP.

Tableau 3. Étudiants choisis pour cette étude, 'Corpus Copenhague'

	Long	Nombre d'appr	Années d'étude	Âge	Période suivie	Occas. d'enreg	Tâches: Int 15-20 min, BD, Vid	Type d'acquisition
LNN								
Gr 1 Débutants	Ro Int1-7, Em 1-4, Pe 1-4, Ja 1-4	4	0	19-30	2 semest.	5-10 1	+	guidée
Gr 2 Lycéens	To Int1-2, Li 1-2, Os 1-2	3	3,5	16-18	2 ans	2 1	+	guidée
Gr 3 Étudiants Universit.	Int 4: Ev, Ma, Yv, Chr, Pe, Le	6	(4,5) – 6	19-26	4-5 semest.	10 1	+	semi-guidée
LN	-	5	-	23-26	-	-	+	

5.4 Hypothèses

Afin de préciser ce qui nous intéresse plus particulièrement dans le rapport entre stades de développements, d'un côté, et séquences préfabriquées de l'autre côté, nous émettons deux hypothèses qui structureront cette étude.

1) Étant donné que la production des locuteurs natifs contient le taux le plus important de langage préfabriqué, nous supposons que plus l'apprenant avance dans son acquisition, et ainsi s'approche de la LC, plus il emploiera de langage préfabriqué.

2) Compte tenu des résultats obtenus dans Forsberg (2005,) nous supposons que les catégories SPILI et SP Autobiographiques sont caractéristiques des stades initiaux, et qu'elles diminuent par la suite. Par contre, la catégorie des SP Lexicales est caractéristique de la production des natifs et constitue le pas le plus difficile à franchir pour les apprenants L2.

6. Les SP et les stades 1-5

Ci-dessous sont présentés sous forme de tableau les résultats de l'étude. Dans les colonnes 3-8 est indiquée la distribution des différentes catégories. La neuvième et dernière colonne contient le résultat d'un tout autre calcul et ne doit donc pas être confondue avec les autres : il s'agit ici du pourcentage de mots appartenant à des SP par rapport aux mots produits au total par un apprenant donné. Jan 1 a produit 327 mots, dont 36 font partie de séquences préfabriquées. Ainsi, nous avons obtenu le pourcentage indiqué dans la dernière colonne, à savoir 11 % de langage préfabriqué dans sa production orale.

Tableau 4. Résultats stades / SP

Stade	Informateur et interview	SP Lexicales	SP Gram	SP Discursives	SPILI	SP Autobio	Total SP	Pourcentage de SP
1. Stade initial	Heidi 1	2	0	3	15	7	27	8 %
	Jan 1	3	1	9	2	2	17	11 %
	Carin 1	0	1	0	0	3	4	4 %
	Clara 1	3	4	8	0	6	21	10 %
	Cia 1	3	2	3	0	4	12	13 %
Total		11 (14 %)	8 (10%)	23 (28 %)	17 (21 %)	22 (27 %)	81	Moyenne: 9,2 %
2. Stade post-initial	Emelie 1	3	4	24	7	4	42	17 %
	Jan 4	3	9	52	4	1	69	21 %
	Oskar 1	1	5	11	0	4	21	11 %
	Pelle 1	4	1	3	5	1	14	7 %
	Pelle 4	2	3	21	0	1	27	12 %
Total		13 (8 %)	22 (13 %)	111 (64 %)	16 (9 %)	11 (6 %)	173	Moyenne : 13,6%
3. Stade intermédiaire								

	Emelie 4	7	4	19	0	5	35	22 %
	Linnea 1	6	4	27	0	4	41	14 %
	Linnea 2	7	3	28	5	3	46	15 %
	Oskar 2	5	4	41	4	3	57	17 %
	Robert 4	9	8	38	4	5	64	19 %
Total		34 (14%)	23 (9%)	153 (63%)	13 (5%)	20 (8%)	243	Moyenne : 17,4 %
4. Stade avancé inférieur	Eva 4	32	10	95	0	6	143	23 %
	Christina 1	57	14	81	0	10	162	28 %
	Christina 4	28	21	83	0	1	133	24 %
	Pernilla 4	21	11	80	0	3	115	22 %
	Yvonne 4	40	19	194	7	1	261	28 %
Total		178 (22%)	75 (9%)	533 (65%)	7 (1%)	21 (3%)	814	Moyenne: 25 %
5. Stade avancé moyen	Lena 4	51	26	243	9	8	337	26 %
	Marie 4	45	21	120	0	2	188	26 %
	Anders 4	40	46	110	0	0	196	24 %
	Malena 4	38	16	124	0	0	178	25 %
	Anita 4	48	17	106	0	0	171	28 %
Total		222 (21%)	126 (12%)	703 (66%)	9 (1%)	10 (1%)	1070	Moyenne : 26 %
Locuteurs natifs	Henry	112	27	134	-	4	277	33 %
	Isabelle	102	26	171	-	1	300	30%
	Jérôme	53	50	143	-	3	249	27 %
	Mélanie	146	56	223	-	3	428	29 %
	Olivier	84	17	118	-	0	219	27 %
Total		497 (34%)	176 (12%)	789 (54%)	0	11 (1%)	1473	Moyenne : 29,2%

6.1 Résultats au niveau des catégories

SP Lexicales

Aux trois premiers stades, il n'y a pas d'évolution claire, mais nous constatons ensuite une augmentation importante du taux SP Lexicales aux stades 4 et 5, qui entre eux ne présentent pas de différence importante. Par contre, on observe une différence très nette entre les stades avancés et les locuteurs natifs. Nous tirons donc la conclusion que c'est la catégorie qui différencie le plus la production des natifs de celle des apprenants.

SP Grammaticales

Cette catégorie est la seule à ne pas vraiment montrer d'évolution nette. Les pourcentages se ressemblent tout au long de l'acquisition, jusqu'à la production des LN. Une explication possible est que cette catégorie comporte des expressions des classes fermées et contient un

nombre limité de séquences. Cela fait qu'elle n'augmente pas considérablement au cours de l'acquisition.

SP Discursives

La catégorie en question présente une évolution très frappante entre le stade initial et le stade post-initial. Cela tient surtout au fait que *c'est* apparaît abondamment au deuxième stade, ce qui a aussi été montré par Bartning (2005). Par la suite, on ne remarque pas de différences importantes sauf entre les apprenants universitaires et les locuteurs natifs. Il se trouve que cette catégorie est moins importante proportionnellement chez les LN que chez les apprenants. Les apprenants semblent suremployer les structures de cette catégorie. Cela s'explique par le fait qu'ils les utilisent à un plus grand degré que les LN, comme une stratégie de production (cf. Hancock 2000)

SPILI

Les pourcentages de cette catégorie connaissent une évolution visible. Au stade initial, le pourcentage s'élève à 21 %, pour ensuite diminuer de façon décisive. Aux stades avancés, cette catégorie n'obtient que des taux infimes. Cette catégorie est par conséquent caractéristique des productions initiales.

SP Autobiographiques

Les SP Autobiographiques montrent une évolution à travers les stades qui ressemble à celle des SPILI : un pourcentage assez important au stade initial (27%) suivi par une diminution, menant à des taux presque inexistantes aux stades avancés et chez les locuteurs natifs. Tout comme les SPILI, cette catégorie pourrait donc être considérée comme caractéristique des stades initiaux.

6.2 Résultats au niveau de la quantité de SP dans la production

Comme nous le postulons dans l'une de nos hypothèses, il est visible que le taux de langage préfabriqué augmente au fil des différents stades : plus l'apprenant avance dans son acquisition, plus il/elle emploiera de langage préfabriqué. La plus grande différence se trouve entre le stade intermédiaire et le stade avancé inférieur, où le pourcentage augmente d'environ sept points de pourcentage. Cela s'expliquerait probablement par le fait que les apprenants au stade avancé inférieur ont tous été exposés à de l'input en milieu naturel, ce qui encouragerait

l'emploi du langage préfabriqué. Or, nous pouvons constater qu'il existe des différences individuelles considérables dans ces pourcentages. Considérons par exemple le stade post-initial, où les pourcentages de SP dans la production orale varient de 7 % à 21 %. Nous effleurons ici la problématique des différences individuelles qui, bien qu'existant dans n'importe quel aspect langagier, semblent plus fortes encore dans le domaine du langage préfabriqué (voir Forsberg 2005, en cours). Pour le montrer, nous pouvons nous référer par exemple à la catégorie de SPILI, dont Heidi produit la majorité des occurrences comptabilisées, alors que les autres apprenants n'en produisent guère. Nous pouvons également considérer la production de l'apprenant débutant Emelie, dont la première interview se trouve déjà au deuxième stade et qu'elle connaît déjà un certain nombre de SP Discursives, notamment « c'est », alors que Heidi ne produit quasiment aucune SP Discursive.

7. Conclusions

Dans cette étude pilote, nous avons tenté de mettre en rapport deux domaines d'étude acquisitionnels, à savoir les stades de développement en morphosyntaxe d'une part, et les séquences préfabriquées d'autre part. Se peut-il que les séquences préfabriquées, tout comme la morphologie verbale, entre autres, suivent un itinéraire acquisitionnel généralisable ?

Par le biais des hypothèses avancées dans 5.4, et leur confirmation, nous avons pu constater que l'usage des séquences préfabriquées évolue dans l'acquisition du français L2. La quantité de SP augmente au cours de l'acquisition et la distribution en catégories change. Or, ce développement correspond-il à des stades bien séparés ou s'agit-il plutôt d'une évolution sous forme de continuum, pour reprendre la distinction faite par Sharwood-Smith & Traugott (2005) ? Pour ce qui est des SP, on peut distinguer des phénomènes initiaux et des phénomènes tardifs, mais étant donné que les différences individuelles sont importantes, il serait plus prudent de parler de continuum au lieu de stades, dans lesquels un phénomène donné précède obligatoirement l'émergence d'un autre. Par exemple, la morphologie verbale montre un développement stricte, alors que le développement de l'usage des SP semble plutôt régi par l'input et le style communicatif de l'apprenant (voir Forsberg, en cours)

Références bibliographiques

- Bartning, I. (2002) : La variété avancée et sa place dans les phases acquisitionnelles d'une langue étrangère, in : Bartning, I., J. Falk, L. Fant, M. Forsgren, R-M. Jacobsson & J. Nystedt. (éds) : *Mélanges publiés en hommage à Gunnel Engwall*. Acta Universitatis Stockholmiensis. Romanica Stockholmiensia 20, Almqvist & Wiksell International, Stockholm, pp. 1–12.
- Bartning, I. (2005) : Une formule bien utile. Le cas de *c'est* en français parlé L1 et L2, in : Riegel, M., C. Schneidecker, P. Swiggers & I. Tamba : *Aux carrefours du sens. Hommages offerts à Georges Kleiber pour son 60^e anniversaire*. Peeters, Louvain
- Bartning, I. & S. Schlyter (2004a) : Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies* 14, pp. 281-199.
- Bartning, I. & S. Schlyter (2004b) : Itinéraires acquisitionnels et stades de développement en français L2 - version étendue. <http://www.fraita.su.se/interfra>.
- Clahsen, H., J.M. Meisel & M. Pienemann (1983) : *Deutsch als Zweitsprache. Der Spracherwerb ausländischer Arbeiter*. Gunter Narr, Tübingen
- Erman, B. & B. Warren (2000) : The idiom principle and the open choice principle. *Text* 20,1, pp. 29-62.
- Forsberg, F. (2005) : « Prêt-à-parler ». *Étude des séquences préfabriquées en français L2 et L1. Aspects comparatifs et développementaux*. Mémoire de phil. lic. Département de français, d'italien et de langues classiques, Université de Stockholm.
- Forsberg, F. (en cours) : *Préfabriquer l'oral. Étude acquisitionnelle et comparative en français parlé L2 et L1*. Thèse de doctorat. Département de français, d'italien et de langues classiques, Université de Stockholm.
- Forsberg, F., V. Hancock & I. Bartning (à paraître) : Le rôle des séquences préfabriquées dans le talk-show. Actes du colloque *Le français parlé des médias*, Stockholm le 8-12 juin 2005.
- Granfeldt, J. (2005) : *Direkt profil* et deux études sur la morphologie verbale et les stades de développement. *Perles* 20, pp. 65-85, Université de Lund. <http://www.rom.lu.se:8080/profil>.
- Granfeldt, J. & S. Schlyter (éds) (2005) : Acquisition et production de la morphologie flexionnelle. *Perles* 20, Université de Lund.
- Hancock, V. (2000) : *Quelques connecteurs et modalisateurs dans le français parlé d'apprenants universitaires*. Thèse de doctorat Département de français et d'italien. Cahiers de la recherche 16, Université de Stockholm.
- Gross, G. (1996) : *Les expressions figées en français* . Ophrys, Paris.

- Klein, W. & C. Perdue (1997) :The Basic variety (or: Couldn't languages be much simpler?). *Second Language Research* 13,4, pp.301-347.
- Myles, F. R. Mitchell & J. Hooper (1998) : Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning* 48,3, pp. 49–80.
- Nattinger, J. & J. DeCarrico (1992) : *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Peters, A.M. (1983) : *The Units of Language Acquisition*. Cambridge University Press, New York
- Pienemann, M. (1998) : *Processability Theory*. John Benjamins, Amsterdam.
- Schmitt, N. & R. Carter (2004) : Formulaic sequences in action, in : Schmitt, N. (éd.) : *Formulaic sequences : Acquisition, processing and use*. John Benjamins, Amsterdam
- Sharwood-Smith, M. & J. Truscott (2005) : Stages or Continua in SLA: A MOGUL solution. *Applied Linguistics* 26,2, pp. 219-240.
- Wray, A. (2002) : *Formulaic Language and the Lexicon*. Cambridge University Press, Cambridge
- Weinert, R (1995) : The role of formulaic language in second language acquisition : a review. *Applied Linguistics* 16,2, pp. 180-205
- Yorio, Carlos (1989) : Idiomaticity as an indicator of second language proficiency, in: Hyltenstam, K. & L.K. Obler (éds) : *Bilingualism across the lifespan*. Cambridge University Press, Cambridge, pp. 55-72
- Ågren, M (2005) : La morphologie du nombre dans le syntagme verbal en français L2 écrit – l'accord de la 3^e personne du pluriel, in : Granfeldt, J. & S. Schlyter (éds) : *Perles (Petites études romanes de Lund)* 20, pp. 131-151.

Appendice

Résultats sous forme d'une échelle implicationnelle qui a été publiée *in extenso* dans Bartning & Schlyter 2004a. Le tableau 5 ci-dessous ne traite que les traits du syntagme verbal.

Tableau 5. Résumé de l'apparition et de l'utilisation des formes dans les deux corpus, basé sur 80 productions orales des deux corpus

Phen/ Enreg	Ca 1 <u>He</u> <u>12</u>	Ro 1	Ja 46 <u>Mt</u> <u>1</u>	Ca 35 <u>Bj</u> <u>1</u>	Ro 23 <u>Pt</u> <u>12</u>	Pe 47 <u>Sr</u> <u>12</u>	Li 13	To 12 <u>Ka</u> <u>1</u>	Ro 45 <u>Sm</u> <u>12</u>	Os 12	Li 12 <u>Mt</u> <u>23</u>	Ro 67 <u>Ka</u> <u>23</u>	Sm 34	Em 47 <u>Sm5</u>	E v 1 2	E 34 <u>Ka</u> <u>45</u>	Yv 14 <u>Ni</u> <u>13</u>	Ch 14 <u>Ni</u> <u>46</u>	Pr 13 <u>Mg</u> <u>14</u>	Le 14 <u>Kn</u> <u>13</u>	Mal 4 <u>Mg</u> <u>59</u>	An 14	Pe V3 14	Ke V3 Int 14	Ida 4 Vi 3	
PC	-	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Sub ordin	-	-	++	++	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Mod +Infin	-	-	-	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Impf	-	-	-	++	++	-	++	++	+	+	++	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Forme finie-c	-	-	-	-	++	++	++	++	++	++	++	++	+	+	+	+	+	+	+	+	+	+	+	+	+	+
nous V-ons	-	-	-	++	+	++	++	++	++	-	+	+	+	+	+	+	++	+	+	+	+	+	+	+	+	+
FutPé	-	-	-	++	++	++	+	+	+	-	+	++	+	+	+	+	+	+	+	+	+	+	+	+	+	+
voudr	-	-	-	-	-	+	-	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+
3ppl ils ont	-	-	-	++	++	++	+	++	++	++	++	++	++	+	+	+	++	+	+	+	+	+	+	+	+	+
Pqpf	-	-	-	-	-	-	-	-	-	-	-	++	-	-	-	-	+	-	-	-	+	+	+	+	+	+
FSmp	-	-	-	-	-	++	-	-	++	++	-	++	-	-	-	-	+	+	+	+	+	+	+	+	+	+
Cond	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	++	+	+	+	+	+	+	+	+	+	+
3 ppl Vlx-nt	-	-	-	-	++	++	++	-	++	-	-	++	-	++	++	++	++	++	++	++	++	++	++	++	++	++
Subj	-	-	-	-	-	++	-	-	++	-	-	++	-	++	++	++	++	-	++	++	++	++	++	++	++	++
Stade	1	1	1	2	2	2	2	2-3	3	3	3	3	3-4	4	4	4	4	5	5	5	5	5	5-6	5-6	5-6	5-6

Légende : Les apprenants soulignés : corpus Lund ; en gras : **non-guidés** ;

Si la valeur d'un trait diffère dans la même case, celle qui se trouve sur la 2^e ligne concerne le corpus Lund.

Abbreviations :

- 'voudrai' = (je) voudrais utilisé comme formule passe-partout (n'est pas compté comme conditionnel) ;

Valeurs : Dans le cas général :

« - » = pas d'occurrences ; « ++ » = la forme apparaît ; « + » = à partir de 75% environ.

Cas spécifiques :

- 'Forme finie c' (=formes finies courtes): « - » 40-65% ; « ++ » 65-90% ; « + » <90% ;