

## *Profils grammaticaux et lexicaux – à la recherche d'un rapport\**

Granfeldt, Jonas, Université de Lund  
Linguistique

**Résumé :** L'article pose la question de savoir s'il existe une corrélation entre le niveau grammatical et le niveau lexical dans la production écrite des apprenants du français. L'hypothèse nulle est que la langue étrangère se développe à tous les niveaux linguistiques à peu près simultanément. Ainsi, un apprenant qui se trouve à un stade grammatical avancé aurait aussi un vocabulaire (un lexique) « riche », « profond » et « varié ». Pour vérifier cette hypothèse générale, nous avons appliqué des mesures lexicales et une mesure grammaticale à 40 textes en français, écrits par deux groupes d'apprenants adultes suédophones à des niveaux linguistiques différents. Chaque apprenant a composé son texte à partir d'un récit en images nommé « Le voyage en Italie » (Ågren, 2005). L'article rend compte des mesures appliquées, explique la préparation des données, la méthode et la démarche d'analyse adoptée. Nous avons réalisé les analyses du vocabulaire avec différents logiciels. Pour l'analyse grammaticale, nous avons utilisé un programme que nous développons nous-même, *Direkt Profil* (Granfeldt, Nugues et al., 2005). Les résultats montrent que si plusieurs mesures sont capables de séparer les deux groupes d'apprenants postulés, il n'existe pas de relation de corrélation simple entre elles. Une mesure de rareté lexicale est la seule qui montre une corrélation importante avec la mesure grammaticale employée.

---

\* Ce travail fut présenté au XVIème *Congrès des Romanistes Scandinaves*, Universités de Copenhague et de Roskilde, 24 - 28 août 2005. Je tiens à remercier Pierre Nugues et Anita Thomas pour leurs commentaires d'une première version de cet article et Malin Ågren de m'avoir autorisé à utiliser son corpus. La responsabilité des lacunes m'incombe évidemment. Les travaux présentés dans cet article bénéficient du soutien du *Conseil de Recherche* (Vetenskapsrådet), contrat numéro (421) – 2004-1674 pour la période 2005-2008.

## 1. Introduction

Dans cet article, nous posons la question de savoir s'il y a une corrélation entre le niveau grammatical et le niveau lexical dans la production écrite des apprenants du français langue étrangère. L'hypothèse nulle doit être que la langue étrangère se développe à tous les niveaux linguistiques à peu près simultanément. Un apprenant qui se trouve à un stade grammatical avancé aurait aussi un vocabulaire (un lexique) « riche » et « profond » et *vice versa*.

Notre objectif trouve sa motivation dans des arguments théoriques et pratiques. Le paragraphe suivant montrera que les théories d'acquisition diffèrent quant à la conceptualisation de la relation lexique-syntaxe dans une perspective de développement. Des recherches empiriques ciblant le développement du lexique et de la syntaxe pourraient contribuer à ces discussions.

Du côté pratique, il y a un besoin de mesures et de critères bien définis et opératoires pour indiquer la compétence de la langue étrangère chez l'apprenant en français. Ici, les intérêts des chercheurs coïncident avec les besoins des enseignants. Nous soulignons que notre intérêt principal est d'identifier ou, si nécessaire, de développer ces mesures sur la base de la production libre pour accéder aux connaissances surtout implicites de la langue étrangère. Il y a maintes raisons de faire une distinction entre connaissance (grammaticale/lexicale) explicite et compétence implicite.<sup>1</sup> En tenant compte des contraintes temporelles, situationnelles et cognitives imposées sur le sujet dans une situation de production (surtout à l'oral mais aussi à l'écrit), nous supposons que c'est d'abord la compétence implicite qui est employée dans la production libre (cependant voir Hedbor, 2005 pour une étude qui suggère qu'il y a des différences individuelles). Pour ce qui est des connaissances grammaticales explicites, il existe déjà un grand nombre de tests (diagnostiques et des tests de langue) qui peuvent servir aux apprenants et aux enseignants (p ex DIALANG pour ne mentionner qu'un des travaux récents<sup>2</sup>).

Toujours du côté pratique, il est bien connu que mesurer ou indiquer le niveau linguistique d'un apprenant d'une langue étrangère n'est pas chose facile. Pour commencer, il semble difficile de trouver une seule mesure d'une compétence linguistique qui fonctionne pour toutes les productions des apprenants, pour tous les types textuels, pour les deux modes de production (écrit / oral), etc. C'est une des raisons pour lesquelles nous avons adopté la démarche de « profils » grammaticaux (et lexicaux) qui permettent d'inclure un grand nombre de mesures pour analyser les productions des apprenants. Dans leur survol de mesures de développement à l'écrit, Wolfe-Quintero et al. (1998) comparent et évaluent une bonne centaine de mesures de fluidité (*fluency*), complexité (*complexity*), et production correcte (*accuracy*) et arrivent à la conclusion que seul un petit nombre d'entre elles répondent aux critères de validité et fiabilité qu'ils proposent. Et même parmi celles identifiées comme les meilleures, les résultats divergent.

Plusieurs mesures de richesse, de diversité et de rareté lexicale ont été proposées et certaines sont adaptées au français, mais les essais de comparaison semblent plus rares, surtout les comparaisons systématiques entre des mesures grammaticales et lexicales. C'est la raison pour laquelle nous effectuons ici une analyse qui inclut trois catégories de mesures – des

---

<sup>1</sup> Pour une introduction succincte à ces deux termes et aux débats qui les entourent, voir l'introduction de Huljstijn (2005) au numéro thématique de *Studies in Second Language Acquisition*.

<sup>2</sup> DIALANG est un système d'évaluation de quatorze langues européennes. Il propose des tests diagnostiques informatisés disponibles à <http://www.dialang.org>.

mesures de quantité, de rareté et de diversité lexicale et de grammaire. Les mesures lexicales que nous employons sont « la diversité lexicale », (Richards & Malvern, 2004) et « le profil lexical » (Laufer & Nation, 1995, adapté au français par Goodfellow, Jones & Lamy, 2002). La mesure morphosyntaxique est le pourcentage de « Formes Verbales Lexicales Conjuguées » dans un contexte pour un verbe conjugué. Cette mesure se développe depuis un certain temps dans différents projets sur le français parlé (résumé dans Bartning & Schlyter, 2004).

## **2. Compétence grammaticale et compétence lexicale dans une perspective de développement**

Dans une perspective de développement, il existe au moins deux positions théoriques concernant la relation entre la compétence grammaticale (désormais CG) et la compétence lexicale (désormais CL). Ces positions, conflictuelles, sont présentées de manière simplifiée dans ce qui suit.

(1) La CG et la CL se développent indépendamment l'une de l'autre avec des liens « faibles »

Cette position pourrait recouvrir certaines approches modulaires à l'acquisition du langage. C'est le cas des premières versions de la grammaire générative, *Principes et Paramètres* et *Gouvernement et Liage* (Chomsky, 1981), où le lexique et la grammaire étaient clairement distincts. L'acquisition de la CG était vue comme un processus indépendant de l'acquisition d'autres domaines linguistiques, par exemple le lexique (cf. la notion de l'*autonomie* de la syntaxe). Dans ce modèle, certains items lexicaux spécifiques servent à « déclencher » la fixation des paramètres, liés à des catégories fonctionnelles, qui, pour la majorité des chercheurs étaient supposées innées ou sujettes à un processus de « maturation ». Dans une perspective de développement, le rapport entre la compétence grammaticale et lexicale dans cette première optique est alors « faible » dans le sens que la CL a besoin de relativement peu de « mots » pour être déclenchée et acquise.

(2) La CG et la CL se développent ensemble – le développement du lexique « véhicule » le développement de la grammaire

Cette position, sous des formes assez diverses, est défendue par des chercheurs de traditions empiriques et cadres théoriques différents (voir Granfeldt, 2003 :40-42 pour une discussion). Cette position est compatible avec le courant de grammaire générative nommé *Minimalisme* (Chomsky, 1995) dans lequel les propriétés de la syntaxe dépendent des propriétés du lexique. Pour Chomsky (1995), les différences entre les langues sont surtout de nature lexicales. En acquisition, cette optique a été appliquée dans différentes *théories constructivistes* dans lesquelles les propriétés des catégories syntaxiques, et dans certains cas, les catégories elles-mêmes, se développent en fonction de l'acquisition des items lexicaux. Le développement de la syntaxe est ainsi véhiculé par l'acquisition du lexique, surtout les mots grammaticaux. Un des modèles actuels propose que l'enfant et l'adulte acquièrent des items lexicaux et les emmagasinent d'abord comme des unités discrètes. Au moment où il y a suffisamment d'items (une masse critique nécessaire), ils sont « inspectés » par des mécanismes d'analogie qui visent à économiser les représentations. L'inspection d'un grand nombre d'éléments peut avoir comme résultat l'émergence d'une règle, capable de remplacer certaines parties des « mots ». Au lieu d'emmagasiner les participes « parlé », « travaillé », « mangé », une règle

d'analogie pourrait remplacer : participe passé = radical + désinence (é).<sup>3</sup>

### 3. Trois groupes de mesures – quantité, diversité et rareté lexicale, grammaire

Le travail le plus important à ce stade est d'identifier les mesures grammaticales et lexicales qui sont fiables et valides pour indiquer le développement de la compétence (implicite) dans ces deux domaines. Il existe un certain nombre de propositions mais, à notre connaissance, il n'a pas eu de tentatives systématiques de corrélation de mesures lexicales et grammaticales. C'est notre question principale pour la présente étude. Il a fallu choisir parmi les mesures possibles. Le premier critère de sélection était l'automatisation des analyses. Pour répondre aux besoins pratiques des enseignants (mais aussi des chercheurs), nous pensons qu'une méthode aussi automatisée que possible est désirable. Le deuxième critère important était que la mesure produisait des hypothèses intuitives pour notre cas. Nous revenons aux indices lexicaux choisis et aux logiciels pour les calculer ci-dessous.

Pour ce qui est des mesures grammaticales en français langue étrangère, il n'existe pas à notre connaissance de mesure morphosyntaxique généralement acceptée, ni un système disponible pour automatiser complètement l'analyse d'un « profil ». Pour cette étude, nous avons eu recours à un logiciel que nous développons au sein de notre projet, nommé *Direkt Profil*. *Direkt Profil* est un analyseur automatique de textes écrits en français comme langue étrangère. Utilisant des techniques du *Traitement Automatique des Langues* (TAL), son but est d'analyser le texte de l'apprenant et de produire un « profil grammatical » de la production (Granfeldt, Nugues et al, 2005). Le système met en œuvre la théorie des itinéraires acquisitionnels en français langue étrangère élaborée par Bartning & Schlyter (2004). Le but ultime de ces recherches et du logiciel est de pouvoir situer une production d'un apprenant sur un « stade de développement ».

#### 3.1 Des mesures de quantité et de longueur moyenne

(a) Le nombre des mots du texte (*mots totaux ou total des mots*). Cette mesure pourrait représenter une première approximation simple pour indiquer le degré d'automatisation de l'apprenant. Étant donné que tous les sujets disposaient d'autant de temps pour écrire leurs textes, les apprenants avec un haut degré d'automatisation écriront plus de mots que les apprenants avec un degré d'automatisation plus basse. Il faut reconnaître que cet indice n'est pas sans problème.<sup>4</sup>

(b) Longueur moyenne des phrases (*LMP*). Cette mesure simple rappelle l'indice longueur moyenne des énoncés (*LME*) utilisée pour donner une première indication du développement de la langue de l'enfant. Plus la *LMP* est élevée, plus les phrases des textes sont longues. On peut faire l'hypothèse simple qu'un apprenant de bas niveau écrira des phrases moins longues

---

<sup>3</sup> Cette présentation est pourtant une simplification insuffisante pour deux raisons. D'abord, nous faisons abstraction d'une discussion si les notions de « radical », « désinence », « règle » etc. sont des concepts psycholinguistiquement valables (la question de la représentation symbolique). Ensuite, nous n'avons pas abordé la discussion de mécanismes de traitement nécessaire pour traiter l'input (oral et écrit) de telle façon que l'identification d'un « radical », une « désinence » etc. soit possible (pour les deux questions, voir Carroll, 2001).

<sup>4</sup> Il est par exemple impossible de tenir compte des mots effacés. Dans des recherches futures, nous nous servirons d'une méthode pour enregistrer les activités du clavier pour combler cette lacune et d'autres.

qu'un apprenant avancé à cause d'une maîtrise et une automatisation de la langue et son système orthographique moins élevée (cf. mesure (a)). Un développement ultérieur ici serait de considérer le T-units (Hunt, 1970). L'unité de base pour calculer le T-unit est la proposition. Le T-unit est défini par Hunt comme [the] "main clause [of a sentence] plus all subordinate clauses and nonclausal structures attached to or embedded in it." (Hunt, 1970: 64-67) . Les recherches de Hunt ont montré que la moyenne du T-unit par phrase est plus fiable que la LMP pour un corpus de textes anglais.

(c) Longueur moyenne des mots (*LMM*). L'inclusion de cette mesure est motivée par plusieurs facteurs. D'abord, il est connu que les mots les plus fréquents sont aussi les mots les plus courts. Cette observation est formulée comme la « loi de Zipf » (Zipf, 1935)<sup>5</sup>. La loi de Zipf est valide pour des mots de 3 caractères et plus longs, mais ici nous avons pris en considération tous les mots. Nous ignorons pour le moment à quel point il y a des vérifications de la loi de Zipf en français. Dans l'hypothèse plausible où l'on acquiert d'abord les mots fréquents, on peut pourtant émettre qu'un apprenant avancé produira plus de mots longs qu'un apprenant débutant. Une LMM élevée serait ainsi l'indication d'un niveau avancé.

### 3.2 Des mesures de diversité et rareté lexicale

(d) La mesure D pour mesurer la *Diversité lexicale* est un développement de l'indice traditionnel *Type-Token Ratio* (TTR). Elle a été développée et testée par un groupe de chercheurs autour de B. Richards et D. Malvern (Malvern, Richards et al., 2004). Le TTR traditionnel est sensible à la longueur des échantillons : pour des raisons logiques, plus un échantillon est long, plus le TTR décroît. Le risque est donc que le TTR est sur-estimé dans de petits échantillons. La mesure D, par contre, est calculée comme la probabilité de variation de l'indice TTR (Durán et al., 2004). Il s'agit d'élaborer un modèle mathématique sur la variation de TTR en fonction du nombre d'occurrences dans l'échantillon. Un logiciel a été conçu pour analyser la valeur D d'un texte ou une transcription. *VocD* choisit de façon aléatoire des sous-échantillons de mots du texte pour ensuite trouver la meilleure correspondance entre le modèle mathématique et les données empiriques. La valeur D représente un paramètre dans l'équation qui correspond au meilleur ajustement. De grandes valeurs de D sont l'indice d'un haut niveau de diversité lexicale et, par hypothèse, un niveau avancé. Pour calculer D d'un échantillon il faut au moins 50 occurrences (mots) analysables.

(e) La mesure de rareté de mots dans l'usage – *Le Profil Lexical*. Cette mesure a été développée pour l'anglais (Laufer et Nation, 1995) et récemment adaptée au français par Goodfellow, Jones & Lamy, 2002 et ensuite par Horst et Cobb (2004). D'après les informations disponibles, la version française n'a pas encore été autant validée que celle anglaise. La mesure ou le « profil » lexical de Laufer et Nation a eu un grand impact sur le domaine de l'acquisition de langue étrangère. Leur analyse classe les familles des mots du texte de l'apprenant en fonction de quatre bandes de fréquence : (1) Les 1 000 mots les plus fréquents, (2) les mots du deuxième millièm, (3) les mots de spécialités, fréquents dans des textes universitaires et (4) les mots les moins fréquents, i.e. ceux qui ne sont pas inclus dans les listes (1) à (3). Laufer et Nation (1995) montrent que cette mesure est capable de discerner entre des apprenants et qu'il y a une corrélation entre la mesure et des résultats d'un « test de vocabulaire » répandu.

---

<sup>5</sup> Dans les mots de Zipf (1935: 38), cité dans Sigurd et al. (2004: 37) : "In the view of the evidence of the stream of speech we may say that the length of a word tends to bear an inverse relationship to its relative frequency. Footnote: Not necessarily proportionate; possibly some non-linear mathematical function"

Par rapport à la mesure de *diversité lexicale* (cf. d *supra*), le *profil lexical* de Laufer et Nation comporte l'avantage qu'il ne se prononce pas seulement sur la diversité lexicale mais aussi sur la rareté des mots utilisés. Laufer et Nation (1995 : 313) argumentent que : « The LFP will discriminate between subjects who use frequent and less frequent vocabulary, not just between those who can or cannot vary their possibly limited vocabulary ».

Pour la présente étude, nous avons utilisé la version du *profil lexical* développé pour le français par Horst & Cobb (2004). L'idée est la même que dans LFP de Laufer et Nation mais les listes sont différentes. D'abord les listes sont basées sur la prose journalistique : les corpus du *Monde* et du *Soir* ont été utilisés pour établir les listes. Ensuite, il n'y a pas de liste de « mots de spécialité » comme dans la version anglaise. La raison pourrait en être que Horst & Cobb (2004) affirment que ce vocabulaire est en plus grande partie incluse dans le deuxième millièmè des mots en français qu'en anglais. Ceci veut dire qu'il n'est pas tout à fait clair qu'une liste séparée pour les mots académiques fréquents en français soit nécessaire. Les listes du logiciel *Vocabprofile* (version française) sont : 1) Les 1 000 mots les plus fréquents, (2) les mots du deuxième millièmè, (3) les mots du troisième millièmè (4) les mots les moins fréquents, i.e. ceux qui ne sont pas inclus dans les listes (1) à (3).<sup>6</sup>

Notre hypothèse pour cet indice est que les apprenants du bas niveau produiront proportionnellement plus de mots fréquents (de la liste (1)). Dans l'étude de Goodfellow, Jones & Lamy (2002), c'était les mots du deuxième millièmè qui étaient ceux qui séparaient le mieux entre les apprenants.

### 3.3 Mesure grammaticale – les Formes Verbales Lexicales Conjuguées (FVLC)

La mesure grammaticale que nous avons utilisée dans cette étude est le taux de *formes verbales lexicales conjuguées* (FVLC) en pourcentage. Elle est calculée ici comme un rapport : le nombre de verbes lexicaux conjugués dans un contexte de verbe conjugué divisé par le nombre total de contextes pour des verbes lexicaux conjugués. La mesure fait partie de l'inventaire de mesures grammaticales de Bartning & Schlyter (2004) mais elle est ici restreinte aux phrases contenant un pronom sujet (pour des raisons pratiques). Il s'agit d'une mesure de production correcte du domaine verbal qui s'est avérée valide pour des données orales et pour différencier entre des apprenants de différents niveaux. Développée pour des apprenants suédophones par Bartning et Schlyter, nous avons montré ailleurs que cette mesure arrivait à différencier les apprenants anglophones avec des années d'études de français différentes (Granfeldt, 2005). Cet indice est intégré dans l'analyse de *Direkt Profil* (version 1.5.2). Ce logiciel est disponible sur <http://www.rom.lu.se:8080/profil>.

## 4. Corpus

La présente étude est exploratoire, ce qui est reflété dans le choix d'un corpus limité. Pour nos analyses, nous avons utilisé un sous-ensemble du corpus CEFLE de Lund (*Corpus Écrit de Français Langue Étrangère*) (voir Ågren, 2005, ce volume). Le CEFLE est constitué de textes de longueur variée, produits par des élèves suédophones de français au lycée de niveaux linguistiques variés et par un groupe de contrôle (des jeunes Français). Au total le corpus contient environ 100 000 mots.

---

<sup>6</sup> Le *Vocabprofile* de Tom Cobb est disponible sur <http://www.lex Tutor.ca/vp>

Pour la présente étude, nous avons choisi d'étudier des textes produits d'après une procédure d'élicitation, nommée *Le voyage en Italie* (cf. Ågren, 2005). Il s'agit d'une série d'images dessinée sans mots qui raconte le voyage en Italie de deux jeunes filles suédoises. La série d'images était conçue pour faire produire des SNs et des SVs au pluriel. Il y a aussi un grand nombre d'objets féminins sur les images pour éliciter des déterminants et des adjectifs au féminin.

Ågren a appliqué un sous-ensemble des critères verbaux de Bartning & Schlyter (2004) (entre autres FVLC) pour classer tous les textes du CEFLE selon le stade de développement. Nous avons choisi des textes des deux niveaux stades 2 et 4, voir tableau 1. Le choix de textes a été réalisé de façon aléatoire.

Tableau 1. *Les données du CEFLE*

stade	Nbr d'apprenants	Nbr d'années d'études de français	Nbr de mots du corpus
2	25	1-3	env. 4 300
4	15	3-4	env. 6 200

## 5. Questions de recherches

Nos questions de recherches dans cette étude sont au nombre de trois :

(a) Quelle *variation intra-groupe* ces mesures produisent-elles ? Nous savons que la variation individuelle est souvent importante dans des données de production en L2. Mais la variation n'est pas forcément aussi importante pour toutes les mesures.

(b) Quelles *différences inter-groupes* ces mesures produisent-elles ? Rappelons que les 40 textes sélectionnés du corpus CEFLE de Lund ont été classés comme étant aux stades 2 (n=25) et 4 (n=15). Pour la classification des critères verbaux de Bartning & Schlyter étaient utilisés (voir *supra*). Est-ce que les autres mesures produisent les mêmes différences entre les stades 2 et 4 ? Cette question porte surtout sur la question plus générale de la validité des mesures lexicales et grammaticales.

(c) Y-a-t-il des *corrélations intra-groupe* pour ces mesures ? Notre question principale est, rappelons-le, de savoir s'il y a une corrélation entre certaines mesures lexicales et la mesure grammaticale utilisée ici.

## 6. Méthode

### 6.1 Préparation des données

La préparation des données avant l'analyse représente une étape importante, susceptible d'influer sur le résultat final. Il s'agit de sélectionner un sous-ensemble des données pour une analyse ultérieure et de les modifier pour optimiser la qualité des résultats. Nous avons préparé les données en deux étapes : (a) sélection d'un sous-ensemble des données et (b) correction orthographique partielle.

(a) Sélection d'un sous-ensemble des données

Nous avons établi des critères d'inclusion et d'exclusion. Pour éviter de *surestimer* la richesse et la diversité lexicale, certains mots ont été exclus des textes. Les catégories des mots exclus sont :

- (1) Les noms propres (ex. *Amanda, Baltazaar, etc.*)
- (2) Les chiffres (ex. *1, 2, 5 etc.*)
- (3) Les mots clairement étrangers (ex. *sightseeing, solnedgång, etc.* )
- (4) Les mots inconnus et éventuellement étrangers (ex. *bold, supé, cèlebite, garde, etc.*)

Les catégories des mots inclus sont :

- (5) Les mots français correctement épelés
- (6) Les mots français incorrectement épelés où la forme correcte est prévisible (ex, *conetre (connaître), pière (pierre) etc.*)

Ces critères ne s'appliquent pas sans difficulté et il existe un certain nombre de cas limites. La plus grande difficulté a été de distinguer entre (4) et (6). Nous avons utilisé un critère phonétique : les mots épelés de telle façon que la prononciation correcte du mot ne changeait pas ont été inclus. Inversement, les mots où l'orthographe proposée changeait clairement la prononciation par rapport à celle du mot ciblé ont été exclus.

(b) correction orthographique partielle

L'orthographe des mots de la catégorie (6) a été corrigée à l'aide de Microsoft Word<sup>TM</sup>. Nous avons considéré cette étape comme nécessaire pour optimiser le résultat des analyses de la mesure « D » et le « Profil Lexical ». Pour celle-ci, il faut rappeler que le *Vocabprofile* de T. Cobb utilise un dictionnaire de formes fléchies. Ce dictionnaire ne reconnaît que des formes correctement orthographiées. Pour éviter de *sous-estimer* la rareté et la diversité lexicale des textes des apprenants, nous avons corrigé l'orthographe des mots de la catégorie (6).

## 6.2 Analyse et évaluation

Une démarche ascendante est sans doute la meilleure pour ce type de recherches qui visent à établir des critères et des mesures de compétence grammaticale et lexicale. C'est aussi la raison pour laquelle nous ne ferons pas de sélection initiale des mesures appliquées ; pour nous ce sont des questions ouvertes et empiriques (a) quelles sont les mesures les plus efficaces et (b) s'il y a des corrélations entre elles.

La méthode consiste à appliquer les indices lexicaux et grammaticaux sur les textes d'apprenants des deux groupes (voir section *Corpus, supra*). Après la préparation des données, nous avons traité les textes avec le logiciel respectif pour obtenir les résultats. Le tableau 2 montre les systèmes respectifs utilisés.

Tableau 2. *Mesures et logiciels utilisés*

	Mesures de quantité et de longueur moyenne			Mesures de diversité et rareté lexicale		Mesure grammaticale
	Mots totaux	LMP	LMM	D	Profil Lexical	FVLC
Logiciel	CLAN : freq	CLAN : wdlen	CLAN : wdlen	CLAN : vocd	Vocabprofile	Direkt Profil

Légende : LMP = Longueur Moyenne des Phrases ; LMM = Longueur Moyenne des Mots ; FVLC = Formes Lexicales Conjuguées

Nous avons utilisé plusieurs logiciels de la boîte à outils de CLAN (MacWhinney, 2000). Plusieurs de ces outils calculent la mesure « total des mots ». Pour analyser la mesure D, il faut adapter le texte à un format CHAT de base (*minCHAT*). Un logiciel de conversion, nommée *Textin*, prend en charge l'insertion des en-têtes et lignes principales nécessaires pour ce format, sans effectuer d'autres changements dans le texte original.

Les résultats ont été transférés d'abord à Excel<sup>TM</sup> et ensuite à SPSS<sup>TM</sup> pour des analyses statistiques. Trois types d'analyses ont été appliqués à chacune des mesures :

(i) des calculs de Moyennes, Médianes, Écarts Types et valeurs extrêmes pour répondre à la question de recherche (a) sur la *variation intra-groupe* ;

(ii) Des analyses de variances (ANOVA) où le facteur indépendant était le stade de développement préalablement établi et le facteur dépendant était la mesure en question. Ces analyses montreront si ces mesures arrivent à distinguer de façon significative les productions des deux groupes d'apprenants (question de recherches (b));

(iii) Des analyses de corrélation (PEARSON) montreront à quel point il y a des corrélations entre les différentes mesures. Ce qui nous intéresse particulièrement ce sont des corrélations entre les mesures lexicales et grammaticales (question de recherches (c)). Une question annexe est de savoir à quel point il y a une corrélation entre les différentes mesures à l'intérieur du groupe lexical.

## 7. Résultats

Les résultats sont encore préliminaires mais ils ne suggèrent pas un rapport simple entre les niveaux grammatical et lexical.

Le tableau suivant montre les moyennes et les écarts types des trois groupes de mesures :

Tableau 3. *Groupes de mesures, Moyennes et Écart Type*

Mesures de longueur moyenne et quantité							Mesures de diversité et rareté lexicale							Mesure Grammaticale.		
S	Mots totaux		LMM		LMP		Valeur D		%K1-fam		%K2-fam		%K3-fam		%FVLC	
	M	ÉT	M	ÉT	M	ÉT	M	ÉT	M	ÉT	M	ÉT	M	ÉT	M	ÉT
2	194	69	4,23	0,17	9,05	2,47	68,4	21,4	89,0	2,2	6,0	3,0	4,9	2,7	90,2	15,0
4	406	213	4,48	0,14	7,79	0,71	89,8	17,2	86,9	5,0	8,5	3,2	4,6	1,9	99,6	1,3

Légende : LMP = Longueur Moyenne des Phrases ; LMM = Longueur Moyenne des Mots ; FVLC = Formes Lexicales Conjuguées ; %K1/2/3-fam = le pourcentage des familles des mots appartenant au 1<sup>er</sup>, 2<sup>e</sup> et 3<sup>e</sup> millièmes des mots fréquents ; M = Moyenne ; ÉT = Écart Type

Notre premier groupe de résultats suit, *grosso modo*, les hypothèses que nous avons faites pour chacune des mesures appliquées (voir section 3 *supra*). Seuls deux résultats vont à l'encontre des ces prédictions sans pour autant être significatifs: (a) La LMP est plus élevée au stade 2 (9,05) qu'au stade 4 (7,79), ( $p = 0,064$ ) contrairement à notre hypothèse sur l'automatisation ; (b) la moyenne des familles du 3<sup>e</sup> millième de fréquence (%K3-fam) au stade 2 est de 4,9% et de 4,6% pour stade 4. Ceci pourrait suggérer que les apprenants du stade 2 utilisent plus de mots assez rares que les apprenants du stade 4. Mais la différence n'est pas significative ( $p=0,622$ , voir tableau 4, *infra*).

Notre deuxième groupe de résultats répond à la question de recherche (a) sur la variation intra-groupe pour ces mesures. Des analyses de la médiane et des valeurs extrêmes, montrent qu'en général, il y a plus de variation intra-groupe chez les apprenants du stade 2 (cf. aussi les valeurs des Écarts Types du tableau 3 *infra*). La valeur D des apprenants du stade 2 varie d'environ 35 à 100 (plusieurs apprenants) alors que la variation du stade 4 va de 65 à 110 (avec un apprenant exceptionnel, no 14, à 125), cf. figure 1. La mesure %K1-familles (figure 2) suit même modèle: Le stade 2 varie de 82% à 96%. Au stade 4, la variation s'étend de 80% à 90% si certaines valeurs exceptionnelles sont mises à part. Notre impression est que la mesure de %K1-familles donne lieu à un peu moins de variation intra-groupe que la mesure D.

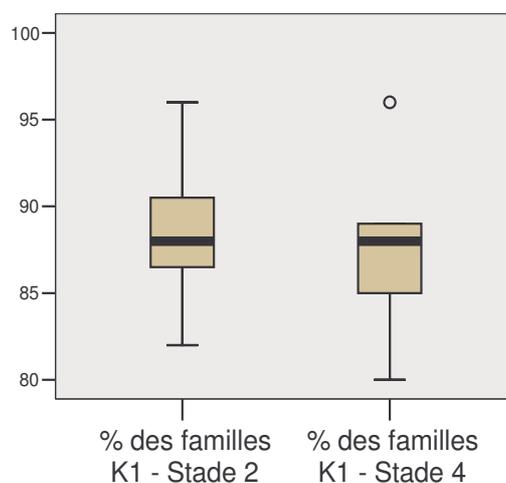
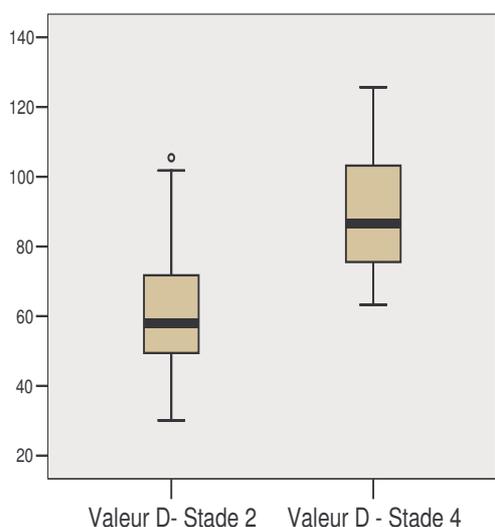


Figure 1. La Valeur D, Médiane et valeurs extrêmes

Figure 2. Profil lexical – le pourcentage de familles K1 – Médiane et valeurs extrêmes

La figure 3 montre que la mesure %FVLC produit un effet plafond au stade 4 : pour cette mesure il n'y a pratiquement pas de variation intra-groupe au stade 4 ce qui suggère que cette mesure est moins utile pour les stades avancés. Nous revenons à cette observation dans le dernier paragraphe.

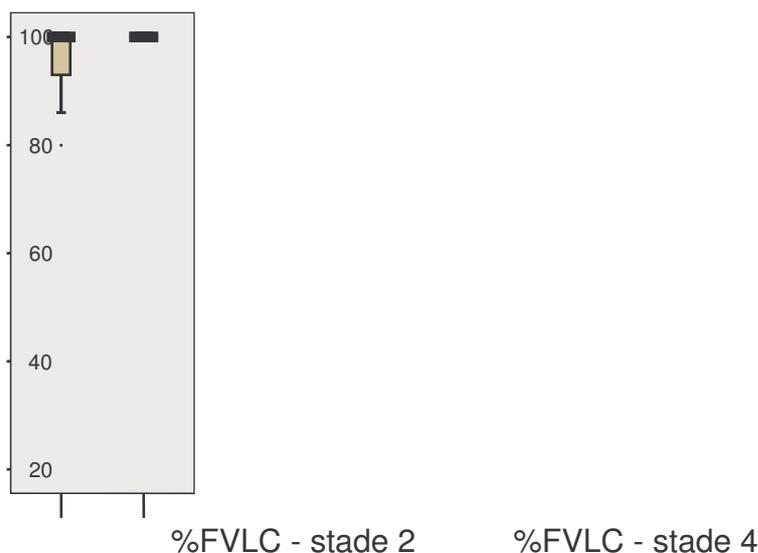


Figure 3. Le %FVLC, Médiane et valeurs extrêmes

Notre troisième résultat montre que cinq des huit mesures produisent une différence significative entre les deux groupes (question de recherches (b)). D'après nos analyses, les mesures de « Mots totaux » et « LMM » produisent des différences très importantes entre les stades 2 et 4 ( $p < 0,001$ ). Les mesures de Valeur D, %K2-fam, et %FVLC sont significatives au niveau inférieur ( $p < 0,05$ ).<sup>7</sup> Nous pouvons interpréter ces résultats comme une confirmation, au niveau du groupe, de la classification initiale en deux stades de développement (2 et 4) : Le niveau linguistique lexical et grammatical des deux groupes est effectivement différent. Des analyses ultérieures, au niveau d'apprenants individuels, sont pourtant nécessaires pour décider les limites de la variation à l'intérieur d'un stade donné.

Tableau 4. *Analyses de variance*

	Longueur moyenne et quantité			Diversité et rareté lexicale			Gramm.	
	Mots totaux	LMM	LMP	Valeur D	%K1-fam	%K2-fam	%K3-fam	%FVLC
F	21,217	23,410	3,642	10,814	2,906	6,131	0,247	5,783
P	<0,001**	<0,001**	0,064	0,002*	0,096	0,018*	0,622	0,021*

Légende : \* = différence significative à  $p < 0,05$  ; \*\* = différence significative à  $p < 0,001$

Pour ce qui est de notre question principale de cette étude qui était, rappelons-le, d'analyser une corrélation éventuelle entre les niveaux lexical et grammatical dans la production des apprenants, les résultats ne s'interprètent pas facilement. Les résultats sont présentés dans les tableaux 5 et 6. Nous limitons la discussion dans cette section aux résultats significatifs. Nous discutons de façon plus générale les différentes mesures appliquées dans la section suivante.

<sup>7</sup> Il n'est pas surprenant que la mesure %FVLC produise cette différence car elle faisait partie des critères de sélection.

Dans le groupe des apprenants du stade 2, quatre mesures montrent des corrélations significatives (cf. tableau 5) : (i) %FVLC et %K1-fam, (ii) le total des mots et la valeur D, (iii) %K1-fam et %K2-fam, et (iv) %K1-fam et %K3-fam. Les deux dernières sont des effets de l'analyse effectuée et seront mises à part. Les mesures de « Mots totaux » et « Valeur D » montrent une corrélation significative faible ( $p=0,026$ ) mais ceci pourrait être dû au fait que les deux mesures se ressemblent.

Le résultat le plus intéressant pour nous est que le taux de FVLC montre une petite corrélation négative avec le taux de %K1-fam, une mesure de rareté lexicale. Ceci veut dire que plus le taux de FVLC est élevé, moins l'apprenant utilise des mots de très haute fréquence (parmi les mille mots les plus fréquents). Ce résultat est d'autant plus intéressant que la mesure de diversité lexicale (D) ne montre pas de corrélation significative. Une possibilité est qu'un vocabulaire varié dans un texte à l'écrit n'est pas corrélé avec le développement grammatical. Une autre possibilité, plus probable, est que le type de tâche utilisée ici n'offre pas suffisamment d'espace à un choix de vocabulaire personnel et ainsi à une variation lexicale individualisée. Pour ce type de mesure, un autre type textuel et une tâche linguistique plus « ouverte » seraient à préférer. Nous y revenons dans le dernier paragraphe.

Si ce résultat qui suggère une corrélation entre la rareté lexicale et un niveau grammatical intermédiaire est valide, il appuie les théories de développement car il suggère un lien entre la fréquence des mots dans l'usage et le développement de la grammaire. Ce lien, qui n'est que vaguement suggéré dans le cadre de cette étude, est discuté dans plusieurs des articles réunis par Granfeldt & Schlyter (2005).

Tableau 5. Niveaux de corrélation entre les mesures - stade 2

	Longueur moyenne et quantité			Diversité et rareté lexicale			Gramm.	
	Mots totaux	LMM	LMP	Valeur D	%K1-fam	%K2-fam	%K3-fam	%FVLC
Mots	1	,019	,267	,445*	-,229	,341	-,087	,187
LMM	,019	1	,060	,330	,042	-,206	,173	-,054
LMP	,267	,060	1	,051	-,222	,186	,043	,346
D	,445	,330	,051	1	,378	-,133	-,308	-,304
%K1-fam	-,229	,042	-,222	,378	1	-,652**	-,589**	<b>-,443*</b>
%K2-fam	,341	-,206	,186	-,133	-	1	-,217	,269
%K3-fam	-,087	,173	0,43	-,308	-	-,217	1	,272
%FVLC	,187	-,054	,346	,304	<b>-,443*</b>	,269	,272	1

Légende : \* = corrélation significative à  $p<0,05$  ; \*\* = corrélation significative à  $p<0,001$

Pour ce qui est des corrélations au stade 4, nous n'en avons pas trouvé, à part celles qui sont les résultats des méthodes de calculs effectués. Pour nous, le résultat le plus important est que la corrélation entre %FVLC et %K1-fam, significative au stade 2, n'est pas significative au stade 4. Ceci pourrait en grande partie être dû à l'effet plafond de la mesure %FVLC. Alors qu'il y a une variation intragroupe pour la mesure %K1 au stade 4 (cf. figure 2, *supra*), la valeur de %FVLC est 100% pour 14 des 15 apprenants du stade 4 (cf. figure 3 et tableau 3, *supra*).

Tableau 6. Niveaux de corrélation entre les mesures - stade 4

	Longueur moyenne et quantité			Diversité et rareté lexicale				Gramm.
	Mots totaux	LMM	LMP	D	%K1-fam	%K2-fam	%K3-fam	%FVLC
Mots	1	-,125	,575*	,339	-,320	,290	,116	,134
LMM	-,125	1	-,352	-	-,201	-,033	,418	,318
LMP	,575*	-,352	1	,264	,062	-,018	-,188	,095
D	,339	-,359	,264	1	-,086	,082	-,106	-,346
%K1-fam	-,320	-,201	,062	-	1	-,865**	-,614*	-,138
%K2-fam	,290	-,033	-,018	,082	-,865**	1	,160	-,035
%K3-fam	,116	,418	-,188	-	-,614*	,160	1	,349
%FVLC	,134	,318	,095	-	-,138	-,035	,349	1

Légende : \* = corrélation significative à  $p < 0,05$  ; \*\* = corrélation significative à  $p < 0,001$

## 8. Résumé des résultats et discussion des sources d'erreurs potentielles

Cette étude a recherché des rapports éventuels entre un certain nombre de mesures lexicales et une mesure grammaticale dans la production écrite de deux groupes d'apprenants de français langue étrangère. Les apprenants avaient été préalablement regroupés dans deux stades de développement grammatical différent, les stades 2 et 4, selon des critères verbaux de Bartning & Schlyter (2004).

Les résultats montrent que certaines mesures produisent plus de variation intra-groupe que d'autres. La « diversité lexicale » est un exemple de mesure qui donne lieu à une variation intra-groupe importante alors que la rareté lexicale, « profil lexical », (p. ex %K1-familles) semble varier moins dans les deux groupes. De façon générale, il y a plus de variation intra-groupe au stade 2 qu'au stade 4. Cinq des huit mesures appliquées produisent une différence significative entre les deux groupes. Nous interprétons ce résultat comme une confirmation du regroupement initial des apprenants en deux stades de développement distincts (Ågren, 2005). La question principale pour notre étude était de savoir s'il y avait une corrélation entre les mesures lexicales et la mesure grammaticale. La réponse n'est pas tout à fait claire. Une seule corrélation significative a été trouvée : au stade 2 une mesure de rareté lexicale (%K1-fam) montre une corrélation négative avec le taux de formes verbales conjuguées (%FVLC). Dans ce qui suit nous essayerons d'identifier des sources d'erreurs potentielles,

Les facteurs les plus important à discuter, en vu d'une étude future, sont de deux catégories :

- (a) Des facteurs reliés à la méthode appliquée ;
- (b) Des facteurs reliés au type de mesure

Dans la première catégorie, on peut discuter des processus de sélection et de préparation des données (cf. section 6.1 *supra*). Nous avons choisi d'exclure certains mots, par exemple des noms propres, des emprunts etc. Plus important nous avons aussi corrigé un certain nombre de

mots mal orthographiés pour qu'ils soient détectés par les logiciels. Ainsi, nous avons neutralisé une différence possible entre les apprenants, le niveau de maîtrise de l'orthographe. Mais il n'est pas probable que ces interventions influent à un haut degré sur les résultats. Des essais préliminaires sur la mesure D montre qu'elle change très peu (quelques dixièmes) en fonction de ces paramètres. En d'autres termes, ces paramètres ne sont pas suffisamment importants pour changer les résultats.

Par contre, le processus de la collecte des données et le type de tâche sont deux facteurs reliés à la méthode susceptibles d'avoir un impact considérable sur les résultats. Pour commencer, la tâche utilisée ici (raconter à l'écrit d'après une série d'images) ne semble pas idéale pour obtenir des variations lexicales. Une tâche « plus libre » dans le sens qu'elle permet à l'apprenant de choisir son propre vocabulaire pourrait donner de meilleurs résultats. On peut rappeler le fait que Laufer et Nation ont élaboré le profil lexical original avec le « composition » comme type textuel. En revanche, il n'est pas certain qu'une tâche plus libre serait possible avec les apprenants de bas niveau linguistique qui souvent ont besoin d'un cadre bien précis pour produire (M. Ågren, communication personnelle).

Pour ce qui est du processus de collecte des données, il serait souhaitable d'avoir plus d'informations sur la production des apprenants. Nous ignorons encore à quel point la mode de production influe sur la mise en œuvre des compétences linguistiques. Notre impression est que la mise en œuvre de la compétence grammaticale, p ex le taux de formes verbales conjuguée, diffère à l'oral et l'écrit chez les apprenants. Mais il n'est pas clair s'il s'agit d'une différence catégorielle où des différences selon un continuum. Une méthode de collecte de données qui enregistre la production des apprenants en ligne peut nous informer sur les stratégies des apprenants (voir Hedbor, 2005 pour une étude qui utilise *ScriptLog*).

Avant d'avoir optimisé la méthode, il est difficile de savoir si nous avons choisi les mesures correctes pour indiquer les niveaux grammatical et lexical des apprenants (catégories (b) ci-dessus). Nous avons déjà pu voir, cependant, que la mesure de formes lexicales conjuguées montre un effet plafond au stade 4 à l'écrit. Ce résultat suggère que cette mesure est moins utile à ce niveau de développement. Nous avons aussi soulevé le fait que la mesure de longueur moyenne de phrase s'interprète difficilement. Il suffit de considérer les trois phrases suivantes pour comprendre qu'il n'y pas de relation simple entre la longueur des phrases et le niveau linguistique :

- (1) Il y a deux femmes et elles sont en train de partir en Italie avec leur voiture. [= 17 mots]
- (2) Il y a deux femmes qui partent en Italie avec leur voiture. [= 12 mots]
- (3) Il y a deux femmes partant en Italie en voiture. [= 10 mots]

La première phrase en est la plus longue mais aussi la plus simple syntaxiquement (coordination, parataxe). La troisième phrase est la plus courte, mais on peut affirmer qu'elle a la syntaxe la plus complexe (hypotaxe, proposition réduite). La réduction ou le compactage des informations est un processus de haut niveau linguistique. Ces exemples montrent de manière simplifiée l'importance de la prise en considération des types de structures grammaticales utilisés par les apprenants lorsqu'il s'agit de définir le niveau linguistique de leurs textes. C'est l'une des pistes de recherches que nous suivrons dans notre projet.

## 9. Références

- Bartning, I. et Schlyter, S. (2004) : Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, 14,3, pp. 281-299.
- Carroll, S. (2001): *Input and Evidence*. John Benjamins, Amsterdam/Philadelphie.
- Chomsky, N. (1981) : *Lectures on government and binding*. Foris, Dordrecht.
- (1995) : *The minimalist program*. MIT Press, Cambridge (Mass.).
- Cobb, T. et Horst, M. (2004). Is there room for an AWL in French?, in : Laufer, B. et Bogaards, P.(éds.) : *Vocabulary in a Second Language: Selection, acquisition, and testing*, pp. 15-38. John Benjamins, Amsterdam/Philadelphie.
- Cobb, T.: *The compleat lexical tutor for data-driven language learning on the web*. [En ligne: disponible à <http://www.lextutor.ca>].
- Goodfellow, R., Jones, G. and Lamy, M.-N. (2002) : Assessing learners' writing using Lexical Frequency Profile. *ReCALL* 14:, pp.129-142.
- Granfeldt, J. (2003) : *L'acquisition des catégories fonctionnelles, Étude comparative du développement du DP français chez des enfants et des apprenants adultes*. Thèse de doctorat, Études Romanes de Lund 67, Institut d'études romanes, Université de Lund, Suède.
- Granfeldt, J. Nugues, P et al. (2005) : Direkt profil : un système d'évaluation de textes d'élèves de français langue étrangère fondé sur les itinéraires d'acquisition. *Actes de la conférence Traitement Automatique des Langues Naturelles, TALN & RECITAL 2005*, Tome 1 – Conférences principales, pp. 113–122, Dourdan, France, 6-10 juin 2005.
- Granfeldt, J. (2005) : Direkt Profil et deux études sur la morphologie verbale et les stades de développement, in : Granfeldt, J. et Schlyter S. (éds.) (2005), pp. 65-85.
- Granfeldt, J. & Schlyter, S. (éds.) (2005) : *Acquisition et production de la morphologie flexionnelle*. *Actes du « Festival de la morphologie »*. Petites Études Romanes de Lund Extra Seriem (PERLES), n° 20. Institut d'études romanes de Lund. Université de Lund.
- Hedbor, C. (2005) Le traitement implicite et explicite de l'inflexion de la 3<sup>e</sup> personne du pluriel du verbe chez les apprenants guidés du français LE, in : Granfeldt, J. et Schlyter, S (éds.) (2005), pp. 151-166.
- Hulstijn, J. (2005) Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition* 27, 2, pp. 129-140.
- Hunt, K.W. (1970) *Syntactic maturity in school children and adults*. *Monographs of the Society For Research in Child Development*, 53 (134). University of Chicago Press, Chicago.
- Laufer, B. et Nation, P. (1995) : Vocabulary Size and Use : Lexical Richness in L2 Written Production. *Applied Linguistics*, 16, 3, pp. 307-322.
- MacWhinney, B. (2000) *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum Associates , Mahwah, NJ.
- Sigurd, B., Eeg-Olofsson, M. et van de Weijer, J. (2004) : Word length, sentence length and frequency – Zipf revisited. *Studia Linguistica* 58, 1, pp. 37-52.
- Zipf, G.K. (1935) (Reprinted 1965) : *The psycho-biology of language*. MIT:Press, Cambridge MA.
- Wolfe-Quintero, K., Shunji, I et Hae-Young, Kim (1998) : *Second language development in writing: measures of fluency, accuracy and complexity*. Technical Report 17. Second Language Teaching & Curriculum Center. University of Hawai'i.
- Ågren, M. (ce volume)

Ågren, M. (2005) : *Développement de la morphologie du nombre en français langue étrangère à l'écrit. Etude transversale*. Mémoire de licentiatexamen. Petites Études Romanes de Lund Extra Seriem (PERLES), n° 21. Institut d'études romanes de Lund. Université de Lund.