



LUND UNIVERSITY

Acquisition et production de la morphologie flexionnelle

Actes du Festival de la morphologie

Granfeldt, Jonas; Schlyter, Suzanne

2005

Document Version:

Förlagets slutgiltiga version

[Link to publication](#)

Citation for published version (APA):

Granfeldt, J., & Schlyter, S. (Eds.) (2005). *Acquisition et production de la morphologie flexionnelle: Actes du Festival de la morphologie*. (PERLES : petites études romanes de Lund. Extra seriem; Vol. 20). Institut d'études romanes de Lund ; Lunds universitet : Romanska institutionen.

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Jonas Granfeldt et Suzanne Schlyter
(éds.)

*Acquisition et production de la morphologie
flexionnelle*

Actes du « Festival de la morphologie », mars 2005 à Lund

Petites études de l'Institut d'études romanes de Lund,
Extra Seriem, n° 20

©2005 Jonas Granfeldt et Suzanne Schlyter

La loi suédoise interdit les copies ou reproductions destinées à une utilisation collective. Toute représentation ou reproduction intégrale ou partielle faite par quelque procédé que ce soit, sans le consentement de l'auteur ou des ses ayants, est illicite et constitue une contrefaçon sanctionnée.

All rights reserved. No part of this text may be reproduced in any form or by any electronic or mechanical means including information storage and retrieval systems without permission in writing from the copyright owner, except by a reviewer who may quote brief passages in a review.

Imprimé en Suède.

ISSN : 1400-1810

Préface

Ce volume est le fruit de discussions à l'occasion de la visite de deux chercheurs français à l'Institut d'études romanes de Lund au printemps 2005 : Cyrille Granget (Université de Paris III) et Pierre Largy (Université de Toulouse le Mirail). Ils ont discuté leurs travaux sur la morphologie avec les doctorantes, les chercheuses et chercheurs du Centre de Langues et Littérature (SOL) de Lund. L'ambiance créative et amicale, ainsi que les festivités qui ont eu lieu à l'occasion de ces visites, nous ont spontanément incités à parler d'un « Festival de morphologie ». Les discussions à partir de nos données respectives nous ont permis de constater que même si nous discutons des travaux en cours – encore qu'imparfaitement au point – la comparaison de nos données s'avérait déjà intéressante et méritait d'être réunie dans notre série de prépublications. Ce volume est également disponible sur <http://www.rom.lu.se/durs/morph.pdf>



Festival de morphologie : de gauche à droite : Cyrille Granget, Angela Byrskog, Suzanne Schlyter, Malin Ågren, Anita Thomas et Cecilia Hedbor



Festival de morphologie, bis ; de gauche à droite et de haut en bas: Pierre largy, Suzanne Schlyter, Malin Ågren, Anita Thomas, Lieve van Espen ; Cecilia Hedbor, Nina Bengtsson, Jonas Granfeldt

Ce volume est aussi le résultat d'un effort commun. C'est pourquoi nous nous remercions mutuellement !

Kerstin Schlyter, la mère de Suzanne Schlyter, a eu l'extrême gentillesse de mettre à notre disposition son appartement, non seulement pour loger les chercheurs invités mais aussi pour y célébrer des fêtes très gaies. Nous lui dédions ce volume.

Table des matières

Jonas Granfeldt & Suzanne Schlyter

Introduction 7

Cyrille Granget

La désambiguïsation des formes verbales en /e/ dans des récits oraux d'apprenants germanophones du français 17

Anita Thomas

Formes ambiguës en –E : input et catégories sémantiques 35

Suzanne Schlyter & Inge Bartning

L'accord sujet-verbe en français L2 parlé 53

Jonas Granfeldt

Direkt Profil et deux études sur la morphologie verbale et les stades de développement 65

Pierre Largy

Approche psycholinguistique de la mise en œuvre de la morphologie flexionnelle du nombre verbal en français langue maternelle 89

Angela Byrskog

Beeinflusst die Muttersprache den Erwerb der Subjekt-Verb-Kongruenz im Plural bei Schülern, die Deutsch als L2 studieren? 97

Cyrille Granget

Développement de l'accord verbal avec un sujet pluriel dans les récits écrits d'apprenants germanophones scolarisés du français 111

Suzanne Schlyter

Formes verbales au pluriel à l'oral plus marquées chez les apprenants guidés que non-guidés ? 125

Malin Ågren

La morphologie du nombre dans le syntagme verbal en français L2 écrit - L'accord de la 3^{ème} personne du pluriel 131

Cecilia Hedbor

Le traitement implicite et explicite de l'inflexion de la 3^e personne du pluriel du verbe chez les apprenants guidés du français LE 151

Introduction

Jonas Granfeldt & Suzanne Schlyter

Université de Lund
Institut d'études romanes
{jonas.granfeldt, suzanne.schlyter@rom.lu.se}

Ce volume de PERLES est consacré à une comparaison de l'acquisition et de la production de la morphologie flexionnelle à travers de différents corpus. Si la morphologie appartient au cœur des recherches linguistiques, elle a toujours dû négocier sa place entre le lexique, la syntaxe et la phonologie. Il est symptomatique que la réorientation la plus importante dans les théories de grammaire formelle cette dernière décennie a touché la place de la morphologie et son rapport avec les autres modules. Dans la tradition dite « Gouvernement et liage » (Chomsky, 1981), la morphologie est absorbée dans la syntaxe. Les « mots » entrent dans la syntaxe sans flexion et les formes fléchies sont ainsi le résultat de la dérivation (*l'approche de l'insertion tardive*). A l'opposé, le *Minimalisme* (Chomsky, 1995) fait l'hypothèse que les mots entrent déjà fléchis dans la syntaxe (*l'approche de l'insertion précoce*). Dans cette optique la syntaxe est déterminée par la morphologie et la computation syntaxique (la dérivation) est véhiculée par les informations grammaticales (les traits) réalisées (ou non) dans les morphèmes flexionnels ou dérivationnels. Pour d'autres encore, la morphologie constitue un module à part entière. C'est le cas dans la théorie de *Morphologie Distribuée* (Halle et Marantz, 1994) où une composante post-syntaxique prend en charge l'épellation de la morphologie sur la base des matrices de traits comme personne, genre, temps etc. (*insertion tardive*). Ces positions opposées traduisent un traitement vacillant de la morphologie en tant que domaine linguistique.

L'autonomie relative de la morphologie et son rapport avec la syntaxe sont aussi des questions pour les recherches en acquisition de langues. S'il est généralement admis que chez le petit enfant la syntaxe et la morphologie flexionnelle s'acquièrent, en principe, ensemble (Meisel, 1994 et d'autres), le lien est beaucoup plus discuté en acquisition d'une langue étrangère à un âge adulte.¹ A titre d'exemple représentatif, certains ont proposé que l'acquisition de la négation dépende de l'acquisition de la morphologie verbale, plus particulièrement de la « finitude » (Perdue et al. 2002). Selon cette optique, la position correcte de la négation s'acquiert si et seulement si l'apprenant a déjà acquis certaines distinctions centrales exprimées par la morphologie verbale flexionnelle (l'opposition entre une forme conjuguée et une forme non conjuguée²).

¹ Nous disons "en principe" parce que certaines interprétations du phénomène dit « Infinitifs de la racine » chez les petits enfants de 2 à 3 ans sont compatibles avec une dissociation de la morphologie et la syntaxe. Nous n'aborderons pas la question ici.

² Ces termes correspondent à la distinction anglosaxonne « [$\pm$ finite forms] ».

D'autres ont argumenté en faveur d'une dissociation entre morphologie et syntaxe. Il a été trouvé que si les formes non conjuguées apparaissent souvent dans les contextes des formes conjuguées (i.e. comme verbe principal de la phrase), l'inverse est beaucoup plus rare (Prévost & White, 2000). En simplifiant, tout se passe comme si les formes non-conjuguées n'étaient « non-finies » qu'en surface. Si les formes conjuguées et non-conjuguées connaissent des distributions différenciées, ceci parle en faveur de l'hypothèse qu'une fonction spécifique est attribuée à chacune de ces formes. Leur distribution ne serait ainsi pas arbitraire. Dans un tel cas de figure, toujours selon Prévost et White, le problème des apprenants n'est qu'un problème de *morphologie de surface*. Ils nomment ainsi leur théorie *Missing Surface Inflection Hypothesis* (Prévost & White, 2000).

Les contributions de ce volume proposent d'aller plus loin dans les recherches sur l'acquisition de la morphologie. Les articles témoignent d'une volonté d'attribuer à l'acquisition de la morphologie verbale flexionnelle une autonomie aussi au niveau des objets de recherche. Cette autonomie est justifiée par l'observation que le développement de la mise en œuvre de la morphologie, p. ex. l'accord sujet-verbe, surtout à l'oral, représente un procès étendu sur plusieurs années d'acquisition. Ceci pose un problème pour une théorie d'acquisition gouvernée par des règles internes qui s'appliquent de façon successive et aveugle. C'est pourquoi plusieurs auteurs de ce volume (Schlyter et Bartning, Thomas, Ågren, Hedbor) adoptent plutôt une autre approche basée sur la fréquence. Comme l'exprime Thomas dans sa contribution : « Dans une approche centrée sur l'**usage de la langue** (*usage-based approach*), l'acquisition d'une langue est étroitement liée aux probabilités d'occurrence mettant en relation forme et fonction (N. Ellis 2002). » Ces auteurs citent aussi Bybee, Andersen, et mentionnent Pinker, pour discuter leurs données. Largy arrive à une conclusion semblable mais par des méthodes expérimentales sur la production par des adultes francophones.

Nous pouvons identifier un certain nombre de questions qui sont communes aux auteurs du présent volume, ce qui nous a permis de les réunir dans trois parties entre lesquelles il y a pourtant une communication importante. Dans ce qui suit nous tenterons de discuter autour de ces questions en comparant les approches et certains résultats des contributions individuelles. Cette introduction se veut donc à la fois une présentation et une petite synthèse de certains résultats émergents.

I Le 'triangle infernal' du français : parler - parlé – parlait

Le principe alphabétique opère un appariement des (groupes de) graphèmes à des phonèmes. Dans le cas idéal, comme le dit Ågren (ce volume), tout phonème correspondrait à un et un seul graphème et inversement. Dans cette perspective le français doit être considéré comme un système opaque avec un certain nombre de « morphèmes silencieux ». A l'oral, pour les formes verbales de la 1^{ère} conjugaison (-er), le résultat en est au moins une triple ambiguïté des formes (/parle/ > *parler* ou *parlé* (ou *parlez*) et selon les prononciations *parlai/s/(en)t*). Ce « triangle infernal » du français, pour emprunter le terme de Largy, pose un certain nombre de questions pour l'étude de la morphologie verbale. Tout d'abord, ces formes en [e] finales demandent une désambiguïsation. Les articles de Thomas et de Granget posent la question de savoir comment désambiguïser ces formes et par quels critères : sémantiques (Thomas) ou distributionnels

(Granget). D'autre part, il faut discuter de leur source. Thomas pose la question de savoir quelle est l'influence des catégories sémantiques, par l'intermédiaire de leur fréquence en usage (input), sur la production des formes verbales par les apprenants.

II Traits morphologiques comme indicateurs de niveau linguistique

Dans la tradition de la grammaire générative, les règles syntaxiques sont interdépendantes. Un paramètre sous-jacent peut contrôler plusieurs règles de surface. Les études sur l'acquisition d'un paramètre spécifique prennent ainsi en considération plusieurs phénomènes de surface.

Différemment, la morphologie a souvent été étudiée dans une perspective plus atomiste, trait par trait, morphème par morphème, marquage par marquage. Les essais de systématisation à un niveau plus général nous semblent peu nombreux du moins pour le français langue étrangère. Ceci est l'un des buts centraux que se sont données Bartning & Schlyter dans leurs travaux sur les « stades de développement » formulés comme « Profils grammaticaux ». Sur la base de différents corpus d'interlangue, ces auteurs ont identifié un certain nombre de phénomènes morphosyntaxiques qui semblent communs aux apprenants suédophones (Bartning & Schlyter, 2004). Avec leurs travaux basés sur des corpus et des analyses détaillées, ces auteurs vont plus loin dans le sens qu'ils montrent que le problème dit général de la morphologie de surface n'est pas aussi général comme la théorie de *Missing Surface Inflection Hypothesis* pourrait le laisser croire. Il s'avère que ce problème touche en effet différemment les sous-catégories linguistiques, souvent en fonction des facteurs externes identifiables, notamment la fréquence de la forme dans l'usage. Seule une comparaison intra- et intercatégorielle permet de relever ces différences subtiles. L'article de Schlyter & Bartning du présent volume sur l'acquisition de l'accord sujet - verbe en est un exemple. Plutôt que de parler d'un problème général de mise en relation forme – fonction (i.e. « mapping ») de traits de personne et nombre sur la morphologie verbale de surface, ils montrent que l'accord sujet – verbe suit un itinéraire d'acquisition spécifique :

1. verbes non-thématiques au singulier : *je suis - il est; j'ai - il a* etc.
2. 1^{ère} personne pluriel : *nous V-ons*
(2'. *je parle* vs. *je parler*)
3. Verbes en -ont à la 3^{ème} personne pluriel : *ils ont/sont/vont/ont*
4. Verbes thématiques à la 3^{ème} personne pluriel, type *ils prennent*

Qui plus est, nous trouvons de grandes similarités entre ce développement, à l'oral, et celui de l'écrit tel que présenté par Hedbor ainsi que par Ågren et par Granfeldt.

Towell (2004) dans son commentaire récent de cette approche dit que :

The challenges for this kind of approach are of two main kinds. First, there are significant methodological barriers to overcome, especially with regard to statistically based comparisons between cohorts of different sizes learning in different contexts. Ensuring that comparisons are justified is a major task. In addition, the items which are highlighted in the proposed orders of acquisition tend to be the ones which have occurred in the data, as opposed to having been elicited by predetermined tightly controlled investigative methods. This rarely gives rise to the number of exemplars which are required for statistical treatment. Second, there is an 'explanatory' challenge. The advantage which the approach has of offering data up for a variety of possible explanations becomes a disadvantage when the information is presented without a supporting theoretical framework. [Towell, 2004: 364]

Concernant le premier point, une réponse partielle pourrait se trouver dans la suite de la contribution de Granfeldt. Le logiciel *Direkt Profil* est un analyseur automatique de textes écrits ou de transcriptions de français langue étrangère. Le système implémente la théorie des itinéraires d'acquisition de Bartning & Schlyter dans sa version actuelle. Grâce à *Direkt Profil* et l'emploi des données écrites, il sera possible de monter encore plus la base empirique pour ces recherches au niveau où les analyses statistiques nécessaires sont possibles. Deuxièmement, la contribution de Granfeldt montre aussi la validité du critère de « Formes Lexicales Conjuguées » vs. « Les Formes Lexicales Non Conjuguées ». Ce critère s'avère valide pour analyser des corpus si différents que le français parlé par des anglophones (corpus de Myles et al.) et le français écrit des suédophones (corpus d'Ågren).

Le deuxième commentaire de Towell (2004) concerne la question de savoir pourquoi certains itinéraires d'acquisition apparaissent. Il semblerait qu'il n'existe pas, à l'heure actuelle, une théorie cohérente en acquisition de langue étrangère pour discuter, avec le niveau de détail nécessaire, l'acquisition successive des marquages morphologiques. Nous sommes d'avis qu'une telle théorie peut émerger sous deux conditions nécessaires et suffisantes : (i) une meilleure compréhension des phénomènes qui vont ensemble en acquisition du français langue étrangère et (ii) une intégration de différentes approches théoriques. Si la morphologie a dû négocier sa place dans les théories linguistiques, c'est aussi parce qu'elle est complexe. Inutile de dire que la théorie qui veut la traiter en acquisition reflétera cette complexité inhérente. La dernière section du présent volume montre bien comment un seul phénomène, à l'occurrence le 3^{ème} personne du pluriel, doit être traité sous des angles différents.

III Un regard sur la 3^{ème} personne du pluriel

Les articles de la troisième partie de ce volume se concentrent donc tous sur la production et l'acquisition de la morphologie de la troisième personne du pluriel. La section est introduite par l'article de Largy sur l'approche psycholinguistique de ce phénomène, où il montre l'intérêt théorique de ce phénomène linguistique, qui à première vue, semble très isolé. Sa conclusion va à l'encontre d'un traitement syntaxique : « Si l'accord en nombre entre le verbe et son sujet est enseigné par le pédagogue comme un principe grammatical, les recherches en psycholinguistique nous montrent que cet accord n'est pas exclusivement régi par des processus syntaxiques mais qu'il s'applique aussi sous l'influence des niveaux phonologique et sémantique de la production langagière. »

Les articles suivants reflètent bien cette mise en perspective mais s'insérant dans un contexte d'acquisition de langue étrangère, ces auteurs sont obligés d'élargir encore le spectre d'influences. Nous identifions, dans ces articles, les facteurs suivants mentionnés comme jouant un grand rôle pour le développement de la marque de la 3^{ème} personne pluriel chez les apprenants :

- (i) la langue maternelle (Byrskog)
- (ii) l'instruction grammaticale (Granget, Schlyter)
- (iii) la régularité et la fréquence des formes écrites (Ågren, Hedbor)

Byrskog, dans son article (en allemand) sur l'acquisition de l'allemand L2, compare des élèves

suédois et italiens à deux niveaux généraux comparables. Elle montre que les italophones maîtrisent nettement mieux la 3^{ème} personne pluriel que les suédophones, et ceci bien que ces derniers aient reçu plus d'enseignement que les élèves italiens. Ces résultats soulignent, selon Byrskog, l'importance des propriétés du système verbal de la langue maternelle.

On pourra éventuellement discuter ici le rôle de l'instruction grammaticale pour son résultat. Celle-ci est possiblement plus intense dans les écoles italiennes que suédoises, dans lesquelles on enseigne aujourd'hui plutôt la communication que la grammaire. Dans son deuxième article de ce volume, Granget propose justement que le bon résultat de ses élèves germanophones serait dû à un enseignement centré sur des formes de l'accord sujet-verbe. Pourtant, en ce qui concerne le pluriel dans le français parlé, Schlyter dans sa comparaison des apprenants guidés et non-guidés, tous suédophones, n'enregistre aucune différence. Ensemble, ces trois articles montrent donc d'une manière intéressante la complexité des recherches en langue étrangère et la nécessité d'une approche multifactorielle à ces questions.

Plusieurs études dans cette section concernent les mêmes formes de la 3^{ème} personne au pluriel dans des corpus différents : Hedbor, Ågren, et Schlyter (ainsi que Schlyter et Bartning) discutent les formes représentant la 3^{ème} personne au pluriel, en faisant une distinction entre d'une part les formes non-thématiques très fréquentes qui terminent par *-ont* : *ont, sont, vont, font* et d'autre part les verbes lexicaux (thématiques et modaux), du type *il prend - ils prennent*. Les études de l'écrit prennent aussi en considération les marques inaudibles à l'oral, donc pour les verbes de la première conjugaison comme *il parle - ils parlent*. Toutes ces études montrent, d'une part, qu'il y a une progression continue dans la maîtrise de ces formes par l'apprenant ; d'autre part que les formes en *-ont* sont acquises plus tôt que les formes des verbes lexicaux. Ce développement est surtout très net dans le corpus écrit, très riche, d'Ågren. Dans la contribution de Hedbor, cette différence est analysée aussi en temps réel, où elle montre un accès immédiat pour les formes '*ont, sont, vont, font*' alors que les verbes lexicaux demandent plus de réflexion et semblent ainsi être traités par une règle plutôt que par un accès lexical immédiat.

Concernant l'écrit et l'oral, nous pouvons observer que les marques morphologiques de la 3^{ème} personne pluriel sont acquises plus tôt à l'écrit. Dans le corpus d'Ågren on observe que, au stade 4, les apprenants fournissent la désinence *-nt* à pratiquement 100%, même si la racine est parfois mal choisie. En revanche, dans les deux corpus de français parlé, Schlyter et Bartning montrent que les niveaux supérieurs, représentant les stades 4 et 5, n'atteignent que 40% et 59%. Granget constate une quasi omniprésence de l'accord au pluriel dans les productions en français de ses apprenants germanophones, et propose une relation avec le contexte scolaire d'apprentissage et l'input écrit. Mais il faut aussi tenir compte des facteurs intralinguistiques. Les résultats de Granget montrent l'influence des propriétés du sujet grammatical. En général, les verbes de 3^{ème} pluriel sont mieux accordés avec les sujets pronominaux. Granfeldt fait la même observation dans son article.

Les résultats de ces comparaisons pointent dans la direction des facteurs qui indiquent une fréquence de la cohésion forme - fonction. Le rapport entre forme et fonction de la 3^{ème} personne pluriel à l'écrit suit parfaitement le principe bi-univoque (Andersen, 1984) alors qu'à l'oral c'est loin d'être le cas. Dans la combinaison « ils ___ent » l'accord est acquis plus tôt que dans les combinaisons avec des SN variés, ce qui suit le même principe que celui proposé pour « nous ___ons » dans Schlyter et Bartning: un cadre invariable avec un 'slot' pour le lexème verbal.

Malgré certaines différences méthodologiques, il n'en reste pas moins qu'en comparant les données de Granget d'élèves germanophones avec celles d'Ågren, on observe une différence nette, de telle sorte que les apprenants suédois développent leur accord au pluriel beaucoup plus tard que les germanophones. Les germanophones étudiés par Granget produisent la désinence du pluriel dans environ 80% des cas, dès le début et dans toutes les classes d'école étudiées, alors que les élèves suédois n'arrivent à ce taux qu'après plusieurs années au lycée. Nous ne savons pas si cette différence très nette est due à l'influence de l'allemand comme langue source, qui marque aussi l'accord sujet-verbe (comme l'italien dans l'étude de Byrskog), ou en revanche, si le type d'instruction dans les pays différents joue un rôle.

Tableau récapitulatif des articles du présent volume

Auteur	Phénomène(s)	Corpus (écrit / oral)	Type d'apprenants / sujets	L1 des apprenants / sujets
Granget I	formes en [e]	oral	guidés (purs)	allemand
Thomas	formes en [e]	oral	guidés	suédois
Schlyter & Bartning	accord sujet – verbe	oral	guidés et non-guidés	suédois
Granfeldt	formes en [e], pluriel	oral / écrit	guidés	anglais / suédois
Largy	accord sujet- verbe, 3 ^{ème} pluriel	écrit (corpus expérimental)	adultes francophones	adultes francophones
Byrskog	accord sujet-verbe, 3 ^{ème} pluriel	écrit allemand L2	guidés	suédois / italien
Granget II	accord sujet – verbe, 3 ^{ème} pluriel	écrit	guidés	allemand
Schlyter	accord sujet –verbe, pluriel	oral	guidés vs. non-guidés	suédois
Ågren	3 ^{ème} pluriel	écrit	guidés	suédois
Hedbor	accord sujet – verbe	écrit	guidés	suédois

Références bibliographiques

- Andersen, R. (1984) « The one-to-one Principle of interlanguage construction ». *Language Learning* 34 : 77-95.
- Bartning, I. & Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies* 14 : 281-299
- Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht : Foris.
- (1995) *The minimalist program*. Cambridge (Mass.) : MIT Press.
- Halle, M. & Marantz, A. (1994) « Distributed morphology and the pieces of inflection ». In K. Hale et S-J. Keyser (éds.), *The view from building 20 : Essays in linguistics in honor of Sylvain Bromberger*, pp. 111-176. Cambridge (Mass.) : MIT Press.
- Meisel, J. M. (1994) « Getting FAT: Finiteness, Agreement and Tense in Early Grammars ». In J. M. Meisel (éd.) *Bilingual First Language Acquisition*, pp. 89-129. Amsterdam/Philadelphia : John Benjamins.

- Perdue, C., Benazzo, S. & Giuliano, P. (2002) « When finiteness gets marked: the relation between morphosyntactic development and use of scopal items in adult language acquisition ». *Linguistics* 40(4) : 849-890.
- Prevost, P. & White, L. (2000) « Missing Surface Inflection or Impairment in Second Language Acquisition? Evidence from Tense and Agreement ». *Second Language Research* 16(2) : 103-133.
- Towell, R. (2004) « Research into the second language acquisition of French: achievements and challenges ». *Journal of French Language Studies* 14 : 357-375.

I

*Le ‘triangle infernal’ du français :
parler - parlé – parlait*

La désambiguïsation des formes verbales en [e] dans des récits oraux d'apprenants germanophones du français¹

Cyrille Granget

Université de Paris III
cgranget@wanadoo.fr

1 Question pratique, enjeux théoriques

La désambiguïsation des formes en [e] soulève un enjeu majeur lorsqu'on travaille sur un corpus d'apprenants en français parlé. En effet, quiconque entreprend de transcrire établit une correspondance entre les formes verbales en [e] dans le français parlé et des formes verbales écrites. Transcrire un corpus oral est toutefois une entreprise paradoxale pour celui dont le questionnement ne porte pas uniquement sur le contenu informationnel du discours mais aussi sur les formes linguistiques et notamment les unités morphologiques et le sens qu'elles véhiculent.

D'abord, comme le souligne Blanche Benveniste (1997 :117), la morphologie verbale est le domaine où les différences entre le français parlé et le français écrit sont les plus grandes. Ainsi, il peut aussi sembler hasardeux, et à plus forte raison dans un corpus d'apprenants, d'établir une correspondance entre une forme verbale orale en [e] et une forme du français écrit. La morphologie verbale du français écrit est beaucoup plus déterminée et riche que la variété morpho phonologique du français oral : elle véhicule davantage d'informations morphosyntaxiques et sémantiques comme la flexion, l'accord en nombre, en genre, le temps, et l'aspect. D'où le fait que le [e] post-verbal peut potentiellement se transcrire de nombreuses façons : -é (participe passé), -er (infinitif), -ez (présent, 2^{ème} personne pluriel) , -ai, -ais, -ait -aient (imparfait, respectivement 1^{ère}, 2^{ème}, 3^{ème} personne singulier, et 3^{ème} personne pluriel, la distinction /e/ et /ɛ/ n'étant pas toujours établie dans les interlangues). Trancher en faveur d'une transcription repose sur l'idée que le locuteur parlant possède une représentation grammaticale du français telle que la donne à voir l'écrit. Or ce qui vaut pour l'adulte lettré n'est peut-être point valable dès lors que l'on considère le jeune apprenant du français langue première ou l'apprenant du français langue seconde en immersion sociale ? L'enfant, ou l'adulte, qui n'est exposé qu'au seul français oral peut-il à partir du seul input déduire que [vole] dans les énoncés a, b et c ci-dessous ne correspond pas à la même forme, mais en (a) à un participe et en (b) et (c) à un infinitif ?

¹ Nous tenons à remercier Suzanne Schlyter de nous avoir incités à entamer cette réflexion sur le corpus et à prendre conscience des intuitions qui guidaient notre démarche de transcription.

- a. il a /vole/
- b. il doit /vole/
- c. /vole/ ça n'se fait pas

Désambiguïser, c'est supposer que le locuteur a une représentation grammaticale de la langue. Or, est-ce possible sans entrer dans le français écrit ? En fait, il est curieux de constater que l'analyse linguistique vise précisément à répondre à cette question qui en réalité surgit bien avant l'analyse, lorsque le matériau est en cours d'élaboration.

Ensuite, il importe de prendre en considération le fait que l'apprenant scolarisé du français langue seconde, contrairement à l'apprenant de langue première, entre non seulement dans le français parlé mais aussi et surtout dans le français écrit dès les premières heures de son apprentissage de la langue étrangère. Ainsi, dans un corpus d'apprenants scolarisés du français L2, une forme verbale en [e] peut être une réalisation oralisée de formes stockées sous leur forme écrite. Dans ce cas, l'oralisation peut s'effectuer selon les règles grapho-phonétiques de la langue première. Ainsi, un morphème graphique en –e peut donner lieu pour un germanophone à une réalisation en [e]. D'une façon générale, le fait que l'acquisition de la phonologie du français est en cours ouvre des possibilités supplémentaires d'interpréter les formes verbales en [e].

Bien que l'ambiguïté des formes en [e] soit patente, la désambiguïstation s'avère sinon nécessaire alors souhaitable pour celui qui souhaite analyser un corpus oral. Cependant, la transcription, ce travail de fourmi apparemment dénué de réflexion, a des implications théoriques non négligeables dans un champ linguistique traversé par des options théoriques fortes, et notamment par la question de l'accessibilité des catégories fonctionnelles de la langue, tels l'inflexion et l'accord, au cours de l'acquisition.

Les grandes hypothèses qui circulent en acquisition de L2 à propos du développement de la morphologie verbale portent en effet sur le fait de savoir si les formes verbales présentes dans les lectures d'apprenants, dont les formes en [e], obéissent à une distribution systématique ou aléatoire (hypothèse de la variété de base, Klein et Perdue 1997), si l'occurrence de formes en [e] en position de verbe fléchi résulte d'un déficit morphologique de surface (*Missing surface inflection*, Lardière 1998, Herschensohn 2001) ou, au contraire, de l'absence de flexion ou de finitude morphologique (Perdue et al. 2002). La transcription des données faisant déjà apparaître des hypothèses concernant la présence des catégories morphosyntaxiques et sémantiques dans les langues d'apprenants, il importe de s'interroger sur les implications théoriques de la démarche de transcription. On peut ainsi penser que les divergences d'interprétation du corpus ESF d'apprenants arabophones du français L2 rencontrées dans les études de Noyau et al. (1995) et Prévost et White (2000) reposent en partie sur une transcription différente du corpus oral initial, transcription phonétique versus orthographique (cf. Schlyter 2003 pour d'autres explications).

Il n'y a donc probablement pas de transcription innocente, fût-elle conçue dans un esprit de fidélité à l'original. Quel que soit le discours de départ, on transcrit à la fois pour rendre compte d'un matériau oral, mais, cette entreprise étant d'emblée paradoxale, le français parlé étant moins déterminé que le français écrit, on peut donner à voir plus que ne donne à voir le matériau initial. C'est pourquoi il importe de s'interroger sur les implications théoriques de cet acte apparemment

banal qu'est la transcription. De façon schématique, deux options sont possibles : soit on considère que le discours de l'apprenant est un flux sonore segmenté que l'on transcrit intégralement phonétiquement en mentionnant les pauses intonatoires, soit on considère qu'il s'agit d'une lecture oralisée du français écrit, auquel cas une transcription orthographique s'impose. Dans le premier cas, la transcription présente l'avantage d'être fidèle au discours initial, mais difficile d'accès et d'utilisation pour l'analyse linguistique selon les catégories de l'analyse morphosyntaxique. Dans le second cas, la transcription orthographique présente l'avantage d'être lisible et prête à l'analyse, mais présente aussi l'inconvénient d'être surinterprétée. Il est fort probable en effet que le transcripateur impose à l'apprenant sa propre segmentation de la langue, travers, assez répandu dans les disciplines analytiques, que souligne par exemple Bourdieu (1994 : 215) et nomme la *scholastic fallacy* ou l'illusion scolastique en référence à la *scholastic view* d'Austin. Peut-on dès lors transcrire à des fins d'analyse sans tomber dans le piège de l'opacité ou de l'illusion ? La transcription idéale serait celle qui rendrait compte des représentations grammaticales de l'apprenant, qui nous indiquerait par exemple que *je ne sais pas* dans les énoncés *il vole le pain il a faim je ne sais pas il est triste* est en fait une formule préfabriquée qui ne suppose pas de l'apprenant qu'il maîtrise à ce stade de l'acquisition, la négation post-verbale, et qui, à ce titre pourrait être transcrite en un bloc graphique. Remarquons toutefois que ce type de transcription ne peut être qu'une transcription après coup qui a lieu après l'analyse détaillée du corpus, après confrontation de plusieurs discours. Nous partons de l'idée qu'il n'y a pas de transcription définitive, que le matériau transcrit est sujet à révision au fil de l'analyse linguistique et selon les domaines que l'on considère, morphosyntaxe, discours ou lexique, d'où l'importance des corpus oraux digitalisés qui permettent de faire se correspondre la transcription et le flux sonore (possibilité offerte notamment par l'éditeur CHAT de CHILDES) et permettent notamment, du fait de la repérabilité des pauses intonatoires, d'avoir accès aux segmentations micro-syntaxiques effectuées par l'apprenant et susceptibles de fournir quelques éclairages sur le processus de production.

Cet article a pour vocation d'exposer les cas ambigus des formes en [e] rencontrés dans notre corpus de récits oraux d'apprenants germanophones scolarisés, d'expliquer les choix de désambiguïsation et de non désambiguïsation que nous avons effectués. Il ne s'agit pas d'édifier un protocole de transcription des données orales mais d'ouvrir la discussion sur les indices pertinents pour désambiguïser les données. Tâche délicate mais nécessaire pour quiconque se lance dans l'aventure de la transcription des données d'apprenants.

2 Les données

2.1 Les récits oraux

Les données considérées sont des récits oraux d'un extrait du film *Les Temps Modernes* de C. Chaplin (1936), dans lequel deux protagonistes sont à tour de rôle sujets et objets de nombreuses péripéties dont certaines nécessitent de référer à des événements passés.

Les narrateurs apprenants, âgés de 12 à 16 ans, sont issus de cinq groupes classes différents et répartis sur quatre niveaux institutionnels comme l'indique le tableau 1 ci-dessous. Ils ont visionné l'extrait de film de 8' puis en ont fait un récit en temps illimité face à un interlocuteur

francophone natif. Les narrateurs ont obtenu une fiche lexicale sur laquelle apparaissent quelques formules préfabriquées, des noms et des verbes à l’infinitif. Le recours à cette fiche est fréquent dans les stades initiaux et plus rares dans les stades avancés.

Tableau 1. *Groupes classes, année d’apprentissage du français et âge*

GROUPE-CLASSE	1 (7C)	2A (8C)	2B (8E)	3 (9C)	4 (10C)
Année d’apprentissage du français	1^{ère} (7 ^{ème})	2^{ème} (8 ^{ème})	2^{ème} (8 ^{ème})	3^{ème} (9 ^{ème})	4^{ème} (10 ^{ème})
Age (en années)	12-13	13-14	13-14	14-15	15-16

Nous disposons au total de 17 récits oraux du même récit de film. Le tableau 2 ci-dessous indique la répartition du nombre de récit par groupe ainsi que l’identifiant des narrateurs-apprenants.

Tableau 2. *Participants au récit de film oral*

Groupe N =	Participants
17	
1 (n = 5)	MAT, MAY, MIK, NIN, SAS.
2’(n = 5)	ASL, BEN, BER, HIL, NAD.
2 (n = 6)	ANT, MES, MIL, NOR, REA, VIO.
3 (n = 4)	DAN, EYS, LAU, REG.
4 (n = 7)	AID, ANN, DAV, GREG, LEA, LEO, SAH

La consigne de narration, qui a été donnée à tous les apprenants, était « Raconte-moi le petit film s’il te plaît ». Elle n’imposait pas d’ancrer le récit au passé. Seuls quelques apprenants parmi les plus avancés le font. Toutefois des formes en [e], auxiliées ou pas, surgissent dans les récits. Il importe de remarquer que cet ancrage optionnel du récit au présent ou au passé rend la désambiguïsation particulièrement difficile. A cet égard, les récits personnels sont moins ambigus car ils font plus clairement référence à une période temporelle donnée par rapport à l’ici-maintenant du narrateur.

2.2. L’exposition aux catégories morphosyntaxiques du français écrit

Les apprenants, scolarisés, ont été observés durant plusieurs mois pendant leurs cours de français. Il a ainsi été possible de reconstituer quels étaient les temps verbaux auxquels ils avaient été exposés en classe, de façon explicite et implicite. Il s’avère que la progression linguistique de l’enseignement est un précieux indice de désambiguïsation (Granget 2004).

Tableau 3. *Exposition quantitative et qualitative des groupes-classes au français au moment du recueil des récits de film.*

Groupes	Exposition * (en heures)	Temps verbaux**	ADVT	Verbes
1 (n=5)	78	PR	alors, encore, puis, et puis, après, depuis, en ce moment, aussi, toujours, seulement, déjà	être, avoir, aller, verbes en –er
2' (n=5)	218	PR		+ prendre, pouvoir, vouloir, acheter, appeler, payer, manger, commencer, prendre, dire, écrire, mettre, savoir, voir, etc.
2 (n=6)	218	PR, PC		+ plaire, courir, etc.
3 (n=4)	338	PR, PC, IMP		
4 (n=7)	458	PR, PC, IMP, PQP		

* Nous avons calculé le nombre de séances ayant effectivement lieu sur la base d'une année scolaire, multiplié ce nombre par 45 mn, puis transcrit le chiffre obtenu en nombre d'heures. Une année de cours correspond sur cette base à 120 h d'exposition au français.

** PR : Présent, PC : Passé Composé, IMP : Imparfait, PQP : Plus-que-parfait

On observe ainsi que l'imparfait est enseigné tardivement. De fait, aucun indice autre ne permet d'avancer que les apprenants en années 1 et 2 emploient un imparfait. Nous avons en conséquence exclu la possibilité que coexistent deux formes du passé (passé composé et imparfait) avant que les deux formes ne surgissent dans l'input. Autrement dit, les occurrences sujet - V[e] dans un contexte de référence au passé précédant l'enseignement de l'imparfait sont considérées comme des participe passé, si l'apprenant par ailleurs produit des formes aux+ V-[e].

3 Les indices de désambiguïsation

Désambiguïser les suffixes en [e], c'est réduire le champ des possibles en optant pour un morphème. Pour désambiguïser une forme en [e], il importe donc de trouver dans le corpus des indices selon lesquels l'apprenant emploie une forme en [e] du fait de contraintes syntaxiques ou à des fins sémantiques distinctives précises.

Les indices qui nous ont permis de désambiguïser sont les suivants :

- Le contexte extra-linguistique, c'est-à-dire l'action du film dont les apprenants font un récit. A quelle réalité réfère le locuteur-apprenant ? Dans quelle relation se situe

l'événement mentionné par le verbe, par rapport à ceux précédemment mentionnés : succession, antériorité ?

- Le co-texte : essentiellement les autres formes verbales, le sujet du verbe, pronom clitique ou syntagme nominal, et les adverbes.
- L'analyse parallèle de récits écrits, élicités à partir du même corpus.
- L'observation des activités institutionnelles qui d'une part permettent de déterminer où se situe l'apprenant dans son apprentissage de la phonétique et de la morphologie du français et d'autre part déclenchent l'émergence de nouvelles formes verbales et ouvrent de nouvelles possibilités d'interpréter les formes verbales en [e].

Il nous semble que les indices varient en fonction des stades ou degrés de morphologisation de l'interlangue mais aussi en fonction des locuteurs, des locuteurs-pairs appartenant au même groupe, maîtrisant par exemple la subordination, pouvant, concernant la morphologie faire des distinctions variables d'un narrateur à l'autre.

Il importe donc de prendre en compte le fait que les connaissances déclaratives (morphologiques surtout) et procédurales (phonologiques par exemple) d'un apprenant évoluent au cours de l'apprentissage. Cette évolution nous conduit à chaque stade à envisager de nouvelles interprétations des formes en [e] et à en éliminer d'autres. Par ailleurs, on observe parfois ce que l'on pourrait appeler un style individuel, le fait par exemple qu'un apprenant n'emploie en position de verbe fléchi que des formes en [e] et des infinitifs irréguliers, saillants à l'oral. Ce cas nous invite à interpréter les formes en [e] du récit comme des infinitifs. Il existe également des invariants intra-groupes qui incitent à privilégier une piste de désambiguïsation. Enfin, il se peut qu'aucun indice ne permette de désambiguïser les formes verbales en [e] ou que des indices différents mènent à des désambiguïses divergentes. Dans ces cas de figure, nous maintenons une transcription phonétique de la forme en [e] qui, loin de ne rien signifier, signifie au contraire que les langues d'apprenants sont des systèmes intermédiaires faisant état de formes idiosyncrasiques résistant aux analyses classiques mais porteuses du sens de la dynamique acquisitionnelle.

4 Des cas de désambiguïsation et de non-désambiguïsation

Désambiguïser les formes verbales en [e] permet d'établir un premier relevé quantitatif des formes verbales du corpus et de relever des grandes tendances de la dynamique acquisitionnelle. Il apparaît ainsi que les apprenants germanophones emploient uniquement le présent dans les stades initiaux puis le passé composé puis l'imparfait. Concernant les formes non fléchies, dont les formes en [e], elles représentent entre 0 et 18 % des formes verbales des apprenants de première et deuxième années mais elles représentent entre 0 et 60 % des formes verbales des apprenants de troisième et quatrième années. En d'autres termes, au fil de l'acquisition, l'écart intra-groupe augmente, en même temps que se diversifient les formes verbales. En effet, c'est également en troisième et quatrième années que le passé composé et l'imparfait se répandent dans les productions des apprenants.

Tableau 4. *Développement de l'emploi des formes verbales dans les récits oraux (en %)**

Groupe N=27	PR	PC	IMP	Vnf	Vlr
1 (n= 5)	90 (80-90)	0	0	8 (0-13)	2 (0-13)
2' (n=5)	70 (50-100)	0	0	12 (0-18)	15 (0-30)
2 (n=6)	80 (70-90)	4 (0-7)	0	5 (0-11)	11 (3-33)
3 (n=4)	60 (50-80)	11 (0-26)	3	13 (3-40)	11 (3-17)
4 (n=7)	60 (20-80)	9 (0-18)	4	25 (0-60)	3 (0-18)

* Les chiffres entre parenthèses indiquent l'écart intragroupe. Le premier chiffre indique le taux d'emploi le plus faible du groupe, le second le plus élevé.

Légende : PR : présent, PC : passé composé, IMP : imparfait, Vnf : verbe non fléchi, Vlr : verbe en langue ressource (allemand ou anglais)

Nous allons désormais présenter une étape de l'élaboration du tableau 4 ci-dessus en regardant de plus près les formes en [e] du corpus. La présentation des cas de désambiguïsation suit la chronologie du développement linguistique en présentant dans un premier temps le cas de formes verbales en [e] employées par des apprenants débutants pour finir par des occurrences dans des récits d'apprenants ayant été exposés près de 450 heures au français et notamment au passé composé, à l'imparfait et au plus-que-parfait.

Année 1

En année 1, les apprenants n'ont été exposés qu'à deux formes du verbe, la forme infinitive et la forme fléchie du temps présent. Ce constat nous incite à penser que la forme verbale *demander* dans le récit de l'apprenant NIN en (1) est la forme infinitive.

(1)

*NIN: Charlie Chaplin et dame um ja@s euh l'occasion <et elle bff> [/] et
il pff est sur une bank oder?

*INV: oui ah sur un banc.

- *NIN: <et **demander**> [/] et demande euh und euh dix+jours+plus+tard euh.
 *INV: hm hm.
 *NIN: le madame euhm euh ja **[se]** [:=? say c'est]
 j'ai une surpra pour toi
 je trouve notre maison.

Nous avons tendance à considérer que la forme infinitive est la forme de base dans les lectures d'apprenants germanophones scolarisés car elle est très présente dans l'input de la classe, que ce soit dans les manuels, au tableau lorsque l'enseignante note un verbe, mais aussi dans les exercices de pratique linguistique. Le terme métalinguistique d'infinitif est connu des apprenants, il est également utilisé pour la description de l'allemand. Sur la fiche lexicale qui est distribuée aux apprenants lors de l'expérience de récit, les verbes apparaissent soit sous forme de *chunks*, soit sous leur forme infinitive. Nous considérons par conséquent que *demander* est un infinitif et nous permettons de l'orthographier en conséquence. Nous en déduisons que l'apprenant dans son processus de production en temps réel mobilise dans un premier temps la forme de base infinitive puis la forme fléchie.

En revanche, nous n'avons pas de solution pour désambiguïser [se] et en proposons entre crochet deux interprétations possibles. D'un côté cette forme verbale ressemble phonologiquement à *c'est* en français, mais d'un autre côté, elle apparaît souvent dans des contextes où elle pourrait avoir le sens de *dire*, comme ci-dessus, ce qui nous a conduit à penser que [se] pourrait être une forme empruntée à l'anglais *to say* et mobilisée en l'absence du verbe *dire*. En effet, dans le récit de NIN ci-dessus, le verbe *dire* n'est pas employé. Cette dernière explication fonctionne dans plusieurs cas mais pas dans tous les cas, ce qui nous a incité à maintenir une transcription intermédiaire. Ceci a également été motivé par la lecture des travaux en acquisition non guidée (Starren 2001) où [se] apparaît comme un marqueur intermédiaire important dans le processus de morphologisation des lectures d'apprenant.

En année 1, on peut également noter que l'apprenant n'a pas toujours maîtrisé la distinction entre [ə], [e] et [ɛ] ou n'arrive pas encore à réaliser ces trois phonèmes de façon distincte comme en témoigne le récit de l'apprenant MIK ci-dessous dans lequel apparaît également [se] à côté de [le], article placé devant les noms *filles* et *garçon*.

(2)

- *MIK: et il eum **lé** fille eu eum **sé** [:=? say c'est] l la fille pas la
 homme eum et eum <lé lé> [/] **lé** file eu elle <donner> [/] **donner**
 pour luiss et <lé eu> [/] **lé** garçon dans le police oui et eu lé
 fille dit ne me reconnaisse vous pas je suis l'homme avec le pain.

En année 1, les apprenants germanophones scolarisés ont bénéficié de 78 heures environ de cours de français. Durant le cours, ils sont certes exposés au français parlé à travers les cassettes audio qui accompagnent le manuel et présentent les dialogues reproduits dans le livre mais aussi à travers le discours de l'enseignante. Cependant cet input oral spontané est doublé d'un input écrit, accessible à tout moment, en classe, mais aussi en dehors de la classe. On peut penser que les apprenants ont en tête la forme écrite des mots, des énoncés, mais pas encore leur réalisation phonétique, ce qui expliquerait qu'en situation de production orale, ils oralisent des mots dont ils ne connaissent pas la réalisation phonétique. Ainsi, ils oraliseraient selon les règles grapho-phonétiques non pas du français mais de l'allemand. Dans cette perspective, le *e* graphique peut-

être prononcé [e] en allemand. On peut également penser que *c'est* est réalisé [se], la réalisation en [ɛ] du *est* graphique ne tombant pas sous le sens. En classe par ailleurs, les apprenants débutants sont très souvent amenés à verbaliser en lisant, lors de la correction d'exercices faits à la maison par exemple. Les activités de production spontanée sont très rares à ce stade de l'apprentissage. Ces habitudes de production orale sont encouragées dans la mesure où les apprenants ont reçu une fiche lexicale sur laquelle apparaissent des verbes parfois inconnus comme *emmener*. Si le sens de ces verbes est accessible du fait de leur équivalent en allemand, leur réalisation phonétique ne l'est pas, ce qui explique des prononciations intermédiaires comme dans le récit de MAY ci-dessous : l'apprenante prononce *emmener* **[ɛ]mener** au lieu de **[ɑ̃]mener**

(3)

*INV: donc il n'a pas d'argent hm hm.

*MAY:ok euh le monsieur n'a pas d'argent
et l'agente de police euh **[ɛ]mener** le monsieur et la madame. (lecture de la fiche)

Année 2, groupe A

Le cas de l'apprenant Ben en (4) est intéressant car les pistes pour désambiguïser sont à la fois nombreuses et hasardeuses.

(4)

*INV: et après?

*BEN: sie@s werden@s erwischt@s irgendwie@s.

*INV: pardon elle vole un pain et?

*BEN: et ## erwischt@s?

*INV: elle est arrêtée.

*BEN: **arrêtée** <et il **aider**> [//] <et il> [//] et elle est à la auto (à réécouter)
de police.

*BEN: ils partent **et après ils poser**.

*INV: oui.

*BEN: à la rasen@s rasen@s?

*INV: euh sur l'herbe.

*BEN: sur <l' arbe > [//] l' herbe **et il voir**.

*INV: oui oui.

*BEN: il voir une famille dans une maison
et <il> [//] ## il pose des questions à la fem.

La première occurrence en [e] ci-dessus, *arrêtée*, est une reprise de l'étayage de l'interlocuteur francophone. Celui-ci répond à la demande de l'apprenant qui veut savoir comment on dit *erwischt* (le participe passé en allemand, homophone du présent, du verbe *erwischen* qui signifie attraper), sous-entendu en français. L'interlocuteur qui a le scénario du film en tête rend compte de la scène par *elle est arrêtée*. Aussitôt, l'apprenant, Ben, reprend *arrêt[e]* et poursuit le récit. On pourrait penser que *arrêt[e]* est un participe passé mais aussi objecter que l'apprenant n'a pas, à ce stade, de catégorie telle que « participe passé ». Il se contente de reprendre le verbe lexical sous la forme saillante dans laquelle il apparaît chez son interlocuteur, natif, sans le morphème préverbal

qu'est l'auxiliaire. Une explication à cette chute morphologique est que le verbe est porteur de l'information qui permet de référer à la situation, de faire progresser le récit et de remplir la tâche narrative.

A la suite de *arêt[e]*, le verbe qu'il emploie aussitôt est également en [e]. Ce phénomène s'observe dans d'autres récits, à des stades différents: une forme en /e/ en appelle une autre. On rencontre également dans le corpus plusieurs cas de verbes nouveaux ou rares qui sont employés à l'infinitif. Une raison de la non flexion de verbes rares est le fait qu'il ne connaissent que cette forme, car c'est la forme de base à l'école, et ne connaissent pas la, ou les, forme(s) fléchi(e)s. Le français montre que la maîtrise de l'infinitif ne présuppose pas la maîtrise de la flexion, sachant la variation lexicale du verbe, par exemple avec un sujet grammatical pluriel. Dans le flot du discours, l'apprenant est incité à opérer des choix économiques. Ainsi, *il voir* répond à la même logique. On peut penser qu'il s'agit d'un infinitif et que *poser* en est un aussi. On peut également penser que *voir*, du fait qu'il s'agit d'un verbe irrégulier moins fréquent dans l'input que *être* et *avoir* par exemple, est employé dans sa forme de base à l'écrit, c'est-à-dire à la forme infinitive.

Il importe à notre sens de prendre en considération que l'apprenant scolaire est susceptible de stocker des informations linguistiques différentes de celles stockées par un apprenant non scolarisé exposé au seul input en contexte de communication ordinaire ou en écoutant simplement le français : sur son lieu de travail, dans ses interactions avec les commerçants, à la télévision, etc. En effet, l'apprenant scolaire apprend des verbes à partir de listes dans lesquelles cette classe lexicale apparaît à l'infinitif. Lorsqu'il fait un test de vocabulaire comme cela se pratique dans le contexte observé, la forme demandée du verbe est la forme infinitive. Même si, par ailleurs, il effectue aussi des exercices dans lesquels il est amené à fléchir ces verbes, en fonction du contexte de l'énoncé, c'est-à-dire à apposer au verbe les marques obligatoires en français tels le nombre, le temps, l'aspect, on peut penser que le contexte scolaire incite l'apprenant à ne retenir que la forme infinitive à partir de laquelle il peut ensuite fléchir le verbe. Ceci fonctionne dans le cas de verbes réguliers puisque la base lexicale du verbe infinitif reste la même en cas de flexion mais cela ne fonctionne pas en cas de verbe irrégulier. Il semble que les apprenants sont conscients du phénomène et ne tentent pas de produire une forme lexicale intermédiaire mais préfèrent employer l'infinitif.

Dans le récit de Hilal ci-dessous, on observe que le verbe *raconter* est employé dans sa forme en [e] que nous avons choisi de transcrire à l'infinitif car nous supposons là aussi que l'apprenante du fait qu'elle n'a pas été exposée au passé composé et ne présente pas par ailleurs de réalisation en [e] de ce type mobilise le verbe à l'infinitif. Notons d'ailleurs qu'elle convoque dans le premier énoncé un verbe en allemand et que ce dernier se trouve à l'infinitif (*sehen*).

(5)

- *HIL: et <e m a madame> [/] une madame e f non <sehen> [/] sehen elle regarde.
- *INV: oui.
- *HIL: et **elle raconter** à policier et eum <elle les filles> [/] la fille va schnell.
- *INV: vite.
- *HIL: em elle va vite.
- *INV: mm.
- *HIL: euh eum et <un un garçon> [/] c'est un garçon.

- *INV: mm.
 *HIL: oui euh non non le blange
 dans la boulangeriee un monsieur
et la madame raconter à monsieur
et le monsieur raconter à policier.

*INV: ah ah.

D'autres verbes dans le récit d'Hilal sont à l'infinitif : *dire* par exemple, mais aussi *voir*. Ces occurrences nous incitent encore à penser que le verbe à l'infinitif est la forme de base en milieu scolaire. Les verbes irréguliers non thématiques apparaissent à plusieurs reprises dans leur forme infinitive. Inversement, quand une forme invariable est répétée dans tous les contextes, il s'agit de la forme infinitive.

(6)

- *HIL: arrête euh le garçon et le policier et la garçon va et la madame
 quand sehen.
 *INV: elle voit.
 *HIL: quand <elle voit> [/] elle voit **et dire** non c'est n'est pas le garçon
 c'est la fille.
 *INV: ah ah.
 *HIL: garçon non le mousieur doul boulangerie va vite **et dire** à la
 policier ce n'est pas le garçon c'est <le fille> [//] la fille.
 *INV: ah ah.

Cependant on observe aussi dans le récit de HIL un phénomène qui rend compte de l'enjeu de la désambiguïisation des formes en [e]. Ci-dessous, comme dans le récit de Ben ci-dessus l'apprenant reprend une forme en [e] donnée par l'interlocuteur natif. Le natif à ce moment-là répond à la demande de lexique de l'apprenant. Celui-ci a entamé un énoncé mais il ne sait pas comment on dit *rauchen* en français. Il interroge son interlocuteur qui répond, en avançant comme il le fait souvent dans les autres récits, la forme infinitive du verbe régulier *fumer*. Dès lors, l'apprenant reprend sa phrase en avançant une traduction de l'auxiliaire allemand *le garçon a* et en ajoutant la forme en [e] ce qui donne *il a fum[e]*. L'observation du processus de production de cet énoncé nous incite fortement à maintenir la forme en [e]. Même si l'énoncé en français parlé est tout à fait approprié au contexte du fait que le participe passé du verbe *fumer* a la même réalisation phonétique que l'infinitif, rien ne nous autorise à penser que l'apprenant a acquis le passé composé. Ici on peut dire que la forme utilisée *est Aux + V[e]*, V[e] étant la forme donnée par le natif. Ce passage montre donc qu'une forme composée a émergé spontanément dans la production de l'apprenant avant même que cette forme n'ait été enseignée.

(7)

- *HIL: oui <et il> [/] et <i em> [//] le policier téléphone et em
 inzwischen@s inzwischen@s?
 *INV: entretemps.
 *HIL: **entretemps le garçon hat em rauchen@s?**
 *INV: **fumer.**
 *HIL: **la garçon a fumi[e] fum[e].**
 *INV: hm hm.
 *HIL: et <le poli e> [/] <le voiture> [//] la voiture.
 *INV: ouais ouais ouais.
 *HIL: le voiture arrive.

Il apparaît dans le récit de BER ci-dessous que le plus important pour l'apprenant, une fois que le décor est planté, est de mentionner les situations successives dans lesquelles sont impliqués les personnages. Quand l'apprenant réfère à une série d'actions dont le caractère sériel est souligné par les adverbes temporels, l'occurrence d'une forme en [e] peut être considérée comme une forme de base. Seulement, étant donné les formes verbales auxquelles l'apprenant a été exposé, nous en déduisons que cette forme fonctionnellement de base, ne visant pas autrement dit autre chose qu'à dénommer l'action, est la forme infinitive.

(8)

*BER: <l'ho> [/] l'homme sè [=? Say, c'est]

je vole le pain

et le police euh bringen@s euh (souffle)

was@s war@s nehmen@s euh ?

*INV: le police prend.

*BER: prend l'homme

et après le madame sè [=? say, c'est]

non c'est euhm le nomme non vole pas le pain le madame vole le pain euh

après le police prend le madame

puis l'homme **manger** au restaurant

il nen <was@s heisst@s bezahlen@s> [/] ah was@s heisst@s bezahlen@s

et il n'en **payer** le gold@s essen@s.

*INV: euh le repas.

*BER: après le police arrive et prend le l'homme.

Année 2, groupe B

En année 2, dans le groupe qui a été exposé au passé composé, on observe également des formes en [e] mais aussi des formes verbales *Aux + V[i]* ou *V[i]* correspondant formellement aux participes passé de certains verbes irréguliers du français parlé, ce qui nous incite à penser que, parmi les formes possibles se cachant derrière les formes en [e], il pouvait se trouver dans ces récits d'apprenants des participes passé, comme c'est le cas ci-dessous. Il apparaît ainsi que l'auxiliaire dans les lectures d'apprenant n'est pas nécessairement accolé au participe passé. Afin de savoir si la forme du verbe en [e] peut être désambiguïsée en un participe ou en un infinitif, le contexte extralinguistique est un précieux indice. Il détermine si l'apprenant fait référence à une situation qui s'inscrit dans la suite des situations précédemment mentionnées ou s'il réfère à une situation précédente. Dans l'exemple ci-dessous l'occurrence de *vol[e]* dans l'énoncé *parce qu'il vol[e]* le pain indique que l'apprenant réfère à une situation précédant la situation mentionnée auparavant *l'agent de police emmène l'homme* (= l'agent de police emmène l'homme). Nous considérons qu'il s'agit d'un participe passé que nous comptons dans la rubrique « verbe non fléchi » mais que nous distinguons dans les analyses qualitatives de la forme infinitive. L'occurrence de *la fille a pris le pain* indique que l'apprenant possède dans son répertoire actif des formes verbales plus complexes que celles mobilisées par des apprenants moins avancés, ce qui ne se traduit pas nécessairement par un emploi adapté en contexte mais témoigne d'une évolution qu'il convient de souligner.

(9)

*NOR: ## ach@s so@s l'agent de police emà lom

parc+que parc+qu' il **voler** le pain.

*INV: ah ah d'accord ouias et après qu'est ce qu'il fait qu'est ce qu'il fait lom?

*NOR: **la fille a pris** le pain
après lom et la fille. bousculent.

Dans le récit de l'apprenant NOR, la forme [se], dont il a été question auparavant, est employée, en (9), avec un auxiliaire dans un contexte où elle prend le sens de dire. Nous maintenons ici l'orthographe intermédiaire [se].

(10)

*NOR: et lom erkannte@s?

*INV: ah il le reconnaît.

*NOR: lom rekeune la femme **et il a [se]** [:=? say] ne me reconnaissez vous
pas je suis lom avec le pain.

%sit: lit

Le cas d'Ant est intéressant. Cette apprenante produit un récit qui témoigne également d'une partie du processus en temps réel de production puisqu'elle emploie ci-dessous une forme composée *Aux + V* et hésite quant à la forme verbale pour finalement s'arrêter sur la forme *prendre*. Nous avons maintenu l'orthographe en *é* car il s'agit de toute évidence d'une forme distincte de l'infinitif et d'une surgénéralisation de la règle abordée explicitement en classe (déduite par les apprenants, énoncée par l'enseignante, lue dans le livre de grammaire et écrite dans le cahier de l'apprenant) selon laquelle le participe passé des verbes en -er se forme en -é. On assiste ici à un processus de flexion par dérivation. Le contexte d'apprentissage et le contexte extralinguistique constituent des indices convergents pour valider le participe passé.

(11)

*ANT: des gens <dit tout> [//] ah ne@ [= non] dit <a le> [/] a le boulanger
que la fême a <porte> [//] <prendre> [//] **prendre** oder@s so@s le baguette
<et> [/] eum et la fff ne@s.

*INV: hmhm.

*ANT: moment@s <le boulanger dit> [//] le boulanger <ça dit> [/] ne@s doch@s.

*INV: oui oui oui.

*ANT: ça dit <a> [/] <a le> [//] a des agents de police
et les agents courent em à la fême.

Année 3

En année 3, on observe des phénomènes, déjà relevés, de reprise comme en (11) ci dessous ou de flexion à partir de la base infinitive du verbe lexical en (12). On observe également de nouveaux phénomènes liés au contexte d'apprentissage. D'une part, le passé composé a été enseigné et il est employé par plusieurs apprenants dans plusieurs contextes. D'autre part, l'imparfait a également été enseigné et il est employé par un apprenant seulement, comme c'était le cas pour le passé composé dans le groupe précédent.

(12)

*INV: les deux discutent.

*LAU: il dit euh
vous ne me reconnaissez pas je suis l' homme avec le pain ?

*INV: ah d'accord.

*LAU: et elle lui **reconnaître** pff.

*INV: oui.

(13)

*LAU: et euh dix+jours+après euh les deux retrouvent
et elle dit
j'ai une surprise pour toi
et les deux <aller et euh> [/] <va > [/] euh **aller** I don't know euh chez son magasin

L'apprenant DAN dans son récit de film ci-dessous emploie à plusieurs reprises des formes en [e] dont certaines sont précédées d'un auxiliaire. Elle dit également *une dame a vu ça*, ce qui témoigne de son usage du passé composé. Les contextes d'occurrence dans le récit montrent que ces formes verbales qui alternent avec des formes au présent peuvent être interprétées comme des références à une situation passée. *volé* fait référence au vol passé et *est arrivé* à la situation accomplie. C'est la raison pour laquelle nous considérons que cette forme en [e] se distingue de la forme non fléchie infinitive telle qu'elle est mobilisée dans les contextes analysés précédemment, c'est-à-dire comme une forme de base employée pour référer à la situation avant d'exprimer une forme fléchie ou comme seule mention du procès. Ici les occurrences des formes en [e] indiquent un contraste dont il est toutefois encore difficile de dire s'il est aspectuel ou temporel.

(14)

*DAN: elle a **volé** des croissants des brot@s [= du pain] <et et> [/]
et vend scheise@s [= merde] weglaugen@s [= s'enfuir]
elle ...
*INV: elle s'enfuit elle part.
*DAN: oui oui et une dame <a> [/] a vu ça
<et> [/] et dit <à la> [/] à la pâtisserie
<cete dame> [/] cette dame <elle a> [/] elle a porte <des> [/] des croissants des brota et ce nom **volé** a lui
et cette dame qu'elle a porte <des> [/] des croissants
elle <crash> [/] crash.
%sit: tape dans les mains
*DAN: avec un autre nomm
et elle tombe à la boden@s
<et euh> [/] et euh <le pâtissière> [/] le pâtissier appelle <la> [/] <la police> [/] le police ja@s
et eum le police <a euh arrivé> [/] **est arrivé**
<et> [/] et eum le nom helfen@s help@s.

Dans l'extrait ci-dessous du récit de Daniela, on rencontre un phénomène nouveau et rare, à savoir un processus de création lexicale à partir de l'anglais *to meet*. Il s'agit d'un phénomène régulièrement observé dans les productions des germanophones. La forme créée est en [e] et nous l'orthographions -er en pensant qu'il s'agit d'un infinitif, c'est-à-dire un verbe lexical qui réfère au procès mais ne porte pas d'information temporelle ou de personne.

(15)

*DAN: et euh il cette nom je crois charlie chaplin
et cette dame **se metter** à la voiture
et charlie chaplin dit
je ce nom qui **aidé** toi
et cette dame <a> [/] a lachelt@s.
*INV: souri.
*DAN: souri ja@s
et charlie chaplin vont discuter avec elle

et le voiture a un crash.

La forme en [e] suivante, *aidé*, est, à notre sens, un participe passé, sans que l'on puisse toute fois déterminer sa fonction sémantique. En effet, elle s'insère dans un énoncé rapporté, introduit par un verbe de dire, et fait référence à une situation passée de même que la forme verbale *a lächelt* qui s'ensuit. Dans le récit de Dan, ces formes verbales précédées ou pas d'un élément morphologique que l'on peut appeler sans ambiguïté un auxiliaire contrastent avec des formes au présent et s'insèrent dans un texte globalement ancré au présent. Si l'on s'en tient au contexte extralinguistique et en l'absence d'éléments lexicaux spécifiant, ce contraste exprimé par la forme composée ou la forme seule en [e] peut tout aussi bien être un contraste temporel, l'apprenant réfère à un événement passé, qu'aspectuel, l'apprenant fait référence à une situation passée sous l'angle de son résultat au moment présent. Les deux lectures de la forme en [e] sont possibles à ce stade du récit. Il est intéressant qu'en français ces deux fonctions, aspectuelles et temporelles, sont encodées dans une même forme, le passé composé.

Certains apprenants comme c'est le cas de Reg ci-dessous emploient plus de formes non fléchies que d'autres. Le fait que ces formes font référence à des événements en série suggèrent qu'elles n'expriment pas de contraste temporel ou aspectuel spécifique et qu'il s'agit de formes de base que nous identifions étant donné le contexte d'apprentissage à des infinitifs. L'occurrence parmi elles de *se mettre* nous conforte dans ce choix.

(16)

*REG: <et il> [/] <et il ah> [/] et il **aller** <au> [/] <en> [/] au magasin.

*INV: hm oui oui.

*REG: et <il> [/] il monge dans le magasin

et mais <il ne>[/] **il ne pas payer** <pour le> [/] pour le menu.

et la police euhm euhm aller aussi?

*INV: ouais ouaisouais.

*REG: <le> [/] le <dans> [/] dans ce euh <was heisst auf xx> [:=g comment on dit dans xx]?

*INV: dans la voirure.

*REG: dans le voiture de police le monsieur et la fille euhm **se mettre**.

*INV: ouais ouais.

L'infinitif non ambigu des verbes irréguliers est donc un indice pour opter pour la forme infinitive quand Vrég-[e].

Dans ce même groupe d'apprenants un seul apprenant, Eys, emploie l'imparfait. A ce stade, l'apprenant réalise différemment un [e] et un [ɛ]. Par ailleurs Eys emploie des formes en [e] précédées de l'auxiliaire, accordé avec le sujet de la proposition, comme on peut le constater ci-dessous. Par conséquent l'occurrence de *c'était* ne peut pas apparaître comme un passé composé défectif. Il s'agit à notre sens d'un imparfait, ce qui est une possibilité étant donné que l'apprenant y a été exposé. Cependant on s'aperçoit dans le contexte du récit que l'imparfait a une valeur d'irréel qui ne correspond pas à sa valeur cible. De la même façon, la première occurrence du passé composé entraîne un conflit sémantique avec les autres éléments de l'énoncé. Ces remarques montrent que la désambiguïsation formelle des formes doit s'accompagner ensuite d'une analyse de leur contenu sémantique qui ne correspond pas nécessairement à celui des formes telle qu'employées par des natifs.

(17)

- *EYS: oui il se sauve
et la femme et Charlie Chaplin **a été** ensemble pour+quek+temps
et <la femme dit non> [/]
charlie chaplin dit qu'il veut être d'avoir une maison ensemble.
*INV: ah d'accord.
*EYS: une maison ensemble
et euhm la femme dit
oui **c'était** vouaim c'était bien
et uehm <dix+jours+après> [/] dix+jours+après la femme venir chez Charlie Chaplin et dit
<j'ai a> [/] **j'ai trouvé** la maison euh
que nous **avons cherché**.

La prononciation de l'apprenant témoigne de réalisations distinctes [e] et [ɛ].

Année 4

En année 4, les formes en [e] sont beaucoup plus présentes dans les récits, la plupart des apprenants employant au moins un passé composé dans leur récit. A ce stade, les passé composé alternent parfois avec des imparfait. Toutefois des ambiguïtés subsistent, mais là aussi, elles peuvent être levées par l'observation attentive de leur contexte d'émergence. Ainsi ci-dessous, l'occurrence de volé dans le récit de Gregor dans un contexte de référence à une situation passée suggère que cela pourrait être une forme contrastive et non de base

(18)

- *INV: alors raconte moi le petit film s'il te plaît je t'écoute.
*GRE: ôfff euh c'est euh.
%sit: tape sur la table deux fois.
*GRE: Charlie Chaplin.
*INV: oui.
*GRE: euhm eu eu eum o <femme> [/] <fèem> [/] fème euh **volé** un pain.
*INV: d'accord.
*GRE: **a** (souligne) **volé** un pain.

A ce stade de l'apprentissage où les apprenants possèdent théoriquement plusieurs formes verbales pour un même verbe lexical et donc de possibilités de contrastes, ils font preuve en situation de production d'une liberté plus grande qu'auparavant dans la sélection et l'organisation des événements. C'est ce qu'on observe dans le récit de l'apprenant AID. Ce récit, étoffé, témoigne d'une alternance entre le passé et le présent. Ces contrastes au niveau morphologique sont accompagnés d'indices linguistiques contrastifs *mais, dans la même temps, dans la même situation, a dit que, dit qui*. Ainsi *parlé* qui apparaît avec un adverbe de simultanéité peut être considéré comme un PC défectif du fait que dans la même situation auparavant corrélait avec une forme composée *Aux + V-[e]*.

Concernant la dernière forme en [e], nous considérons qu'il s'agit aussi d'un participe parce qu'aucune forme infinitive irrégulière suggère qu'il pourrait s'agir d'une forme de base et d'autre part parce qu'une analyse détaillée des récits montre que la clôture d'une scène est exprimée de façon récurrente par un participe ou un passé composé chez les apprenants employant par ailleurs cette forme dans leur récit.

(19)

*AID: <c'est un l'histoire> [//] <c'est un> [//] c'est <en fille> [//] un fille
et elle a beaucoup de faim
et elle **a volé** <en> [//] un pain
et eum elle **a trouvé** monsieur charlie chaplin
et euh mais il a une femme
et il parle avec le boulangereur
mais boulangeur **é** dit
que la fille **a volé** le pain
dans+la+meme+temps l'agent de police **a gardé** le fille
et monsieur chaplin dit euhm
qui il euh il y a euh il **il a volé** le pain
et c'est ça pour quoi le police veut eum partir avec le monsieur
mais dans+la+meme+situation la femme qui euh **parlé** avec <le boulangeur> [//]le boulangeur
dit aussi
qu'il n'a pas le monsieur
c'est la fille
et monsieur chaplin peut partir
et le police **allé** avec la fille dans un voiture.

5 Conclusion

Notre démarche de désambiguïsation a consisté à observer des phénomènes spécifiques à chaque stade sachant que les apprenants ont été exposés à un input de plus en plus riche, contenant nouveaux verbes, ainsi que de nouvelles formes verbales de verbes déjà connus. Cet enrichissement de l'input étant susceptible de se traduire par l'emploi de nouvelles formes, semblables à la langue cible ou idiosyncrasiques, nous avons d'abord observé globalement si les récits témoignaient de phénomènes réguliers. Ensuite nous avons observé récit par récit où apparaissaient les formes en [e] et si des contraintes morphosyntaxiques ou sémantiques étaient susceptibles de nous renseigner sur ce phonème.

La principale difficulté de désambiguïsation des formes verbales en [e] dans les récits oraux d'apprenants germanophones porte sur celles qui apparaissent isolément, sans élément pré-morphologique, en position de verbe fléchi. La désambiguïsation a mis en évidence trois types de formes verbales en [e]. Premièrement il peut s'agir d'une forme de base, c'est-à-dire, étant donné le contexte scolaire, d'un infinitif. Deuxièmement, il peut aussi s'agir d'une forme contrastive à valeur temporelle ou aspectuelle que nous avons transcrite sous la forme d'un participe passé, sachant qu'elle survient dans des récits d'apprenants ayant été exposés au passé composé et en faisant usage dans leur récit. Enfin, la forme verbale en [e] peut-être une forme idiosyncrasique, auquel cas nous avons maintenu dans la transcription l'écriture phonétique.

On rencontre également dans les récits les plus avancés des occurrences de l'imparfait. Cependant, à ce stade du développement linguistique, il ne s'agit plus de formes verbales en [e] mais de formes en [ɛ] dénuées d'ambiguïté formelle.

Soulignons enfin que la distinction principale entre « forme de base » et « forme contrastive » a été élaborée à partir de ce corpus de récits de film oraux et en tenant compte du facteur externe de l'enseignement.

6 Références bibliographiques

- Blanche-Benveniste, C. (1997) *Approches de la langue parlée en français*. Paris : Éditions Ophrys.
- Bourdieu, P. (1994) *Raisons pratiques. Sur la théorie de l'action*. Paris : Editions du Seuil
- Granget, C. (2004) *L'acquisition par des apprenants germanophones de la référence au passé en français : un cas d'école*. Thèse de doctorat. Université de Paris III.
- Herschensohn, J. (2001) « Missing inflection in second language French : accidental infinitives and other verbal definites ». *Second Language Research* 17: 273-305.
- Klein, W. & Perdue, C. (1997) « The Basic Variety (or Couldn't natural languages be much more simpler ?) ». *Second Language Research* 13 : 301-347.
- Lardièrre, D. (1998) « Dissociating syntax from morphology in a divergent L2-end state grammar ». *Second Language Research* 14 : 359-375.
- Perdue, C., Benazzo S. & Giuliano P. (2002) « When finiteness gets marked : the relation between morphosyntactic development and use of scopal items in adult language acquisition. *Linguistics* 40(4) : 849-890.
- Schlyter, S. (2003) « Development of verb morphology and finiteness in children and adult acquiring French. In Dimroth C. StarrenM. *Information structure and the Dynamics of Language Acquisition*, pp. 15-44. Amsterdam/Philadelphia : John Benjamins.
- Starren, M. (2001) *The second time*. Utrecht: LOT.

Formes ambiguës en *-E* : input et catégories sémantiques

Anita Thomas

Institut d'études romanes
Université de Lund
anita.thomas@rom.lu.se

1 Introduction

Cette étude s'inscrit dans la recherche sur l'acquisition du français langue étrangère par des adultes suédophones et plus précisément du développement et de l'utilisation de la morphologie verbale dans l'interlangue des apprenants. Elle est avant tout centrée sur la production des verbes en *-er* (*parler, habiter* ...) et plus particulièrement sur l'utilisation des formes du type *je parlE*. Ce genre de formes pourrait correspondre à différentes formes verbales dans l'interlangue des apprenants : un infinitif *-er*, un participe passé *-é*, une 2^e personne du pluriel *-ez*, ou encore une forme de l'imparfait *-ais/ait/aient* (voir aussi Granget, ce volume). Les formes du type *parlE* comprennent donc à la fois les sons finaux /e/ et /ɛ/, la distinction entre ceux-ci n'étant pas toujours faite par des apprenants en début d'acquisition. La graphie *-E* sera utilisée ici pour garder l'ambiguïté de la fonction visée.

L'objectif de cette étude est d'une part de voir dans quels contextes ce genre de formes en *-E* apparaissent dans l'interlangue des apprenants et d'autre part de voir si l'input auquel les apprenants sont exposés favoriserait l'émergence de telles formes.

Les sujets de cette étude sont des apprenants débutants du français, ce qui permet de faire l'hypothèse qu'ils se trouvent à un stade dans lequel ils soumettent le système verbal de la langue française à un processus de simplification. Ce processus pourrait être basé sur une utilisation pragmatique de la langue et influencé par la distribution des formes verbales trouvées dans l'input.

2 Considérations théoriques

Bartning et Schlyter (2004) proposent que les formes du type *je parlE* correspondent à une forme non-finie (longue) utilisée en alternance avec des formes finies (courtes) du type *je parle*, et que ces formes soient utilisées de manière invariable dans différents contextes en début d'acquisition. A partir des données de leurs corpus longitudinaux, elles ont pu ensuite observer "[qu'i]l y a un développement graduel vers une distinction de telle sorte que les 'formes finies courtes' sont utilisées dans les contextes finis et les formes non-finies dans les contextes non-finis" pour finalement arriver à un stade où "les formes sont utilisées en accord avec la langue cible." (Bartning et Schlyter 2004 : 285). La question qui se pose alors est de savoir pourquoi les

apprenants développent la morphologie verbale vers la langue cible.

Les tenants d'une approche plus pragmatique de l'acquisition d'une langue proposent de faire le lien entre le développement de la morphologie verbale et la structuration des énoncés de l'apprenant (Klein et Perdue 1992). Selon eux, la fonction du verbe exprimé par l'apprenant est avant tout de porter un contenu lexical, le contexte temporel étant marqué par d'autres moyens, par exemple par des adverbes de temps, des prépositions ou encore par l'organisation discursive du récit. Voici un exemple tiré du corpus de cette étude :

*CAR: autrefois euh je travaille à Paris.

Selon cette approche, les stades proposés sont les suivants (Perdue, Benazzo et Guiliano 2002 :857¹ :

Stade NUO, *Nominal Utterance Organisation*, les phrases sont organisées sans verbe exprimé ;

Stade IUO, *Non Finite Utterance Organisation*, l'organisation des phrases est centrée sur un verbe non-fini ;

Stade FUO, *Finite Utterance Organisation*, le verbe est conjugué.

Cette vision du développement de la morphologie verbale met en valeur le fait que l'apparition du verbe au stade IUO permet à l'apprenant d'exprimer tout un champ sémantico-temporel, même s'il ne marque pas encore la distinction morphologique *fini* vs *non-fini*. Le passage progressif à un stade dans lequel le verbe est produit dans des formes finies (FUO) est caractérisé par le besoin de l'apprenant d'exprimer les différentes relations temporelles entre le temps des événements et celui de la narration de manière plus nuancée et plus indépendamment de la structure syntaxique (Perdue, Benazzo et Guiliano 2002 :861 ; Klein 1989 :145ss).

La sur-utilisation de certaines formes – par exemple de la forme non-finie – est donc possible parce que le contexte temporel est exprimé par d'autres moyens. Cette approche pragmatique de l'interlangue des apprenants est intéressante parce qu'elle met en valeur le fait que la *forme* verbale puisse jouer un rôle mineur par rapport au contexte entier dans lequel elle est exprimée, l'élément déterminant pour la communication étant les données lexicales du verbe.

Influence de l'input

Dans une approche centrée sur l'**usage de la langue** (*usage-based approach*), l'acquisition d'une langue est étroitement liée aux probabilités d'occurrence mettant en relation la répartition forme / fonction (N. Ellis 2002).

Language learning is the associative learning of representations that **reflect the probabilities of occurrence of form-function mappings**. Frequency is thus a key determinant of acquisition because “rules” of language, at all levels of analysis [...] are **structural regularities** that emerge from learners' lifetime **analysis of the distributional characteristics of the language input**. Learners have to *figure* language out.” (Ellis 2002 : 144 – c'est moi qui souligne)

¹ Pour une description plus détaillée de ces stades voir par exemple Klein (1989 :124ss) ou Klein et Perdue (1992 :302).

Ainsi, même si les verbes en –er se forment à partir de règles régulières, celles-ci ne deviennent automatiques qu’après que l’apprenant ait eu l’occasion d’analyser le fonctionnement régulier de ce système. Selon Ellis, il y a donc des effets de fréquence non seulement sur les verbes irréguliers, mais aussi sur les verbes réguliers, du moins en début d’acquisition (Ellis 2002 : 154)². Selon cette approche, la production de formes verbales du type *je parlE* serait donc liée à la fréquence et à la distribution de cette forme dans la langue à laquelle les apprenants sont exposés.

Principes opératoires

R. Andersen a proposé un certain nombre de principes (*Operating principles*, cf. R. Ellis 1994 : 380-381 ; Mitchell & Myles 1998 : 75-77 ou *Acquisitional Principles* Andersen 1993 : 327ss) permettant de rendre compte du lien entre l’input et la production des apprenants. Ces principes représentent un certain nombre de mécanismes cognitifs sous-jacents permettant aux apprenants de résoudre, dans un premier temps, les difficultés liées à l’absence de correspondance univoque entre les formes et leurs diverses fonctions.

En début d’acquisition, deux principes semblent particulièrement actifs. Le premier, appelé *one-to-one principle*, pose l’hypothèse que les apprenants commencent par simplifier le système d’utilisation des verbes en attribuant une forme invariable claire par sens. Pour cela les apprenants choisiront une forme saillante, ayant un sens prototypique jusqu’à ce qu’ils réussissent à percevoir une nouvelle nuance. Des formes du type *je parlE* pourraient avoir un tel sens, notamment dans des contextes du passé, le passé composé comme l’imparfait se terminant le plus souvent par un son –E.

Le deuxième principe est celui des tendances distributionnelles - *distributional bias* - qui complète le premier. Par *tendances distributionnelles*, on entend la répartition de l’input entre un certain nombre de catégories (existantes ou émergentes). Il s’agit donc de la fréquence avec laquelle une certaine fonction est réalisée par une certaine forme. Le principe se définit de la manière suivante :

Si les formes X et Y peuvent apparaître dans les deux contextes A et B, mais que les tendances distributionnelles font percevoir que X apparaît surtout dans A et Y dans B, alors les apprenants vont restreindre l’utilisation de X pour le contexte A et de Y pour le contexte B. (Mitchell & Myles 1998 : 76)

² Pinker (1991, 1998) ainsi que d’autres tenants du dual system proposent deux mécanismes cognitifs distincts pour le traitement des verbes réguliers et irréguliers. Selon lui, seuls les verbes irréguliers seraient soumis à des effets de fréquence dans l’input.

Tableau 1. *Principe des tendances distributionnelles*

	Contexte A	Contexte B
	<u>forme X</u> et forme Y	forme X et <u>forme Y</u>
exemples locuteurs natifs	<u>elle habite</u> / elle a habité	il tombe / <u>il est tombé</u>
exemples interlangue	elle a habite	il tombE

Légende : la forme soulignée est la forme choisie selon le principe des tendances distributionnelles.

Pour pouvoir observer de telles tendances, il est nécessaire de classer les données dans des contextes du type *A* et *B* du tableau 1. Pour cela j'utiliserai d'une part la référence au présent et au passé et d'autre part les catégories sémantiques. Ces catégories permettent de classer les verbes selon leur aspect inhérent, à savoir selon les caractéristiques temporelles inhérentes au contenu lexical du verbe. Un tel classement pourrait être pertinent du moment où l'on considère que les apprenants débutants s'appuient sur un usage très pragmatique de la langue. Un tel cadre devrait aussi permettre, dans une étude ultérieure, d'observer un développement spécifique ou non de la morphologie verbale dans ces catégories.

Les catégories sémantiques

Les catégories sémantiques, telles qu'elles sont utilisées aujourd'hui dans le cadre de la recherche sur l'acquisition des langues, ont été systématisées par le philosophe Vendler (1957). Les noms donnés à ces catégories varient d'un auteur à l'autre, selon les traits sémantiques qu'il veut mettre en avant. J'utiliserai ici les mêmes termes que R. Andersen (1993) :

Auteur	Noms des catégories de verbes			
R. Andersen (1993)	Etat	Activité	Téliques	Ponctuels
Vendler (1957)	Etat	Activité	Accomplissement	Achèvement
Platzack (1979)	Etat	Processus illimités	Processus limités	Événements ponctuels

Pour pouvoir classer les formes verbales dans les différentes catégories sémantiques, il faut considérer un certain nombre de traits (Cf. Platzack 1979 ou Vet 1980 pour une présentation complète de ces traits). Voici, à titre d'exemple la présentation de trois traits sémantiques : le changement, la durée et la limitation ainsi que la "divisibilité".

Tableau 2. Répartition de trois traits sémantiques

Traits	Etat	Activité	Téliques	Ponctuels
changement	- changement statique	+ changement dynamique	+ changement dynamique	+ changement dynamique
durée		+ durée	+ durée	- durée
limitation		- limitation pendant	+ limitation en	+ limitation à
“divisibilité”		+ divisible faire des gâteaux	- divisible faire un gâteau	

La catégorie des verbes d’état est caractérisée par l’absence de changement, contrairement aux trois autres catégories et ils expriment donc une forme de permanence.

J’habite à Lund. J’aime la France.

Les traits de durée et de limitation permettent de faire la différence entre les verbes d’activité, téliques et ponctuels.

- A : Il a fait un doctorat pendant 2 ans. (Le doctorat n’est pas terminé.)
T : Il a fait un doctorat en 2 ans. (Le doctorat est terminé.)
P : *Il a fait un doctorat à 15h.
- A : *Il est tombé de l’échelle pendant trois heures.
T : *Il est tombé de l’échelle en trois heures.
P : Il est tombé de l’échelle à 15h.

Finalement le trait de “divisibilité” permet de faire la différence entre les verbes d’activité et les verbes téliques. Il s’agit de voir si l’objet a une quantité spécifique ou non. Pour les verbes d’activité, la quantité n’est pas spécifiée, l’objet est illimité, alors que pour les verbes téliques on a un *produit*, une quantité spécifique, un objet limité :

- A : Je lis des livres. Le dimanche je fais des gâteaux.
T : J’ai lu trois livres. J’ai fait un gâteau.

Ainsi, lorsque les verbes *faire* ou *lire* sont suivis de l’article partitif, ils correspondent à des verbes d’activité et lorsqu’ils sont suivis d’une quantité spécifique, ils correspondent à des verbes téliques.

Comme leur nom l’indique, les verbes ponctuels décrivent une situation ou un événement momentanés. Il n’y a donc pas de durée et le cadre est limité :

- P : Le livre est tombé.

Comme on l’aura peut-être déjà senti à partir des exemples, chaque catégorie sémantique est exprimée de préférence par certains temps des verbes (prototypique). Ainsi le présent est typiquement le temps des verbes d’état et l’imparfait pour les verbes d’état au passé. Les verbes téliques et ponctuels sont le plus souvent au passé et comme l’action est limitée dans le temps, ils sont le plus souvent exprimés au passé composé (Bardovi-Harlig 2000 : 425 ; Salaberry 1998).

3 Objectifs et méthode

L'objectif de cette étude est de voir dans quels contextes des formes du type *je parlE* sont produites par les apprenants. Pour cela j'utiliserai les *catégories sémantiques* comme cadre d'analyse de base, ce qui signifie que toutes les formes verbales seront classées selon ces catégories. Un tel classement pose évidemment de nombreuses difficultés, en particulier lorsqu'il s'agit d'énoncés produits par les apprenants. Comment classer le verbe *regarder* (activité) lorsqu'il est utilisé dans le sens de *voir* (télique/ponctuel) ? Cette étude se base sur un classement qui est donc relativement fragile et qui doit être considéré avec une certaine prudence, même s'il peut montrer un certain nombre de tendances.

Pour pouvoir mettre en relation une possible influence de l'input et la production (pragmatique et guidée par des principes menant à une simplification du système verbal) des apprenants, je commencerai par une analyse de la production verbale de l'intervieweuse du corpus, en particulier au niveau de la répartition de temps des verbes entre les différentes catégories sémantiques. Cette analyse devrait permettre de mettre en valeur des différences et des similarités entre les différentes catégories et des contextes plus ou moins propices à l'apparition de formes du type *je parlE*.

La définition de ce qu'est un *input* est une question débattue (Carroll 2001 : 33, note 8). Le terme d'*input* est emprunté à la littérature sur l'intelligence artificielle, donc au domaine de l'informatique (Carroll 2001 : 8). Mais dans un environnement biologique et notamment dans le cadre de l'acquisition d'une L2, 'ce que les apprenants entendent / perçoivent / traitent de la langue cible' est nettement plus complexe à définir. L'objectif de cet article n'est pas d'entrer dans la discussion sur cette question, c'est pourquoi l'analyse de 'l'input', sera ici limitée à une analyse des données produites par l'intervieweuse du corpus. Cet input n'est qu'une petite partie de tout ce que les apprenants entendent / perçoivent / traitent de la langue cible. En effet tous ont accès à un usage natif de la langue lors de leurs séjours en pays francophone ou lorsqu'ils regardent des films. Tous ont également accès à la langue utilisée par leurs collègues apprenants. Les raisons pour lesquelles j'ai choisi ici d'analyser les données de l'intervieweuse sont les suivantes :

- Premièrement l'intervieweuse étant aussi l'enseignante des apprenants, le contact avec sa manière de parler est régulier, dans la durée et probablement chargée d'un certain lien affectif. En ce sens 'l'input' apporté par l'intervieweuse pourrait avoir un certain poids.
- Une deuxième raison pour l'analyse de la production verbale de l'intervieweuse est que sa manière de converser lors des interviews reflète assez bien sa manière de converser en classe. Il est ainsi possible de calculer la fréquence de certaines prédications verbales et par là même la possible fréquence des réponses entendues en classe, qui renforcent ou atténuent la fréquence de certaines données de l'input.

L'analyse des données dans cette étude se fera en trois étapes. Tout d'abord, comme je l'ai déjà mentionné, je présenterai les données de la production des verbes de l'intervieweuse, puis la production des apprenants avec l'objectif de voir si leur production de formes du type *je parlE* est lié à l'expression d'un contexte particulier ou propre à un certain contexte sémantique. Finalement je présenterai une comparaison des formes les plus fréquentes de verbes en *-er* trouvées chez les apprenants et l'intervieweuse. La dernière section sera consacrée à une discussion.

4 Corpus

Le corpus est constitué de 20 apprenants adultes suivant des cours du soir de français avec moi. Ils ont été interviewés pendant environ 10 à 15 minutes. Quelques élèves ont été enregistrés plusieurs fois (une fois par semestre) durant 20 à 30 minutes. Environ la moitié des apprenants ont appris le français à l'école dans leur jeunesse, l'autre moitié a commencé dans le cadre de ces cours. Les plus débutants ont un semestre de français derrière eux (un peu plus 24h de cours) et les plus avancés quatre semestres (un peu plus de 96 heures). Tous les apprenants ont le suédois comme L1, sauf une qui a le norvégien. Ils ont tous l'anglais comme L2 et la plupart connaissent également l'allemand.

Lors de la première série d'interviews, les apprenants ont su à l'avance que je leur demanderais de se présenter et de parler de leurs loisirs. Ils ont tous déjà eu ce genre de conversations en classe. Avec la plupart nous avons aussi parlé de la France et du français. À la fin de chaque interview, je leur ai demandé de lire un texte et parfois la conversation a repris après la lecture. Lors des deuxièmes et troisièmes interviews le cadre de discussion était plus libre. Les apprenants ayant plus de difficultés ont notamment pu choisir de parler à partir d'une image représentant une personne.

5 Résultats

5.1 Production de l'intervieweuse

Les données pour la production verbale de l'intervieweuse correspondent aux énoncés produits dans seulement sept conversations du corpus, qui avaient été analysées dans le cadre d'une étude précédente (Thomas 2004). Ces conversations sont néanmoins représentatives de la production de l'intervieweuse dans la plupart des conversations enregistrées.

Le tableau 3 ci-dessous, présente la répartition des formes verbales au présent, au passé et à l'infinitif dans les différentes catégories sémantiques. Il montre donc quel est le temps le plus utilisé dans chaque catégorie avec des exemples de quelques verbes fréquents. Ces données comprennent les verbes en –er ainsi que les verbes irréguliers (y compris les verbes *être*, *avoir* et les verbes modaux comme *pouvoir* et *vouloir*). Les expressions comme *c'est*, *est-ce que*, *il y a* ainsi que les verbes *être*, *avoir* et *aller* dans leur fonction d'auxiliaires ont été exclus du calcul. L'objectif est de mettre en évidence la distribution des temps des verbes, y compris de l'infinitif, entre les catégories sémantiques.³

³ La catégorie passé comprend l'imparfait et le passé composé. Au total seulement 8 formes de l'imparfait ont été produites, 6 dans les verbes d'état et une chaque fois pour les verbes d'activité et téléiques. Des 12 formes du passé des verbes d'état, 6 sont des imparfaits.

Tableau 3. Répartition des temps des verbes, pour l'intervieweuse, dans les catégories sémantiques

Etat	Activité	Téliques	Ponctuels
présent (88.5%)	présent (62.3%) (passé 15.1% infinitif 22.6%)	présent (38.9%) (passé 30.6% infinitif 30.6%)	infinitif (47.1%) (passé 41.2%)
être, avoir, connaître, pouvoir, vouloir, aimer, habiter	écouter, étudier, jouer, manger, parler, regarder, travailler, utiliser, voyager, faire, lire	aller, apprendre, comprendre, lire, venir, voir, déménager, gagner	commencer, demander, finir

Les verbes d'état ne comprennent que peu de verbes réguliers en *-er* et de nombreux verbes irréguliers, contrairement aux verbes d'activité qui contiennent un grand nombre de verbes en *-er* et quelques verbes irréguliers. Les verbes téliques contiennent environ un quart de verbes en *-er* et le verbe *aller* correspond également à un quart des occurrences. Pour les verbes ponctuels, qui sont très peu nombreux, les verbes en *-er* correspondent à 65 % des occurrences.

Le temps le plus utilisé par l'intervieweuse est nettement le présent, qui est le temps le plus fréquent pour toutes les catégories excepté celle des verbes ponctuels. Pour les verbes d'état, le pourcentage de verbes au présent est particulièrement haut avec 88.5 % et c'est aussi la catégorie qui contient le plus de formes du présent (53 % des formes verbales au présent sont des verbes d'état). Pour les verbes téliques, on peut constater que les formes verbales se répartissent entre le présent, le passé et l'infinitif à environ un tiers pour chacun.

Un calcul de la distribution des formes verbales *entre* les différentes catégories, permet de voir que la catégorie des verbes d'activité contient le plus grand nombre d'occurrences à l'infinitif (49 % des formes à l'infinitif sont des verbes d'activité) ainsi que du passé (37 % des occurrences du passé sont des verbes d'activité). La catégorie des verbes téliques contient un tiers des formes produites au passé (34 %) et à l'infinitif (30 %).

On peut donc penser que les verbes d'activité et les verbes téliques comprennent un grand nombre de formes verbales se terminant par un son *-E* pour les verbes de la catégorie *-er* étant donné que le passé (composé) et l'infinitif sont fréquents. Un calcul du pourcentage de finales en *-E* pour les verbes en *-er* (sans *aller*) chez l'intervieweuse montre effectivement un tel résultat comme on peut le voir dans le tableau 4 ci-dessous⁴ :

Tableau 4. Répartition des finales en *-E* des verbes en *-er* chez l'intervieweuse

Etat	Activité	Téliques	Ponctuels
58.7 % (27/46)	80.9 % (93/115)	81.8 % (18/22)	90.9 % (10/11)

Légende : les données entre parenthèses correspondent au nombre de formes se terminant par le son *-E* sur la somme des occurrences des formes de verbes en *-er*.

Bien que la catégorie des verbes d'état se distingue des trois autres, on peut tout de même constater que le pourcentage de formes verbales se terminant par un son en *-E* est élevé. Étant

⁴ Chez l'intervieweuse les formes en *-E* correspondent à tous les sons [e] et [ɛ], donc y compris la deuxième pluriel.

donné que l'intervieweuse vouvoie ses élèves, un grand nombre de contextes du présent sont également réalisés par une forme se terminant en *-E* (voir aussi le tableau 7 plus loin).

La production de l'intervieweuse présente donc une utilisation des verbes qui pourrait se répartir entre les verbes d'état vs les autres. Les verbes d'état sont produits au présent et contiennent moins de formes qui se terminent par un son en *-E* que les autres catégories. Pour les autres catégories, les verbes sont produits à différents temps et se terminent fréquemment par un son *-E*.

A la différence des verbes d'activité et téléiques, les verbes ponctuels ne sont que rarement produits au présent. Finalement on pourrait aussi voir une distinction entre les verbes d'activité et les autres en considérant le nombre de verbes en *-er* différents, nombreux pour les verbes d'activité et moins nombreux dans les autres catégories.

5.2 La production des apprenants

L'objectif de cette partie est de voir la distribution des formes ambiguës du type *je parlE* chez les apprenants. L'analyse se limitera donc à la production des verbes en *-er* dans les contextes du présent et du passé dans les différentes catégories sémantiques pour voir dans quels contextes les formes en *-E* apparaissent.

Au total, les contextes du présent et du passé ont été marqués 83 fois par une forme du type *je parlE*. Ces formes se répartissent de manière relativement équitable entre les contextes du présent ($47/83 = 57\%$) et du passé ($36/83 = 43\%$). Il est alors intéressant de voir comment ces formes sont produites dans les différentes catégories sémantiques et si ce sont les mêmes contextes qui sont marqués par de telles formes dans toutes les catégories sémantiques.

Le tableau 5 ci-dessous présente la répartition des formes du type *je parlE*. Le pourcentage indique le nombre de verbes produits à une forme en *-E* par rapport au total des formes verbales produites dans les contextes du présent respectivement du passé⁵.

Tableau 5. *Pourcentage de formes du type je parlE dans les contextes du présent et du passé chez les apprenants*

Contexte	Présent	Passé	Total 2
Etat	2 % (3/124)	- (0/8)	2 % (3/132)
Activité	32 % (38/117)	34 % (23/67)	33 % (61/184)
Téliques	33 % (4/12)	70 % (7/10)	50 % (11/22)
Ponctuels	29 % (2/7)	38 % (6/16)	35 % (8/23)
Total 1	18 % (47/260)	36 % (36/101)	23 % (83/361)

Légende : Le total 1 est la somme des formes en *-E* sur le total des contextes du présent, respectivement du passé. Le total 2 est la somme des formes en *-E* sur le total des contextes du présent et du passé de la catégorie concernée.

⁵ A noter que les formes comme vous marchez dans un contexte du présent ont été comptées comme correctes, par contre dans un contexte du passé vous marchE entre dans le calcul des verbes en *-E*.

Ce tableau permet de voir que les contextes du présent et du passé sont exprimés par des formes en *-E* dans 23 % des cas. Sur tous les contextes du présent, 18 % sont exprimés par une telle forme, pour les contextes du passé, il s'agit de 36 %. Il est donc plus fréquent qu'un contexte du passé soit marqué par une forme en *-E* qu'un contexte du présent.

Il peut être intéressant de noter que les données de ce corpus diffèrent légèrement de celles de Bartning et Schlyter (ce volume). En effet, le pourcentage de formes finies s'élève ici à 82 %, ce qui est un peu plus élevé que leurs données pour les débutants. Cette différence pourrait être due à l'absence des occurrences des verbes irréguliers dans les calculs du tableau 5.

Comme on peut le voir dans le tableau 5, les verbes d'état sont très rarement réalisés par une forme en *-E*. Il s'agit de trois formes du verbe *habiter* produits dans un contexte du présent, par un même apprenant :

- (1) *LUC: et **je habitE** à www.
- (2) *LUC: euh un fille et une euh <elle vont> [/] **elle** euh **habitE** dans www.
- (3) *LUC: oui oui , **un écrivain** euh **habitE** en www.

Pour les verbes d'activité, les formes du type *je parlE* correspondent à un tiers des verbes produits : 32 % des formes produites dans un contexte du présent et 34 % des contextes du passé. Le pourcentage de formes en *-E* est donc très semblable pour les deux contextes. C'est la catégorie qui contient le plus de ces formes (73 % des formes en *-E* sont des verbes d'activité), réparties sur 16 verbes différents, mais c'est aussi la catégorie qui contient le plus d'occurrences. Voici quelques exemples :

- (4) *Contexte du présent*
 *INT: mais qu' est ce que vous faites alors maintenant que vous êtes retraité?
 *BER1: euh [/] euh **je travailleE** à la euh jardin .
- (5) *Contexte de l'imparfait (SAN parle de son défunt mari)*
 *SAN: et **il travailleE** beaucoup euh dans Amérique , Africa , Japon .
- (6) *Contexte du passé composé*
 *CAR: euh **il** euh **mangE**, mangE à le restaurant .

Pour les verbes téliques, 50 % des verbes sont exprimés par une forme en *-E*, 70 % des contextes du passé et 33 % des contextes du présent. Contrairement au verbes d'activité, ces formes ne se répartissent pas équitablement entre les deux contextes temporels, mais servent avant tout à l'expression du passé. Voici quelques exemples :

- (7) *Contexte du présent*
 *SEB: un peu un peu euh euh euh quand **je achetE** à la kiosque .
- (8) *Contexte du passé composé*
 *PIE: oui oui oui **nous visitE** un un vin farm un vin vigne vigne vin .

Pour les verbes ponctuels, les formes en *-E* sont un peu plus fréquentes dans les contextes du passé que du présent, mais il n'y a que très peu d'occurrences (8 formes pour 5 verbes différents).

- (9) *Contexte du présent*
 *PIE: **commencE** , football **commencE** .
- (10) *Contexte du passé composé*
 *JAC: +, **je <rencontre> [/] rencontrE** ma femme .

Les formes en *-E* sont donc avant tout produites dans les catégories des verbes d'activité, téliques

et ponctuels. Pour les verbes d'activité, on trouve un pourcentage semblable de ces formes dans les contextes du présent et du passé, alors qu'elles servent plus à marquer le passé pour les verbes téléquies et ponctuels.

5.2.1 Exemples d'utilisation de verbes en –er dans leurs formes longues et courtes

Le tableau 6 ci-dessous présente quelques verbes en –er fréquemment utilisés par les apprenants tous contextes confondus⁶, donc y compris les contextes de l'infinitif et du futur proche. Par forme longue j'entends les formes du type *parlE* (y compris les formes *vous parlez*⁷) et par forme courte celles du type *parle*. Ces formes sont classées dans les deux colonnes indépendamment du contexte dans lequel elles ont été produites. Chaque colonne contient donc des formes correctes et incorrectes. Ce tableau permet de mettre en évidence une possible relation entre la forme la plus fréquemment utilisée et les différentes catégories sémantiques, abstraction faite du temps des verbes.

Tableau 6. *Présentation de quelques verbes en –er dans leurs formes longues et courtes chez les apprenants*

	verbe	nombre de formes courtes	nombre de formes longues	Σ	forme la plus fréquente	%
Etat	aimer	37	0	3	aime	10
				7		0
	habiter	49	9	5	habite	84
				8		
	penser	13	0	1	pense	10
				3		0
	préférer	8	0	8	préfère	10
						0
Activité	étudier	0	20	2	étudiE	10
				0		0
	jouer	2	13	1	jouE	87
				5		
	parler	19	14	3	parle	58
				3		
	regarder	5	11	1	regardE	69
				6		
	travailler	44	18	6	travaille	71
				2		
	voyager	6	10	1	voyagE	63
				6		
	acheter	1	5	6	achetE	83
	commencer	2	5	7	commencE	71
Téliques/Ponctuels	rencontrer	1	3	4	rencontrE	75
	visiter	0	11	1	visitE	10
				1		0

Le verbe *habiter* est le seul verbe d'état en –er qui soit produit dans une forme longue comme on l'a vu plus haut (exemples (1) à (3) ci-dessus) . Pour ce verbe, il est par contre intéressant de constater que la forme du présent est sur-utilisée sept fois dans des contextes demandant une forme longue *habité*, *habiter* ou *habitait* :

⁶ Les formes marquées avec –ons de la première pluriel n'entrent pas dans les calculs.

⁷ Le corpus ne contient que 2 formes de la 2e pluriel, l'autre étant *vous marchez*.

- (11) *Contexte pour l'imparfait*
 *INT: vous étiez dans un hôtel ou est ce que vous avez euh +...
 *CAR2: nej <je> [/] **j' habite** à hôtel à www .
- (12) *Contexte pour le futur proche*
 *JUL: euh **nous allons habite** en euh Perpignan ou euh +...
- (13) *Contexte pour le passé composé*
 *JAC: <j' ai habite euh en> [/] **je ne habite** en Suède pour quinze vingt ans .

Les autres verbes d'état sont par contre produits à une forme courte à 100%. Ces verbes sont produits uniquement dans des contextes du présent et à chaque fois correctement au niveau du choix temporel (*je aime, j'aime, tous aiment, ma femme aime ; je pense, elle pense ; je préfère, nous préfère*).

Pourtant la forme courte n'est pas exclusivement réservée aux verbes d'état, puisque les verbes d'activité *parler* et *travailler* sont également plus fréquents à une forme courte chez les apprenants. Pour le verbe *travailler* on trouve une forte alternance dans l'utilisation des formes en *-E* et des formes courtes. Sur les 62 occurrences de ce verbe, la forme *travaille* exprime 11 contextes du présent et la forme *travaille* 12 contextes demandant une forme longue, dont quatre pour l'infinitif. Ce verbe semble donc particulièrement difficile à produire correctement.

- (14) *PIE: beaucoup , <j' adore> [/] j' adore **travaille** .
 (15) *RUT: euh j' aime de <travailler> [/] **travaille** dans le jardin .⁸

Les autres verbes d'activité sont tous plus souvent produits à une forme longue, mais il y a plus de variation au niveau du pourcentage de la forme la plus fréquente (de 58 à 100 %).

Pour les verbes téléquies et ponctuels, les verbes sont le plus souvent produits à une forme longue et ce à plus de 70 %.

5.2.2 Influence des formes fréquentes des verbes en *-er* chez l'intervieweuse

On pourrait ensuite se demander s'il y a un lien entre les formes verbales les plus fréquentes chez l'intervieweuse (une partie importante de l'input) et les formes verbales les plus fréquentes chez les apprenants pour les verbes en *-er*. Le tableau 7 ci-dessous présente les formes de verbes les plus fréquentes (en occurrences) chez l'intervieweuse et chez les apprenants (voir aussi tableau 6). Les données sont tirées du corpus entier et classées par catégories sémantiques.

⁸ Dans les cas où deux formes de verbes ont été produites, seule la deuxième a été comptée.

Tableau 7. Formes de verbes les plus fréquentes de l'intervieweuse et des apprenants

	Intervieweuse	formes courtes -e / -es / -ent	-ez	-er	-é	Apprenants
Etat	aimE	11	73	0	3	aime
	habite	22	7	2	0	habite
	pense	11	3	0	0	pense
	préfèrE	0	5	0	1 ⁹	préfère
Activité	étudiE	1	11	1	11	étudiE
	jouE	10	17	3	2	jouE
	parlE	10	26	10	4	parle
	regardE	1	24	3	4	regardE
	travailE	10	17 ¹⁰	4	2	travaille
	voyagE	3	9	9	6	voyagE
	achetE	0	0	1	1	achetE
Téliques/Ponctuels	commencE	1	0	3	5	commencE
	(rencontre	2	1	0	1)	rencontrE
Activité/Téliques	visitE ¹¹	0	0	5	1	visitE

Légende : les formes courtes correspondent à la 1^{ère}, 2^e et 3^e du présent singulier ainsi qu'à la 3^e pluriel du présent ; –ez : correspond à la 2^e du pluriel (au présent sauf pour travailler voir note 10) ; –er : correspond à l'infinitif ; –é : correspond au participe passé (et donc à un contexte du passé composé).

Commentaires sur le tableau 7 :

- A l'exception des verbes *habiter*, *penser* et *rencontrer* les verbes de l'intervieweuse sont plus fréquents à une forme longue (voir aussi tableau 4).
- La forme longue la plus fréquente chez l'intervieweuse est celle de la 2^e pluriel, contrairement à ce qui avait été prévu plus haut (cf. tableau 4 et le commentaire qui le précède).
- Aucun verbe en –er est utilisé au singulier de l'imparfait par l'intervieweuse. Le seul verbe en –er à être produit à l'imparfait est le verbe *travailler* à la 2^e pluriel.
- Les verbes de l'intervieweuse sont de préférence produits au présent pour les verbes d'état et d'activité. Le verbe *étudier* est aussi fréquent dans la forme du participe passé et le verbe *commencer* est le seul verbe qui soit le plus fréquemment utilisé dans la forme du participe passé.
- On peut voir que la forme la plus fréquente chez l'intervieweuse et chez les apprenants est la même pour la plupart des verbes, sauf quatre : les verbes *aimer* et *préférer* sont plus fréquents à la forme longue chez l'intervieweuse et à la forme courte chez les apprenants (cf aussi tableau 6 pour les pourcentages des apprenants) alors que les verbes *parler* et *travailler* sont plus fréquents à la forme longue chez l'intervieweuse et à la forme courte chez les apprenants. Les apprenants ne semblent donc pas simplement reprendre les formes les plus fréquentes de l'intervieweuse.

6 Discussion

⁹ La forme *préféré* est aussi utilisée 7 fois comme adjectif.

¹⁰ Y compris 3 formes de l'imparfait.

¹¹ L'intervieweuse utilise 5 fois l'expression *rendre visite* qui n'est pas comptée ici.

L'analyse de la production verbale de l'**intervieweuse** avait pu mettre en évidence trois caractéristiques spécifiques à certaines catégories sémantiques, que l'on peut résumer comme suit :

- a) verbes d'état vs les autres catégories : les verbes d'état sont clairement produits au présent et les formes finales en *-E* correspondent à 60 % des occurrences des verbes en *-er*. Pour les autres catégories on trouve des occurrences du présent, mais aussi un certain nombre de formes au passé et à l'infinitif. Dans ces catégories, les formes verbales se terminant en *-E* constituent plus de 80 % des occurrences des verbes en *-er*.
- b) verbes d'activité vs les autres catégories : les verbes d'activité contiennent un grand nombre de verbes en *-er* différents, contrairement aux autres catégories, qui n'en contiennent que peu (cf tableau 3).
- c) verbes ponctuels vs les autres : les verbes ponctuels ne contiennent pratiquement pas de verbes au présent, alors que le présent est le temps le plus fréquent à l'intérieur des autres catégories.

Il semblerait donc qu'on puisse voir quelques éléments propres aux catégories sémantiques, mais ces éléments pourraient aussi être le fait d'un usage individuel de la langue par l'intervieweuse. On peut notamment souligner le fait que la production des verbes téliques ne correspond pas à un usage prototypique, qui serait celui du passé composé. L'analyse des formes fréquentes de verbes en *-er* de l'intervieweuse avait aussi montré que seul le verbe *commencer* était plus fréquent à ce temps, les autres étant plus fréquents à l'infinitif (cf. tableau 7).

L'analyse de la production des verbes en *-er* des **apprenants** a montré que les formes en *-E* étaient plus fréquentes dans des contextes du passé que du présent, mais qu'elles marquent aussi un tiers des contextes du présent des catégories activité, téliques et ponctuels (cf. tableau 5).

La production des apprenants montre elle aussi une opposition verbes d'état vs les autres, dans la mesure où cette catégorie ne contient quasiment pas de formes en *-E*. Les trois formes en *-E* de cette catégorie ont été produites par un seul apprenant, ce qui pourrait correspondre à une spécificité individuelle.

On a aussi pu constater que la forme la plus fréquente des verbes d'état chez les apprenants était une forme du type *aime*, donc courte, alors que pour les verbes téliques et ponctuels il s'agissait d'une forme du type *commencE*, donc longue. Pour les verbes d'activité, on a pu observer que certains verbes étaient de préférence à une forme longue et d'autres à une forme courte et que ces "formes préférées" étaient utilisées dans tous les contextes - présent, passé et infinitif.

Finalement une présentation des formes les plus fréquentes chez l'intervieweuse a montré que les formes en *-E* correspondaient souvent à la 2^e personne du pluriel. Une comparaison entre les formes fréquentes de l'intervieweuse et des apprenants a montré qu'il n'y avait pas de correspondance unilatérale ; 4 des 14 verbes présentés n'ont pas la même forme la plus fréquente.

En résumé, on peut dire que la production verbale à l'intérieur des verbes d'état est différente des autres catégories, aussi bien chez l'intervieweuse que chez les apprenants. Ces verbes étant de préférence exprimés dans un contexte et une forme du présent, l'input (intervieweuse +

apprenants) de formes courtes est plus fréquent dans cette catégorie que dans les autres. La catégorie sémantique qui semble poser le plus de difficultés est celle des verbes d'activité. Elle semble contenir un plus grand nombre d'occurrences de verbes en –er, produites aussi bien dans des contextes du présent que du passé, mais aussi dans des constructions à l'infinitif. Pour les verbes téléquies et ponctuels, il est intéressant de constater que les formes en –E sont plutôt produites dans des contextes du passé et qu'il y a peu d'occurrences de verbes en –er (cf. tableaux 4 et 5).

On peut alors se demander si la variation dans la production des verbes d'activité est liée à des éléments spécifiques de cette catégorie ou au fait qu'un plus grand nombre de verbes utilisés met mieux en évidence certaines difficultés. Ou dans le même ordre d'idée, peut-on observer un lien entre la production des formes en –E et les catégories sémantiques ou le problème est-il lié à la production de chaque verbe individuellement, notamment pour des raisons phonologiques ?

Une analyse de la production des verbes irréguliers pourrait être intéressante pour pouvoir formuler des hypothèses sur la forme visée par les apprenants. Les formes du type *je parle* correspondent-elles à une forme par défaut, par exemple l'infinitif ou pourraient-elles être vue comme un participe passé ou un imparfait notamment dans les contextes du passé ? Pourrait-on dire que la "forme par défaut" n'est pas la même selon la catégorie sémantique dans laquelle elle est produite ?

Cette étude, basée sur un petit nombre de données, a permis de voir que la classification en catégories sémantiques de la production des verbes en –er et notamment des formes en –E, permettait d'apporter un éclairage spécifique sur le système interlangagier des apprenants. Ces catégories semblent donc pertinentes même pour l'analyse de formes produites dans un contexte du présent.

7 Références bibliographiques

- Andersen, R. W. (1993) « Four operating principles and input distribution as explanations for underdeveloped and mature morphological systems ». In Hyltenstam K. & Viberg Å (éds.) *Progression and regression in language. Sociocultural, neuropsychological, and linguistic perspectives*, pp. 309-339. Cambridge : Cambridge University Press.
- Bardovi-Harlig, K. 2000. « Tense and Aspect in Second Language Acquisition : Form, Meaning, and Use ». *Language Learning*, Vol 50, supplément 1.
- Bartning, I. & Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies* 14:281-299
- Carroll, S. (2001) *Input and Evidence. The raw material of Second Language Acquisition*. Amsterdam/Philadelphia: John Benjamins.
- Ellis, N. C. (2002) « Frequency Effects in Language Processing. A Review with Implications for Theories of Implicit and Explicit Language Acquisition ». *Studies in Second Language Acquisition* 24 : 143-188.
- Ellis, R. (1994) *The study of second language acquisition*. Oxford : Oxford University Press.
- Granget, C. (ce volume) « La désambiguïsation des formes verbales en [e] dans des récits oraux d'apprenants germanophones du français ».
- Klein, W. (1989) *L'acquisition de langue étrangère*. Paris : Armand Colin.

- Klein, W. & Perdue, C. (1992) *Utterance Structure. Developing grammars again*. [SiB 5] Amsterdam/Philadelphia: John Benjamins.
- Mitchell, R. & Myles, F. (1998) *Second Language Learning Theories*. London : Arnold
- Perdue C., Benazzo S. & Guiliano P. (2002) « When finiteness gets marked: the relation between morphosyntactic development and use of scopal items in adult language acquisition ». *Linguistics* 40-4:849-890.
- Pinker, S. (1998) « Words and rules ». *Lingua* 106 : 219-42.
- Pinker, S. (1991) « Rules of Language ». *Science* 253 : 530-535.
- Platzack C. (1979) *The Semantic Interpretation of Aspect and Aktionsarten. A Study of Internal Time Reference in Swedish*. Studies in Generative Grammar 8. Dordrecht : Foris.
- Salaberry, R. (1998) « The development of Aspectual Distinctions in L2 French Classroom Learning ». *Canadian Modern Language Review* 54(4) : 508-542.
- Thomas, A. (2004) *Les formes verbales chez des apprenants adultes débutants: input et catégories sémantiques. Mémoire de 80 points*. Institut d'études romanes, Université de Lund.
- Vet, C. (1980) *Temps, aspects et adverbes de temps en français contemporain. Essai de sémantique formelle*. Genève : Droz.
- Vendler, Z. (1957) « Verbs and Times ». *Psychological Review* 66 :143-60.

II

*Traits morphologiques comme indicateurs de
niveau linguistique*

L'accord sujet-verbe en français L2 parlé

Suzanne Schlyter

Université de Lund
Institut d'études romanes
suzanne.schlyter@rom.lu.se

et

Inge Bartning

Université de Stockholm
Département de français et d'italien
inge.bartning@frait.su.se

1 Introduction

Comment rendre compte du développement de l'accord sujet-verbe chez des apprenants adultes L2 ? Une approche, dans le cadre de la grammaire générative, est que l'apprenant 'acquiert' ou 'lexicalise' la Catégorie Fonctionnelle qui concerne l'inflection, IP (=Inflection Phrase), ou plus précisément, la catégorie AgrS (Agreement-Subject). Une telle approche a été reliée à l'acquisition de la morphologie correspondante (voir Clahsen et al 1996 pour l'allemand, Meisel 1994 pour le français L1). Une telle approche postule que cette acquisition se fait assez rapidement, et que tous les traits qui sont liés à cette catégorie fonctionnelle surgissent à peu près simultanément. Comme Meisel (1994) l'a montré, cela semble être le cas pour les enfants qui acquièrent le français comme langue maternelle.

En ce qui concerne une langue acquise comme L2, surtout à l'âge adulte, une telle analyse est pourtant plus problématique, et la question est loin d'être résolue. Certains proposent un développement similaire aux enfants, de telle sorte qu'il y a une relation entre l'acquisition de la morphologie verbale (notamment les formes finies) et la syntaxe comme par exemple le placement de la négation ou des adverbes (Perdue et al 2002). On propose dans ces cas que les catégories fonctionnelles soient 'acquises', ou accessibles par l'apprenant, dans l'ordre de bas en haut : VP > IP > CP (Klein & Perdue 1997, Myles 2005, Hawkins 2005).

Une conception différente est celle de la MSIH (Missing Surface Inflection Hypothesis) d'après laquelle la syntaxe et la morphologie sont nettement séparées, la syntaxe étant accessible dès le début mais la morphologie constituant un développement à part. Granfeldt et Schlyter défendent depuis un certain temps une position de ce type, ce qui est surtout dû aux différences très nettes que nous avons observées entre les enfants et les adultes (Granfeldt 2000, 2003 ; Schlyter 2003a, 2005).

Une discussion détaillée de ces questions est fournie par Hawkins (2005 :248), où en

discutant les contre-exemples de la MSIH il envisage, entre autres :

« whether the apparent counterexamples can be explained as the effects of something outside of the operation of UG (for exemple frequency of forms in the input);”.

Nous ne voulons pas entrer dans ces détails, mais montrer le caractère extrêmement lent et graduel de l'accord sujet-verb en français parlé, un développement qui nous semble compatible avec la fréquence des formes/fonctions en question dans l'input.

2 Nos deux corpus

Nos données proviennent de deux grands corpus constitués de la production orale 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 1a-b ci-dessous.

Tableau 1a. *Le corpus Lund*

	Long/ Trans	Nombre d'appr	Années d'étude de français	Âge	Période suivie	Occas. d' enreg	Tâches Int, BD, Vid	Type d'acquis ition
LNN								
Adultes séjournant en France	long	8	quelques semaines ou 0	20-50 env	2 à 15 mois	2 à 5	interv, récits, traduct	non- guidée
Étudiants d'université	long	7	6 ans de lycée	19-22	5 à 12 mois	4 à 9	interv, récits, traduct	guidée
Débutants Corpus Thomas	trans/ long	5	6-8 mois	35-70	-	1 à 2	interv	guidée

Tableau 1b. *Le corpus InterFra, Stockholm*

	Long/ Trans	Nombre d'appr	Années d'étude	Âge	Période suivie	Occas. d'enreg	Tâches Int, BD, Vid	Type d'ac- quisition
LNN								
Gr 1 Déb	a) long b) trans	a) 6-10 b) 18	0	19-30	2 semest.	5-10 1	+	guidée
Gr 2 Lyc	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 Fut prof	Long	6	7-8	23-34	2 semest. +stage	4	+	semi-guidée
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; Déb=vrais débutants à l'université, Lyc= lycéens, Étudiants=étudiants universitaires, du 1er au 4e semestre, Fut.prof= étudiants futurs professeurs; Int=Interview, BD=récits de bandes dessinées, Vid=récits de films vidéo muets. Les chiffres en gras indiquent le nombre d'apprenants examinés en particulier pour cette étude.

Dans Bartning et Schlyter (2004), nous proposons, à partir des données de différents types de ce corpus, les six stades de développement suivants pour les apprenants suédophones de français L2, adultes et adolescents:

- 1: Initial
- 2: Post-initial
- 3: Intermédiaire
- 4: Avancé bas
- 5: Avancé moyen
- 6: Avancé haut

Ces stades sont basés sur le développement d'un grand nombre de différents phénomènes linguistiques tels que l'accord sujet-verbe (qui sera donc indiqué plus en détail ici), Temps-Mode-Aspect, négations, pronoms d'objet, cliticisation des pronoms sujets et articles, l'accord en genre, subordination et 'empaquetage' etc. Dans Schlyter (2003b) nous avons aussi traité l'amalgame des prépositions avec l'article.

3 Le développement de l'accord sujet-verbe des verbes non-thématiques

Dans les discussions sur l'accord sujet-verbe, on distingue souvent les verbes thématiques (= lexicaux, pleins, de contenu) des verbes non-thématiques, ces derniers ayant très peu de contenu propre mais figurant surtout comme copules et auxiliaires (*être, avoir, aller* et les verbes modaux du type *vouloir, pouvoir* etc.). Les verbes non-thématiques semblent avoir un accord sujet-verbe plus facile à acquérir, ou qui en tout cas est acquis très tôt (Parodi 2000).

Cela est le cas aussi dans notre corpus. Les débutants, au stade 1, semblent avoir surtout une forme invariable et manquent d'opposition entre personnes, même au singulier :

j # je # j'E un petit peu fatigué ; oui c'est vrai* (Henry 1) Stade 1

*INT: oui d'accord, vous avez des enfants?

*CAR: *euh oui [/] euh je [/] euh # [/] euh [/] euh je # [/] je a, nej...* Stade 1

Comme Henry ne connaît pas encore l'expression *j'ai peur*, la forme *j'E* est à interpréter comme *j'est fatigué* ou *je fatigué*. En tout cas cela indique qu'il ne connaît pas la forme *je suis*.

On observe pourtant que très tôt, déjà au stade 2, ces verbes figurent dans au moins deux formes différentes qui signalent des personnes différentes :

je suis musicien ; il est français (Björn 1) Stade 2
j'ai pas beaucoup d'amis ici ; elle a fait le gâteau (Björn 1)

Chez Robert 3, on voit encore une petite hésitation concernant l'accord, qui semble pourtant sur le point d'être maîtrisé :

je suis eh étudiant de français ; elle suis eh elle est tou- toulousien (Rob 3)
demain j'ai un examen ; mais étudiant de Informatique (I :mm) a été
beaucoup de travail (Rob 3) Stade 2

Tableau 2. être/avoir/aller avec opposition entre 1^{ère} et 2^{ème} / 3^{ème} ps sing,
 Corpus Lund

	1p sg je suis	2/3p sg tu es/ c'est/ il est	1p sg j'ai	2/3p sg tu as/ il a	1p sg je vais	2/3p sg tu vas/ il va	stade B&S
Caroline	+	+	-	+	-	-	1
Henry 1-2	-	+	+	-	-	-	1
Martin 1	-	+	+	+	+	-	2
Martin 2-3	-	+	+	+	+	+	3
Petra 1	+	+	+	+	+	-	2
Petra 2	+	+	+	+	+	+	2-3
Lisa 1	+	+	+	-	-	-	2
Lisa 2	+	+	+	+	+	-	2
Sama 1	-	+	+	+	+	+	3
Sama 2	+	+	+	+	+	+	3

Tableau 2. être/avoir/aller avec opposition entre 1^{ère} et 2^{ème} / 3^{ème} ps sing,
 Corpus InterFra

	1p sg je suis	2/3p sg tu es/ c'est/ il est	1p sg j'ai	2/3p sg tu as/ il a	1p sg je vais	2/3p sg tu vas/ il va	stade B&S
Rob 1-2	+	+	+	+	-	-	1-2
3-6	+	+	+	+	+	-	2-3
6-7	+	+	+	+	+	+(ça va, vont)	3
Pelle 1	-	+	+	-	+	-	1
2	+	+	+	+	+	-	2
3-4	+	+	+	-	- /+	-	2
5-7	+	+	+	+	+	-	2
Emelie 1-2	+/-	+	+	+/-	-	-	2
3-5	+	+	+	+/-	+	-	3
6	-	+	+	+	-	+	3-4
7	+	+	+	+	+	+	3-4

4 Le développement des ‘formes finies’ et ‘formes non-finies’ des verbes thématiques

Les verbes thématiques, en revanche, se développent plus lentement. Nous présentons ici plus en détail les données qui sont à la base de la description succincte Bartning et Schlyter (2004).

Nous avons étudié ici l'opposition entre les deux formes de base (formes par défaut, d'après Prévost 2004) du type *il parle* vs *il parler/parlé*, utilisées pour un verbe fini, donc avec un sujet explicite. Ces deux formes semblent être les deux premières à être utilisées par les apprenants, surtout les apprenants non-guidés. Une telle opposition entre une forme finie

simple - non conjuguée pour la personne - et une autre forme, non-finie, ressemble aussi au système suédois, dans lequel il n'y a pas d'accord sujet-verbe mais en revanche une différence très nette entre la forme du présent et la forme infinitivale. Nous appelons ici la forme du type *il parle* (ainsi que *elle sort, je vois* etc.) la 'forme finie' et la forme du type *il parler/é* (ainsi que *elle sorti(r), je voir/vu* etc.) 'forme non-finie'. Nous avons écarté les formes qui réfèrent au passé et qui peuvent donc être considérés comme des imparfaits (cf. Granget dans ce volume), et nous ne prenons pas en considération, pour le moment, les formes marquées du type *nous parlons, vous parlez* ou *ils viennent*, que nous allons étudier séparément (voir ci-dessous).

Nous pouvons observer un développement graduel et très lent du taux des formes finies (ainsi définies) chez les apprenants, que ce soit chez des apprenants guidés (G) ou non-guidés (NG).

Tableau 3a. *Formes finies vs formes non-finies dans le Corpus Lund.*

Groupe ou appr/enreg		mois d'exposition	présent, formes finies	présent, formes non-finies	% de formes finies
Caroline	G	6 (24heures)	7	5	58%
Henry 1-2	NG	< 3	26	15	64%
Björn 1	NG	3	23	16	58%
Björn 2		5	24	15	62%
Karl 1	NG	8	53	34	61%
Karl 2		10	48	17	74%
Karl 3		12	75	20	79%
Karl 4		14	13	5	72%
Karl 5		16	81	11	88%
Martin 1	NG	7	57	3	95%
Martin 2		14	32	7	82%
Martin 3		16	20	4	83%
Knut 1-2	NG	20-22	78	7	92%
Knut 3		35	80	1	99%
Lisa 1	G	école 6 ans	44	15	75%
Lisa 2		+ 1 mois	70	11	86%
		univ			
Lisa 3		+ 2 "	35	8	81%
Lisa 4		+ 3 "	89	25	78%
Sama 1	G	école 6 ans	40	2	95%
Sama 2		+ 1 mois	38	10	79%
		université			
Sama 3		+ 2 «	59	5	92%
Sama 4		+ 3 "	74	11	87%
Sama 5		+ 4 "	23	1	96%
Nina 1	G	école 6 ans	22	1	96%
Nina 2		+ 1 mois	29	3	90%
		université			
Nina 3		+ 2 "	59	1	98%
Nina 4		+ 3 "	52	3	90%
Nina 5		+ 4 "	16	0	100%
Maguy 4	G	école 6 ans	40	0	100%
		+ 4 mois			
		université			
Maguy 9		+9 mois	49	0	100%
		université			

Tableau 3b. *Formes finies/non-finies, corpus Lund*

Apprenants	formes finies, %
Moins de 9 mois d'exposition	66%
Groupe intermédiaire	80%
Plus de trois ans d'exposition dans le pays ; ou après l'école et un semestre d'université	94%

Nous trouvons un développement très semblable dans le corpus InterFra:

Tableau 3c. *Formes finies/non-finies des débutants pendant deux semestres, corpus InterFra, étudiants guidés*

Apprenant/groupe	présent, forme finie	présent, forme non- finie	Total	% de formes finies
ROBERT Int1(oct)	8	5	13	62%
Int 2,BD1 (nov)	22	11	33	67%
Int 3,Bi1 (nov)	28	4	32	87%
Int 4,Vi1 (dec)	36	5	41	88%
Total 1 ^{er} semestre	94	25	119	79%
Int 5 (feb)	36	3	39	92%
Int 6,BD2 (mar)	37	3	40	92%
Int 7,Bi2 (avr)	20	2	22	91%
Total 2e semestre	93	8	101	92%
CARIN Int 1 (feb)	5	7	12	42%
Int 2, BD1 (mar)	9	3	12	75%
Int 3, Bi1 (avr)	16	6	22	73%
Int 4 (mai)	12	2	14	86%
Int 5,Vi1 (jui)	14	3	17	82%
Total 1er semestre	56	21	77	73%
JAN Int 1 (feb)	5	-	5	100%
Int 2,BD1 (mar)	8	-	8	100%
Int 3,Bi1 (avr)	11	4	15	73%
Int 4 (mai)	11	1	12	100%
Int 6,Bi2 (août)	13	4	17	76%
Total 1er sem	37	8	45	82%
PELLE Int 1 (oct)	6	2	8	75%
Int 2, BD1 (nov)	17	7	24	71%
Int 3,Bi1 (dec)	11	3	14	79%
Int 4,Vi1 (feb)	12	7	19	63%
Total 1er sem	46	19	65	71%
Int 5,BD2 (mar)	15	2	17	88%
Int 6, Bi2 (avr)	15	4	19	79%
Int 7 (mai)	15	4	19	79%
Int 8,Vi2 (jui)	32	3	35	91%
Total 2e sem	77	13	90	86%
VERA Int 1 (oct)	12	-	12	100%
Int 2,BD1 (nov)	14	2	16	87%
Int 3,Bi1 (dec)	13	1	14	93%
Int 4,Vi1 (feb)	10	12	22	45%
Total 1er sem	49	15	64	77%
Int 5, BD2 (mar)	14	1	15	94%
Int 6, Bi2 (avr)	11	4	15	73%
Int 7, Vi2 (mai)	8	8	16	50%
Int 8 (jui)	10	2	12	83%
Total 2e sem	43	15	58	89%
EMELIE Int1 (oct)	8	1	9	89%
Int 2, BD1 (nov)	12	1	13	92%
Int 3,Bi1 (nov)	13	2	15	87%
Int 4,Vi1 (dec)	14	1	15	93%
Total 1er sem	47	4	52	92%
Int 5 (feb)	14	1	15	93%
Int 6, BD2 (mar)	25	-	25	100%
Int 7, Bi2 (avr)	29	1	30	97%
Vid 2 (jui)	4	-	4	100%
Total 2e sem	72	2	74	97%

Tableau 3d. *Formes finies/non-finies dans 3 groupes d'InterFra*

Apprenants/ groupes	présent fo finies	Présent fo non-fi	% de fo finies
InterFra 6 débutants Int 1-4, 1er sem	329	92	78%
InterFra 8 lycéens Int 1-2	464	78	86%
InterFra 4 appr univ Int 1-4	1633	1	100%

On peut voir que ce développement est lent et graduel (cf. par ex. Karl 1-5, Lund et Robert Int 1-7, InterFra). Si on avait eu à faire avec une acquisition paramétrique, dans le sens de White (2003 : 179), on se serait attendu à un développement plus soudain, illustré par un palier au moment du début de l'accès à la catégorie fonctionnelle AgrS (ou IP, plus généralement). Il semble aussi difficile d'en rendre compte en postulant un parallélisme entre ce développement morphologique et un développement syntaxique.

5 Développement du pluriel (1p et 3p) marqué à l'oral

Notre troisième étude ici concerne le développement des formes au pluriel : 1p pl (*nous parlons*), et 3p pl, où nous distinguons d'un côté les verbes *ils ont/sont/vont* et de l'autre les verbes lexicaux (thématiques, mais y inclus les modaux) dans lesquels la marque du pluriel est audible, comme *ils prennent, ils sortent, ils veulent, elles peuvent* (*Vlex-nt* dans le tableau 4). Notre but, dès le départ, était de voir si l'on pouvait observer un ordre d'acquisition entre ces formes et aussi par rapport aux autres formes étudiées, et si dans ce cas on pouvait trouver un développement semblable dans nos deux corpus indépendants. Cela semble effectivement le cas :

Tableau 4a. *Accord sujet-verbe au pluriel 1ère et 3e personne, Corpus Lund.*

Apprenants/ groupes	<i>nous</i> <i>Vons</i> correct	<i>nous</i> <i>V-</i> incorr	% <i>nous</i> <i>Vons</i>	<i>ils -ont</i> correct	<i>ils -ont</i> incorr	% <i>ils</i> <i>-ont</i>	<i>Vlex-nt</i> corr	<i>Vlex-</i> incorr	% Vlex-nt
Moins de 9 mois d'exposition	23	10	70%	8	4	67%	1	5	17%
Groupe intermédiaire	94	18	84%	27	32	46%	2	12	14%
Plus de 3 ans d'expo dans le pays ; ou après un semestre d'université	42	7	86%	53	16	77%	8	12	40%

Légende: *nous Vons* correct: l'accord au pluriel 1^{ère} personne marqué;

nous V-incorr : « « « non marqué ;

ils -ont correct/ incorr : « « 3^{ème} « avec les verbes en *-ont* marqué/non-marqué

Vlex-nt corr : « « 3^{ème} « « verbes lexicaux marqué ; etc.

Le corpus InterFra (de Inge Bartning) montre le même ordre de développement:

Tableau 4b. *Accord sujet-verbe au pluriel 1ère et 3e personne, corpus InterFra*

Apprenants / groupes	<i>nous</i> <i>Vous</i> correct	<i>nous V-</i> incorr	% <i>nous</i> <i>Vous</i>	<i>ils -ont</i> correct	<i>ils -ont</i> incorr	% <i>ils</i> <i>-ont</i>	<i>Vlex-nt</i> corr	<i>Vlex-</i> incorr	% Vlex- nt
6 Débutants, 1 ^{er} sem									
Total Int, BD, Vid	42	13	76%	27	13	67%	3	8	27%
8 Lycéens									
Total Int, BD, Vid	44	18	71%	21	39	35%	1	14	7%
4 Appr univ									
Total Int, BD, Vid	15	1	94%	205	28	88%	74	51	59%
5 doct+3 appr av BD, Vid seulement	+			+			106	16	87%

Légende: voir 4a

Nous pouvons en déduire un ordre d'acquisition (concernant aussi bien l'apparition que la productivité de la forme), qui est le même dans les deux corpus :

1. Verbes non-thématiques au sing : *je suis - il est; j'ai - il a* etc.
2. Plur de la 1^{ère} personne : *nous V-ons*
(2'. 'Formes finies' vs 'formes non-finies' : *je parle* vs. *je parler*)
3. Verbes en *-ont* à la 3p plur : *ils ont/sont/vont/font*)
4. Verbes lexicaux à la 3p plur, type *ils prennent*).

Il est difficile de placer exactement l'acquisition des 'formes finies' (telles que définies ici). Dans Bartning et Schlyter (2004), nous considérons différemment deux cas de morphologie : pour la première, ayant une alternance entre une forme radicale (ou défaut, comme *parl-*) et une forme marquée (comme *nous parlons*) nous considérons 75% de contextes marqués correctement comme une utilisation productive de la forme. Le même taux ne peut pas être utilisé là où deux formes plus ou moins égales alternent (cf. le genre), comme c'est le cas pour les deux formes de base *parle - parler/é*. (Ici nous avons 90% comme taux productif.) Intuitivement, nous dirons que la distinction 'formes finies/non-finies' est productive peu après, ou simultanément avec, celle de la 1p pl *nous parle/nous parlons*, mais l'ordre est incertain et pourra encore être changé.

La première opposition de notre séquence, celle entre *je suis - il est* etc., est en place très tôt. On l'observe après quelques mois d'exposition au français, ce qui correspond au Stade 2 selon Bartning et Schlyter (2004). Les oppositions 2 et 3 sont en place un peu plus tard, dans les stades intermédiaires. Au Stade 4, les apprenants n'utilisent plus de formes non-finies à la place des formes finies. La dernière opposition, celle des verbes thématiques au pluriel, ne s'établit qu'après une longue période d'exposition. Ce ne sont que les apprenants très avancés : ceux qui ont étudié le français pendant 3-4 semestres à l'université (après 5 ans au lycée) ou ceux qui ont passé plusieurs années dans un pays francophone, qui maîtrisent l'accord à la 3^{ème} personne au pluriel en français parlé.

Ce développement extrêmement étendu et graduel semble difficile à décrire à l'aide d'une fixation du paramètre AgrS, si une telle fixation doit se passer à un moment précis. Il nous faudrait pourtant étudier plus en détail si les deux formes les plus régulières, 'formes finies/non-finies' et 'nous V-ons' pourraient éventuellement constituer le noyau de la 'finitude' et représenter la Catégorie Fonctionnelle AgrS. Nous n'allons pourtant pas aborder

cette question ici.

6 Une approche associative basée sur la fréquence de la forme dans l'input

Pour rendre compte du développement de ces différents types de suffixes flexionnels représentant l'accord sujet-verbe en français parlé, nous voudrions esquisser un développement dans un autre cadre que celle de la syntaxe générative. Nous supposons que l'apprenant a accès à la syntaxe, de par sa langue maternelle, comme le proposent p.ex. Prévost & White (2000), ou Granfeldt (2001, 2003). Le développement de la morphologie flexionnelle, en revanche, se passe séparément, influencé par la fréquence de la forme/fonction dans l'input. Nous nous posons donc dans l'optique de Bybee (1991), Nick Ellis (2002) etc. en ce qui concerne la morphologie flexionnelle, voir aussi Bartning (1998 ; à paraître). Nous ne voulons pourtant pas prendre position sur la divergence entre ces auteurs et le système 'Duel' (Pinker 1998), selon lequel une des formes serait régulière, et donc acquise par le système des règles alors que les autres formes (avec leur fonction) seraient acquises comme des lexèmes et ainsi sensibles à la fréquence dans l'input. La distinction entre ces deux théories n'est pas possible à faire avec les données spontanées que nous avons, mais exige une expérimentation plus poussée.

Nous présentons ici comme simple ébauche comment nous pensons procéder, et il est évident qu'il faudra, par la suite, trouver des arguments plus importants pour une telle approche, avec des études sur la fréquence de ces formes/fonctions dans les corpus de français oral et dans les discours adressés aux apprenants.

1. Les formes *je suis - tu es - c'est ... j'ai - tu as - il y a, je vais - tu vas* au singulier (et dans un moindre degré : *il/elle/on va*), sont des combinaisons très fréquentes, qui ont une grande 'token frequency' (voir Bartning, à paraître) et aussi une grande 'transition frequency' entre le sujet et le verbe. Ces combinaisons sujet-verbe sont donc acquises comme n'importe quels lexèmes fréquents, très tôt.

2. Le suffixe *-ons* dans *nous parlons, nous travaillons* etc., très régulier, figure avec un grand nombre de verbes, toujours de la même manière systématique (donc une relation bi-univoque), avec une grande fréquence d'occurrences (token frequency). On peut s'attendre à ce que la combinaison *nous avons* soit la première à être acquise et qu'elle servira de modèle aux autres verbes, pour lesquels la désinence sera segmentée (par la combinaison *nous __ons*, avec un 'slot' pour la racine du verbe). Cela prédit une acquisition assez précoce.

3. Le choix entre les formes que nous appelons ici 'formes finies' et 'non-finies' semble être lié à une forme invariable pour chaque verbe, par exemple *il parle - *pour parle*, mais **je étudier - pour étudier*. Il est possible que la forme *parle* figure souvent, dans l'input de ces apprenants, dans des contextes du type *tu parles français ?*, *moi je ne parle pas suédois* etc. En revanche, *étudieE* figure peut-être plus souvent dans des contextes comme *Qu'est-ce que tu vas étudier ? / as étudié ?* Avec le temps, ces patterns de fréquence dans l'input peuvent former des catégories telles que celles étudiées par Anita Thomas, ce volume. C'est là des études qui restent à faire. Ici, nos études sur corpus rejoignent celles expérimentales de Largy et de son équipe, voir Largy, ce volume.

4. Le pluriel 3p semble être relativement rare dans l'input, comparé aux mêmes verbes au singulier. On prédit donc une acquisition assez tardive. Les formes *ont/sont/vont* doivent être acquises plus tôt que les verbes thématiques (*prennent, sortent*) étant donné la fréquence de ces premiers verbes. Une telle approche est suggérée par Hedbor, ce volume. (En revanche, les données de Ågren, ce volume, montrent que même ces verbes peuvent servir de base à une segmentation en verbe+ le suffixe *-ont*.)

5. Le pluriel des verbes thématiques (lexicaux) est rarement marqué dans le français parlé. L'accord se fait le plus fréquemment (token frequency) sans suffixe ajouté, p.ex. *il parle - ils parlent* (voir Bartning à paraître). Ce type d'accord est facilement sur-étendu à d'autres verbes, ce qui entraîne possiblement un manque de marque de pluriel pour les verbes comme *viennent, mettent* etc. aussi. Il est possible que ce système oral se développe avec le temps, au fur et à mesure que les apprenants entendent plus de français parlé, alors que les débutants à l'école ont un input surtout écrit. Nous voyons ici une certaine différence entre les apprenants guidés et non-guidés, ainsi que la courbe en U que nous entrevoyons dans nos données, voir Schlyter, ce volume.

Ce ne sont là pourtant que quelques résultats préliminaires, qui devront être étudiées par des tests à plus grande échelle dans des études ultérieures.

7 Références bibliographiques

- Andersen, R. W. (1993) « Four operating principles and input distribution as explanations for underdeveloped and mature morphological systems ». In Hyltenstam K. & Viberg Å. (éds) *Progression and regression in language. Sociocultural, neuropsychological, and linguistic perspectives*, pp. 309-339. Cambridge University Press.
- Bartning, I. 1998. « Procédés de grammaticalisation dans l'acquisition des prédications verbales en français parlé ». *Travaux de linguistique* 36 : 223-34.
- Bartning, I. (à paraître) « Late morphosyntactic and discourse features in advanced/very advanced L2 French – a view towards the end state ». In Habermatzl, S. (éd.) *Processes and Outcomes. Explaining Achievement in Language Learning*. Berlin: Mouton de Gruyter.
- Bartning, I. & Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies* 14 : 281-299.
- Bybee, J. (1991) « Natural morphology: the organization of paradigms and language acquisition ». In T. Huebner & C. Ferguson (éds), *Crosscurrents in second language acquisition and linguistic theories*, pp. 67-91. Amsterdam/Philadelphia : John Benjamins.
- Clahsen, H., Eisenbeiss, S. & Penke, M. (1996) « Lexical Learning in Early Syntactic Development ». In H. Clahsen (éd.), *Generative Perspectives on Language Acquisition: Empirical Findings, Theoretical Considerations, and Crosslinguistic Comparisons*, pp. 129-159. Amsterdam/Philadelphia : John Benjamins.
- Ellis, N. C. (2002) « Frequency Effects in Language Processing. A Review with Implications for Theories of Implicit and Explicit Language Acquisition ». *Studies in Second Language Acquisition* 24 : 143-188
- Granfeldt, J. (2000) « The acquisition of the Determiner Phrase in bilingual and second

- language French ». *Bilingualism : Language and Cognition* 3(3) : 263-280.
- Granfeldt, J. (2003) *L'acquisition des catégories fonctionnelles*. Thèse de doctorat, Institut d'Études Romanes, Université de Lund.
- Hawkins, R. (2004) « The contribution of the theory of Universal Grammar to our understanding of the acquisition of French as a second language ». *Journal of French Language Studies* 14 : 233-255.
- Klein, W. & Perdue, C. (1997) « The Basic Variety (or: Couldn't natural languages be much simpler?) ». *Second Language Research* 13(4): 301-347.
- Meisel, J. M. (1994) « Getting FAT: Finiteness, Agreement and Tense in Early Grammars ». In J. M. Meisel (éd.) *Bilingual First Language Acquisition*, pp. 89-129. Amsterdam/Philadelphia : John Benjamins.
- Myles, F. (2005) « The emergence of Morpho-syntactic Structure in French L2 ». In J.M. Dewaele (éd.) *Focus on French as a Foreign Language: Multidisciplinary Approaches*, pp. 36-62. Clevedon : Multilingual Matters.
- Parodi, T. 2000. « Finiteness and verb placement in second language acquisition ». *Second Language Research* 16(4) : 355-381.
- Perdue, C., Benazzo, S. & Giuliano, P. 2002. « When finiteness gets marked: the relation between morphosyntactic development and use of scopal items in adult language acquisition ». *Linguistics* 40(4) : 849-890.
- Pinker, S. (1998) « Words and rules » *Lingua* 106 : 219-42
- Prevost, P. & White, L. (2000) « Missing Surface Inflection or Impairment in Second Language Acquisition? Evidence from Tense and Agreement ». *Second Language Research* 16,2: 103-133.
- Prévost, P. (2004) « Morphological variation in early adult second language French. A cross-sectional study ». *EUROSLA Yearbook* 4 : 147-175
- Schlyter, S. (2003a). « Development of verb morphology and finiteness in children and adults acquiring French » In Dimroth, C. & Starren, M. (éds.), *Information structure, linguistic structure, and the dynamics of learner language*, pp. 15-45. Amsterdam/Philadelphia : John Benjamins (Studies in Bilingualism).
- Schlyter, S. (2003b) « Stades de développement en français L2 : Exemples d'apprenants suédophones, guidés et non-guidés, du "Corpus Lund" ». Disponible sur www.rom.lu.se/durs/STADES_DE_DEVELOPPEMENT_EN_FRANCAIS_L2.PDF
- Schlyter, S. (2005) « Adverbs and functional categories in L1 and L2 acquisition of French ». In Dewaele, J.M. (éd.), *Focus on French as a Foreign Language: Multidisciplinary Approaches*, pp. 36-62. Clevedon : Multilingual Matters.

Direkt Profil* et deux études sur la morphologie verbale et les stades de développement

Jonas Granfeldt

Université de Lund
Institut d'études romanes de Lund
jonas.granfeldt@rom.lu.se

Résumé

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 de produire une évaluation du stade de développement des apprenants sous la forme d'un profil grammatical. *Direkt Profil* réalise une analyse des phrases fondée sur des *itinéraires d'acquisition*, i.e. des phénomènes morphosyntaxiques locaux liés à un développement dans l'apprentissage du français. Cet article commence par constater combien les ordinateurs ont changé les méthodes des recherches en linguistique appliquée. Il présente ensuite le système *Direkt Profil* dans sa version actuelle et discute son apport aux domaines du présent volume par l'intermédiaire de deux études préliminaires sur la morphologie verbale chez des apprenants anglophones (données orales) et suédophones (données écrites) de français. L'article se termine par une discussion des possibilités qui se présentent aux chercheurs s'intéressant aux recherches sur l'acquisition de la morphosyntaxe du français.

* Je tiens à remercier Pierre Nugues pour ses 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

Nous sommes nombreux à avoir réalisé à quel point l'informatique peut être un outil essentiel aux recherches en linguistique appliquée. Les avantages les plus évidents de l'ordinateur, par rapport aux analyses manuelles, sont la rapidité et l'efficacité du traitement. En fait, il est aujourd'hui difficile de trouver des domaines des recherches où les outils informatiques ne jouent pas un rôle central, que ce soit dans les processus de collecte des données, du stockage, du traitement et de l'analyse. C'est la démonstration que l'informatique s'étend aujourd'hui bien au-delà de son propre champ.

Mais si les linguistes utilisent couramment des logiciels dans leurs analyses, ils ont moins l'habitude d'en développer eux-mêmes. Cet article présente un outil de recherche que nous développons dans un projet pluridisciplinaire.¹ Il s'agit d'un analyseur automatique appliqué à la morphosyntaxe en français langue étrangère (FLE). Avec cet outil, nommé *Direkt Profil*, nous abordons, entre autres, les questions suivantes :

- (a) Est-ce que l'acquisition du français par des adultes passe par des stades d'acquisition possibles à définir en termes morphosyntaxiques ?
- (b) Serait-il possible de dresser un « Profil » linguistique grammatical d'un texte produit spontanément en se basant sur des critères morphosyntaxiques développés sous forme d'*itinéraires d'acquisition* (voir *infra*) ?
- (c) Quelle serait la relation entre ce « Profil » basé sur la production en ligne et libre fondé sur des itinéraires d'acquisition et un Profil basé sur des tests « traditionnels » fondé sur des indicateurs pragmatiques qui relèvent de l'usage de la langue (cf. *Cadre Commun de Référence*, Conseil d'Europe)

L'idée de concevoir le système de *Direkt Profil* est née en l'an 2000 lors des préparatifs du congrès suédo-danois *Mål, Medel och Mätningar i språkämnena*. Ce congrès avait pour but de faire l'état des lieux sur les méthodes d'enseignement et d'évaluation des langues étrangères dans les deux pays ainsi que de présenter les recherches faites sur le développement des langues étrangères (surtout le français) en Suède. Cette idée est née grâce à la conjonction de deux facteurs:

- (i) Les recherches sur les *itinéraires d'acquisition* en FLE étaient arrivées à un point où les premières synthèses étaient possibles. Les deux équipes suédoises travaillant dans ce domaine, celle d'Inge Bartning à Stockholm et celle de Suzanne Schlyter à Lund, avaient produit des résultats préliminaires sur les caractéristiques morphosyntaxiques de différentes variétés de français langue étrangère parlé (par des suédophones). Voir Bartning, 1997, Schlyter, 2003, et aussi Granfeldt, 2003. Ces recherches ont été récemment réunies dans un article commun Bartning & Schlyter (2004).
- (ii) Les techniques computationnelles devenaient suffisamment fiables, accessibles et puissantes pour qu'il soit possible d'envisager le développement d'un outil spécifique pour l'analyse morphosyntaxique du FLE. D'abord, les logiciels dans la boîte à outils de CHILDES

¹ A l'origine, le projet est le résultat d'une collaboration entre l'auteur et Suzanne Schlyter de l'Institut d'études romanes de l'université de Lund (LU) et Pierre Nugues de l'Institut d'informatique de Lund (LTH).

(MacWhinney, 2000), originalement conçus pour les recherches sur l'acquisition de langage chez l'enfant, s'étaient aussi avérés très utiles pour les recherches sur l'acquisition d'une langue étrangère, à l'oral ou à l'écrit. Deuxièmement, le *Rapid Profile* de Manfred Pienemann montrait l'utilité d'intégrer les acquis de recherches sur l'acquisition des langues étrangères pour évaluer le niveau linguistique des apprenants. Ce système est développé pour l'allemand et l'anglais.²

1.1 Avant Direkt Profil – CLAN et ses limites

Dans notre domaine, le système de CHILDES (MacWhinney, 2000) est un bon exemple d'outil qui a profondément changé les méthodes et l'analyse des recherches. Nos recherches partagent avec la description linguistique, le recours aux corpus, oraux et écrits, pour l'établissement des objets de recherches. Le système de CHILDES met à notre disposition un certain nombre d'outils pour la gestion, le traitement et l'analyse des corpus. Ce système, très répandu, comprend une boîte à outils de CLAN (*Computerized Language Analysis*) comportant un certain nombre de programmes de calcul de fréquence, de concordances, etc. Il permet de formuler des requêtes générales à l'aide d'expressions régulières (ou expressions rationnelles) pour parcourir de façon automatique le corpus. Le but des programmes est de trouver des morceaux de textes qui valident le modèle de la requête (un « match ») et les présenter à l'utilisateur. Le programme *Combo* permet la combinaison d'expressions régulières et des opérateurs booléens, SUIVI PAR (^), OU(+) et SAUF(!) pour mieux préciser les requêtes. On peut aussi recourir à des « listes » de mots spécifiques. Si nous sommes intéressés par l'accord à la 3^e personne pluriel des pronoms sujets (les sujets clitiques, « Scl »), une requête *Combo* pour retrouver ces occurrences serait :

(1) liste_Scl^*^*ent !liste_(-V) (« chercher les séquences ou un des mots de la liste pronoms sujets clitiques précède directement ou pas un mot qui se termine en –ent SAUF ceux qui se trouvent dans la liste liste_(-V] »)

(Notons que la liste de non-verbes doit aussi être produite manuellement après une phase d'extraction de tous les mots du corpus.)

Par rapport à nos besoins, les programmes de CLAN présentent un certain nombre de limitations. Premièrement, la solution locale nécessaire pour désambiguïser certaines formes (en passant par une liste d'exclusion-inclusion par exemple pour les formes non-verbales) représente une méthode simple et suffisante pour les cas où les textes à traiter sont connus et stockés, permettant ainsi un traitement préalable.³ Ce n'est pas notre cas. À long terme, nous souhaitons traiter un corpus (ici un texte) « en direct » où l'entrée du système ne vient pas du chercheur mais de l'apprenant, le sujet lui-même. Deuxièmement, en CLAN les analyses se font toutes, à quelques exceptions près, une par une. Cette démarche ne suffira pas pour les recherches sur des Profils grammaticaux qui sont caractérisés par un large inventaire de catégories différentes. Nous avons besoin de lancer en cascade plusieurs requêtes à la fois. Finalement, nous avons aussi besoin de faire des requêtes plus détaillées visant par exemple certains groupes de mots (GN, GV, etc.) ou même certaines catégories de verbes. De plus, il

² <http://groups.uni-paderborn.de/rapidprofile/>

³ Après le développement et l'intégration des programmes MOR (pour l'analyse des parties des discours, %mor :) et de POST (pour la désambiguïstation de la ligne %mor) cette démarche n'est plus nécessaire. Mais MOR ne pourra pas effectuer les analyses de dépendances qui nous intéressent (voir *infra*).

faudra aussi traiter certaines dépendances et des accords nominaux et verbaux, etc. Ces derniers types d'analyses rendent nécessaire de passer par une étape préalable d'étiquetage des parties de discours.

Pour résumer : par rapport aux analyses de CLAN, *Direkt Profil* nous permet de :

- Traiter des corpus plus larges
- Traiter et analyser des textes et des corpus inconnus, en direct, sans présence d'un expert-chercheur
- Lancer en cascade un ensemble de requêtes
- Spécifier des requêtes plus précises qui se réfèrent à des catégories lexicales et grammaticales et aux phénomènes d'accord verbal et nominal.

De plus, l'architecture du *Direkt Profil* nous permet aussi de :

- Modifier les catégories d'analyse (à un coût temporel moyen) après évaluation
- Rajouter des catégories d'analyses au fur et à mesure (et permettre à d'autres chercheurs à en rajouter).

1.2 Le fonctionnement de Direkt Profil

Direkt Profil fonctionne en mode client-serveur où le serveur réalise l'annotation d'un texte et le client, intégré à un navigateur, prend en charge l'affichage et l'interaction. Aucune installation n'est nécessaire. Il suffit d'aller à la page Web <http://www.rom.lu.se:8080/profil> pour se connecter.

L'interface du programme est conçue pour que les chercheurs (ou les enseignants ou les apprenants) puissent copier des textes écrits par des apprenants (ou des transcriptions de langue orale) et les soumettre à l'analyseur. Après la connexion, l'utilisateur se trouve face à un écran d'analyse où il écrit ou colle un texte.

Figure 1. *Interface de Direkt Profil 1.5.2 – Écran d'analyse.*

L'utilisateur appuie sur le bouton « Analyser le texte » et le programme identifie les structures

caractéristiques du développement grammatical du français. Après quelques secondes, le système affiche les résultats à un écran :



Figure 2. Interface de Direkt Profil 1.5.2 – Écran de résultats.

L'écran de résultats est divisé en deux parties. Dans la partie gauche, les structures détectées sont colorées selon des codes que nous avons conçus au préalable. La partie droite énumère les catégories d'analyses et le nombre d'occurrences de chaque catégorie. Grâce à un script Java, il est possible de cliquer sur le nombre d'occurrences pour faire apparaître une liste contenant les structures correspondantes dans le texte analysé. Cette fonction est nommée « Montrer occurrences » dans le système.

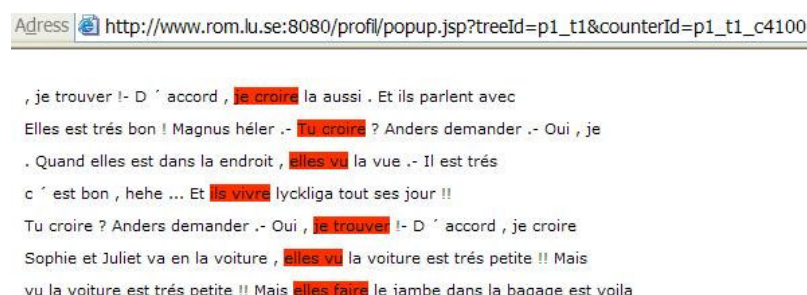


Figure 3. Interface de Direkt Profil – « Montrer occurrences » (ici formes verbales non-conjuguées)

2 Deux exemples d'analyses avec Direkt Profil

Dans cette section, nous présenterons deux études préliminaires qui visent les buts suivants : (a) tester les stades d'acquisition de Schlyter (2003) sur deux corpus différents, un corpus oral d'apprenants anglophones de français (Myles et al., 2002, *Linguistic Development Corpus*) et un corpus écrit (Ågren, 2005, le corpus *CEFLE*) et (b) donner un exemple de l'usage de *Direkt Profil* dans ces recherches.

Le tableau de Schlyter (2003) comprend plus de 25 catégories dont certaines proviennent de Granfeldt (2003). Nous travaillons actuellement à la réalisation d'un système qui prennent en compte tous ces phénomènes. Dans la version actuelle (1.5.2), il détecte surtout des catégories reliées au groupe verbal (finitude, temps et mode). L'analyse est limitée à des phrases avec un pronom sujet. La détection et l'analyse de sujets nominaux sont prévues pour la version 2.0. Pour la présente étude, les critères verbaux du Tableau 1 seront pris en considération.

Tableau 1. *Critères verbaux choisis pour l'évaluation des données (de Schlyter, 2003)*

ID.	Stades	1	2	3	4	5	6
A.	% de formes conjuguées (<i>je parle</i>) vs. (* <i>je parlE</i>) (% occurrences) (FLC) ⁴	50% - 75%	70-80%	80-90%	90-98%	+	+
B.	1-2-3 pers. singulier (<i>être/avoir</i>) est, a, va	formule sans opposition: j'ai/ c'est	opposition (ex. <i>j'ai</i> vs. <i>il a</i>)	erreurs isolés (ex. * <i>je va</i> * <i>je a</i>)	+	+	+
C.	1 ^{re} pers. pluriel (V- <i>ons</i>) (% correct)	-	70-80%	80-95%	erreurs dans des constr. complexes	+	+
D.	3 ^e pers. pluriel (<i>ont, font, sont, vont</i>) (% marquées)	-	* <i>ils est</i>	50%	60-80%	+	+
E.	Sujet + <i>viennent, veulent, prennent</i>	-	- (ex. * <i>ils prend</i>)	occs. isolées de V-(<i>n</i>)ent	50%	problèmes encore	+
F.	Mod+Inf/ FutProche FutSimple	-	+/-	Productif AG : FutS cas isolés	Fut Simple +/-	Fut Simple plus productif	+
G.	Imparfait	-	cas isolés, incorr	Impf mais souvent Prs pour Impf	<i>était, avait</i> Opp. PC/imp	<i>était, avait, modaux</i>	Impf avec diff V
H.	Conditionnel	formule <i>voudrais</i>	d :o	d :o	+/-	Cond plus productif	+
I.	Plus-que-parfait	-	-	-	apparaît utilis.incert	Ppf OK	+

Légende : + = productif ; - = n'apparaît pas ; +/- = commence à apparaître

Ce tableau est soumis à une révision continue et les taux sont encore approximatifs. Une des raisons en est que ces catégories et les détails de chaque catégorie sont le résultat d'une analyse de quelques 20 apprenants. Un des objectifs principaux de notre projet est d'évaluer ces critères sur un nombre plus large d'apprenants. Nous commençons ce travail dans la section suivante.

2.1 L'analyse du corpus oral de Myles et al. (Linguistic Development Corpus)

Il s'agit d'un corpus transversal de 20 apprenants formels anglophones de trois années scolaires. Les apprenants ont été enregistrés à l'oral. Au total, le corpus contient 60 apprenants. Les élèves avaient entre 13 et 16 ans et ils ont réalisé quatre tâches différentes :

⁴ Il y a une petite différence dans le calcul de cette catégorie ici et celui effectué par Schlyter (2003). Pourtant, cette différence n'empêche pas les comparaisons envisagées ici.

conversation autour d'une série de photos, une tâche interrogative, un récit après un film, une description d'une personne connue (pour faire produire la négation).

Pour la présente étude, nous avons retenu l'intégralité des tâches afin d'optimiser le nombre d'occurrences. Étant donné que *Direkt Profil* v.1.5.2 n'analyse que des groupes verbaux avec un pronom sujet, ce choix nous était nécessaire. Nous avons retenu les transcriptions de 38 apprenants. Les informations sur le corpus choisi se trouvent dans le Tableau 2.

Tableau 2. *Les données de LDC (Myles et al.)*

Nbr d'apprenants	Année scolaire	Nbr d'années d'études de français	Nbr de fichiers	Nbr de mots du corpus
19	9	3	76	env. 20 000
19	11	5	76	env. 25 000

2.1.1. Questions de recherche

Le corpus LDC de Myles et al. est un corpus transversal qui présente des données d'apprenants de trois années successives dans l'école anglaise : les années 9, 10 et 11. Nous testerons l'hypothèse selon laquelle les apprenants de l'année 9 avec 3 ans d'études de français se trouvent sur un stade moins avancé que les apprenants de l'année 11 avec 5 ans d'études de français.

Nous partirons d'un critère de la grille de Schlyter (Phénomène A du Tableau 1), le taux de *Formes Lexicales Conjuguées* (FLC) dans un contexte où le verbe doit être conjugué. La première question de recherche est de savoir si cette mesure parvient à distinguer significativement les deux groupes d'apprenants. Ensuite, les mesures de FLCs seront informellement testées contre les 8 autres indicateurs du système de finitude et du système tempo-aspectuel pour vérifier la corrélation entre les critères (les phénomènes B à I du Tableau 1). La deuxième question de recherche est ainsi de savoir s'il y a une corrélation entre le taux de FLCs et les autres critères.

2.1.2 Méthode

Les transcriptions ont été téléchargées du site Web de *flloc* (*French Learner Language Oral Corpora*)⁵. À l'origine, les transcriptions contenaient les énoncés des apprenants et de l'expérimentateur. À l'aide du logiciel *Combo* de *CLAN*, nous avons d'abord enlevé les énoncés de l'expérimentateur (en utilisant une combinaison des options +d et -t, voir MacWhinney, 2000). Le résultat de cette opération sert d'entrée à *Direkt Profil* dont voici un extrait :

(2) Extrait des données d'entrée soumises à *Direkt Profil*

[...]

*48N: elle a la cheveux blonde et longue .

*48N: mmm <elle aime manger> [/] elle n' aime pas manger les glaces .

*48N: et elle va # um # [%eng: horseriding] [+eng] .

*48N: elle aime faire@g de@g l' équit(ation) l' équitation@g .

*48N: et # euh elle n' aime pas regarder à le [?] télé .

*48N: um <elle est> [/] um elle habite avec Guy Ritchie .

⁵ <http://www.flloc.soton.ac.uk/index.html>

*48N: et ils sont très um famous@a +/.

*48N: +, célèbre@g .

*48N: [%eng: yeh] [+eng] .

Ce type de texte contient un assez grand nombre de codes, commentaires, séquences en anglais, etc. qui ne sont pas intéressants pour nos analyses et qui risquent de perturber les résultats d'un étiqueteur « ordinaire ». Mais grâce à notre système d'étiquetage « local », *Direkt Profil* ne tient compte que des structures qui sont pertinentes pour les itinéraires d'acquisition en excluant le reste. Pour un système qui utilise une approche globale à l'étiquetage, ce genre de texte ne serait pas possible à traiter (voir Granfeldt et al., 2005 pour une discussion de ces différentes approches).

2.1.4 Les résultats

Après l'analyse, *Direkt Profil* produit l'écran de résultats suivant :

Année 9 – 3 ans d'études de français	Année 11 - 5 ans d'études de français
Code couleur Regarder le résultat en mode non connecté Coloration de l'analyse active Ôter coloration de l'analyse active ▼ Statistiques des phrases, des mots et des expressions figées Nombre de mots ab 28099 Nombre de phrases ab 7513 Longueur moyenne des phrases ab 3 Expressions figées ab 112 ▼ Analyse des groupes verbaux y inclut des pronoms Phrases sans verbe ab 394 Verbe sans accent ab 42 ☐ Total: Verbes conjugués et non conjugués ab 807 Verbes non conjugués ab 167 ☐ Total: Verbes conjugués ab 640 ☐ Total: Temps et modes simples (précoces) ab 636 ☐ Total: Verbes au Présent ab 573 ☐ Total: Verbes au Passé composé ab 81 ☐ Total: Verbes modaux auxiliaires + infinitif ab 3 ☐ Total: Verbes à l'Imparfait ab 5 ☐ Total: Verbes au Futur simple ab 0 ☐ Total: Temps et modes complexes (tardifs) ab 1 ☐ Total: Verbes au Plus-que-parfait ab 0 ☐ Total: Verbes au Conditionnel ab 1 Être / avoir conjugués à d'autres temps ab 2	Code couleur Regarder le résultat en mode non connecté Coloration de l'analyse active Ôter coloration de l'analyse active ▼ Statistiques des phrases, des mots et des expressions figées Nombre de mots ab 26161 Nombre de phrases ab 5855 Longueur moyenne des phrases ab 4 Expressions figées ab 278 ▼ Analyse des groupes verbaux y inclut des pronoms Phrases sans verbe ab 229 Verbe sans accent ab 25 ☐ Total: Verbes conjugués et non conjugués ab 1208 Verbes non conjugués ab 94 ☐ Total: Verbes conjugués ab 1114 ☐ Total: Temps et modes simples (précoces) ab 1111 ☐ Total: Verbes au Présent ab 964 ☐ Total: Verbes au Passé composé ab 115 ☐ Total: Verbes modaux auxiliaires + infinitif ab 19 ☐ Total: Verbes à l'Imparfait ab 31 ☐ Total: Verbes au Futur simple ab 12 ☐ Total: Temps et modes complexes (tardifs) ab 4 ☐ Total: Verbes au Plus-que-parfait ab 1 ☐ Total: Verbes au Conditionnel ab 3 Être / avoir conjugués à d'autres temps ab 1

Figure 4. Parties droites des écrans de résultats pour le LDC (Myles et al.) Année 9 et Année 11.

Après une inspection détaillée des résultats globaux à l'aide de la fonction « montrer occurrences », nous avons pu écarter certaines occurrences analysées incorrectement. Le résultat par rapport aux critères d'évaluation choisis (Tableau 1) est présenté *infra*.

Tableau 3. *Le résultat du Linguistic Development Corpus (Myles et al.)*

ID	Critères	Année 9 - 3 ans d'études	Année 11 – 5 ans d'études
A	Formes lexicales conjuguées vs. non-conjuguées : (% occurrences) (FLC)	formes lexicales non conjuguées / total formes lexicales = 172 / 661 = 26 % de formes non-conjuguées et 74 % de FLCs	formes lexicales non conjuguées / total formes lexicales = 71 / 712 = 10 % de formes non-conjuguées et 90% de FLCs
B	1-2-3 pers singulier, <i>est, a, va</i> (<i>con/aux</i>)	acquis (<i>j'ai</i> vs. <i>il a</i>)	acquis (<i>j'ai</i> vs. <i>il a</i>)
C	<i>nous Vous</i> accord marqué (% occs.)	-	-
D+E	<i>ont font sont et Suj+ viennent, veulent, prennent</i>	-	3 occurrences correctes sur 5 = 60%
F	Mod+Inf/ FutProche	-	7 Mod + Inf / 8 Fut Proche et 12 FutSimple
G	Imparfait	Cas isolés avec <i>être / avoir</i> non productif	Productif avec <i>être / avoir</i> . Cas isolés des verbes lexicaux
H	Conditionnel	1 occurrence	3 occurrences (= <i>j'aimerais aller</i>)
I	Plus-que-parfait	-	1 occurrence

Une analyse de variance (ANOVA) suivi d'un Tukey's Parwise Comparison (95,0% Simultaneous Confidence Intervals) montre que le taux de FLCs distingue significativement l'Année 9 de l'Année 11 ($F(4,937)$, $p = 0,0089$). Pour une liste d'exemples de formes lexicales non conjuguées voir l'Appendice, Tableau A. Pour les résultats individuels des apprenants, se reporter à l'Appendice, Tableau B.

Après 3 ans d'études de français, ces apprenants produisent à l'oral autour de 74% de FLCs dans des contextes où les verbes sont normalement conjugués (écart type de 17,8%). Ce taux de FLCs les situe entre les stades 1 et 2. La maîtrise de la distinction en personne au singulier des verbes avoir et être (B) et les quelques occurrences d'imparfait avec les mêmes verbes (G) sont d'autres critères qui sont compatibles avec le Stade 2. Il n'y avait pas de contexte pour vérifier les critères (C) et (D+E) du Tableau 1.

Après 5 ans d'études de français, ces apprenants produisent 90% de FLCs (écart type de 10%). Ce taux les situe entre les Stades 3 et 4. Les autres critères compatibles avec le Stade 3 incluent la présence de Mod+Inf et de Futur Proche (F) et les quelques exemples de Conditionnel (H). Ceci pointe vers le Stade 4 ainsi que la production de l'Imparfait avec *être* et *avoir*, les quelques exemples d'Imparfait avec verbes lexicaux, la présence Futur simple (mais souvenons-nous qu'il s'agit d'apprenants formels chez qui cette construction apparaîtrait déjà au Stade 3 selon le Tableau 1).

Nous concluons que le taux de FLC distingue les deux groupes de manière significative. En

plus, cette mesure montre une corrélation préliminaire, dans les deux groupes, avec la maîtrise des personnes au singulier (verbes non-lexicaux) et avec l'apparition de l'imparfait.

Ces résultats suggèrent que le taux de FLCs chez les apprenants anglophones du français ne semblerait pas plus élevé que chez les apprenants suédophones. Schlyter a élaboré les stades d'acquisition pour les apprenants suédophones à l'oral. Si une corrélation entre plusieurs critères apparaît aussi à l'intérieur d'un stade donné pour des apprenants anglophones, ceci serait un argument en faveur de la généralité de certains itinéraires d'acquisition. De plus, ce serait un argument contre l'influence morphologique de la L1 sur certains aspects de la production en L2. Prévost (2004) émet l'hypothèse que le taux très élevé de Formes Lexicales Non Conjuguées (FLNCs) chez les apprenants anglophones serait lié au fait qu'en anglais la finitude n'est pas marquée morphologiquement sur les verbes lexicaux. En revanche, cette distinction est marquée en suédois et l'on s'attendrait alors à une différence entre les deux groupes d'apprenants. Ces résultats préliminaires ne confirment pas cette hypothèse. Plusieurs réserves sont pourtant nécessaires dont la plus importante est que nous n'avons analysé que les sujets pronominaux dans cette étude.

2.2 L'analyse du corpus écrit de Ågren – Le corpus CEFLE

Il s'agit d'un corpus de textes écrits par 85 apprenants suédophones du français langue étrangère (« Corpus Écrit de Français Langue Étrangère de Lund ». Ågren 2005). En outre, il comprend des textes écrits par 22 jeunes Français. Le corpus contient environ 100 000 mots. Les textes qui le composent sont des récits de longueur et de niveaux variés. Ågren a classé tous les apprenants selon les stades de Bartning & Schlyter (2004).

Pour la présente étude, nous avons retenu les textes des apprenants des stades 2 et 4 sur le « Voyage en Italie » (cf. Ågren, ce volume).

Tableau 4. *Les données du CEFLE*

Stade	Nbr d'apprenants	Nbr d'années d'études	Nbr de mots du corpus
2	25	1-3	env. 5 800
4	15	3-4	env. 6 600

2.2.1. Question de recherche

Ågren a classé les apprenants de son corpus selon les critères verbaux de Schlyter & Bartning (2004). Les critères de nombre (cf. D+E dans le Tableaux 1) n'ont pas été inclus (cf. Ågren, ce volume). À l'aide de *Direkt Profil*, nous évaluerons si le taux de *Formes Lexicales Conjuguées* (FLC) montre une corrélation avec les critères du nombre. Deuxièmement, il faut noter qu'Ågren a établi le stade de chaque apprenant en prenant en considération les phrases avec des sujets nominaux et pronominaux. Ici, suite aux limitations actuelles de *Direkt Profil*, nous n'analyserons que les phrases ayant un sujet pronominal. Une différence de taux de FLC dans notre analyse, par rapport au classement d'Ågren, pourrait ainsi être liée à un traitement différent de ces deux types de sujets par les apprenants.

2.2.2 Méthode

Nous avons d'abord sélectionné les textes sur « Voyage en Italie » des apprenants au Stades 2 et 4 dans CEFLE. Nous les avons fait passer dans *Direkt Profil*. Pour écarter les occurrences incorrectement analysées, nous avons utilisé la fonction « montrer occurrences ».

2.2.3 Les résultats

Le Tableau 5 présente les résultats globaux :

Tableau 5. *Résultats du CEFLE*

ID	Critères	Stade 2	Stade 4
A	Formes conjuguées vs. non-conjuguées : (% occurrences) (FLC)	formes lexicales non conjuguées / total formes lexicales = 24 / 224 = 9,6% de formes non-conjuguées et 90,2% de FLCs	formes lexicales non conjuguées / total formes lexicales = 1 / 293 = 0,4 % de formes non-conjuguées et 99,7% de formes conjuguées
B	1-2-3 pers singulier, <i>est, a, va</i> (cop/aux)	acquis	acquis
C	<i>nous Vous</i> accord marqué (% occs.)	9 sur 10 = 90%	très peu d'occurrences
D	<i>ont font sont et vont</i>	54 sur 71 = 76%	142 sur 146 = 97%

Plusieurs informations sont accessibles à partir de ce Tableau 5. D'abord un ANOVA suivi d'un Tukey's Parwise Comparison (95,0% Simultaneous Confidence Intervals) montre que la différence entre les Stade 2 et 4 du taux de FLCs (A) est significative ($F(5,75)$, $p = 0,022$). Par contre, il n'y aucune différence significative entre les groupes pour les formes correctes en *V-ont* (D) ($F(2,60)$, $p = 0,116$). Pour les résultats individuels des apprenants, se reporter à l'Appendice, Tableau C. Un test de corrélation linéaire (PEARSON) montre qu'il n'y a pas de corrélation significative entre le taux de FLCs et les formes correctes en *V-ont* (D) au Stade 2 ($p=0,835$). Au Stade 4, par contre, cette corrélation (A et D) est significative ($p=0,040$).

Deuxièmement, nous observons que pour le Stade 4 le moyen du taux de FLCs ne correspond pas avec ce Stade. Il est plus élevé et indiquerait plutôt un Stade 5 (cf. Tableau 1). La même chose peut être dit sur le Stade 2. Ceci peut être dû au fait que l'analyse ne prend en considération que les sujets pronominaux. Les résultats suggéreraient alors que les FLNCs sont pratiquement absentes avec les sujets pronominaux au Stade 4 et que les FLNCs qui restent sont produites avec des sujets nominaux.

Troisièmement, par rapport aux autres critères sur le nombre, nous observons que les deux groupes ont acquis l'opposition de personne au singulier (B). Pour ce qui est du critère (D), Schlyter propose pour l'oral une proportion d'environ 50% d'accord correct sujet-verbe avec *ont, sont, font, vont* au Stade 2. À l'écrit, les apprenants de ce stade arrivent à le marquer à 76%. Il se peut qu'il s'agisse ici d'une différence réelle entre les deux modes de production. Cette hypothèse est soutenue par le fait que les apprenants du Stade 4 se trouvent à 97% d'accord correct sujet-verbe dans cette catégorie (D), à comparer avec les 60-80% proposé pour Stade 4 à l'oral.

Pour terminer cette section nous mentionnons brièvement une autre interprétation possible qui dirait que les propriétés du sujet influent sur l'accord verbal. Souvenons-nous que nous n'avons pris en considération ici que les sujets pronominaux. Pour le *V-ont* il y a, en principe, cinq configurations possibles (abstraction faite de la négation et des clitiques) :

- (3) ils/elles ____ont
- (4) [NP_plur]____ont
- (5) [[NP_sing] et [NP_sing]] ____ont
- (5') [[NP_sing] et [NP_plur]] ____ont
- (6) [[NP_plur] et [NP_sing]] ____ont

Dans le premier cas, le verbe est précédé par le pronom sujet qui ne varie qu'en genre. Pour le verbe à droite, il s'agit donc d'un contexte bien défini qui semble correspondre à un « slot » possible. Les entrées de « ils » / « elles » sont annotées avec les traits [3^{ème} personne] [-singulier] dans le lexique mental.¹ Ces traits sont donc disponibles toute suite après l'accès lexical. Sous l'hypothèse que le nombre est lexical et non pas syntaxique (i.e. il y a des entrées lexicales distinctes pour « chat » et « chats ») la configuration (4) pourrait être traitée selon la même procédure. Les deux configurations suivantes (5 et 5') avec deux NPs coordonnés où le premier est doté du trait singulier représentent des cas qui demandent au moins une opération supplémentaire. L'apprenant doit d'abord passer par un calcul interne du nombre du NP sujet complexe pour ensuite procéder à un processus d'accord interphrastique avec le verbe qui suit. Notre observation est simplement qu'avec les SNs sujets, il existe plus de variations dont l'apprenant en planifiant/produisant (ou traitant) son énoncé doit tenir compte. Le résultat en est peut-être que ces contextes ne s'apprennent pas à la définition d'un « slot » pour le verbe. En tout cas, ce que nos résultats montrent, c'est que la configuration (3) est différente et elle semble dans ce sens proche au contexte « nous ____ons » discuté par Schlyter & Bartning (ce volume). Au niveau d'analyse où nous sommes actuellement avec *Direkt Profil*, nous ne pouvons pas distinguer (4) et (5) mais cette analyse sera possible dans l'avenir.

Pour terminer, nous remarquons que ces configurations peuvent servir de base pour vérifier le lieu du trait de nombre. Si le nombre est lexical, les configurations (3) et (4) seront traitées ensemble et contrasteront avec (5). Par contre, si le nombre est syntaxique et l'on accepte l'hypothèse vraisemblable que les pronoms sont dotés de leur trait de nombre dans l'entrée lexicale, (3) et (4) sont des configurations distinctes.

2.4 Résumé des résultats

Il faut encore souligner que ces deux études sont préliminaires. Nous en retenons surtout trois résultats. D'abord, le taux de FLCs est, à l'oral et à l'écrit, une mesure valide pour définir des niveaux linguistiques et distinguer les apprenants. Sur les données orales des anglophones, cette mesure arrivait à distinguer les apprenants d'année 9 des apprenants d'année 11. Sur les données écrites du CEFLE, cette mesure résulte en une différence significative lorsqu'on compare les apprenants des Stades 2 et 4. Dans cette étude, la différence entre les Stades 2 et 4 en taux de V-ont correct n'était pas significative à l'écrit.

Deuxièmement, nous pouvons inférer de ces analyses que le taux de FLCs est vraisemblablement plus élevé avec un sujet pronominal qu'avec un sujet nominal. Ceci semblerait surtout être le cas dans les stades plus avancés.

Troisièmement, il n'y avait pas corrélation significative entre les taux de V-ont et de FLCs au Stade 2 à l'écrit alors que cette corrélation était significative au Stade 4.

¹ Les noms communs et les noms propres ont tous la 3^{ème} personne et si nous la traitons comme la personne par défaut, « ils » / « elles » auront les traits [-personne] [-sing]. Nous pensons que cette dernière est une annotation psycholinguistiquement plus correcte.

3 *Direkt Profil* et les recherches à venir

Dans cette section, nous présentons d'abord quelques détails du système pour ensuite développer quelques réflexions sur l'apport de *Direkt Profil* par rapport aux autres projets de recherches présentés dans ce volume.

3.1 Le moteur analyseur, les règles et le lexique

Le système *Direkt Profil* est modulaire et se compose de trois parties : un moteur d'analyse, un ensemble de règles et un lexique des formes fléchies. Le moteur d'analyse lit les règles que nous avons écrites manuellement. Dans la version actuelle, il y a trois ensembles de règles : le premier segmente le texte en mots, le deuxième identifie les expressions figées (en s'appuyant sur une liste que le chercheur peut modifier), le troisième ensemble étiquette les parties des discours et annote les groupes syntaxiques. Le moteur s'appuie sur un lexique de formes fléchies provenant de l'ABU.² Nous avons modifié le lexique original de plusieurs façons et nous l'avons reformaté en XML. Nous avons par exemple ajouté des informations sur la racine de chaque mot afin de pouvoir faire des recherches portant sur les radicaux des verbes. Dans l'état actuel, une entrée du dictionnaire pour un verbe à la forme suivante :

```
(7)
<wordform entry="voyaient" lemma="voir" pos="ver">
  <stem>voy</stem>
  <feature>iimp:pl:p3</feature>
</wordform>
```

Les règles représentent des structures partielles de groupe et sont divisées en une partie <condition> et une partie <action>. La partie condition contient les paramètres de recherche. Il peut s'agir d'un lemme, d'une expression régulière ou d'une classe de flexion. Le moteur parcourt le texte et applique les règles à l'aide d'un arbre de décision. Il exécute la partie condition pour identifier les séquences de mots contigus. Chaque règle produit un résultat positif (« *match* ») ou négatif (« *no match* »). Ils sont exécutés suivant le résultat de la condition et ont pour effet d'annoter le texte, de compter le nombre d'occurrences du phénomène et d'enchaîner une autre règle. En parcourant les nœuds de l'arbre, le moteur mémorise les règles qu'il a parcourues sur son chemin ainsi que les résultats des parties condition de ces règles. Quand il arrive à un nœud terminal, le moteur applique les parties action de toutes les règles.

(8) Exemple d'une règle :

```
<rule id="p1_t1_r300_verb_lex_imperfect">
  <description>Check if the verb has imperfect tense.</description>
<search framesize="0">
  <inflection category="verb">
    <tense value="imperfect"/>
  </inflection>
</search>
<action>
```

² <http://abu.cnam.fr>

```
<match nextrule="p1_t1_r350_accord_lex_imperfect">
  <incrcounter value="p1_t1_c5430"/>
  <incrcounter value="p1_t1_c5400"/>
  <incrcounter value="p1_t1_c5000"/>
</match>
  <nomatch nextrule="p1_t1_r310_verb_lex_conditional"/>
</action>
</rule>
```

La partie <search> de la règle utilise ici deux informations qui se trouvent dans le dictionnaire : la partie du discours (pos = « ver ») et l'information temporelle de la forme décrite dans l'attribut <feature>.

3.2 Développements possibles

L'avantage du système modulaire est qu'il permet d'apporter des modifications aux trois ensembles des règles et au lexique sans changer le moteur d'analyse. Ceci veut dire, en principe, qu'un chercheur qui s'intéresse à une question spécifique pourrait développer des nouvelles règles et, si nécessaire, rajouter des informations dans le lexique. Par rapport aux domaines de recherches présentés dans ce volume, nous voyons un certain nombre de possibilité de ce genre :

1. Plusieurs contributions de volume (Ågren, Thomas, Hedbor) concernent le marquage de la 3^e personne au pluriel. Pour chaque verbe du lexique, les informations de personne et de nombre sont déjà spécifiées et il suffira d'écrire quelques nouvelles règles du type <condition> / <action> pour avoir un résultat (et un compteur) séparé de cette catégorie.

2. Thomas (ce volume) s'intéresse à la relation entre la sémantique prototypique d'un verbe et la production des apprenants. Sa première hypothèse est que les verbes, selon leur aspect inhérent (Aktionsart), sont utilisés suivant des schémas spécifiques au niveau de la forme. Sa deuxième hypothèse est que les apprenants dans leur propre production refléteront ces schémas. Ainsi, les verbes d'état par exemple seront plus, et peut-être plus vite, utilisés dans leur forme conjuguée (forme courte), comme un reflet de cette entrée.

Si on attribue aux verbes du lexique des informations sur leur catégorie sémantique typique (état, activité, etc.), ce serait ensuite possible de prendre en compte ces informations dans la formulation des règles. Il faut dire que la modification du lexique (ici le rajout d'un trait supplémentaire (« feature »)) peut aussi se faire de manière automatique. Dans la version actuelle, il est d'ailleurs déjà possible de faire passer des corpus écrits et oraux et avoir une liste détaillée de formes non-conjuguées (voir Tableau A, de l'Appendice).

3. Hedbor (ce volume) utilise *ScriptLog* pour son analyse de la production écrite des apprenants. Ce logiciel peut produire des fichiers « rtf » qu'il est possible de charger ensuite dans *Direkt Profil*. Nous voyons plusieurs interactions intéressantes entre les deux systèmes. Si les sujets produisent leur texte dans *ScriptLog*, il peut ensuite être analysé dans *Direkt Profil*. Le chercheur peut ensuite revenir au *ScriptLog* pour des informations plus détaillées sur la production d'un exemple spécifique, par exemple en utilisant la fonction « Transition time String Context ».

4. Ågren étudie le marquage du pluriel dans le groupe nominal et le groupe verbal. La version

suivante de *Direkt Profil* inclura une analyse des accords de nombre et genre à l'intérieur du groupe nominal ce qui permettra des analyses à grande échelle de cette catégorie.

4 Conclusion et travaux futurs

Dans cet article, nous avons présenté le système de *Direkt Profil*. Il s'agit d'un analyseur morphosyntaxique pour le FLE. Nous avons d'abord présenté le fonctionnement général du système. Dans la section 2, nous avons présenté les résultats préliminaires de deux études sur la morphologie verbale des apprenants où nous avons utilisé *Direkt Profil* pour nos analyses. La première porte sur des données orales des apprenants anglophones (corpus de Myles et al.) et la seconde sur des données écrites des apprenants suédophones (corpus de Ågren). Par ces deux analyses, nous avons voulu montrer que notre système peut être utilisé pour les deux types de production (l'oral et l'écrit). Pour un résumé des résultats de ces études, voir la section 2.4, *supra*.

Par rapport à nos questions de recherches générales (cf. Introduction, *supra*), les travaux futurs se centreront d'abord sur un élargissement des critères détectés par le système. Nous estimons que vers la fin de 2005, le système détectera un nombre suffisant de phénomènes pour pouvoir commencer à traiter la définition des « Profils » grammaticaux. La démarche que nous voulons suivre pour ce travail utilisera des techniques d'apprentissage automatique à partir des textes et des mesures de régularité statistique.

5 Références

- Bartning, I. (1997) « L'apprenant dit avancé et son acquisition d'une langue étrangère ». *Acquisition et Interaction en Langue Étrangère* 9 : 9-50.
- Bartning, I. et Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies*, 14:3:281-299.
- 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 » TALN 2005, 6-11 juin. Dourdan, France.
- MacWhinney, B. (2000) *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Myles, F. et al. (2002) « Corpus d'interlangue française : outils de gestion et d'analyse » Journée d'étude de l'ATALA : Constitution et exploitation de corpus du français parlé Paris, France.
- Prévost, P. (2004) « Morphological variation in early adult second language French. A cross-sectional study ». *EUROSLA Yearbook* 4 : 147-175.
- Schlyter, S. (2003, Ms) « Stades de développement en français L2. Exemple d'apprenants suédophones, guidés et non guidés, du Corpus Lund ». Institut d'études romanes de Lund. Disponible sur www.rom.lu.se/durs/archive.htm.
- Ågren, M. (2005, Ms) « L'accord morphologique du nombre dans la phrase nominale. Une

étude sur l'acquisition du français L2 écrit »: Institut d'études romanes de Lund.

Tableau A. Exemples de FLNC chez Année 9 (gauche) et Année 11 (droite) dans LDC (Myles et al.)

i cream?]. gla@ g. il ne manger pas de gla.[% eng: how : " E_10_LN_2002_01_29WAV" _550293_559190. um. um elle jouer tennis. elle boit thé. um um et le[?]. monstre vrai ok?. il aimer coca cola et+/.+. fruits. ehm right pose xx dan or something]. je danser. je ne pas à la piscine la télé. je écouter le musique. je jouer mon ami. um[% eng: ne pas jouer au golf. um um elle ne pas de fumer[?]. um elle boit de thé . il est euh très intéressant. um je rester@ f en maison les vacans j football. um je regarder la télé. je écouter le musique. je jouer mon ami . elle boit@ g le thé. elle ne regarder pas la télévision.[% eng: what ne chanter pas. il nager.< il ne regarder télévision pas> [/]/[% eng: 'est que> qu'est que on prendre la photo?. ok um xx. le weekend je jouer au football. um je regarder la télé. je écouter le musique télé. elle manger de fruits. um elle ne pas jouer au golf. um um elle ne fumer pas. il ne chanter pas. il nager.< il ne regarder télévision pas] le. sept de février. um je acheter le seamonkeys@d. xx pour au fruits. il ne fumer pas. il ne chanter pas. il nager.< il at the weekend]. à le weekend je jouer au football. um je regarder la pas joue dans la piscine. um oh elle ne pas monter de la gare.[% eng: um . elleelle ne pas regarder la télé. elle manger de fruits. um elle ne pas . j'habite à Sandford. euh nous aller le maison de grand mère. euh s ride?]. um. xx. elle monter le cheval um[% eng: what dans le oui d'accord. um ils prendre le bateau+/. mmm.+ um en au vin. il manger au fruits. il ne fumer pas. il ne chanter pas. euh j'aime regarder la. télé je jouer au piano. um.[% eng: . um il nene bu au vin. il manger au fruits. il ne fumer pas @ d dans le lac. um. il dit c' est le monster@d french?].+. oh le jardin. um il nene bu au vin. il manger au fruits journaliste prené des photos. les autres enfants ils regarder la télé les regarder le monstre. um what's running running?].< elle ne s [/]/ elleelle courir. um elle ne pas joue dans work so should I say+]. ok elle ne secréter@ v[?] pas.[% eng: . il ne pas mange la gla. il ne pas regarder la télévision. il aime le football [% eng: I know this word] ils visiter une ville le mère. regarder les	nouvelle année? oh um nous allons euh nous rester à la maison. ok. mais lac. vrai@ g. ok. elle jouer au tennis. elle elle elle aime pas boit le vin. et elle joue elle ne jouer pas de golf. um elle est la photo? ah oui. euh oui je@ n aller Espagne dans octobre le euh[% eng courts? um qu'est qu' elle le cheveux frisés?[% eng: pardon?]. . je ne fumer la cigarette xxx. je@ n jouer le en(fant) enfants[*]]. comment@ g est@ g il@ g habillé@ g? ok où est le il vrai. il[?] football. euh il jouer xx Manchester United.[% eng: bien nous nageons dans le le mer et nous jouer le ping pong et le foot et je ne sais pas jouer au golf. je ne fumer la cigarette xxx. je@ n près de Sandford. eh non eh eh je vu eh à les vac(ans) ne sais pas. non. euh oui l'aider à la maison euh je fais passer um il ne pas le vin rouge. il ne pas le manger le gla. il ne pas le eng: and riding a horse?]. je monter@ g sur[?] le@ g il ne il ne um chanter pas. je ne jouer pas de basket. il um il m(ère) le mère et enfant qui surprise. beaucoup de monde prennent des photos pas ordinateur. je@ n chanter. je je ne pas rester le maison. je ne sais pas pas l'alcool[*]. et euh il passer l' aspirateur[% eng: um . il ne mange pas de glas. il ne il ne um chanter pas. je ne jouer pas de um. je ne sais pas ordinateur. je@ n chanter. je je ne pas rester le what's smoking in french?]. je ne pas fumer@ g. la mange orange. . ok um elle est une chanteuse. elle jouer le tennis. elle a une enfant singer]. elle chanteuse@g. elle jouer au tennis. euh il elle n avec ma amie. euh dans le weekend je@ n euh jouer au basket ou aller skating@ v les pommes[*]. pommes@g. je@ n chanter.. um. je ne sais je jouer@ p au foot. et je jouer au hockey. non. ah Sou elle ne boit pas le vin. um elle manger une pomme. um. ok um m(a) ma famille. euh ils habiter euh dans euh à Bradford. dans . il mange beaucoup de fruits. et il ne regarder pas la télévision. et il ne
---	--

Tableau B. *Les résultats individuels de LDC (Myles et al.)*

Groupe Année 9 – 3 ans d'études (N=19)			
Appr.	Verbes lexicaux (total)	Formes lexicales non- conjuguées	%FLC
1	36	16	55,56%
2	60	24	60,00%
3	21	13	38,10%
4	58	12	79,31%
5	41	9	78,05%
6	43	7	83,72%
7	60	23	61,67%
8	30	1	96,67%
9	50	12	76,00%
10	50	25	50,00%
11	12	2	83,33%
12	51	18	64,71%
13	15	1	93,33%
14	37	1	97,30%
15	9	1	88,89%
16	12	4	66,67%
17	12	2	83,33%
18	39	0	100,00%
19	25	1	96,00%
Total / Moyen	661	172	73,97%

Groupe Année 11 – 5 ans d'études (N=19)			
Appr.	Verbes lexicaux (total)	Formes lexicales non- conjuguées	% FLC
1	40	3	92,50%
2	19	3	84,21%
3	22	1	95,45%
4	43	3	93,02%
5	63	9	85,71%
6	29	2	93,10%
7	31	5	83,87%
8	58	1	98,27%
9	67	0	100,00%
10	28	5	82,14%
11	60	4	93,33%
12	37	2	94,59%
13	42	3	92,86%
14	14	4	71,43%
15	68	2	97,06%
16	27	9	66,66%
17	24	5	79,17%
18	26	9	65,38%
19	14	1	92,86%
Total / Moyen	712	71	90,02%

Tableau C. *Les résultats individuels de CEFLE*

Stade 2 (N=25)	Verbes lexicaux	Formes lexicales non- conjuguées				
Appr.	(total)		% FLC	V-ont	*V-ont	%V-ont
1	4	1	75,00%	2	1	66,6%
2	8	0	100,00%	3	0	100,00%
3	7	1	85,71%	9	0	100,00%
4	10	0	100,00%	1	0	100,00%
5	15	3	80,00%	3	0	100,00%
6	10	0	100,00%	5	4	56%
7	8	0	100,00%	0	0	0,00%
8	12	0	100,00%	0	0	-
9	7	1	85,71%	2	1	66,6%
10	4	0	100,00%	2	0	100,00%
11	7	0	100,00%	2	0	100,00%
12	4	0	100,00%	2	0	100,00%
13	16	0	100,00%	5	0	100,00%
14	8	0	100,00%	0	3	0,00%
15	4	0	100,00%	2	0	100,00%
16	10	5	50,00%	1	0	100,00%
17	23	7	69,56%	0	0	-
18	4	1	75,00%	2	0	100,00%
19	13	0	100,00%	0	0	-
20	8	0	100,00%	2	2	50,00%
21	9	2	77,77%	2	1	66,6%
22	6	0	100,00%	0	0	-
23	8	0	100,00%	1	3	25,00%
24	12	0	100,00%	3	1	75,00%
25	7	3	57,14%	5	1	83,33%
Total / Moyen	224	24	90,24%	54	17	54,97%

Stade 4 (N=15) Appr.	Verbes lexicaux (total)	Formes lexicales non- conjuguées	%FLC	V-ont	*V-ont	%V-ont
1	13	0	100,00%	9	0	100,0%
2	19	1	94,74%	5	1	83,33%
3	14	0	100,00%	4	0	100,0%
4	13	0	100,00%	8	0	100,0%
5	9	0	100,00%	5	0	100,0%
6	21	0	100,00%	17	0	100,0%
7	16	0	100,00%	10	0	100,0%
8	47	0	100,00%	22	2	90,90%
9	16	0	100,00%	6	0	100,0%
10	10	0	100,00%	7	1	87,50%
11	21	0	100,00%	11	0	100,0%
12	12	0	100,00%	7	0	100,0%
13	19	0	100,00%	4	0	100,0%
14	36	0	100,00%	17	0	100,0%
15	26	0	100,00%	10	0	100,0%
Total / Moyen	292	1	99,65%	142	4	97,45%

III

Un regard sur la 3^{ème} personne du pluriel

Approche psycholinguistique de la mise en œuvre de la morphologie flexionnelle du nombre verbal en français langue maternelle

Pierre Largy

Université de Toulouse le Mirail (France)
Laboratoire P.C.S. –
largy@univ-tlse2.fr

1 Introduction

Cet article propose un survol d'une série de recherches conduites sur la production écrite de l'accord en nombre sujet/verbe en français langue maternelle par des adultes ayant une pratique régulière de l'écrit. Ces recherches ont principalement opposé deux situations, la production écrite de la troisième personne du singulier (vs. la production écrite de la troisième personne du pluriel) de verbes dits du "premier groupe" (i.e., verbes réguliers avec un infinitif en –er) au présent de l'indicatif. Elles considèrent ainsi contraster la production d'un accord verbal au singulier (avec sujet nominal: "Le chat miaule" ou pronominal: "Il chante") avec la production d'un accord verbal au pluriel ("Les chats miaulent" ou "Ils chantent"). Il convient toutefois de garder à l'esprit que les résultats ne valent que pour le cas de la troisième personne. En effet, dans le cas de la première ou de la deuxième personne, l'opposition singulier/pluriel se traduit par une distinction audible des flexions verbales du nombre (respectivement: –e vs. –ons ["je chante" vs. "nous chantons"], –es vs. –ez ["tu chantes" vs. "vous chantez"]), ce qui facilite leur production écrite. Ce n'est pas le cas pour la troisième personne où l'opposition singulier/pluriel est inaudible: "il chante" se prononce exactement comme "ils chantent". Ainsi, opposer le marquage du singulier au marquage du pluriel dans le cas de la troisième personne, c'est se donner l'occasion d'étudier la production écrite de marques flexionnelles du nombre, dans une situation où la production des flexions ne peut s'expliquer par un simple phénomène de transcription phonème/graphème. Autrement dit, c'est se donner l'occasion d'étudier les processus à l'origine du marquage du nombre verbal, indépendamment d'un simple processus de guidage phonologique.

Les travaux rapportés ci-après exploitent le caractère silencieux de l'opposition de nombre (singulier vs. pluriel) des flexions verbales de la troisième personne (–e vs. –nt), d'une part pour mettre en évidence que la production écrite de l'accord verbal fait l'objet d'un monitoring de la part du scripteur, d'autre part pour mieux appréhender le phénomène d'encodage grammatical (*grammatical encoding*: Levelt, Roelofs & Meyer, 1999), c'est-à-dire les processus conduisant à la sélection d'une flexion particulière.

2 Aspects méthodologiques

Les recherches sur la production écrite de l'accord sujet/verbe partagent une même méthodologie. Conduites dans une perspective de psychologie expérimentale, elles analysent les performances orthographiques de sujets adultes sains et tentent ainsi de caractériser les procédures expertes de mise en œuvre de l'accord en nombre sujet/verbe. Ces recherches concernent la production de phrases isolées et reposent sur un paradigme expérimental d'induction d'erreurs. Il est demandé aux participants de rappeler par écrit des phrases qu'ils viennent juste d'entendre. En faisant varier la structure syntaxique des phrases et les conditions dans lesquelles les scripteurs sont amenés à les produire (par exemple, en leur imposant la gestion d'une tâche secondaire parallèlement au rappel écrit d'une phrase), on génère des erreurs d'accord verbal dans la production écrite de ces adultes lettrés. La distinction de différentes catégories d'erreurs et l'analyse des conditions d'occurrence de ces erreurs, permettent de faire des inférences sur la nature des processus impliqués dans la production de la morphologie flexionnelle du nombre verbal.

3 Le paradigme de la double tâche

L'expérience de Fayol et Got (1991) ouvre la voie d'une longue série de recherches sur la production écrite de l'accord sujet/verbe en français langue maternelle. Son paradigme expérimental dit "de la double tâche" sera largement exploité dans le cadre de ces recherches. Il consiste en un rappel écrit de phrases énoncées par l'expérimentateur, sous deux conditions différentes. Soit le participant transcrit la phrase sans aucune autre contrainte, soit il la transcrit tout en gérant parallèlement une autre tâche (e.g., mémoriser une série de mots unisyllabiques en vue d'un rappel ultérieur). Les résultats de Fayol et Got mettent en évidence que des adultes lettrés commettent, lorsqu'ils sont placés en situation de double tâche, des erreurs d'accord verbal qu'ils ne commettent pas lorsqu'ils doivent simplement s'acquitter du rappel de phrases. Ces erreurs consistent à accorder le verbe avec un nom qui n'est pas son sujet grammatical, mais qui se trouve immédiatement placé devant le verbe (e.g., "Le chien des voisins arrivent"). L'utilisation de phrases au syntagme nominal complexe (de type: Nom1 de Nom2 Verbe), dans lesquelles les auteurs faisaient varier expérimentalement le nombre de Nom1 et de Nom2, a permis de confirmer l'existence d'accords par proximité.

Bock et Miller (1991) mettent en évidence des accords par proximité à l'oral en anglais (voir aussi: Bock & Cutting, 1992; Bock & Eberhard, 1993). Pour ce faire, elles font entendre à des adultes un début de phrase (un syntagme nominal: e.g., "The key to the cabinets") et leur demandent de le répéter puis de compléter librement la phrase. Elles analysent ensuite le nombre d'erreurs d'accord verbal produites en fonction de la structure du syntagme nominal. En effet, comme dans l'expérience de Fayol et Got (1991), quatre syntagmes différents étaient utilisés, en fonction du nombre respectif de Nom1 et de Nom2: Singulier Singulier: "The key to the cabinet"; Pluriel Pluriel: "The keys to the cabinets"; Singulier Pluriel: "The key to the cabinets"; Pluriel Singulier: "The keys to the cabinet". Ici encore c'est lorsque Nom1 et Nom2 diffèrent en nombre que les erreurs d'accord verbal surviennent. Toutefois, si ces travaux sur l'oral confirment l'existence d'erreurs d'accord par proximité, ils révèlent des taux d'erreurs nettement plus faibles que ceux rapportés à l'écrit. Deux explications peuvent être fournies à cet écart de performances entre l'oral et l'écrit (mais ici, aussi entre l'anglais et le français), d'une part l'existence [à l'oral en

anglais] (vs. l'absence [à l'écrit en français]) d'une information phonologique permettant de distinguer le singulier du pluriel verbal, d'autre part le fait que Fayol et Got, mais pas Bock et Miller, recourent au paradigme de la double tâche. En effet, de nombreux autres résultats (e.g., Fayol, Largy & Lemaire, 1994) sont venus confirmer que les erreurs d'accord verbal sont significativement plus fréquentes en présence qu'en l'absence d'une tâche secondaire. Fayol et al. (1994) en ont conclu que la production de l'accord verbal s'exerçait en présence d'un monitoring susceptible de détecter un risque d'erreur, et qu'une surcharge cognitive momentanée empêchait ce monitoring de s'exercer. La présence d'une tâche ajoutée mobiliserait des ressources attentionnelles qui dès lors feraient défaut au fonctionnement de cette instance de contrôle de la production écrite, laissant apparaître des erreurs d'accord verbal.

4 Arguments en faveur d'un accord verbal contrôlé

De nombreux résultats sont venus confirmer l'hypothèse d'une erreur d'accord verbal due à une surcharge cognitive. Ainsi, Largy, Chanquoy et Fayol (1993) ont répliqué l'effet de la tâche ajoutée avec des phrases présentant une structure en double pronom: Pronom1 Pronom2 Verbe. L'erreur la plus fréquente consiste à accorder le verbe sous l'influence du nombre du pronom préverbal (e.g., "L'enfant émet des sons. Il les répètent") et ne s'observe qu'en présence d'une tâche secondaire. Un autre argument tient au fait que le nombre d'erreurs d'accord par proximité augmente avec la complexité de la tâche secondaire. Ainsi, Largy, Fayol et Lemaire (1996) observent significativement plus d'erreurs d'accord verbal lorsque les participants doivent maintenir en mémoire des listes de cinq mots que des listes de trois mots. Hupet et al. (1996), mais aussi Fayol et al. (1994), ont confirmé l'implication d'une surcharge cognitive dans la survenue d'erreurs d'accord en reproduisant ce type de résultats avec des tâches secondaires de natures différentes (i.e., mémorisation de non-mots; dénombrement de clics; addition de chiffres). Ce rapport entre efficacité du contrôle de l'accord et coût cognitif de la tâche secondaire a été confirmé par l'analyse des performances des participants à la tâche secondaire. Par exemple, Largy et al. (1996) ont montré que la performance à la tâche secondaire était affectée par la complexité de l'accord à produire: les participants rappelaient en moyenne significativement plus de mots (de la liste de cinq mots à maintenir en mémoire) lors du rappel de phrases où Pronom1 et Pronom2 étaient de même nombre (tous deux singulier ou tous deux pluriel) qu'ils n'en rappelaient lorsque Pronom1 et Pronom2 étaient de nombre différent.

En résumé sur ce point, la production de la morphologie du nombre verbal fait l'objet d'un contrôle cognitivement coûteux (pour plus de détails, cf. Largy, 2003, 2004). Une surcharge cognitive au moment de la production de l'accord expose le scripteur au risque de transcrire graphiquement une flexion verbale inappropriée, automatiquement activée. Des recherches ont ensuite exploité ce résultat en utilisant le paradigme de la tâche ajoutée comme le moyen d'observer quel(s) morphème(s) flexionnel(s) étai(en)t initialement activé(s) avant que n'opère une instance de contrôle corrective. En effet, lors de la production d'un texte, les morphèmes flexionnels transcrits graphiquement peuvent avoir fait l'objet d'une vérification avant même leur transcription. On n'est donc pas certain que le morphème flexionnel transcrit soit celui qui a été initialement activé. En empêchant l'instance de contrôle d'opérer, par adjonction d'une tâche secondaire, on est à même d'observer, dans la production écrite des participants, le morphème initialement activé. Le paradigme expérimental de la tâche ajoutée a dès lors été utilisé pour

étudier la façon dont la morphologie flexionnelle du nombre verbal était activée. C'est en ce sens que Largy et al. (1996) ont considéré que ce paradigme "ouvrait une fenêtre sur les processus syntaxiques".

5 De l'étude de l'encodage grammatical

D'après Levelt (1989, Levelt et al. 1999), un message est élaboré conceptuellement, avant d'être formulé linguistiquement puis articulé. Largy (2003), adaptant le modèle de production du langage oral de Levelt au cas de la production écrite, postule qu'après les étapes d'élaboration conceptuelle et de formulation linguistique, le message peut également être transcrit graphiquement. Selon Levelt, l'étape de formulation linguistique consiste en un accès lexical en deux temps: le temps de l'encodage grammatical, suivi du temps de l'encodage phonologique. L'encodage grammatical permet la récupération des lemmas (lemma: entité abstraite correspondant aux propriétés sémantiques et syntaxiques d'un mot), puis l'encodage phonologique donne accès aux lexèmes (lexème: entité abstraite correspondant aux propriétés phonologiques d'un mot). L'encodage grammatical correspond à la phase de l'accès lexical où les mots sont sélectionnés et leurs relations syntaxiques spécifiées. C'est donc au cours de l'encodage grammatical que le verbe va être sélectionné et que seront précisés ses paramètres diacritiques (i.e., il sera spécifié en mode, temps, personne et nombre; Levelt et al., 1999). On peut considérer que c'est à cet instant que le morphème flexionnel du nombre est activé.

Un débat s'est instauré sur la nature des propriétés du nom susceptibles d'attribuer sa valeur au paramètre diacritique "nombre verbal" (Bock et al., 2001). De nombreuses recherches sur l'oral, inspirées du paradigme de Bock et Miller (1991) ont contribué à ce débat. Dans ce cadre déjà, Bock et Miller (1991) ont testé dans quelle mesure des caractéristiques conceptuelles typiques des noms "sujets" possédées par le nom local (i.e., placé juste avant le verbe) pouvaient influencer l'accord verbal. Elles ont ainsi étudié l'impact du caractère animé (vs. inanimé) et du caractère concret (vs. abstrait) de Nom2. Leurs résultats ne révèlent aucun effet de ces deux variables. Pas plus que les résultats de Bock et Eberhard (1993) ne sont parvenus à mettre en évidence l'effet d'une catégorie particulière de noms (dits "noms collectifs"; e.g., une armée, une forêt, une équipe, etc.) sur le taux d'erreurs d'accord verbal. En effet, avec ces noms employés syntaxiquement au singulier mais désignant un ensemble, une collection d'êtres ou d'objets, les auteurs étaient en mesure d'attendre un effet de la pluralité sémantique. Or, ici encore, les auteurs ne rapportent aucune différence dans l'effet d'attraction (accord par proximité) entre les noms locaux collectifs (grammaticalement singuliers) et les noms locaux individuels (i.e., un soldat, un arbre, un joueur, etc.). Cet ensemble de résultats tend à valider l'hypothèse "minimale" (ou "hypothèse du tout syntaxique"). En effet, tout semble se passer comme si la sélection du morphème flexionnel du nombre verbal ne se faisait que sur la base d'une information syntaxique, c'est-à-dire indépendamment de l'effet des traitements conceptuels (cf. les expériences citées supra) et morphophonologiques (cf. Bock & Eberhard, 1993, autre expérience) de la phrase. Or, les travaux sur la production écrite de l'accord verbal vont permettre de remettre en cause une telle conception, en montrant que l'activation de la morphologie flexionnelle du nombre est sensible aux dimensions sémantique et phonologique de la production langagière.

6 De la "perméabilité" de la syntaxe

Les principaux débats observés en psycholinguistique sur la question de l'accord verbal peuvent être rassemblés sous l'étude de deux phénomènes: le phénomène de "proximité" et celui de "perméabilité".

Le phénomène d'accord par proximité est connu de longue date. Il a permis en son temps d'opposer différentes conceptions explicatives de l'accord verbal. Jespersen (1924) s'est référé au concept d'énergie mentale (*mental energy*) pour rendre compte du fait que plus le verbe est éloigné linéairement de son sujet dans la phrase, plus le risque d'un non marquage en nombre du verbe est élevé. Woodworth (1938) a évoqué le concept d'"effet d'atmosphère" (*atmosphere effect*) pour expliquer l'impact du nombre d'un mot sur l'accord des mots suivants. Actuellement, les hypothèses rivales sont celles renvoyant à la conception d'un accord soit linéaire (*linear proximity*), soit hiérarchique (*syntactic proximity*: hypothèse du tout syntaxique). Alors que les travaux de Bock et de ses collaborateurs (cf. supra) plaident en faveur d'un accord hiérarchique, ceux de Fayol, Largy et de leurs collaborateurs, conduits sur la production écrite de l'accord en français, tendent à valider la conception d'un accord linéaire.

Un premier argument en faveur de la linéarité tient en l'existence d'erreurs du type "Le loup voit des moutons. Il les attrapent" (Fayol et al., 1994). Dans ce cas, la seule présence préverbale d'un mot au pluriel suffit à activer le morphème flexionnel verbal du pluriel, même lorsque ce mot n'appartient pas au syntagme nominal (ce qui va clairement à l'encontre de l'hypothèse hiérarchique). Un second argument réside dans l'existence d'erreurs typiques de phrases au sujet inversé. Ces phrases sont rares en langue mais grammaticalement correctes. Hupet et al. (1996) rapportent les exemples suivants: "Dans les trains passent le contrôleur" et "Dans la cabine remonte les mineurs", où l'accord verbal ne semble pas régi par le sujet grammatical de la phrase mais dans la linéarité sous l'influence du nom préverbal.

Le phénomène de perméabilité de l'accord, mis en évidence par les travaux sur la production écrite de l'accord verbal, est un argument de plus à l'encontre de l'hypothèse du tout syntaxique. La sélection du morphème flexionnel du nombre verbal ne relèverait pas exclusivement de processus syntaxiques mais dépendrait aussi des aspects phonologiques et sémantiques de la production du langage.

L'incidence de la phonologie sur l'accord verbal est mise en évidence par Largy et Fayol (2001). Ces auteurs ont conduit une expérience dans laquelle des adultes entendent des phrases à l'imparfait de l'indicatif et doivent les rappeler par écrit après les avoir transformées au présent de l'indicatif. Ces phrases sont du type "Nom1 de Nom2 Verbe" où Nom1 et Nom2 sont tantôt de même nombre, tantôt de nombre différent. Les verbes sont de deux types, soit du premier groupe (e.g., La mère des garnements arrivait [arriver]), soit d'un autre groupe (e.g., La mère des garnements surgissait [surgir]). Ainsi, quel que soit le verbe, lorsque la phrase est fournie à l'imparfait aucune information phonologique ne permet de différencier la flexion verbale de la troisième personne du singulier de celle de la troisième personne du pluriel (e.g., arrivait/arrivaient; surgissait/surgissaient). Par contraste, au présent, seuls les verbes du premier groupe préservent auditivement l'ambiguïté (e.g., arrive/arrivent). Les autres verbes présentent une terminaison susceptible de guider phonologiquement l'accord (e.g., surgit/surgissent; /i/ vs. /is/).

Les auteurs mettent en évidence qu'en présence d'une tâche ajoutée les erreurs d'accord par proximité sont significativement plus nombreuses sur les verbes dont les flexions (singulier/pluriel) n'étaient pas phonologiquement différenciables que sur les autres verbes. Ils montrent ainsi que la sélection de la morphologie flexionnelle du nombre est sensible à une information de type phonologique.

L'incidence du niveau sémantique sur l'accord verbal est mise en évidence à partir d'une catégorie d'erreurs d'accord particulière, celle consistant à accorder le verbe comme s'il s'agissait d'un nom, c'est-à-dire en lui associant au pluriel la marque du pluriel des noms: -s (e.g., "Le facteur prend des enveloppes. Il les timbres"). Ces erreurs s'observent en situation de double tâche, en présence de verbes homophones de noms (e.g., le verbe "timbrer", homophone du nom "un timbre" lorsqu'il est conjugué au présent de l'indicatif : "il timbre" ; "ils timbret"). Largy et al. (1996) ont construit des phrases susceptibles d'amorcer sémantiquement le candidat nominal en situation d'homophonie verbe/nom. Procédant ainsi, ils ont mis en évidence que lorsque l'homophone nominal était amorcé par le contexte sémantique de la phrase (e.g., "Le jardinier sort les légumes et il les asperge") les erreurs d'accord en -s (i.e., ... il les asperges) étaient significativement plus nombreuses que lorsqu'il ne l'était pas (e.g., "L'éléphant voit les clowns et il les asperge"). Ce résultat a récemment été confirmé par Largy et Fayol (soumis) en distinguant la situation des homophones verbe/nom jugés proches (e.g., brosse / une brosse) ou éloignés sémantiquement (e.g., asperger / une asperge). Ce faisant, les auteurs ont montré que la sélection de la morphologie flexionnelle du nombre semble s'opérer via la sélection de la racine du mot. Ce résultat confirme que l'accord verbal est "perméable" à une information de type sémantique.

L'ensemble des résultats rapportés sur la production écrite de l'accord verbal établit clairement que la sélection de la morphologie flexionnelle du nombre verbal ne dépend pas exclusivement de processus syntaxiques mais s'avère également sensible aux aspects sémantiques et phonologiques de la production écrite. Récemment, les travaux sur l'oral sont venus confirmer une telle conception (e.g., Schriefers & Hartsuiker, 2001).

7 Conclusion

Si l'accord en nombre entre le verbe et son sujet est enseigné par le pédagogue comme un principe grammatical, les recherches en psycholinguistique nous montrent que cet accord n'est pas exclusivement régi par des processus syntaxiques mais qu'il s'applique aussi sous l'influence des niveaux phonologique et sémantique de la production langagière.

Les résultats rapportés à l'écrit permettent d'envisager que la sélection d'une flexion en vue de sa transcription graphique se fasse sur le critère de son niveau d'activation par rapport à celui d'autres flexions concurrentes. L'influence des niveaux évoqués supra sur la sélection d'une flexion erronée peut être expliquée dans le cadre de modèles connexionnistes (cf. Dell, Chang & Griffin, 1999). Ceux-ci fournissent un support intéressant pour se représenter un phénomène de compétition entre la flexion appropriée et la (ou les) flexion(s) concurrente(s). Ils permettent d'envisager l'existence de différentes sources d'activation (syntaxique, conceptuelle/sémantique, morphophonologique, etc.) dont les effets peuvent se sommer et conduire à la sélection d'une flexion particulière. Ces modèles ont également des implications temporelles et peuvent fournir une explication alternative à celle d'un contrôle de la production écrite. En effet, selon de tels

modèles, les erreurs et leur absence peuvent être expliquées en termes de temps et/ou de ressources disponibles à l'accomplissement du processus d'activation. Les erreurs tiendraient au fait que ce processus n'a pas pu atteindre son terme et qu'une flexion erronée, provisoirement plus activée que la flexion exacte, a été sélectionnée.

Toutefois les résultats ici rapportés ne valent que pour des adultes ayant déjà une longue pratique de la langue écrite, chez qui la fréquence d'association entre certaines racines et certaines flexions à une incidence sur la façon dont ils produisent désormais l'accord verbal.

8 Références bibliographiques

- Bock, J.K., & Cutting, J.C. (1992) « Regulating mental energy: Performance units in language production » *Journal of Memory and Language* 31 : 99-127.
- Bock, J.K., & Eberhard, K.M. (1993) « Meaning, sound, and syntax in English number agreement » *Language and Cognitive Processes* 8 : 57-99.
- Bock, J.K., Eberhard, K.M., Cutting, J.C., Meyer, A.S., & Schriefers, H. (2001) « Some attractions of verb agreement » *Cognitive Psychology* 43 : 83-128.
- Bock, J.K., & Miller, C.A. (1991) « Broken agreement » *Cognitive Psychology* 23 : 45-93.
- Dell, G.S., Chang, F., & Griffin, Z.M. (1999) « Connectionist models of language production : Lexical access and grammatical encoding » *Cognitive Science* 23 : 517-542.
- Fayol, M., & Got, C. (1991). « Automatisation et contrôle dans la production écrite » *L'Année Psychologique* 91 : 187-205.
- Fayol, M., Largy, P., & Lemaire, P. (1994). « When cognitive overload enhances subject-verb agreement errors. A study in French written language » *Quarterly Journal of Experimental Psychology* 47A : 437-464.
- Hupet, M., Schelstraete, M.A., Demaeght, N., & Fayol, M. (1996) « Les erreurs d'accord sujet-verbe en production écrite » *L'Année Psychologique* 96 : 587-610.
- Jespersen, O. (1924). *The philosophy of grammar*. Allen and Unwin, London.
- Largy, P. (2003) « Du contrôle de l'orthographe grammaticale. 1re partie: du contrôle pré-graphique » *Le Langage et l'Homme* 38 : 139-152.
- Largy, P. (2004) « Du contrôle de l'orthographe grammaticale. 2ème partie: du contrôle post-graphique » *Le Langage et l'Homme* 39 : 7-21.
- Largy, P., Chanquoy, L., & Fayol, M. (1993) « Automatic and controlled writing: Subject-verb agreement errors in French native speakers » In G. Eigler & T. Jechle (Eds.), *Writing: Current Trends in European Research*, pp. 109-120. Freiburg: Hochsch.-Verl.
- Largy, P., & Fayol, M. (2001) « Oral cues improve subject-verb agreement in written French » *International Journal of Psychology* 36 : 121-132.
- Largy, P., & Fayol, M. (soumis) « All homophones are not alike : On the impact of the semantic proximity of verb/noun homophones on agreement errors »
- Largy, P., Fayol, M., & Lemaire, P. (1996) « The homophone effect in written French : The case of verb-noun inflection errors » *Language and Cognitive Processes* 11 : 217-255.
- Levelt, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge MA: M.I.T. Press.
- Levelt, W.J.M., Roelofs, A., & Meyer, S. (1999) « A theory of lexical access in speech production » *Behavioral and Brain Sciences* 22 : 1-75.
- Schriefers, H., & Hartsuiker, R. (2001) « Encodage grammatical et production du langage :

Largy

l'accord sujet-verbe » *Psychologie Française* 46 : 43-54.

Woodworth, R.S. (1938). *Experimental psychology*. New York : Holt

Beeinflusst die Muttersprache den Erwerb der Subjekt-Verb-Kongruenz im Plural bei Schülern, die Deutsch als L2 studieren?

La langue source influe-t-elle sur l'acquisition de l'accord sujet-verbe au pluriel chez les lycéens qui étudient l'allemand L2 ?

Angela Byrskog

Universität de Lund
Institut d'allemand
angela.byrskog@tyska.lu.se

Résumé

Le but de cette étude a été de montrer comment la langue source influe sur la morphologie de la troisième personne du pluriel de l'allemand langue cible. Nous avons choisi de comparer le suédois et l'italien comme langue source, parce que l'italien montre l'accord qui nous intéresse, et distingue six formes de personne du verbe, alors que le suédois n'a qu'une seule forme pour toutes les personnes, donc ne marque pas cet accord. Nous avons étudié deux classes de suédophones et deux classes d'italophones qui apprennent l'allemand. Les résultats obtenus montrent qu'il a été plus facile pour les italophones de marquer l'accord sujet-verbe que pour les suédophones, bien que les suédophones aient reçu le double nombre de cours d'allemand, et que les italophones aient un niveau linguistique général identique ou même un peu plus bas que les suédophones, jugé à partir d'autres phénomènes.

1 Fragestellung

Diese Arbeit hat zum Ziel, den Einfluss der Muttersprache auf den Zweitspracherwerb zu untersuchen. Näher bestimmt: Die Aneignung der Subjekt-Verb-Kongruenz (SVK) der 3. Pers. Plural im Deutsch-L2-Erwerb wird analysiert. Gibt es Transfer, oder, mit anderen Worten: Beeinflusst die Muttersprache den Deutsch-L2-Erwerb der 3. Pers. Plural? Die SVK im Plural allgemein hat sich als eine schwer zu meisternde Struktur erwiesen für Gymnasiasten¹ mit Schwedisch als L1. Fälle wie **Die Kinder spielt*; **die Mütter weint* etc kommen häufig vor. Besonders auffallend im Deutschunterricht an den schwedischen Gymnasien ist, dass dieselben Schüler, die die oben zitierten Konstruktionen produzieren, die Kasus Kongruenzen oft korrekt markieren, wie z. B.: *Heidi ist in der Stadt*; (Dativ nach Präposition); *Es gibt einen Mann*

¹ In diesem Artikel sind mit "Schüler" nicht nur ausschließlich männliche Schüler gemeint, sondern sowohl Schüler als auch Schülerinnen. Dasselbe gilt auch für die Benennung "Lerner" und "Muttersprachler".

(Akkusativ nach Verb). Die schwedische Sprache besitzt weder SVK im Plural noch Kasuskongruenz. Dennoch bereitet ausgerechnet die SVK den schwedischen Schülern besondere Probleme. Ist es möglich, dass die schwedische Muttersprache dieser Schüler mit dem Deutsch-L2-Erwerb interferiert? Eine Möglichkeit, um diese Untersuchung durchzuführen wäre demzufolge, wenn man diese schwedischen L2-Lerner mit einer Gruppe von Lernern vergleichen könnte, bei denen die Muttersprache dieselbe, spezifisch zu untersuchende syntaktische Eigenheit (SVK im Plural) wie das Deutsche besitzt, um zu sehen, ob die in der Muttersprache gelernte Struktur der SVK im Plural einen Einfluss hat auf den Erwerb derselben Struktur in der L2. Eine dieser Sprachen ist das Italienische. Somit wäre eine vergleichende Untersuchung von Material interessant, das teils von schwedischen, teils von italienischen Muttersprachlern produziert wird, die die deutsche Sprache als L2 erwerben.

Die Hypothese dieser Arbeit lautet demzufolge: Wenn es Transfer gibt, sollte schwedischen Muttersprachlern der Erwerb eines Parameters, der in der schwedischen Sprache nicht vorhanden ist, schwerer fallen als italienischen Muttersprachlern, da die Letzteren diesen Parameter in ihrer Muttersprache schon gesetzt haben. Es wird also angenommen, dass die in der L1 erworbenen Parameter der italienischen Schüler ihnen beim L2-Erwerb behilflich werden.

2 Theorie

Die verschiedenen Transfer-Theorien im Rahmen der generativen Syntax gehen davon aus, dass der L2-Lerner einen direkten Zugang zur Universal Grammar (UG) besitzt, (Mitchell & Myles 2004) und dass er sich die L2 auf die gleiche Weise wie die Muttersprache aneignet. Beobachtete Unterschiede und Abweichungen werden darauf zurückgeführt, dass Erwachsene andere Bedürfnisse haben als Kinder und eine kognitive Reife besitzen, die ein Kind nicht haben kann.

2.1 UG und L2: die Relevanz der UG in Bezug auf den L2-Erwerb

Es wurde in den siebziger Jahren festgestellt, dass verschiedene Ähnlichkeiten bestehen zwischen dem L1 und dem L2 Erwerb. Es wurde gezeigt, dass auch im L2-Erwerb gewisse "Routen" durchgemacht werden, die mehr oder weniger fixiert sind. Beispiele dafür sind Studien über den L2-Erwerb durch Meisel, Clahsen & Pienemann (1983). Diese Erwerbsrouten nahm man als Zeichen dafür, dass die L2-Erwerbler durch interne Prinzipien gesteuert werden, die in Vielem von der Muttersprache unabhängig sind. Diesen Wortfolgestudien folgten weitere Untersuchungen wie z.B. über den Erwerb von Negationen in Englisch als L2, die auch in definierten Stadien vorkamen (Ellis 1994 u.a.) Dieselben Beobachtungen wurden in Bezug auf den Deutsch-L2-Erwerb bei Negationen gemacht. (Mitchell & Myles 1998). Sämtliche erwähnte Untersuchungen und Forschungsergebnisse scheinen auf überzeugende Weise einen systematischen L2-Erwerb vorzuweisen, wobei dieser Erwerb unabhängig von der Muttersprache geschieht und viel Ähnlichkeit mit dem L1-Erwerb aufweist.

Wenn, wie oben erwähnt, auch im L2-Erwerb festgestellte Sequenzen zu finden sind, kommt die Frage auf, welche Rolle die UG im Zweitspracherwerb spielen kann. Verschiedene Hypothesen existieren hierzu, anbei zwei davon:

2.1.1 Full access/ full transfer hypothesis

Der L2-Lerner hat direkten Zugang zur UG, und eignet sich die L2 auf der Basis der Muttersprache an. Beobachtete Unterschiede werden auf unterschiedliche kognitive Reife und unterschiedliche Bedürfnisse des Lernalters zurückgeführt, d.h.: ein Kind hat weder die kognitive Reife noch dieselben Bedürfnisse eines Erwachsenen. Verteidiger dieser Ansicht sind z.B. Schwartz & Sprouse (1996).

2.1.2. Partial access hypothesis

Gewisse Aspekte der UG stehen dem L2-Lerner noch zur Verfügung, andere aber nicht mehr. Zum Beispiel können gewisse Prinzipien für den Lerner noch zugänglich sein, die Parameter aber nicht mehr. Pienemann (1998) und Håkansson, Pienemann & Sayehli (2002) zeigen in ihrem empirischen Material anhand der **Processability Theory**, dass Transfer nur dann erfolgt, wenn der Lerner die zu lernende Struktur verarbeiten kann.

Nachdem Schwartz & Sprouse Transfer zu untersuchen begonnen haben, hat dieses Gebiet der Linguistik in den letzten Jahren einen Aufschwung erlebt (Gass & Selinker 2001). Die Diskussion eines Morphologietransfers wurde weiter von P. Prévost (Eurosla-Konferenz, 2004) anhand von untersuchten Deutsch-L2 Lernern mit teils Spanisch, teils Englisch als L1 aufgenommen. Prévosts Dateien weisen auf einen Einfluss der Muttersprache auf den Erwerb der L2 hin. (Prévost 2004).

3 Korpus

3.1 Untersuchungstyp

Diese Untersuchung basiert auf einer pseudolongitudinalen, vergleichenden Studie des L2-Subjekt-Kongruenzerwerbs innerhalb des Pluralkontextes von 16-18-jährigen Gymnasiasten mit teils Italienisch, teils Schwedisch als L1.

Der Vergleich von Deutsch L2-Lerner, bei welchen die SVK in ihrer Muttersprache **vorkommt**, hier bei italienischen Muttersprachlern, mit Deutsch L2-Lernern, bei denen die Subjekt-Verb-Kongruenz in ihrer Muttersprache **nicht** vorkommt, hier bei schwedischen Deutsch L2-Lernern, wird hier ausgeführt.

Der Erwerb weiterer, **nur** für die deutsche Sprache charakteristischer syntaktischer Phänomene sollte später weiter bei den Italienern sowohl als auch bei den Schweden untersucht und verglichen und der SVK im Plural gegenübergestellt werden, da bei einer solchen Struktur beide L1-Sprachen nicht interferieren können. Eine solche Struktur ist die Kasuskongruenz in der deutschen Sprache.

3.2. Nähere Beschreibung der untersuchten Klassen:

Die untersuchten Schüler wurden aus Gymnasialklassen in Malmö, Schweden, und Rom, Italien ausgesucht. Vom St. Petri Gymnasium in Malmö wurden folgende Schüler untersucht:

Stufe 2: 14 schwedische Schüler, Alter: 17 Jahre, Deutsch-Unterrichtsstunden: 125. Die

Klasse der zweiten Stufe erhielt den Namen *Malmö 2*.

Stufe 4: 17 schwedische Schüler, , Alter: 18 Jahre, Deutsch-Unterrichtsstunden: 566. Die Klasse der 4. Stufe erhielt den Namen *Malmö 4*.

Vom V. Arancio Ruiz Gymnasium in Rom wurden folgende Schüler untersucht:

Stufe 2: 15 italienische Schüler, Alter: 16 Jahre, Deutsch-Unterrichtsstunden: 160. Die Klasse der zweiten Stufe erhielt den Namen *Roma 2*.

Stufe 4: 22 italienische Schüler, Alter: 18 Jahre, Deutsch-Unterrichtsstunden: 262, die Klasse der 4. Stufe erhielt den Namen *Roma 4*.

Um einen so gerechten Vergleich wie möglich zu erreichen, wurde der erhaltene Deutschunterricht in Anzahl Stunden angegeben, wobei mit einer gewissen Variation von Schule zu Schule gerechnet werden muss, die hier nicht aufgefangen werden kann. Da die Schüler aus verschiedenen Grundschulen stammen, ist es unmöglich, ohne einen ungeheuren Zeitaufwand die exakte Anzahl Unterrichtsstunden eines jeden Schülers zu errechnen, deswegen wurde, nach verschiedenen Nachforschungen und einer Analyse der von den Schülern eingefüllten Fragebogen, während der Grundschule ein realistischer Schnitt von 80 Sprach-Unterrichtsstunden pro Jahr angenommen.

Die gesamte Anzahl Deutschunterricht-Stunden der verschiedenen Klassen wurde wie folgt ausgerechnet:

Schwedische Schüler:

1-6- Klasse: 80 Std./Jahr Sprachunterricht pro Fremdsprache.

7.-9. Klasse: 120 Std./Jahr Sprachunterricht pro Fremdsprache.

Gymnasium: 1. Jahr: Englisch 86 Std., weitere Sprachen 84 Std.

2. Jahr: Englisch 86 Std., weitere Sprachen 84 Std.

Italienische Schüler:

Ein Schnitt von 75 Sprach-Unterrichtsstunden pro Jahr wurde für die italienischen Schüler berechnet. Ein wichtiger Faktor, der bei den italienischen Schülern zu größeren Unterschieden in der Anzahl Unterrichtsstunden beigetragen hat, wurde von den verschiedenen Schulreformen in Italien verursacht, die auf die Schüler unterschiedlichen Einfluss ausgeübt haben in Bezug auf die Anzahl des Sprachstunden-Unterrichts.

Roma 4: Sämtliche untersuchten Schüler haben vier Jahre lang Latein studiert (dieselbe Anzahl Unterrichtsstunden wie Deutsch, also 262 Std.). Dasselbe gilt auch für Roma 2, die 160 Std. Latein studiert haben. In Italien wird Latein nicht als Fremdsprache angesehen, deswegen hat keiner der Schüler die Sprache auf dem Fragebogen erwähnt, sondern nur geschrieben, ob er/sie Englisch und/ oder Französisch studiert hat.

Nähere Beschreibung der untersuchten Klassen

❖ Malmö 2-1:	9 Schwedische SchülerInnen, Stufe 2, <i>obligatorischer</i> Unterricht
❖ Malmö 2-2:	5 schwedische SchülerInnen, Stufe 2, <i>freiwilliger</i> Unterricht
	Alter: 17; Unt. Std.: 125
Roma 2:	15 italienische Schülerinnen, Stufe 2
	Alter: 16; Unt. Std.: 160

❖ Malmö 4:	17 schwedische SchülerInnen, Stufe 4; Alter: 17 Jahre; Unt. Std.: 566
Roma 4:	22 italienische SchülerInnen, Stufe 4; Alter: 18 Jahre; Unt. Std.: 262
Stufe 2:	29 SchülerInnen (15 lt., 14 Schw.)
Stufe 4:	39 SchülerInnen (22 lt., 17 Schw.)
Total:	37 lt. 31 Schw. = 68 SchülerInnen

Es wurden 37 italienische und 31 schwedische Schüler untersucht, die totale Anzahl untersuchter Schüler beträgt 68.

2.3 Auswahl der Schüler, Vorbereitung zur Materialerfassung

Das St. Petri Gymnasium in Malmö, Schweden wurde gewählt, weil dort die größte Anzahl an Deutsch-L2-Schüler mit Schwedisch als L1 zu finden war. In Malmö wurde uns zu Beginn der Untersuchung eine Unterrichtsstunde zur Verfügung gestellt, in welcher ich den Schülern eine kurze Einführung in die Spracherwerbsforschung gab. Dabei wurde nichts über die SVK im Plural geäußert. Das Ziel dieser Introduction war einerseits, die Schüler für meine Untersuchung zu gewinnen, da ich nur an freiwilligen Teilnehmern interessiert war. Andererseits wurde ihnen nahegelegt, dass es in der Sprachwissenschaft keine "Fehler" gibt, sondern dass ihre Abweichungen von der Norm von Interesse sind, da diese Abweichungen Wichtiges über den Erwerb der L2 aussagen können. Den Schülern wurde versichert, dass das Resultat der Untersuchung von ihren respektiven Lehrerinnen weder gesehen noch bewertet würde. Auch wurde ihnen erklärt, dass ihre Resultate mit einem Kode versehen würden, der ihre Identität schützt. Dies nicht nur, um eine eventuelle Angst vor Konsequenzen zu verhindern, die sich negativ auf die Noten der Schüler auswirken würden, sondern auch, um sie auf eine so spontane Produktion wie möglich vorzubereiten. In Malmö wurden die Schüler der Stufe 2 und Stufe 4 von zwei unterschiedlichen Lehrerinnen unterrichtet.

Für die Untersuchung der italienischen Schüler wurde das V. Arangio Ruiz Gymnasium in Rom gewählt, weil die geographische Nähe und somit der Einfluss deutschsprachiger Länder auf den

Deutscherwerb vermieden werden sollte, was bei einem Gymnasium in z.B. Mailand nicht möglich gewesen wäre. Auch in Rom wurde den Schülern, aus den oben erwähnten Gründen, vollkommene Anonymität zugesichert. In Rom wurden Stufe 2 und Stufe 4 von derselben Lehrerin unterrichtet.

4 Material, Methode und Arbeitsweise zur Materialerfassung

Die Hauptuntersuchung bestand darin, den Deutsch L2-Schülern mit den zwei verschiedenen Muttersprachen Deutsch und Italienisch eine Aufgabe zu geben, die analysierbares, vergleichbares Material hervorbringen würde. Wegen verschiedener technischer Begrenzungen war die Produktion eines schriftlichen Korpus der Produktion eines mündlichen vorzuziehen. Das Problem des *avoidance phenomenon* (Schachter, 1974) wurde hiermit ebenso vermieden, da das Material zu einer Produktion zwingt, die die zu analysierende syntaktische Funktion erstellt. Ein weiterer Vorteil der gewählten Methode bestand darin, dass ein durch freies Schreiben quantitativ und qualitativ mangelhaftes Untersuchungsmaterial verhindert wurde. Dabei war es wichtig, so weit wie möglich homogene Gruppen zu untersuchen, die sich nicht zu sehr voneinander unterscheiden in Bezug auf erhaltene Unterrichtsstunden und weiteren Faktoren, die die Untersuchung beeinflussen könnten. Trotz längerer Nachforschung hat es sich jedoch als unmöglich erwiesen, schwedische und italienische Schüler zu finden, die dieselbe Anzahl Unterrichtsstunden erhalten haben. Aus diesem Grund wurde der Entschluss gefasst, Klassen zu untersuchen, bei denen die Anzahl Deutsch-Unterrichtsstunden keinen Nachteil für die aufgestellte Hypothese bedeuten würde. Weiter wurde beschlossen, wenigstens beim Material, das den Ausgangspunkt dieser Untersuchung ausmacht, eine so homogene Form wie möglich zu erarbeiten. Ein Material wurde hergestellt, das jedem Schüler dieselbe Möglichkeit erbot, ein schriftliches Produkt zu erzeugen, das vom gleichen Ausgangspunkt aus analysiert werden konnte.

Das den Schülern ausgeteilte Arbeitsmaterial ist ein Heft mit 46 nummerierten Schwarz-Weiß-Fotografien aus dem Film *Heidi*, der auf dem gleichnamigen Buch der Schweizerin Johanna Spyri basiert. Die Schwarz-Weiß-Fotografien wurden einem Filmheft entnommen, das dem vorgewiesenen Film entspricht und so ausgewählt, dass sie die Schüler ermuntern sollten, Sätze mit dem Subjekt abwechselungsweise im Singular und im Plural zu schreiben. Dies, um bei der Produktion und später bei der Analyse des Materiales eine Differenzierung Singular/Plural in der Verbanwendung zu ermöglichen. Um allen untersuchten Schülern eine gemeinsame Grundkenntnis des Testmaterials zu sichern, wurde ihnen eine schriftliche Zusammenfassung über das Buch in ihrer Muttersprache verabreicht. Der Film wurde den Klassen in deutscher Sprache vorgewiesen, um alle zu untersuchenden Schüler so weit wie möglich mit denselben Voraussetzungen vorzubereiten in Bezug auf die Handlung, den Ablauf der Geschichte, der Personengalerie usw.

Für die Untersuchung stand die Zeitspanne von 50 Minuten zur Verfügung, wobei die für die praktische Organisation angewendete Zeit (ca. 25-30 Minuten) vor und nach der schriftlichen Produktion berücksichtigt werden muss. Nach dem Erteilen der Anweisungen, Austeilen der Fotohefte, einloggen der Computer usw., wurden die Schüler gebeten, zu jedem Photo einen Satz zu schreiben, in welchem mindestens ein Subjekt und mindestens ein Verb vorkommen sollten.

Die Sätze wurden mit derselben Nummer des zu beschreibenden Bildes versehen. Weiter sollten bei beendeter Arbeit die geschriebenen Sätze als Dokumente benannt, gespeichert und auf eine Diskette kopiert werden. Die effektive Schreibzeit nach all diesen Vorbereitungen betrug höchstens 20-25 Minuten für 46 Sätze.

Ein Vorteil der knapp zur Verfügung stehenden Zeit war, dass das Geschriebene ein der sprachlichen Produktion so naheliegendes Resultat wie möglich erzeugt, da für die Schüler ein langes Überlegen nicht möglich war (Krashen 1982). Während die Schüler schrieben, gingen die betreffenden Lehrerinnen und ich in der Klasse umher und ermunterten sie, das aufzuschreiben, was sie konnten und sich nicht um den Rest zu kümmern. Selbstverständlich durften den Schülern weder Beispiele noch Wörter suggeriert werden.

5 Empirische Analyse

5.1. Klassendurchschnitt der nicht-zielsprachlichen Subjekt-Verb-Kongruenzen im Pluralkontext, sowie allgemeine vergleichende Daten.

Insgesamt wurde die schriftliche Produktion von insgesamt 68 Schüler untersucht, 37 mit Italienisch, 31 mit Schwedisch als L1, wobei jeder 46 Sätze produzieren sollte, die mindestens ein Subjekt und mindestens ein Prädikat enthalten sollten. Diese Sätze wurden auf die unten aufgezeichneten Faktoren untersucht und miteinander verglichen, wobei der Durchschnitt pro Klasse ausgerechnet wurde.

Um einen besseren Vergleich zwischen den italienischen und den schwedischen Klassen zu ermöglichen, wurde der Durchschnitt allgemeiner Faktoren wie Anzahl Sätze (1), Verbtypen (2), Nebensätze (6), Anzahl korrekte Nebensätze (7) und Anzahl Wörter (8) dem Durchschnitt der Faktoren gegenübergestellt, die im empirischen Material untersucht werden, näher bestimmt: unmarkierte Verbtypen (3), Anzahl unmarkierte Pluralendungen (4) und Minimale Paare (5).

Die Spalten sind wie folgt zu deuten:

ID = Identität des untersuchten Schülers/Schülerin.

R2: Rom, Stufe 2; **M2:** Malmö, Stufe 2; **R4:** Rom, Stufe 4; **M4:** Malmö, Stufe 4.

1. Sätze = Durchschnitt der Anzahl geschriebener Sätze (von 46 möglichen) pro Klasse (misst die *Fluency*, d.h., wie fließend die Schüler die deutsche Sprache beherrschen).

2. Verbtypen = Durchschnitt Anzahl angewendeter Verbtypen pro Klasse, z.B.: Der Lerner hat *gehen*, *essen*, *spielen*, *sprechen* etc angewendet, nicht nur *ist* oder *diskutiert*. Dies deutet auf das Niveau des Wortschatzes hin.

3. % unmarkierter Verbtypen im Plural = Durchschnitt Anzahl Verbtypen ohne SV-Kongruenz im Plural gibt Auskunft darüber, ob der Schüler nur bei *essen* das Plural nicht markiert, oder bei mehreren Verben (gehört zum Kern des Untersuchten).

4. Anzahl unmarkierte Pluralendungen = Durchschnitt gesamter Anzahl SVK im Plural ohne Pluralmarkierung pro Klasse, der eigentliche Kern dieser Untersuchung.

5. Minimale Paare = Durchschnitt Anzahl minimaler Paare pro Klasse. Ein minimales Paar wird wie folgt definiert: Wenn ein Verb mindestens einmal im Singular und einmal im Plural vorkommt, wird es als minimales Paar bezeichnet: Das Kind **geht**, die Mädchen **gehen**. **Geht/gehen** = minimales Paar. Es dient als Kriterium dafür, dass der Lerner das Verb im Singular und im Plural angewendet hat und somit dessen Konjugation im Singular und im Plural beherrscht.

6. N-Sätze = Durchschnitt Anzahl Nebensätze: Gibt Auskunft über den Grad der Komplexität der Sprache.

6 a. N-Sätze korrekt = Prozentsatz korrekter Satzwortfolge pro Klasse. Nach Pienemanns *Processability Theory* (Pienemann 1999) kann man die Satzwortfolge nicht meistern, wenn man die SVK im Plural nicht meistert, da die SVK im Plural auf Stufe 4 der Entwicklungshierarchie steht, die Satzwortfolge jedoch auf Stufe 5. Dieser Punkt bedarf weiterer Untersuchungen.

7. Wörter = Durchschnitt Anzahl Wörter pro Satz pro Klasse (Mean length of utterance).

Die Faktoren 1. (Anzahl Sätze), 2. (Verbtypen), 6. (Nebensätze), 6a. (Anzahl korrekte Nebensätze) und 7. (Anzahl Wörter) dienen dazu, ein allgemeines Bild des untersuchten Schülers in Bezug auf dessen Fähigkeit zur Produktion der deutschen Sprache zu geben. Dies, um das Niveau der verschiedenen Sprachproduzenten miteinander vergleichen zu können.

Klassendurchschnitt der nicht-zielsprachlichen SVK im Pluralkontext versus allgemeine vergleichende Daten

ID	1 Sätze	2 Verbtypen	Klassen- Schnitt			6 N-Sätze	6a N-Sätze % korrekt	7 Wörter/ Satz
			3 % unmark. Verbtypen	4 Anz. un- mark. Pl.	5 Min.Paare			
R2	38.7	11.4	0.7	0.2	2.9	-	-	4.5
M2	40.1	10.9	1.6	2.8	1.4	-	-	4.4
R4	37.4	13.1	0.3	0.4	3.1	0.9	86%	5.4
M4	39.6	18.9	2.5	3.6	1.9	1.0	33%	5.2

5.1.1 Analyse der Resultate

Der Klassendurchschnitt der Anzahl Sätze (1), Anzahl Nebensätze (6) und Anzahl Wörter (7) weisen unbedeutende Unterschiede auf zwischen den Schülern mit Italienisch und den Schülern mit Schwedisch als L1. Die schwedischen Lerner jedoch erhöhen die Anzahl angewendeter Verbtypen von 10,9 auf 18,9 zwischen Stufe 2 und Stufe 4, also beinahe auf das Doppelte, was auf einen bedeutenden Fortschritt in Bezug auf die Entwicklung des Wortschatzes deutet, Lerner

mit Italienisch als L1 jedoch nur von 11,4 auf 13,1, was auf eine deutlich geringere Erweiterung des Wortschatzes hinweist. Dies deutet auf einen wichtigen Unterschied hin in Bezug auf die Entwicklung des Wortschatzes zwischen den beiden Sprachgruppen. Die Testpersonen mit Italienisch als L1 schrieben jedoch bei 86% der Fälle eine korrekte Nebensatzwortfolge, im Gegensatz zu den Schülern mit Schwedisch als L1, die nur eine Ziffer von 33% erreichten. Eine korrekte Nebensatz-Wortfolge deutet auf einen höheren Niveau des Spracherwerbs hin. Dieser Faktor bedarf, wie schon erwähnt, weiterer Untersuchung.

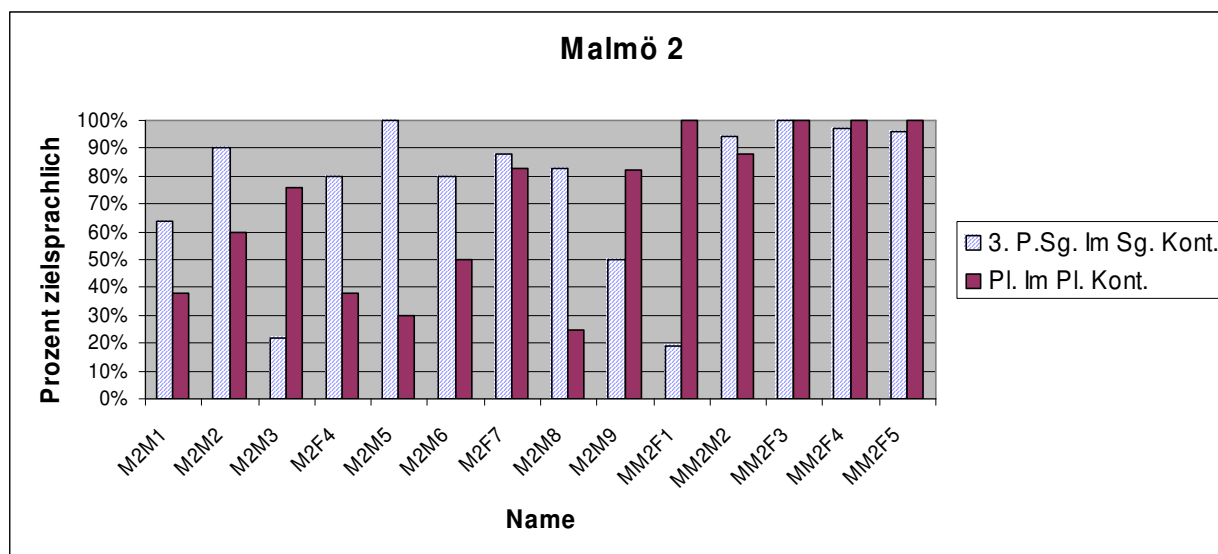
SVK im Plural: Der größte Unterschied besteht in der Anzahl verpasster, beziehungsweise korrekter Markierungen der SVK im Plural (Spalten 3 und 4). Die Klassen der 2. Stufe: Die Schüler mit schwedischer Muttersprache haben im Schnitt 2,8 Verben unmarkiert gelassen im Pluralkontext, im Gegensatz zu den Schülern mit italienischer Muttersprache, die durchschnittlich nur 0,2 Verben ohne SVK im Plural ließen. Dasselbe Phänomen erscheint in den Klassen der Stufe 4, wo die italienischen Muttersprachler nur 0,3 Verben im Pluralkontext unmarkiert ließen, im Gegensatz zu den schwedischen Muttersprachlern, die im Schnitt 2,5 Verben im Pluralkontext unmarkiert ließen. Beispiel: **Die Kinder geht; *Heidi und der Großvater lacht.*

5.1.2 Tabellen SVK im Plural versus 3. Pers. Singular.

Dieser Unterschied in der Pluralmarkierung des Verbs zwischen Deutsch-L2-Lernern mit teils schwedischer teils italienischer Muttersprache wird anhand folgender Tabellen individuell für jeden untersuchten Schüler gezeigt. In den folgenden Dateien wurden die korrekten Verbendungen der 3. Person Singular den korrekten Verbendungen der 3. Person Plural gegenübergestellt. Dies aus folgendem Grund: Wenn der Schüler die 3. Person **Singular** korrekt markiert (z.B. *Heidi geht*) weiß man mit größerer Sicherheit, dass die 3. Person **Plural** korrekt gedeutet werden kann (z. B. *Heidi und Peter gehen*). Wenn der untersuchte Schüler aber für die 3. Person Singular **Heidi gehen* schreibt, was sogleich auch die Infinitivform des Verbs sein kann, gibt es keinen Grund zu glauben, dass er die 3. Person Plural beherrscht, da man keinen Beweis dafür hat, dass *Peter und Heidi gehen* hier wirklich als Pluralform des Verbs gemeint ist. Folglich ist eine korrekte Markierung der 3. Person **Singular** der sicherste Hinweis dafür, dass die vom untersuchten Schüler produzierte, korrekt **aussehende** Markierung der 3. Person **Plural** auch wirklich als Plural gemeint war, und nicht ein infinites Verb darstellt. Diese Kontrollfunktion erlaubt eine bedeutend genauere Analyse des produzierten Materials.

Malmö 2/1-2/2

Die folgende Tabelle zeigt die Resultate zwei verschiedener schwedischer Klassen, die beide auf derselben Stufe des Deutschunterrichts sind. Die Schüler unter 2-1 hatten obligatorischen Unterricht, diejenigen mit der Bezeichnung 2-2 jedoch hatten freiwillig Deutschunterricht gewählt.



Beispiele zur Resultaterläuterung:

Die Daten der Schülerin MM2F1 zeigen deutlich, dass sie keine korrekte Endung in der 3. Person Singular (gestreifte Kolonne) aufweist. Die dunkle Kolonne, die eine 100%-ig korrekte Markierung der 3. Person Plural aufweist, ist demzufolge irreführend. Die Test-Person schrieb, außer in drei Fällen, durchgehend Sätze wie **Heidi sehen die schöne blumchen* ** Die Frau sehen* etc. Dies hat zur Folge, dass die Kolonne, die den Prozentsatz korrekter Subjekt-Verb-Kongruenzen Im Plural zeigt, eigentlich unkonjugierte Infinitivformen des Verbs auffängt. In den von ihr scheinbar syntaktisch korrekt produzierten Sätzen *Heidi und die Frau stehen*; *Heidi und die gross Man sprechen* hat diese Schülerin einfach das Infinitiv des Verbs geschrieben, da sie es nicht konjugieren konnte.

Der Schüler mit der Bezeichnung M2M5 hat in 16 Sätzen das Verb *ist* verwendet, was die 100%-ige Zielsprachlichkeit der 3. Person Sg. erklärt. Er verwendete aber auch dasselbe Verb in Fällen wie ** Heidi und der Mann ist neugierig*. Die SVK im Plural markierte er jedoch nur in 30% der Fälle korrekt.

Weiter kann konstatiert werden, dass nur eine Schülerin von insgesamt 14 untersuchten Test-Personen eine zielsprachliche Form der SVK im Plural von 100% erreicht, zwei ca. 95 %, zwei ca. 80%, zwei 74%, drei 60%, die anderen liegen unter 40% oder noch darunter.

Beachtenswert ist weiter die Tatsache, dass die freiwilligen Deutschschüler bessere Resultate erzielt haben als diejenigen, die ein obligatorisches Sprachjahr wählen müssen (vgl. Krashen 1982, *The Affective Filter Hypothesis*). Andererseits kann man argumentieren, dass man freiwilligen Sprachunterricht nicht wählen würde, wenn man nicht wüsste, dass man leicht Sprachen lernt.

Roma 2

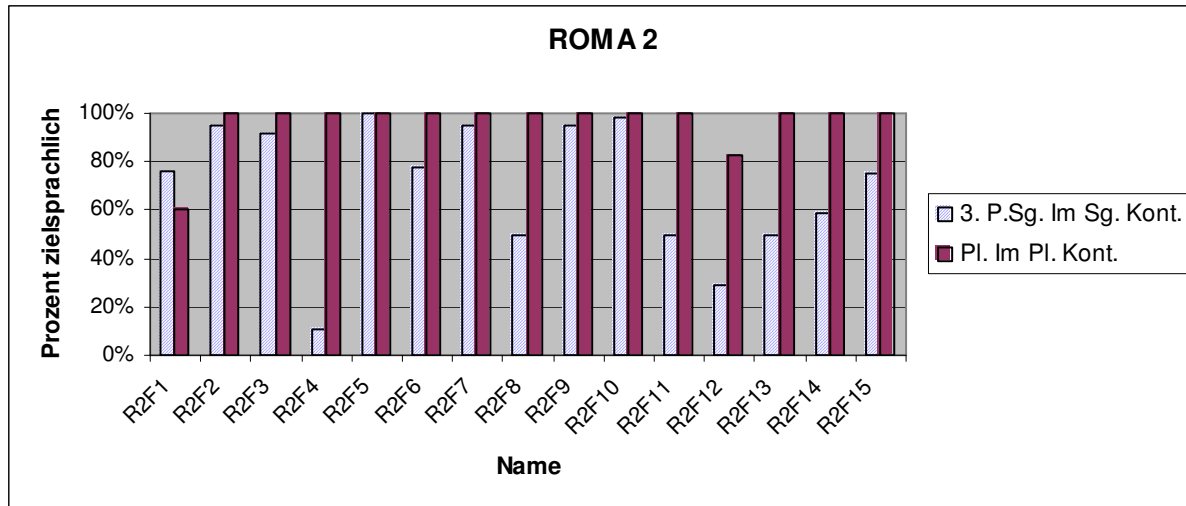
R2 und M2 sowohl als auch R4 und M4 sind an den Faktoren 1. (Anzahl Sätze), 2. (Verbtypen), 6. (Nebensätze) und 7. (Anzahl Wörter pro Satz) durchaus vergleichbar, was zeigt, dass die schwedischen Klassen nicht generell ein niedrigeres Niveau aufweisen als die italienischen.

Die Schüler mit italienisch als L1 weisen eine größere Anzahl korrekter SVK im Plural auf als diejenigen mit schwedischer Muttersprache: Zwei Testpersonen erreichten eine zielsprachliche

Beeinflusst die Muttersprache den Erwerb der Subjekt-Verb-Kongruenz?

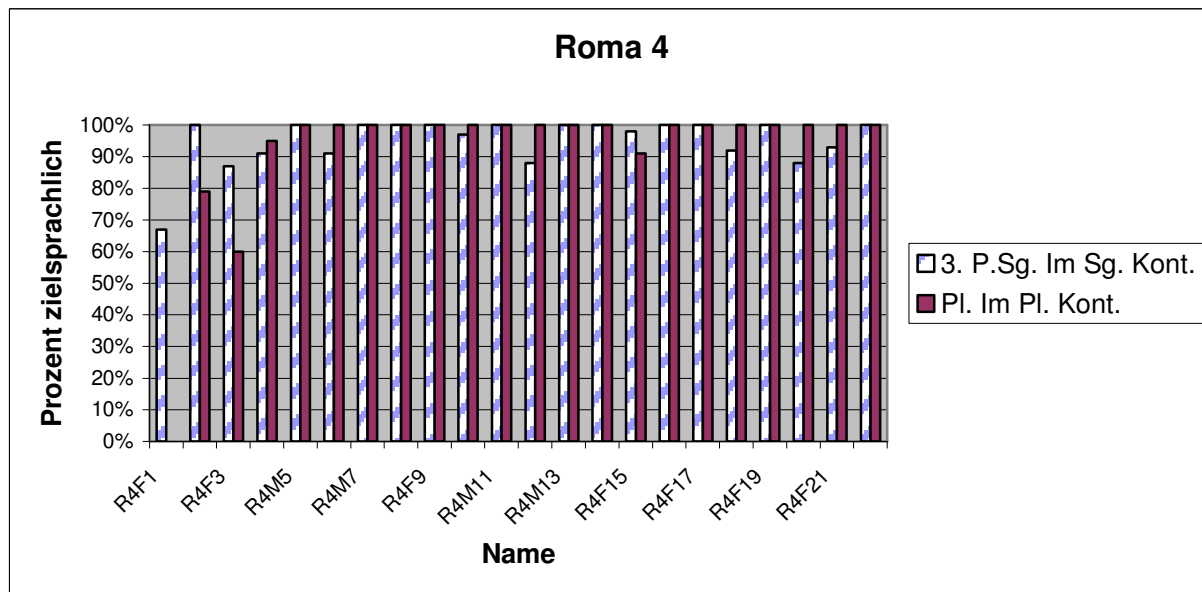
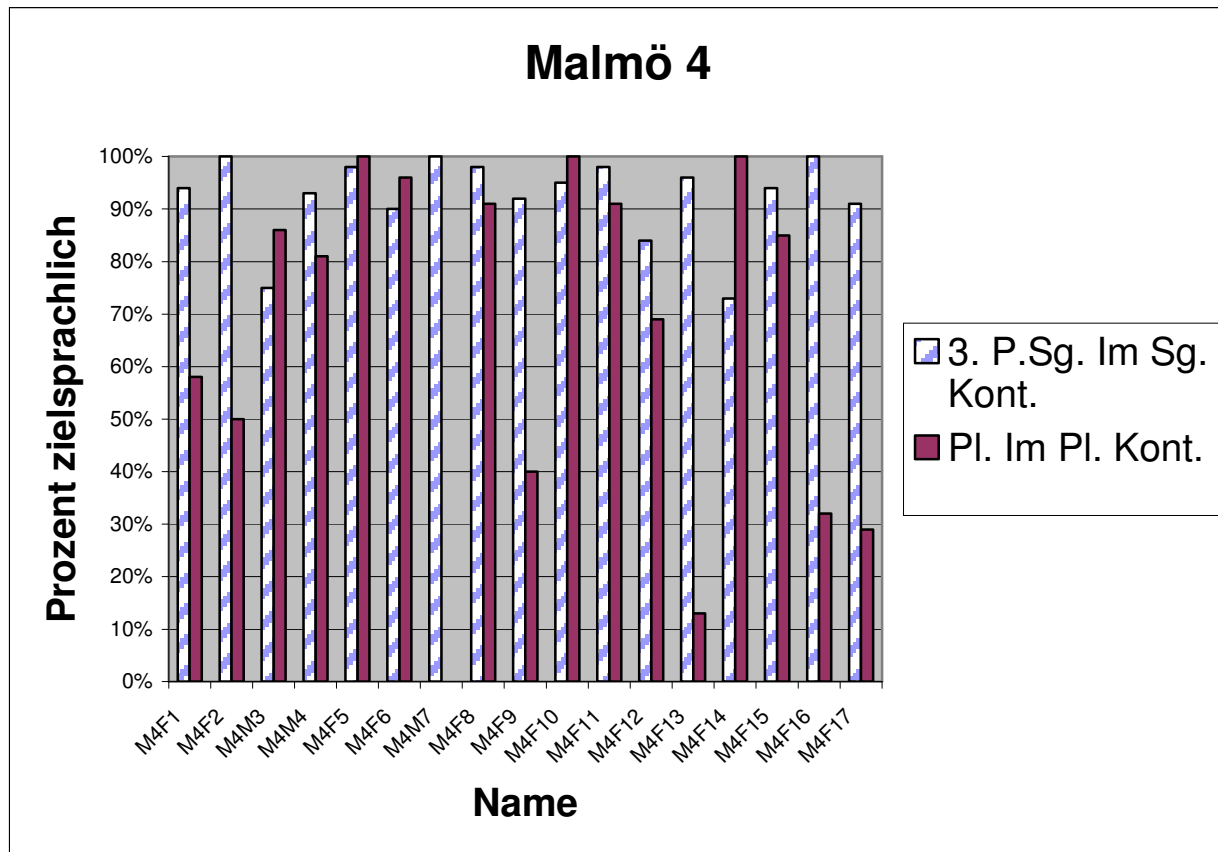
Produktion von 100%, zwei je 93 %. Weitere 3 Schüler je 80%, je ein Schüler 70% und 60%, zwei weitere 50%.

Auch hier kann man bei den vier Schülern R2F4 und R2F13-15 die 100%ig korrekte Plural-SVK nicht als solche betrachten, sondern als nicht-konjugierte Verben, da die zielsprachliche 3. Pers. Sg.-Markierung sehr schwach ausgefallen ist.



Malmö 4 vs. Roma 4

Ein noch größerer Unterschied besteht zwischen der Produktion von Malmö Stufe 4 verglichen mit Roma Stufe 4. Ein einziger Informant erreichte eine zielsprachliche Markierung von ca. 98%, im Vergleich zu Roma 4, wo 5 Informanten eine 100%-ige, und weitere 8 Informanten eine 90%-ige Zielsprachlichkeit erreichten. Außer einer Schülerin, die allgemein große Schwierigkeiten mit der deutschen Sprache hatte, fielen nur 2 Roma 4-Schüler unter die 80%-Linie.



6 Zusammenfassung und Schlussfolgerung

Die Durchschnittsziffern der verschiedenen Klassen unter 1. (Anzahl Sätze), 2. (Verbtypen), 6. (Nebensätze), 6a. (Anzahl korrekte Nebensätze) und 7. (Anzahl Wörter) dienen wie gesagt dazu,

ein allgemeines Bild des untersuchten Schülers in Bezug auf dessen Fähigkeit zur Produktion der deutschen Sprache zu geben. Außer in der Produktion einer zielsprachlichen SVK-Anwendung, die die italienischen Schüler deutlich besser beherrschen, sind die schwedischen Deutsch-L2-Lerner durchaus vergleichbar mit den italienischen Deutsch L2-Lernern. Es scheint, dass Transfer von der Muttersprache für die Resultate der italienischen Schüler eine Rolle spielt.

Weitere mögliche Vergleichsfaktoren wären, wie erwähnt, die zielsprachliche/nicht zielsprachliche Kasusendung nach Verben und Präpositionen, die Kasus verlangen, da solche Konstruktionen (außer bei gewissen Pronomen) weder in der italienischen noch in der schwedischen Sprache vorkommen. Da die korrekte Anwendung der Kasusformen eine in verschiedenen Schritten zu prozessierende Funktion ist, kann sie in Bezug auf die Einstufung der Sprachkompetenz der Schüler von Interesse sein. Vorläufige Ergebnisse deuten darauf hin, dass die schwedischen Klassen eine bessere Kasusbeherrschung als die italienischen aufweisen. Dies sollte jedoch genauer überprüft werden. (vgl. Byrskog, unter Bearbeitung).

Der markanteste Unterschied in Bezug auf erhaltene Deutschstunden besteht zwischen Malmö Stufe 4 und Rom Stufe 4, da die schwedischen Schüler beinahe doppelt so viele Deutsch-Unterrichtsstunden erhielten als die italienischen. Was den Lateinunterricht betrifft, liegt auch ein großer Unterschied vor. 420 Stunden Latein sind für die italienischen Schüler das erreichte Pensum, wenn sie auf Stufe 4 sind, die Schweden hingegen haben kaum Lateinunterricht erhalten. Latein ist eine Sprache mit einer reichen Verbmorphologie, und inwiefern dies auf die Sensibilisierung der Schüler auf das Markieren der SVK im Plural einwirkt, müsste weiter untersucht werden.

Der untersuchte Korpus zeigt deutlich, dass die beiden Rom-Klassen die 3. Person Plural besser markieren, obwohl die Malmö-Klassen die doppelte Anzahl Unterrichtsstunden erhielten. Die Untersuchung zeigt weiter, dass die Rom-Klassen nach den Faktoren Fluency (1) und Wortschatz (2) auf einem niedrigeren Niveau als Malmö stehen. Die Tatsache jedoch, dass R4 beim Faktor 6a (Nebensatzwortfolge) scheinbar bessere Resultate erreicht, muss weiter untersucht werden.

7 Bibliographie

- Ellis, R. (1991) *Understanding Second language Acquisition*. Oxford University Press.
- Håkansson, G., Pienemann, M. & Sayehli, S. (2002) « Transfer and Typological Proximity in the Context of Second Language Processing ». *Second language Research* 18(3) : 250-273.
- Krashen, S. (1982) *Principles and practice in Second Language acquisition*. Oxford: Pergamon.
- Meisel, J.M., Clahsen, H. & Pienemann, M. (1983) *Deutsch als Zweitsprache: der Spracherwerb ausländischer Arbeiter*. Tübingen: Narr.
- Mitchell, R. & Myles, F. (2004) *Second language Learning Theories*. Oxford : Oxford University Press Inc.
- Pienemann, M. (1999) *Language processing and Second Language Development, Processability Theory*. Amsterdam/Philadelphia: John Benjamins.
- Prévost, P., (2004) « Morphological variation in early adult second language French. A cross-

- sectional study » *EUROSLA Yearbook* 4 : 147-175.
- Schachter, J. (1974) « An Error in Error Analysis ». *Language learning* 205-214.
- Schwartz B. & Sprouse, R. (1996) « L2 Cognitive States and the Full Transfer/Full Access Model » *Second language Research* 12(1) : 40-72.

Développement de l'accord verbal avec un sujet pluriel dans les récits écrits d'apprenants germanophones scolarisés du français

Cyrille Granget

Université de Paris III
cgranget@wanadoo.fr

Introduction

Comme le notent Granfeldt et Schlyter en introduction de ce volume, le développement de la morphologie verbale est au cœur des discussions menées actuellement en acquisition de langue seconde. Une question majeure porte notamment sur le moment d'émergence de l'accord verbal dans les langues d'apprenants et sur le sens que l'on doit y accorder. Le marquage de l'accord est-il graduel et fonction du développement également graduel de la syntaxe (Vainikka et Young-Scholten 1996, Paradis et al. 1998) ou doit-on au contraire considérer que le marquage variable de l'accord ne reflète pas l'absence d'une structure syntaxique mais traduit des difficultés de réalisation de surface auxquelles sont confrontés les apprenants de L2 (Prévost et White 2000).

Si l'on s'intéresse au marquage de l'accord verbal en français, il apparaît que les marques spécifiques attestant de l'accord sujet-verbe sont différentes à l'écrit et à l'oral. A l'oral, l'accord du verbe avec un sujet pluriel 3^{ème} personne n'a pas de réalisation distincte de l'accord avec un sujet singulier pour certains verbes réguliers comme *parler* tandis qu'il se traduit par une variation lexicale du verbe, audible, pour d'autres verbes irréguliers, comme par exemple *prendre*.

(1)

parler		prendre	
je	}	je	
tu		tu	[pra ~]
l'homme		l'homme	
l'homme et la fille			l'homme et la fille

A l'écrit en revanche, la variation ou l'absence de variation lexicale sur le verbe se double d'une marque graphique en position suffixale.

(2)

parler	prendre
l'homme parle	l'homme prend
l'homme et la fille parlent	l'homme et la fille prennent

En conséquence, cette marque graphique -nt pourrait suffire à indiquer le marquage de l'accord, sans pour autant s'accompagner d'une variation lexicale du verbe. Ainsi, il est plus délicat de déduire des marques d'accord du français parlé, notamment dans le cas de verbes réguliers, que du français écrit. Il s'ensuit également que des apprenants ayant été exposés au français écrit et des apprenants n'y ayant pas été exposés sont susceptibles d'avoir une représentation différente de l'accord verbal en français : variation lexicale versus variation flexionnelle, éventuelle, plus dérivation flexionnelle. Tandis qu'à l'écrit la dérivation flexionnelle est une marque obligatoire de l'accord pluriel quelque soit le type de verbe, la variation lexicale n'affecte que quelques verbes.

Dans le cas d'apprenants scolarisés et par conséquent familiarisés avec l'écriture, la production de récits écrits offre plusieurs possibilités d'attester du marquage de l'accord et du développement de cette catégorie à différents stades de l'apprentissage du français. A partir d'un corpus de récits de film écrits, produits par des apprenants germanophones scolarisés de quatre niveaux différents, nous nous sommes demandés comment se développe le marquage de l'accord verbal en français des stades débutants (+ 70 heures d'exposition en classe) à intermédiaires (+ 450 heures d'exposition en classe), et quel est le rôle des contraintes internes liées au français et des contraintes externes liées à la fréquentation de la classe de langue.

Nous allons principalement tester deux hypothèses concernant le développement, à l'écrit, du marquage de l'accord verbal avec le sujet pluriel 3^{ème} personne:

Hypothèse 1 : Le développement du marquage de l'accord s'explique par les propriétés du sujet pluriel (pronom clitique, syntagme nominal, sujet élide) et du verbe lexical ([+/-] régulier, [+/-] fréquent dans l'input).

Hypothèse 2 : Le développement du marquage de l'accord verbal s'explique par l'enseignement, à la fois par un volume d'input oral réduit et par la priorité accordée à l'accord graphique en milieu scolaire.

2 Données et méthode

Afin d'expliquer le développement de l'accord du verbe avec un sujet pluriel, nous nous appuyons sur deux types de données. D'une part, nous avons analysé le développement de l'accord verbal avec un sujet pluriel 3^{ème} personne à partir d'un corpus de 95 récits de film écrits produits par des apprenants se situant à différents stades d'apprentissage et d'exposition au français scolaire. Ce corpus de récits écrits présente 289 énoncés dans lesquels le sujet de l'action ou de la situation est un sujet pluriel 3^{ème} personne. Ces énoncés se répartissent selon les groupes de la façon suivante :

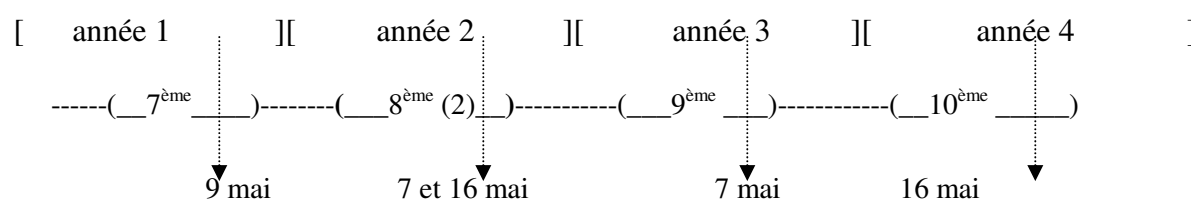
Tableau 1. Répartition des énoncés au pluriel selon les groupes

Groupe	1	2A	2B	3	4
N=95	n=10	n=23	n=18	n=18	n=26
Nbr énoncés pluriel	17	51	51	39	131
Total=289					

D'autre part, nous avons analysé, en amont de la tâche de récit, la place accordée à l'apprentissage de l'accord sujet-verbe dans le contexte des classes de français qui sont, pour ces apprenants, le seul lieu d'exposition au français. Ces analyses reposent sur un deuxième corpus d'interactions en classe de langue.

Le schéma ci-dessous indique les dates de recueil des données de récit de film et les périodes d'observation en classe :

(3)



Les périodes signalées entre crochets correspondent aux quatre années d'apprentissage du français. Les périodes entre parenthèses correspondent pour chaque année ou niveau institutionnel (7^{ème}, 8^{ème}, 9^{ème}, 10^{ème}) à la période au cours de laquelle nous étions en classe, c'est-à-dire une partie de l'année, entre décembre/ janvier (selon les classes) et fin juillet de l'année scolaire 2000/2001. Le chiffre 2 sur la deuxième ligne signale que nous avons observé deux classes de 8^{ème} (2A et 2B) en deuxième année d'apprentissage du français. Sur la troisième ligne, nous indiquons la date de l'année 2001 à laquelle nous avons réalisé l'expérience des récits de film.

Le support utilisé pour les récits est un extrait des *Temps Modernes* de C. Chaplin. Cet extrait met en scène plusieurs protagonistes qui agissent seuls ou de concert, selon les circonstances, ce qui favorise la référence à plusieurs personnes et la réalisation dans le récit d'un sujet grammatical pluriel. Nous avons mené cette expérience telle qu'elle a été menée dans le projet ESF (Perdue 1993) auprès de travailleurs adultes immigrés vivant en France. Après avoir regardé l'extrait des *Temps Modernes*, les apprenants ont disposé de trente à trente-cinq minutes pour écrire le récit de ce film muet. Les apprenants sont restés dans la salle de projection en présence de leur enseignant. Ils disposaient tous d'une aide à la conception du récit, une fiche lexicale, à laquelle ils pouvaient se référer s'ils le souhaitaient. Sur cette fiche figuraient des expressions préfabriquées comme *c'est l'occasion de fuir*, *ils se retrouvent* ainsi que des verbes à la forme infinitive comme *emmener*, *être par terre*, *avoir faim*. L'emploi de *ils se retrouvent* dans les récits écrits n'a pas été pris en compte dans le calcul des marques de l'accord verbal avec le sujet. L'enquêteur a expliqué

aux apprenants qu'il s'agissait d'une expérience dont le résultat lui était destiné et qu'ils ne seraient pas notés.

En ce qui concerne la méthode de comptage des marques d'accord verbal, nous avons considéré qu'en cas de double marquage de l'accord du verbe avec un sujet pluriel en français écrit cible, c'est-à-dire variation lexicale du verbe et suffixation en *-nt* pour les verbes pleins irréguliers, le marquage graphique *-nt* de l'accord constituait une marque suffisante d'accord. Les marques graphiques d'accord du pluriel 3^{ème} personne que nous avons retenues sont *-nt*, *-s*, et *-ez* (une seule occurrence). Autrement dit, nous n'avons pas considéré que la variation lexicale, dans le cas des verbes irréguliers, était un signallement nécessaire pour attester de la présence de l'accord. Ainsi, nous observons à l'écrit de nombreuses occurrences de marquage morphologique du nombre sans modification lexicale du verbe, ou alors variation idiosyncrasique, tels *ils fuïres*, *ils fuïsent*, *la fille et l'homme mets*, *ils voirent*, *l'agent et la fille allent*. Ces exemples témoignent d'un marquage prioritairement morphologique de l'accord verbal.

Lorsque les apprenants germanophones désignent au moins deux personnages dans leur récit, notons qu'ils emploient dans la quasi-totalité des cas un sujet grammatical pluriel : *ils*, *les persons*, *la fille et l'homme*, *le deux*. Cependant, on observe également l'emploi, à cette fin, de *tout le monde*, sujet grammatical singulier. Ces cas minoritaires (2 occurrences sur un total de 289 références à au moins deux personnages) n'ont pas été pris en compte. Nous avons ainsi retenu l'ensemble des énoncés dans lesquels les apprenants font référence à plus d'un personnage et observé si le verbe est accordé avec le sujet grammatical pluriel qui fait référence à un référent pluriel. Nous avons éliminé du comptage les énoncés dans lesquels le verbe n'est pas en français mais en allemand ou en anglais.

3 Le développement de l'accord verbal avec un sujet pluriel dans les récits.

Avant d'aborder le cas particulier de l'accord verbal avec un sujet pluriel, nous allons observer le développement de l'accord sujet-verbe, quel que soit le sujet.

3.1. L'accord sujet-verbe

Si on calcule un taux d'accord sur la base des verbes fléchis du corpus, c'est-à-dire la part des verbes accordés avec le sujet de l'énoncé dans les énoncés comportant un verbe fléchi, on obtient les résultats qui apparaissent dans le tableau suivant :

Tableau 2. *Moyenne (M) et écart type (s) du Taux d'accord sujet-verbe en fonction du groupe*

GROUPE	1	2'	2	3	4
N = 95	n = 10	n = 23	n = 18	n = 18	n = 26
M (en %)	97	94	91	95	95
s	0,065	0,067	0,057	0,061	0,059

Quel que soit le groupe observé, on remarque un taux d'accord élevé et d'une étonnante stabilité au cours du développement de langue. Environ 95 % des verbes fléchis sont accordés avec le sujet de la proposition dans laquelle ils s'insèrent. L'accord est donc en place et ce, au moins dès le 8^{ème} mois de l'apprentissage, c'est à dire après environ 78 heures d'exposition scolaire au français. De plus, les performances des apprenants au sein d'un même groupe sont fort homogènes comme en témoigne l'écart-type, faible d'un groupe à l'autre. En raison de la tâche, le taux d'accord élevé traduit surtout la régularité de l'accord avec la troisième personne du singulier et avec la troisième personne du pluriel qui sont les personnes grammaticales les plus fréquentes dans le récit de film.

3.2 Le développement de l'accord verbal avec un sujet pluriel

Le taux d'accord du verbe avec un sujet pluriel a été calculé sur l'ensemble des verbes en position de verbe fléchi dans l'énoncé, que ce verbe soit fléchi ou pas.

Tableau 3. *Taux d'accord avec un sujet pluriel en fonction du groupe*

GROUPE	1	2A	2B	3	4
N = 95	n = 10	n = 23	n = 18	n = 18	n = 26
Tx Acc (en %)	88	80	84	69	80

On observe que le taux d'accord d'un groupe à un autre est également élevé et relativement stable, même si les apprenants du groupe 3 marquent moins l'accord avec un sujet pluriel, en moyenne, que les apprenants des autres groupes.

4 Les contextes d'occurrence du non-marquage de l'accord verbal au pluriel

Le développement de l'accord sujet-verbe s'inscrit dans le développement lexical, syntaxique et discursif des langues d'apprenants. On peut penser que le développement de l'accord du verbe avec le sujet pluriel est lié à la fois au développement des moyens de référer aux personnes

(pronom clitique, syntagme nominal ou sujet élidé) et au développement du lexique verbal et précisément à la mobilisation d'un répertoire lexical de plus en plus diversifié. Ces observations permettraient de nuancer l'observation précédente selon laquelle l'accord verbal avec un sujet pluriel est un phénomène élevé et stable d'un groupe à un autre, autrement dit au fil de l'appropriation du français.

4.1 Développement des moyens de référer à plusieurs personnages

Si l'on observe l'évolution des moyens de référer à plusieurs personnages dans les récits de film écrits, on constate effectivement que ces moyens évoluent d'un groupe à un autre.

Tableau 4. *Répartition de l'ensemble des sujets pluriels (pronom clitique, syntagme nominal (SN), sujet élidé) en fonction des groupes (en %)*

	Pronom clitique	SN	Sujet élidé
1	71	29	0
2A	55	41	4
2B	54	42	4
3	51	44	5
4	46	45	9

A partir du tableau 4 ci-dessus, on observe que les apprenants débutants (groupe 1) emploient deux fois plus de pronoms clitiques que de SN pour référer à plusieurs personnages. Cette tendance s'atténue sans jamais toutefois s'inverser dans les stades ultérieurs. Dans les groupes 2A et 2B, 4 % des énoncés sont construits avec un sujet élidé. Dans le groupe 4, le plus avancé, les sujets élidés représentent une part plus importante, 9 % des énoncés « pluriel ». Hormis ces cas de figure, les apprenants du groupe 4 font référence à plusieurs personnages autant, en moyenne, au moyen d'un pronom clitique que d'un syntagme nominal.

A stimulus constant, les moyens de référer à un sujet pluriel évoluent d'un groupe à un autre, au fil de l'exposition croissante au français en classe de langue. L'évolution de la distribution des pronoms clitiques est à mettre en rapport avec le développement discursif et syntaxique des langues d'apprenants. Les sujets élidés, apparaissant dans des propositions coordonnées, traduisent une intégration syntaxique croissante des énoncés dans le récit.

Notons, parmi les SN employés par les apprenants germanophones, la présence d'une catégorie intéressante, le duel. L'expression *le deux* ou équivalente, comme par exemple *les deux persons*, est fréquemment employée par les apprenants germanophones du français pour référer à deux personnes. Cet usage est frappant et évolue de la façon suivante d'un groupe à l'autre.

Tableau 5. *Part de le deux ou équivalent parmi les sujets pluriels (en %)*

	le deux
1	0
2A	4
2B	12
3	18
4	14

L'expression du duel n'est pas mobilisée dans le premier groupe mais dans les suivants, selon des proportions variables mais non négligeables. On peut penser qu'il s'agit là d'un transfert de la notion de duel, exprimée dans la langue source par *beide*, et pour laquelle il n'existe pas, en français, de catégorie morphosyntaxique équivalente. Cette expression du pluriel témoigne d'un transfert d'une catégorie sémantique grammaticalisée dans la langue première.

4.2 Accord verbal et types de sujet

Pour notre propos, une question intéressante consiste à se demander si le taux d'accord varie selon le type de sujet. Le tableau suivant présente la répartition des formes verbales non accordées avec un référent pluriel selon la nature de ce dernier.

Tableau 6. *Avec quel sujet les verbes ne s'accordent-ils pas ? (en %)*

	Clitique	SN	élide
1	0	100	0
2A	40	60	0
2B	37,5	50	12,5
3	23	62	15
4	27	58	15
Total	29	59	12

Sur l'ensemble du corpus, les verbes s'accordent moins en moyenne avec un syntagme nominal qu'avec un pronom clitique. Dans le groupe 1 par exemple, alors que les apprenants font majoritairement référence à plusieurs personnages au moyen de pronoms clitiques, les cas de non-accord se rencontrent uniquement avec un syntagme nominal en position sujet. En d'autres termes, la propension à ne pas accorder le verbe avec le sujet pluriel est d'autant plus forte que le sujet est un syntagme nominal. Toutefois, le tableau montre également qu'il n'y a pas de corrélation absolue entre la présence d'un SN et la chute morphologique puisque dans les groupes 2 à 4, même si la plupart des verbes non accordés sont précédés d'un sujet sous forme de syntagme nominal, un part non négligeable de ces verbes sont précédés d'un pronom clitique ne sont pas immédiatement précédés d'un sujet (sujet élide). Il n'est donc pas exclu que le type de sujet influence l'accord verbal mais il ne s'agit pas d'une contrainte absolue. D'autres facteurs explicatifs de la chute morphologique peuvent être envisagés.

4.3. Accord verbal et verbes lexicaux

Une autre hypothèse explicative de la chute, globalement résiduelle, de l'accord verbal avec la

3^{ème} personne du pluriel consiste à avancer que certains verbes irréguliers ou rares dans l'input sont plus fréquemment dénués de marques d'accord. Le tableau ci-dessous rend compte des verbes qui, pour chaque groupe, ne présente pas de marquage suffixal en –nt ou -s.

Tableau 7. *Verbes non accordés selon les groupes*

1	regarde, crie
2A	est, regarde, emence, fait, prendre (2), être (2), fuir, apporte
2B	etre, habit, est, quitter, courir, rever, mett, trouve
3	parle (3) , rencontre (2), tomber, voit, être, fuir, fuite, raconte, prendre, bavarde
4	fuir, va, ressemble, emmene (3) , se retrouve, faire (2), emmener, voise, se trouve, parti, aller, parle, trove, alle, être, est, fuir, courir se rencontre, a, prend, parler

D'après le tableau ci-dessus, il est vrai que l'on rencontre parmi les verbes non accordés de nombreux verbes irréguliers (*faire, prendre, mett, voit*) ou rares, voire absents, dans l'input de la classe de langue (*courir, fuir, emmene*). Cependant, on rencontre également des verbes réguliers (*regarde, crie, se rencontre*) ou fréquents dans l'input de la classe de langue (*est, fait, alle*). Cette deuxième hypothèse, selon laquelle les propriétés du verbe [+/- régulier], [+/- fréquent] pourraient influencer l'accord verbal, n'est donc pas entièrement satisfaisante.

Remarquons par ailleurs qu'il y a deux types de verbes non accordés : les verbes autrement accordés et les verbes non fléchis.

Nous allons à présent aborder notre deuxième hypothèse selon laquelle le développement du marquage de l'accord verbal pluriel dans les récits écrits est dû au contexte d'apprentissage scolaire.

5 L'enseignement de l'accord sujet-verbe en classe

Comme nous le mentionnions en introduction, les apprenants scolarisés bénéficient d'un environnement d'apprentissage et d'exposition au français susceptible d'influencer leur marquage de l'accord graphique.

En effet, dans le contexte de l'apprentissage scolaire tel que nous l'avons observé, c'est-à-dire une exposition hebdomadaire au français de 3 heures en moyenne, dans des classes de 15 à 30 apprenants, le volume d'input oral est fortement réduit par rapport à l'acquisition de L1 par l'enfant ou de L2 par l'adulte en immersion sociale. Cette réduction de l'input oral est compensée en classe de langue par un input écrit véhiculé par les manuels, les exercices, les énoncés et explications au tableau, recopiés dans le cahier, etc. C'est d'ailleurs à partir de cet input écrit et de toutes les traces écrites des activités didactiques que l'apprenant va apprendre ses leçons de français en vue des activités d'évaluation de ses performances.

Un corollaire de cet input écrit consiste en un focus formel sur la grammaire du français langue étrangère. En effet, l'écriture de la langue permet d'attirer l'attention de l'apprenant sur les marques morphologiques et graphiques de l'accord, tel qu'il s'exprime dans le français écrit par exemple. Nous allons présenter quelques séquences de cours de 1^{ère} année afin d'illustrer nos

propos et de montrer la priorité accordée à l'accord verbal avec le sujet dans le contexte du collège berlinois observé.

Prenons la séance où les apprenants de première année sont pour la première fois exposés au verbe *faire* dans l'input. Ils ont commencé les cours à la fin du mois d'août et cette séance a lieu le 16 mars. L'enseignante a commencé une nouvelle leçon en présentant le lexique à partir d'un poster sur lequel se trouvent dessinés tous les nouveaux items de la leçon. Après que les apprenants se sont entraînés à prononcer le lexique nouveau, l'enseignante leur demande d'ouvrir leur livret de grammaire à la page 29 (*Grammatisches Beiheft, Découvertes*) et de lire individuellement et en silence la conjugaison du verbe *faire*...

Dans son livret de grammaire, l'apprenant peut lire les choses suivantes :

Das Verb « faire »/ Le verbe « faire »*

Auch **faire** ist ein unregelmässiges Verb/ *faire est aussi un verbe irrégulier*

Faire		machen	
je	fais	ich	mache
tu	fais	du	machst
il		er	
elle	fait	sie	macht
on		man	
nous	faisons	wir	machen
vous	faites	ihr	macht/Sie machen
ils		sie	machen
elles	font		

* les phrases en italique sont une traduction qui ne figure pas dans l'original.

Entre la conjugaison du verbe *faire* et celle du verbe équivalent en allemand, *machen*, figure la transcription phonétique du français graphique. Par ailleurs, en face de *vous faites* et *ils/elles font* figure un petit personnage auquel on prête le commentaire suivant (nous traduisons) «Fais bien attention à ces deux formes ! Risque d'erreur ! »

Puis les apprenants ferment leur livre à la demande de l'enseignante et entonnent ensemble la conjugaison du verbe *faire* : « faire je fais tu fais il fait nous faisons, vous faites, ils font ». Enfin, chaque apprenant recopie cette liste dans son cahier. A cette activité vient se greffer une autre activité directement liée : les élèves font un exercice qui se trouve dans leur manuel (*Découvertes 1*, p.47 : 2b). Ils doivent compléter des phrases par le verbe *faire*, ce qui suppose de l'accorder avec le sujet :

b faire

Ahmed bavarde avec Julie et Isabelle.

Ahmed : Vous ... une promenade dimanche ?

Julie : Oui, nous ... une promenade dans la forêt de Fontainebleau.

A Fontainebleau papa et maman ... souvent des promenades.

Papa demande toujours : « Julie, on ... une promenade ? » Ah, je déteste ça !

Ahmed : Moi, dimanche, je ... des démonstrations de slalom.

Isabelle : Tu ... ça où ?

Ahmed : Devant la tour Montparnasse.

in *Découvertes 1*, p.47 : 2b

Ils corrigent l'exercice à voix haute et font l'exercice suivant qui consiste à compléter les phrases par *ont sont* ou *font*.

c ont/ sont/ font

Les parents d'Ahmed M. et Mme Saïd, ... une épicerie dans la rue Daguerre. Ils ... à Paris depuis 20 ans. Ils ... l'épicerie depuis 8 ans. Ahmed a une sœur : Yasmina. Ahmed et Yasmina ... des copains dans la rue Daguerre.

Les copains de la rue Daguerre ... toujours ensemble et ... des tours dans le quartier.

A la maison, Ahmed et Yasmina ... souvent la cuisine pour la famille.

in *Découvertes 1*, p. 48 : 2c

Après l'exercice précédent qui attirait l'attention des apprenants sur la flexion du verbe et l'accord avec le sujet, cet exercice fait appel au sens de *faire*, en contraste avec d'autres verbes non pas sémantiquement mais formellement proches. On voit se dessiner une ligne programmatique: d'abord focaliser l'attention des apprenants sur l'apparence formelle du verbe et son accord avec le sujet, soit la grammaticalité de la phrase puis, enfin, sur le sens de la phrase. La séance de 45 minutes se termine avec l'attribution de devoirs pour la prochaine séance. L'enseignante demande aux apprenants d'apprendre, outre le vocabulaire nouveau, la conjugaison du verbe *faire* avec toutes les personnes.

Lors de la séance suivante, le 19 mars, l'enseignante revient à la conjugaison du verbe *faire* : « *Le verbe faire, qui dit toutes les formes ?* » Deux apprenants, David et Martin, récitent la conjugaison puis l'enseignante demande à Kirstin d'aller au tableau écrire la conjugaison du verbe pendant que les autres l'écrivent sur leur cahier une nouvelle fois. Lorsque Kirstin s'est arrêtée d'écrire, l'enseignante s'adresse à elle en remarquant: « Na, Kerstin, glaubst du, dass die Endungen immer gleich sind ? » „*Kerstin, crois-tu vraiment que les terminaisons sont toujours les mêmes ?*». David est alors chargé d'aller au tableau afin d'aider Kerstin, soit de corriger les marques d'accord verbal qui ne sont pas correctes.

Le jour suivant, le 20 mars, les apprenants écoutent la cassette de la leçon. Sur la cassette, on entend le récit d'une soirée chez les Noblet, la famille dont il est question dans le manuel d'apprentissage du français. Le verbe *faire* apparaît dans le récit qui commence ainsi : « Le club de la tour Montparnasse/ chez les Noblet/ un autre soir // Nathalie et Luc font la cuisine ».

L'enseignante interrompt la cassette et pose une question en reprenant le verbe *faire* :

ENS : que font Nathalie et Luc ? Martin
MAR : ils dans la cuisine
ENS : attention, Mike
MIK : ils font la cuisine

La cassette reprend : « Le club de la tour Montparnasse/ Chez les Noblet/ un autre soir// Nathalie et Luc font la cuisine// Après le repas/ la famille reste à table //Luc montre des photos ». A nouveau l'enseignante pose une question avec le verbe *faire*, mais dans une question avec un sujet singulier et non plus pluriel :

ENS : alors, un autre soir chez les Noblet, que fait la famille après le repas ? (pause, face au silence de la classe, l'enseignante répond à sa propre question) après le repas la famille reste à table, la famille reste à table et que fait Luc ? Kerstin
KER : Luc montre des photos
ENS : oui

Le récit reprend avec un dialogue entre les personnages. Le verbe *faire* apparaît cette fois accordé avec la 1^{ère} personne du pluriel: « -regardez on est sur les photos [...] là nous faisons des démonstrations de Slalom ». Comme précédemment il s'ensuit un dialogue entre l'enseignante et certains apprenants afin de vérifier leur compréhension.

Ens : qu'est ce qu'ils font ?
x : ils font une démonstration de Slalom
Ens : oui et qui fait la démonstration ?
x : Luc
Ens : luc avec qui ?
y : Luc avec Ahmed
Ens : oui et avec les autres copains et ce sont des...ils ont des patins à roulettes alors ce sont des...
z : patineurs !
ens : oui

Ces exercices exposent les apprenants à différentes formes du verbe *faire* et permettent à certains de les utiliser en situation de production orale.

Nous pourrions bien entendu étendre cette description des activités en classe centrées sur l'apprentissage de l'accord verbal avec les différentes personnes du français à d'autres séances et observer combien les enseignants insistent sur l'accord sujet-verbe (cf. Granget 2004). Les activités d'apprentissage ne sont pas seulement des activités orales mais aussi écrites. Bien que le type d'exercice varie, l'objectif reste constant : assimiler la variation morphologique sur le verbe en fonction du sujet de la phrase. Pour le devoir sur table qui clôturera la leçon, les apprenants devront à nouveau apprendre la conjugaison du verbe *faire*, car des exercices porteront sur ce point.

Il nous semble difficile d'ignorer l'impact de ces activités spécifiques au milieu scolaire sur le développement des langues d'apprenants scolarisés. Ces exercices conjugués à des pressions institutionnelles fortes conduisent les apprenants à s'approprier la forme du verbe *faire* accordé avec la 3^{ème} personne du pluriel. L'usage régulier de ce verbe dans des activités de commentaire

de texte, d'images ou autres, en classe, peut également assurer la fixation de cette forme dans l'interlangue des apprenants.

6 Résumé et discussion

Nous mentionnions en introduction de ce texte qu'une question majeure en acquisition portait notamment sur le moment d'émergence de l'accord verbal dans les langues d'apprenants et sur le sens que l'on doit y accorder. Nous avons montré que le marquage de l'accord verbal avec un sujet pluriel est précocement réalisé dans les récits écrits d'apprenants germanophones scolarisés. Les récits écrits en français L2 nous montrent que les verbes sont porteurs d'une marque indiquant le nombre de personnes impliquées dans la situation à laquelle ils réfèrent. Les récits écrits montrent également que la réalisation de cette notion d'accord est prioritairement morphologique et s'exprime par ajout du suffixe *-nt* ou *-s*. Cette dernière marque montre que même en l'absence d'un marquage cible, l'apprenant peut exprimer la notion sémantique de pluriel. Ce résultat confirme par ailleurs les observations de Agren (ce volume) issues de récits écrits d'apprenants suédophones. Les apprenants germanophones, mais aussi suédophones, réalisent plus rarement l'accord de façon lexicale, par variation du radical du verbe, dans le cas des verbes pleins irréguliers, comme en témoignent les occurrences de *les deux prennent la fuite*. Or seule la réalisation lexicale de l'accord est audible, tandis que la réalisation morphologique, comme le souligne très bien Agren (ce volume), est silencieuse. Pour autant une réalisation silencieuse de l'accord à l'oral ne nous semble pas remettre en cause l'acquisition de l'accord. Ce résultat est intéressant dans la mesure où le développement du marquage temporel des énoncés par les apprenants non scolaires d'une L2 est décrit par Perdue et al. (2002) comme le passage d'une finitude sémantique à une finitude morphologique. Autrement dit, les apprenants emploient d'abord des moyens lexicaux pour marquer la référence temporelle puis, dans un deuxième temps, des moyens morphologiques. Il apparaît ici, concernant l'accord du verbe avec le sujet, que le parcours inverse vaut pour les apprenants germanophones scolarisés : ils ont d'abord recours à un marquage morphologique de l'accord verbal, puis, plus tard, à un marquage non seulement morphologique mais aussi lexical, en ce qui concerne du moins les verbes pleins irréguliers. Seuls quelques apprenants parmi les plus avancés de notre corpus emploient *ils prennent* par exemple.

Au delà des observations, comment expliquer cet accord sujet-verbe quasi-systématique et précoce chez les apprenants germanophones scolarisés du français, contrairement au « caractère extrêmement lent et graduel de l'accord sujet-verbe en français parlé » observé à partir des productions des apprenants suédophones (Bartning et Schlyter, ce volume) ? Outre les conditions de production (écrit versus oral), nous avons proposé deux hypothèses interprétatives. La première hypothèse consiste à expliquer l'absence de marquage résiduel par des propriétés des verbes et des sujets. Il apparaît cependant que l'absence de marquage est un cas résiduel qui survient avec tous types de sujet et tous types de verbes. Si certaines tendances se dégagent, elles ne peuvent toutefois expliquer à elles seules des difficultés de marquage du pluriel. Notre deuxième hypothèse n'explique pas les cas résiduels de non accord mais fournit une explication plausible à la présence précoce et stable de l'accord dans les récits d'apprenants germanophones scolarisés. Nous avons montré que l'environnement d'apprentissage et les conditions d'exposition au français pouvaient constituer une explication à cet usage abondant des marques de l'accord pluriel, dans la mesure où les « conjugaisons » des verbes sont enseignées précocement de façon

formelle à travers de nombreux exercices de mémorisation et d'application de ces connaissances déclaratives.

On peut donc penser que la catégorie de l'accord est une catégorie présente initialement dans l'interlangue de l'apprenant germanophone du fait qu'elle est enseignée et apprenable. Nous souhaiterions toutefois nuancer cette affirmation. En effet, nous n'avons guère abordé le rôle de la L1 dans cet article si ce n'est en relevant la présence de la catégorie du duel. Le recours à des expressions comme *le deux* pour référer à deux protagonistes du film dans le récit de fiction nous montre que les apprenants cherchent à exprimer en L2 une catégorie sémantique grammaticalisée dans la langue première. Ce procédé nous informe du rôle non négligeable de la langue source et précisément des catégories sémantiques grammaticalisées en langue source comme le duel, certes, mais aussi le nombre. Il est clair que les propriétés de l'accord morphologique en allemand devraient également être prises en compte dans l'analyse des données en français L2. Le marquage du pluriel s'exprime en effet en allemand au niveau du verbe (cf. Byrskog, ce volume). Ce fait est susceptible de favoriser le marquage de l'accord en français dès les débuts de l'apprentissage.

En conclusion, une explication à la présence précoce de l'accord avec la troisième personne du pluriel dans les récits écrits des apprenants germanophones scolarisés pourrait donc être que l'accord sujet-verbe est précocement et durablement enseigné en classe de langue et qu'il est apprenable, sachant la L1.

7 Références bibliographiques

- Granget C. (2004) *L'acquisition par des apprenants germanophones de la référence au passé en français : un cas d'école*. Thèse de doctorat, Paris III.
- Herschensohn J. (2001) « Missing inflection in second language French : accidental infinitives and other verbal deficits ». *Second Language Research* 17 : 273-305.
- Paradis J., Le Corre M., Genesee F. (1998) « The emergence of tense and agreement in child L2 French ». *Second Language Research* 14 : 227-255.
- Perdue C. (1993) *Adult Language Acquisition : Cross LInguistic Perspectives, volume I : Field Methods*. Cambridge University Press.
- Perdue C, Benazzo S., Giuliano P. (2002) « When Finiteness gets marked : the relation between morphosyntactic development and use of scopal items in adult language acquisition ». *Linguistics* 40 : 4, 849-890.
- Prévost P., White J. (2000) « Missing Surface Inflection or Impairment in second language acquisition ? ». *Second Language Research* 16 : 103-133.
- Vainikka A., Young-Scholten M. (1996) « Gradual Development of L2 Phrase Structure ». *Second Language Research* 12 : 7-39.

Manuel et livret de grammaire :

Découvertes 1. Série verte. Klett Verlag. Stuttgart, Düsseldorf, Leipzig. 1995.

Découvertes. Grammatisches Beiheft. Série verte. Klett Verlag : Stuttgart/Düsseldorf, Leipzig, 1995.

Formes verbales au pluriel à l'oral

- plus marquées chez les apprenants guidés que non-guidés ?

Suzanne Schlyter

Université de Lund
Institut d'études romanes
suzanne.schlyter@rom.lu.se

1 Introduction

Une question importante dans la recherche sur l'acquisition d'une L2 est de savoir dans quelle mesure l'instruction grammaticale joue un rôle pour la production libre, écrite ou orale. Nous savons que non seulement les règles fournies par l'enseignant jouent un rôle mais aussi l'input et l'intake de la L2. Dans Schlyter (2000) je donne des arguments en faveur d'un développement très semblable chez les apprenants guidés et non-guidés, établi à partir de leur production orale spontanée, concernant plusieurs phénomènes linguistiques. En revanche, Granget (2004) propose que le bon résultat des apprenants germanophones concernant la morphologie verbale, et notamment l'accord sujet-verbe, peut être dû à leur instruction formelle. Nous voulons ici comparer les apprenants suédophones guidés et non-guidés en ce qui concerne leur maîtrise de ces formes, surtout la 1ère et 3ème personne au pluriel.

2 Le Corpus FIFI

Nous avons comparé ainsi les marques de pluriel du verbe chez les apprenants non-guidés et guidés du Corpus Lund FIFI (Formell och Informell Fransk-Inläring) dans leurs conversations avec un francophone. Les apprenants non-guidés ont appris le français pendant un séjour à Paris, mais n'ont pas suivi de cours de français. Pour des détails sur les apprenants, voir p.ex. Granfeldt (2003). Les apprenants guidés étant généralement un peu plus avancés que les non-guidés, nous avons dans nos calculs enlevé l'apprenant guidé Magyu, nettement plus avancé que les autres, et l'apprenant Henry, nettement moins avancé. Il nous reste ainsi des apprenants de niveau intermédiaire (stade 2, 3 et 4 d'après Bartning & Schlyter 2004).

Nous avons relevé les cas où l'apprenant vise, sans ambiguïté, un sujet au pluriel, et nous en avons repéré les marques (ils sont, les filles ont, elles prennent, la fille et le garçon réussissent). Seuls les cas où il y a une marque audible sur le verbe à l'oral ont été pris en compte, ce qui exclut par exemple les filles parlent. Nous distinguons (à l'instar de Hedbor et de Ågren, ce volume) d'une part les quatre verbes, non-thématiques de haute fréquence, qui au pluriel terminent par -ont (ont, sont, vont, font), d'autre part les verbes thématiques (Vlex dans les tableaux). La différence

entre ces verbes est discutée aussi dans Schlyter et Bartning, ce volume.

Les tableaux ci-dessous indiquent les cas où le verbe est employé, par ces deux types d'apprenants, avec un sujet pluriel, nous ou la troisième personne (ils, etc.).

Tableau 1. *Accord au pluriel sujet-verbe : Apprenants NON-guidés, corpus Lund*

Appr., enregistr	nous Vons corr	nous V- incorr	ils -ont correct	ils - ont incorr	ils Vlexnt corr	ils Vlex- incorr	Stade
Björn 1-2	4	3	-	-	-	2	2
Sara 1	6	2	1	-	-	-	2
Sara 2	7	-	2	1	-	1	2
Petra 1-2	-	-	3	1	-	1	2
Martin 1	-	-	-	-	1	-	2
Ma 2-3	-	-	3	-	-	-	3
Karl 1	6	5	2	1	-	1	2
Karl 2	9	1	2	6	-	4	3
Karl 3	7	1	3	1	-	-	3
Karl 4	16	-	-	1	-	-	4
Karl 5	2	1	1	3	-	-	4
Johan 1	-	-	2	-	-	1	3?
Johan 2	4 (1 type)	-	8	1	-	-	3?
Johan 3	4	2	4	1	-	2	3?
Johan 4	3	-	-	2	-	1	3?
Knut 1	3	-	-	-	-	5	4
Knut 2	3	-	6	3	-	-	4
Knut 3	5	-	1	5	1	2	4
Total	79 (84%)	15	38 (59%)	26	2 (10%)	20	

Tableau 2. *Accord au pluriel sujet-verbe : Apprenants guidés, corpus Lund*

Appr., enregistr.	nous Vons corr	nous V- incorr	ils -ont correct	ils - ont incorr	ils Vlexnt corr	ils Vlex- incorr	Stade
Lisa 1	8	3	-	1	-	-	2-3
Lisa 2	11	6	-	2	-	-	2-3
Lisa 3	12	-	-	3	-	1	3
Lisa 4	5	2	-	6	-	3	3
Sama 1	2	-	-	5	-	1	3
Sama 2	5	1	-	1	-	-	3
Sama 3	-	-	3	3	-	-	3
Sama 4	1	1	3	3	-	4	3
Sama 5	1	-	-	-	-	-	4
Ylva 1	2	1	-	-	-	-	3?
Ylva 2	3	-	3	2	-	-	3?
Ylva 3	2	-	2	-	-	-	3?
Ylva 4	5	-	2	-	-	-	4?
Ylva 5	-	-	7	-	-	1	4?
Ylva 6	9	-	1	-	-	-	4?
Nina 1	14	1	3	-	1	1	4
Nina 2	21	1	3	1	1	1	4
Nina 3	7	1	3	2	-	-	4
Nina 4	6	1	3	(2)	1	-	5?
Nina 5-6	9	1	-	-	-	-	5
Total	123 (87%)	19	33 (52%)	31	3 (20%)	12	

Résumé

Appr., enregistr.	nous Vons corr	nous V0 incorr	ils -ont correct	ils - ont incorr	ils Vlexnt corr	ils Vlexnt incorr	Pluriel tot marqués
Non guidés	84% (79)	(15)	59% (38)	(26)	9% (2)	(20)	47%
Guidés	87% (123)	(19)	52% (33)	(31)	20% (3)	(12)	46%

Nous pouvons observer que le taux total de marquage correct de la 3e personne pluriel dans chacun des groupes d'apprenants, guidés et non-guidés, est pratiquement identique : 47% pour les non-guidés et 46% pour les guidés. On peut aussi observer un parallélisme presque parfait pour l'accord avec nous (83 vs 87%).

On peut éventuellement proposer que les non-guidés maîtrisent mieux les verbes non-thématiques de haute fréquence (token frequency) ont, sont, vont, alors que les guidés maîtrisent mieux les thématiques (verbes pleins), moins fréquents : veulent, prennent etc. Les données sont trop rares pour offrir des résultats pertinents, mais la question mériterait d'être étudiée : cela pourrait

signifier que les apprenants non-guidés, qui ont beaucoup plus d'input, sont plus sensibles à ces formes fréquentes, et les acquièrent par association (implicitement), alors que les apprenants guidés, qui ont davantage accès à la langue écrite - où le pluriel des verbes lexicaux est régulièrement marqué (type frequency) - seraient un peu plus sensibles à ce type d'accord (voir aussi Ågren, ce volume, pour un raisonnement semblable). Un raisonnement de ce type semble être renforcé par les données de Hedbor, ce volume.

3 Corpus FIFI et Corpus Interfra

Dans Bartning & Schlyter (2004) nous comparons nos deux corpus. Il y a donc encore un corpus d'apprenants guidés avec lequel les apprenants non-guidés peuvent être comparés.

Nos résultats communs concernant le pluriel est résumé dans le tableau suivant.

Tableau 3. *Résumé de l'apparition et l'utilisation des formes dans les deux corpus*

Phen / Enr.	J1 Cl	C a1 R 1 H e1 2	Ja 4 6	M t 1	C a3 5	B j1	R o 2 3	Pt 1 2	P e4 7	Sr 1 2	Li 1 3	T o 1 2 K a 1 2	R o 4 5 S a 1 2	O s1 2	Li 1 2	M t2 3	R o 6 7	K a2 3	S a 3 4	E m 4 7 S a5	E v 1 2	E 3 4	K a4 5	Y v 1 4 N i1 3	C l 4 N i4 6 M 1 4	L e1 4	K n 1 3
nous V- ons	-	-	+	-	+	+	+	-	+	+	+	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
ils ont	-	-	+	-	+	-	+	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Vlex- nt	-	-	-	+	+	-	+	-	+	-	-	-	-	-	-	+	-	-	-	+	+	+	-	+	+	+	+
Stad	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	5	5	4

Légende : Les apprenants soulignés en gras : non-guidés ; Si la valeur d'un trait diffère dans la même case, celle qui se trouve sur la 2e ligne concerne le corpus Lund. Valeurs : « - » = pas d'occurrences ; « + » = la forme apparaît ; « + » = utilisation à partir de 75% environ.

Le même ordre d'acquisition et le même manque de différence entre les apprenants guidés et non-guidés est visible ici. On observe pourtant aussi - mais les données sont à préciser - que les apprenants guidés débutants Ca, Ro, et Pe, au Stade 2 du corpus InterFra, produisent quelques cas isolés de 3e p pl (aussi bien du type 'ils -ont' que Vlex-nt), ce que ne font pas les non-guidés du corpus Lund. Il semble (mais les cas sont rares) que ces formes disparaissent par la suite (au Stade 3), avant qu'elles ne réapparaissent au stade 4 et surtout au stade 5. Il peut donc s'agir d'un développement en forme de U. Se peut-il que les apprenants guidés apprennent d'abord, à l'école, certains cas d'accord, qu'ils sont capables de mémoriser quand ils n'ont pas encore beaucoup de français à leur disposition, mais qu'ils perdent par la suite ? Ou bien, c'est la langue parlée, avec son système moins marqué pour le pluriel (forme unique parle/parlent) qui prend la relève, quand les apprenants sont plus avancés et entendent plus de français oral ? Voir aussi Ågren, ce volume,

pour une discussion de ces phénomènes.

4 Résultats préliminaires

- Corpus FIFI, Lund : aucune différence entre les apprenants guidés et non-guidés en ce qui concerne le pluriel.
- Corpus Lund + Corpus InterFra : davantage de marques au début chez les guidés, mais éventuellement une courbe en U.

Nous avons donc obtenu des résultats assez différents de ceux proposés par Granget (2004). D'où vient la différence ? Il est bien possible, comme le propose Granget, que ces apprenants guidés et non-guidés se comportent d'une manière assez semblable, vu que se sont de jeunes étudiants qui avaient déjà appris l'anglais à l'école, et qui donc sont capables d'utiliser des stratégies semblables. Les convergences observées au niveau de leurs stratégies interactionnelles (voir Bozier 2005) pointent dans la même direction. Les deux types d'apprenants communiquent à l'oral, mais pas tellement au-dessus de leurs moyens linguistiques, étant donné qu'ils ont souvent recours à l'anglais quand la communication devient trop compliquée.

Toujours est-il que les apprenants guidés et non guidés développent un marquage verbal de l'accord au pluriel très semblable et suivent, qualitativement, le même itinéraire.

5 Références bibliographiques

- Bartning, I. et Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies*, 14:3:281-299.
- Bozier, C. (à paraître septembre 2005) *De la sollicitation à la séquence potentiellement acquisitionnelle*. (titre provisoire) Thèse de doctorat, Université de Lund.
- Granget, Cyrille (2004) « L'acquisition par des apprenants germanophones de la référence au passé en français : un cas d'école ». Thèse de doctorat, Paris III.
- Schlyter, S. (2000) « Acquisition du français parlé - une comparaison entre apprenants guidés et non-guidés ». In Andersen, L. H. et Hansen, A. B. (éds.) *Le français parlé. Actes du colloque international*, Université de Copenhague, du 29 au 30 octobre 1998. Museum Tusculanum Press : Copenhagen.

La morphologie du nombre dans le syntagme verbal en français L2 écrit - L'accord de la 3^{ème} personne du pluriel

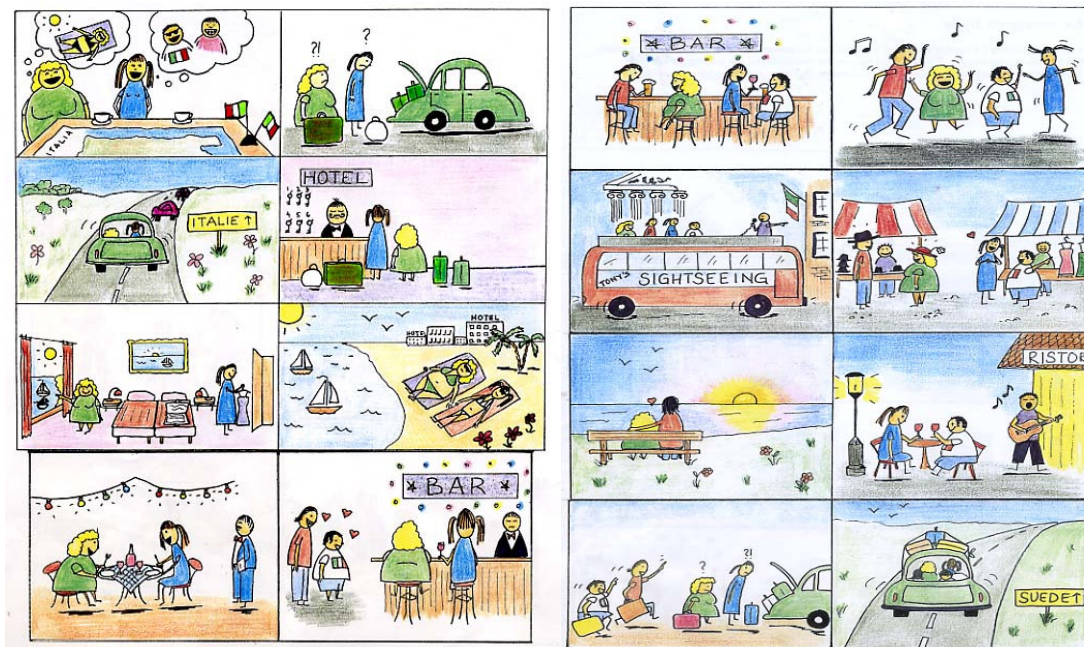
Malin Ågren

Institut d'études romanes
Université de Lund, Suède
malin.agren@rom.lu.se

1 Introduction

Les différences importantes entre le français parlé et le français écrit sont à l'origine de bien des difficultés dans l'acquisition du français écrit, aussi bien pour les enfants francophones (L1) que pour les apprenants de français langue étrangère (L2). Cet article propose une analyse de la production écrite des apprenants de français L2 dans le domaine de l'accord verbal en nombre de la 3^{ème} personne du pluriel (désormais 3pl), cet accord étant le plus souvent de nature silencieuse. L'article met l'accent sur le développement de l'acquisition de la morphologie verbale de la 3pl et discute les facteurs qui influencent la mise en œuvre de la morphologie muette de l'écrit en français L2. Basés sur la série d'images *Le voyage en Italie*, les deux textes ci-dessous, écrits en français L2, illustrent la différence très nette du niveau linguistique général entre un apprenant au stade initial d'acquisition (Amie) et un apprenant à un stade avancé (Daisy).

LE VOYAGE EN ITALIE



Le texte d'Amie

Karin et sa sœur Josefine, Karin a robe vert et Josefin robe bleu. A tableau, deux drapeau, deux tasse de café. Karin est une grosse fille et Josefine une jeune fille. Karins sac a provision, les bagages, et un voyageur vert. Karin et Josefin aller de l'Italie. Trois voyages, un vert, deux rouge. Quatre fleurs. Quatre bagages, Six clefs. Karin et Josefine parler de l'homme, un chambre sil vous plaît, - oui, combien? -deux. Karin et Josefine et la chambre, la plage et mer sa extraordinert. trois robes, un bleu, un vert et un lila. AHHH la plage, sa mer et l'hotel, trois fleurs de rouge. Karin et Josefin se faire tres fatigue. Se manger, une vin blanch, les escargots, un biff tartare, et mousse au chocolat. Karin et Josefine voix le meny. LA BAR...Karin et Josefin boire rouge vin et l'eau, Trois l'hommes, deux parler avec Karin et Josefin. OHh lala, J'etaime J'etaime. Ils parlent la nuit. Ils dansent all nuit, tralala. Un maison de orange, la tonys l'autobus de marron, cinq personnes et un drapeu de l'Italienne.

Extrait du texte de Daisy

Il était une fois 2 femmes qui s'appelaient Sylvie et Simone. Ils avaient travaillé toutes ses vies pour pouvoir faire un voyage en Italie. Sylvie, une dame assez grosse avec les cheveux blondes qui aimait se dresser en verte, voulait partir en Italie pour se reposer au soleil à la plage est devenu bronzée. Simone, une fille mince avec des cheveux bruns et longues qui aimait le couleur bleu, voulait y aller pour faire la connaissance des beaux mecs italiens. Une belle journée, ils sont parties. Ils vont voyager par la voiture verte de Sylvie, (naturellement). Quand ils étaient en train de mettre ses pinales à la voiture, c'était très dût parce que ils avaient beaucoup de valises. Mais une heure plus, ils ont laissé la Suède pour aller en Italie par l'autoroute. Après deux jours, ils sont arrivées en Italie. Ils veulent vivre à l'hôtel "Magnifico" et ils ont reçu la clé pour leur chambre du receptioniste, Carlo. La chambre était merveilleuse. Elle donnait sur la plage, et on pouvait voir l'eau bleu et le soleil qui brillait. Sylvie admirait la vue autant que Simone s'installait dans la chambre. Ils avaient seulement une chambre mais il y avait deux lits très confortables. Sylvie et Simone était très jolies pour avaient trouvé leur paradis. Sylvie et Simone s'amusaient beaucoup. Elles passaient ses jours à la plage, au soleil. Ils ont vu des voiles sur l'Océan et des oiseaux au ciel. C'était fantastique, un vrai paradis! ...

Contrairement au français oral, le français écrit est caractérisé par une morphologie riche. Une caractéristique de la morphologie propre à l'écrit est qu'elle est souvent redondante. Dans le cas régulier, le nombre à l'écrit est marqué morphologiquement par un –s dans le syntagme nominal (SN) et par un –nt¹ dans le syntagme verbal (SV). Ces morphèmes sont, à l'exception de rares cas de liaisons, silencieux à l'oral. Il a été constaté que ces phénomènes sont parmi les plus difficiles et les plus longs à maîtriser lorsque l'on apprend à écrire en français L1 (Totereau, Thévenin et Fayol, 1997 ; Jaffré et David, 1999).

Dans une étude préalable sur l'acquisition de la morphologie du nombre dans la phrase nominale (Ågren, 2005ms), il a été constaté que les apprenants suédophones de français L2 marquent très tôt le nombre sur le nom et le pronom à l'écrit. Déjà à un stade précoce, de façon générale caractérisé par une activité morphologique plutôt faible, les apprenants marquent le nombre sur le nom à 88 % des contextes au pluriel. Il en est tout autrement avec l'accord en nombre de l'adjectif qui, au contraire, est un phénomène tardif et graduel, qui n'est produit de façon relativement correcte qu'à un stade avancé. Le marquage du nombre sur l'article semble acquis à un stade intermédiaire.

Nous avons proposé que l'acquisition de la morphologie silencieuse du nombre en L2 dépende de divers facteurs, entre autre la sémantique, la syntaxe, la fréquence et les contraintes cognitives².

¹ Dans cet article, nous allons faire référence au suffixe de la 3^{ème} personne du pluriel (3pl) comme le morphème –nt, selon la norme dans la littérature sur l'acquisition du français écrit L1. Nous sommes pourtant conscients du fait que dans l'enseignement du français L2 en Suède, la plupart des manuels et des professeurs se réfèrent au morphème –ent dans le contexte de la 3pl.

² Le transfert de l'anglais est un autre facteur qui a été mentionné pour expliquer le marquage extrêmement précoce du morphème du nombre sur le nom. Il est certainement question d'un facteur qui aide à l'automatisation du –s en français L2 mais qui ne suffit pas en lui-même pour expliquer le processus d'acquisition.

Selon Bybee (1985), la pertinence sémantique joue un rôle important en morphologie. Cette idée est également soutenue par des recherches sur le français L1 écrit (Totereau, 1998 ; Fayol, 2003, entre autres). Ainsi, un morphème est pertinent pour un radical si son contenu sémantique affecte de façon directe le contenu sémantique du radical. Puisque le nombre influence le contenu sémantique du nom (et du pronom) de façon beaucoup plus prononcée que celui de l'adjectif ou de l'article, nous proposons que ce facteur explique en partie le marquage du nombre des différents éléments du SN en acquisition du français L2 écrit.

Concernant l'accord en nombre du verbe, il s'avère qu'en acquisition du français L1 écrit, cet accord est acquis plus tardivement que l'accord en nombre du nom et l'adjectif. En s'éloignant de la sphère nominale, la marque du pluriel change. Le *-s*, jusque là omniprésent, est accompagné par un *-nt* dans le SV. Comme l'expriment Jaffré et David (1999 : 18) : « Cette dissymétrie est à l'origine de bien des difficultés dans la maîtrise de l'orthographe du nombre ». L'on a pu constater qu'il existe une phase où le morphème du pluriel du SN, le *-s*, a tendance à être transféré au SV de la façon suivante, ils joues (Totereau, Thévenin et Fayol, 1997, Totereau et al 1998).

Le but de cet article est de décrire l'acquisition et la maîtrise de l'accord en nombre du SV en français L2 dans une perspective développementale. Le nombre n'ayant pas d'influence directe sur le verbe, nous nous attendons à ce que l'accord en nombre se fasse plus tardivement dans le SV que dans le SN, tout comme en français L1. Un autre facteur qui parle en faveur d'une acquisition tardive et graduelle de l'accord en nombre du SV est l'inexistence de flexions verbales en personne et en nombre dans la langue maternelle des apprenants suédois.

2 Le corpus CEFLE³ de Lund

La présente étude sur l'accord verbal du nombre dans l'acquisition du français L2 écrit se base sur un corpus de 85 apprenants formels de français au lycée en Suède. Les apprenants se trouvent à des niveaux différents dans l'acquisition du français, allant des débutants aux apprenants plutôt avancés. A ce matériel s'ajoute un groupe contrôle de 22 lycéens français.

Les apprenants suédois ont été suivis pendant une année scolaire et ont écrit quatre tâches à deux mois d'intervalle. Les textes ont été écrits sur ordinateur et il s'agissait de raconter des histoires à partir de séries d'images ou d'écrire des récits plus personnels du genre : *Un souvenir de voyage* ou encore *Moi, ma famille et mes amis*. Les tâches ont été construites pour encourager une écriture aussi libre et spontanée que possible, tout en assurant un cadre commun à tous les apprenants.

3 Méthode et but

L'analyse sur laquelle se fonde cet article est une étude transversale de quatre groupes d'apprenants de français L2 et un groupe contrôle français. Les textes des apprenants suédois et des scripteurs francophones sont basés sur une série d'images appelée *Le voyage en Italie*.

³ Le corpus CEFLE – Corpus Écrit de Français Langue Étrangère (de Lund).

En ce qui concerne les apprenants formels d'une langue étrangère, il est connu que la longueur des études est un facteur peu informatif du niveau linguistique des apprenants (voir par exemple Bardovi-Harlig, 2000 ou Ellis, 1994). Pour cette raison, nous avons opté pour un regroupement des apprenants selon leurs connaissances linguistiques générales. La division en stades 1, 2, 3 et 4 des apprenants suédois est ainsi fondée sur des critères qui distinguent des phases acquisitionnelles selon le modèle de Bartning et Schlyter (2004). Le regroupement selon le niveau linguistique des apprenants a ensuite été vérifié par un test de niveau. Le nombre d'apprenants à chaque stade acquisitionnel est visualisé dans le tableau 1.

Tableau 1. *Le nombre d'apprenants de français du corpus par stade acquisitionnel*

	Stade 1	Stade 2	Stade 3	Stade 4	Contrôle
	Initial	Post-initial	Intermédiaire	Pré-avancé	Français
Gr.1 (début.)	4	4	1	-	-
Gr. 2 (1 an)	3	12	5	-	-
Gr. 3 (3 ans)	3	13	12	2	-
Gr. 4 (4 ans)	-	-	11	15	-
Total	10	29	29	17	22

Légende : La répartition des apprenants en groupes selon le nombre d'années d'études de français se voit sur l'axe vertical et les stades de Bartning & Schlyter sont exposés sur l'axe horizontal. En bas du tableau on trouve le nombre total d'apprenants pour chaque stade acquisitionnel. C'est ce regroupement qui formera la base de notre analyse.

Puisque l'étude transversale a originalement été réalisée pour une analyse du syntagme nominal, les critères sur lesquels se base la division du matériel en stades sont essentiellement des critères verbaux : l'organisation de la phrase⁴, l'omission du verbe, la finitude (avec les sous-catégories : formes non-finies dans un contexte fini, l'emploi d'une forme de base fondée sur la 3^{ème} personne du singulier (3sg), l'accord en personne), le temps et l'aspect. Nous avons souhaité ainsi éviter les analyses circulaires. En entrant maintenant dans l'analyse de la sphère verbale, nous sommes conscients des possibles problèmes méthodologiques. Pour éviter la circularité des arguments, nous n'allons pas commenter les critères verbaux mentionnés ci-dessus dans notre analyse des verbes, mais uniquement mettre l'accent sur le nombre dans la phrase verbale, plus particulièrement sur l'accord à la troisième personne du pluriel (3pl).

Les questions auxquelles nous souhaitons répondre sont les suivantes : Quand les apprenants en français L2 écrit commencent-ils à faire l'accord verbal en nombre ? Existe-t-il une différence entre les divers groupes de verbes en ce qui concerne l'accord en nombre ? Y a-t-il une tendance de surrégularisation du marquage du nombre ? Existe-t-il un transfert du morphème du nombre nominal, le –s, au SV ?

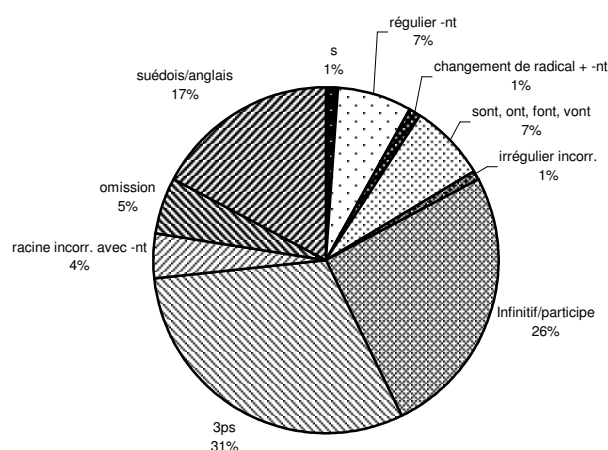
⁴ L'organisation de la phrase se réfère aux capacités des apprenants d'organiser la phrase selon une structure nominale (sans verbe) ou verbale (non-fini ou fini) : NUO = Nominal Utterance Organisation, IUO = Infinite Utterance Organisation et FEO = Finite Utterance Organisation. Ces structures reflètent le niveau linguistique des apprenants dans les variétés Pre-Basic, Basic et Post-Basic de l'étude ESF (Klein et Perdue 1992 et 1997).

4 Analyse des données

4.1 Accord verbal en nombre - un développement graduel

En tenant compte de tous les contextes⁵ de la 3pl dans l'étude transversale, il est possible de voir un fort développement de l'accord verbal en nombre au cours de l'acquisition du français L2 écrit. Les figures 1 et 2 montrent les proportions du marquage ou du non-marquage du nombre dans la SV aux stades 1 et 4. Faute de place, nous avons choisi le point de départ et le point final de notre étude pour bien illustrer les changements qualitatifs qui ont lieu au cours de l'acquisition du français L2 écrit.

Figure 1: Le nombre dans le syntagme verbal (3pl) Stade 1
Contextes au total: 100



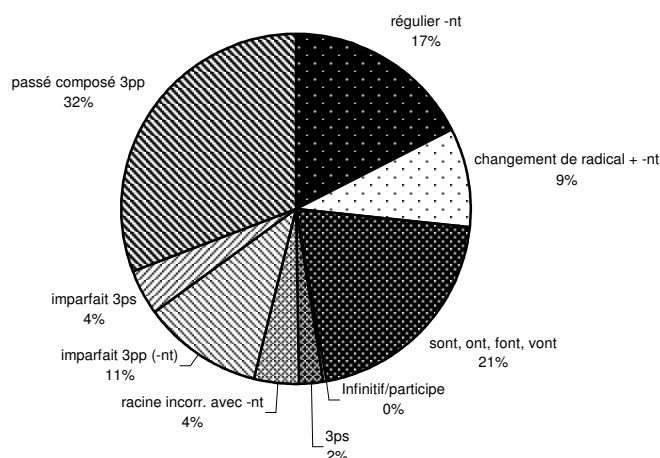
Au stade 1, les apprenants respectent assez rarement l'accord en nombre. Comme l'indique la figure 1, les apprenants conjuguent les verbes correctement à la 3pl dans 15 % des cas (7 % régulier -nt, 7 % *sont, ont, vont, font* et 1 % changement de radical + -nt). Le plus souvent, ils écrivent à l'infinitif (26 %) ou à la 3sg (forme de base : 31 %). Nous soupçonnons que les difficultés d'accord s'expriment également par des omissions de verbes au pluriel ou par des formes verbales en anglais ou en suédois.

Au stade 4 au contraire, comme l'illustre la figure 2, à peu près la moitié des contextes au présent correspond à des formes correctes à la 3pl (17 % réguliers -nt, 9 % changement de radical + -nt et 21 % *sont, ont, vont, font*). Il est très rare que les apprenants omettent l'accord en nombre dans un contexte au présent (2 %) ou au passé (4 %). Le taux d'omission de l'accord verbal en nombre du stade 4 correspond tout à fait au niveau du groupe contrôle français.

⁵ Par contexte à la 3pl, nous entendons toutes les fois où les scripteurs visent un verbe à la 3pl selon leur propre façon de s'exprimer, c'est-à-dire quand le sujet de la phrase est à la 3pl : ils/elles..., les filles..., Marie et Sophie... .

Figure 2: Le nombre dans le syntagme verbal (3pl) Stade 4

Contextes au total: 462



Da

ns les sections suivantes nous allons analyser l'accord verbal en nombre à tous les stades de notre corpus ainsi que chez le groupe contrôle afin d'évaluer le changement qualitatif qui semble avoir lieu en acquisition du français L2 écrit.

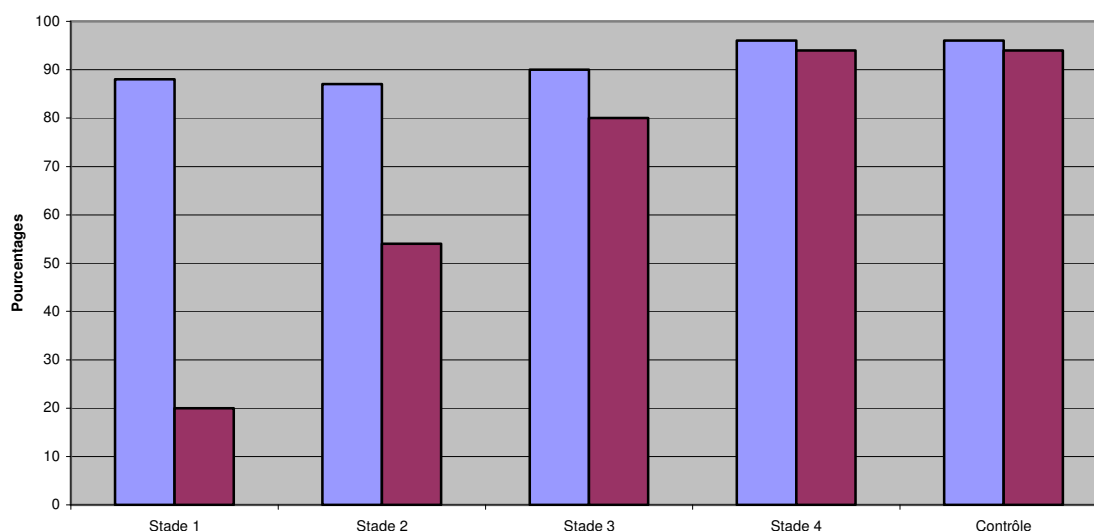
4.2 Le nombre marqué – le verbe par rapport au nom

Le voyage en Italie est une série d'images où le pluriel est explicitement élicité. Il y a plusieurs personnes qui agissent dans l'histoire et il est impossible de la raconter sans être confronté à des contextes pluriels. Néanmoins, au début de l'acquisition, les apprenants semblent parfois réussir à éviter le pluriel. Il est possible d'interpréter les stratégies de deux apprenants au stade 1 comme des efforts pour éviter les contextes du pluriel. Andreas emploie le sujet « le famille » suivit d'un verbe au singulier ou non-conjugué dans tous les contextes : *Le famille voyager depuis quatre semaine après*. Angela choisit d'éliminer une personne de l'histoire. Par conséquent, ce n'est qu'une fille qui agit dans son texte : *Pascale va a de chaud et belle Italie...*

Nous observons que le marquage de la 3pl est un phénomène qui est acquis et maîtrisé graduellement (ou en étapes) en acquisition du français L2 par des suédophones. Ce fait est reflété dans le diagramme 1 qui montre la proportion de verbes accordés à la 3pl par rapport aux noms accordés au pluriel⁶. Dans une étude antérieure, nous avons observé que le nombre était marqué sur le nom déjà à un stade précoce et que, par conséquent, le développement restait plutôt faible entre les stades 1 à 4. A l'instar de la théorie de Bybee (1985), et l'étude de Totereau (1998) et d'autres, sur l'importance de la relation entre sens et forme, nous nous sommes attendus à ce que le marquage du nombre aux noms précède largement celui des verbes. Nos données confirment cette hypothèse comme le montre le diagramme 1 ci-dessous.

⁶ Dans ce diagramme, il est question d'un marquage du nombre dans un sens très vaste. Des moyens divers sont acceptés comme étant des marques du nombre au verbe : au présent -nt, -s, -ont, sont/ont/font/vont, partirent, etc.

Diagramme 1: Le marquage du nombre sur le nom (clair) et le verbe (foncé)



Il est à noter qu'au stade 4, les apprenants ont atteint le niveau du groupe contrôle quant à l'accord du nombre verbal. Or, nous allons voir dans une analyse plus fine, qu'il se cache des différences qualitatives entre les apprenants avancés et les scripteurs Français en ce qui concerne le marquage du nombre sur le verbe.

4.3 Pluriel des verbes réguliers en -er

Les apprenants de français L2 sont confrontés à une grande variété de formes dans le paradigme des verbes réguliers en -er. Plusieurs phénomènes peuvent être considérés à ce propos : la différence du marquage morphologique des personnes grammaticales en français écrit par rapport au suédois, la différence entre le français oral et le français écrit ainsi que le rôle de l'enseignement dans l'acquisition de la morphologie de l'écrit.

En suédois, ni la personne grammaticale ni le nombre ne sont marqués morphologiquement sur le verbe. Toutes les personnes ont par conséquent la même forme: *jag spelar, du spelar, han spelar, vi spelar, ni spelar, de spelar*. De ce fait, de grandes différences entre langue source et langue cible se présentent aux apprenants. En raison de ce décalage, le paradigme verbal au présent des verbes réguliers en -er est répété en classe dès le début de l'enseignement. En ce qui concerne la 3pl, le morphème des verbes réguliers présente l'avantage de correspondre à une relation bi-univoque. La règle est ainsi enseignée et focalisée: **Si** verbe en -er et 3pl, **alors** -nt (parler + -nt = parlent).

Or, non seulement les apprenants rencontrent une morphologie verbale très riche mais ils sont également confrontés aux différences entre l'oral et l'écrit qui caractérisent la morphologie des verbes réguliers en -er. Dans l'enseignement traditionnel du français en Suède, l'input écrit joue un rôle important et les apprenants sont confrontés à une double tâche : mettre en œuvre le paradigme verbal des personnes et ne pas prononcer les morphèmes flexionnels du nombre de l'écrit. Ils doivent aussi apprendre à faire figurer à l'écrit des graphèmes qui ne se prononcent pas.

Notons que du point de vue du sous-système de l'écrit, les conditions d'acquisition entre les apprenants de français écrit L1 et L2 sont essentiellement différentes. Les apprenants du français écrit L1 partent d'un input oral massif et d'un système linguistique oral bien en place alors que les apprenants suédophones sont simultanément confrontés à un input écrit et oral dès les premiers contacts avec la langue. En conséquence, les apprenants suédois pourraient effectivement être moins affectés par le côté silencieux du français écrit que les enfants francophones apprenant à écrire. Une telle différence a déjà été observée pour les pronoms (Ågren, 2005ms) dans une comparaison entre les adolescents Français du groupe contrôle et les apprenants avancés de notre corpus.

En ce qui concerne l'accord verbal en nombre des verbes réguliers en -er, nous nous attendons donc, premièrement, à ce que le morphème -nt de la 3pl se présente de manière isolée déjà au stade 1. Cette hypothèse se base sur les effets de l'enseignement de la morphologie de l'écrit. Deuxièmement, à cause des différences principales d'input en L1 et L2 et les tendances observées pour les pronoms, nous nous attendons à ce que le stade 4 souffre moins que le groupe contrôle quant à la différence entre *ils parlent* et *il parle*. Le tableau 2 ci-dessous montre comment l'accord en nombre des verbes en -er augmente au fur et à mesure de l'acquisition.

Tableau 2. *Les verbes réguliers en -er à la 3pl (-nt) : contextes, occurrences, types et pourcentages corrects versus occurrences et pourcentages incorrects (en gris)*

Stade	Cont. 3pl. -er	Occurrences correctes	% corr.	Types Corr.	O.c./appr.	Inf .	% inf.	3sg.	% 3sg.
1	35	7	20 %	6	0,7	13	37 %	15	42 %
2	95	51	54 %	20	1,8	20	21 %	24	25 %
3	165	117	71 %	36	4,0	12	7 %	36	22 %
4	84	76	90 %	40	4,5	2	3 %	6	7 %
Contr.	150	133	89 %	78	6	-	-	17	11 %

Légende : O.c/appr. = occurrences correctes par apprenant en moyenne,

Inf. = infinitif à la place d'un verbe en -er conjugué à la 3pl,

3sg. = la 3^{ème} personne du singulier remplace la 3^{ème} personne du pluriel.

En prenant en considération, de façon générale, le rôle de l'enseignement pour les apprenants de notre corpus et le rôle de l'écrit dans cet enseignement, il n'est pas étonnant de constater que déjà au stade 1, un certain nombre d'apprenants marque l'accord verbal des verbes réguliers en -er avec le morphème -nt, comme dans les deux exemples: *Ils parlent la nuit. Ils dansent all nuit, tralala* (20 %). Cette constatation correspond à nos hypothèses. Or, le marquage du nombre est loin d'être conséquent. N'oublions pas que dans la majorité des contextes à la 3pl, le verbe se présente sous forme non-finie, *ils parler ou ils parlé* (37 %) ou sous forme de la 3sg, *ils parle* (42 %) à ce stade précoce, même lorsqu'il est précédé par un pronom.

Au stade 2, 17 apprenants (sur 29) marquent au moins une fois le nombre de façon normative aux verbes en -er. Comme l'indique le tableau 2, le nombre d'occurrences et de types de verbes à la 3pl a augmenté par rapport au stade précédant. À ce stade intermédiaire, 8 apprenants ont pourtant choisi de s'inclure dans l'histoire (je, nous), ce qui explique pourquoi il n'y a pas de contexte de

la 3pl chez ce groupe. Dans de nombreux contextes à la 3pl, le morphème –nt manque au verbe, qui est alors exprimé soit par la forme non-finie (21 %) soit par la forme 3sg (25 %).

Au stade 3, l'accord en nombre des verbes en –er est plus systématique. Selon le tableau 2, les occurrences et les types ont encore augmenté. Sur un total de 29 apprenants, 23 marquent de façon normative la 3pl aux verbes en –er. Ceux qui ne le font pas écrivent soit en 1pp (nous), soit au passé composé/imparfait. Il existe de moins en moins de formes à l'infinitif (7 %) ou à la 3sg (22 %) dans un contexte à la 3pl. Il est à remarquer que les formes erronées à la 3sg persistent plus longtemps que celles à l'infinitif. Nous en tirons la conclusion que les apprenants maîtrisent la finitude avant l'accord en nombre à l'écrit (cf. Schlyter et Bartning, ce volume).

Parmi les 17 apprenants au stade 4, seuls 10 écrivent leurs textes au présent, ce qui explique pourquoi les occurrences de verbes en –er conjugués en nombre sont moins nombreuses qu'au stade 3. Le nombre de types différents est pourtant un peu plus élevé qu'au stade précédent. L'absence d'accord est extrêmement rare. Le groupe contrôle ressemble au stade 4 dans le sens où l'accord en nombre à la 3pl est bien maîtrisé. Pourtant, le nombre de types de verbes conjugués à la 3pl est deux fois plus grand pour les scripteurs francophones (78 versus 40). Cette diversité lexicale au niveau des verbes s'explique par le fait que les jeunes Français ont tendance à employer une plus grande variété de verbes en –er pour éviter la répétition ou pour ajouter des nuances, comme *se régaler, discuter, déguster, loger, se précipiter, débarquer, embarquer, accompagner*, etc., là où les apprenants suédois, même ceux à un niveau avancé, se servent de verbes habituels comme *s'amuser, parler, manger, habiter*, etc. ou de *sont/ont/vont/ont* (voir section suivante).

Pour conclure, il est à noter que le groupe contrôle et le stade 4 se distinguent l'un de l'autre dans la mesure où les apprenants avancés font moins d'erreurs du genre *elles danse* que les scripteurs francophones en même temps qu'ils gardent quelques restes d'un infinitif à la place d'un verbe conjugué à la 3pl (*Sylvie et Virginie bronzer*), ce qui n'existe pas dans le groupe contrôle. Ceci pourrait effectivement être dû au fait que les adolescents francophones sont plus influencés par un input oral dans lequel le nombre reste muet que les apprenants de français L2.

Une dernière remarque sera consacrée à la vulnérabilité de l'accord en nombre silencieux qui existe également chez les scripteurs experts. Plusieurs chercheurs ont étudié le caractère cognitivement coûteux de l'orthographe silencieuse du français écrit (Fayol et Bourdin, 1994 ; Fayol, Largy et Lemaire, 1994 ; voir aussi Largy, ce volume) et en tirent la conclusion que même des scripteurs adultes éduqués commettent des erreurs d'accord en nombre sous des conditions dans lesquelles les contraintes cognitives de la tâche à accomplir sont trop lourdes. Vu que le groupe contrôle français de notre corpus comprend 22 adolescents qui n'ont pas encore terminé leur formation au lycée, il est probable que le côté vulnérable de l'accord en nombre se manifeste également dans notre corpus. En examinant les 20 contextes où les scripteurs français omettent le nombre sur le verbe, trois caractéristiques spécifiques reviennent à plusieurs reprises (le même phénomène est observé chez les apprenants au stade 4). D'abord, une partie des omissions apparaissent lorsque deux verbes sont coordonnés et que le deuxième verbe se trouve éloigné du sujet pluriel de la phrase (Pienemann, 1998). Ce fait semble peser lourdement sur la mémoire de travail. Un pronom objet peut également être inséré entre le sujet et le verbe, avec le même effet. Dans plusieurs exemples, ces deux facteurs coopèrent :

Ernest : *Ils dansent toute la nuit et s'amuse beaucoup.*

Eugène : *Elles trouvent une petite pizzeria et s'y installe.*

Ernest : *Le lendemain elle vont au bar et y rencontre deux jeunes hommes...*

Daniel : *Elles les adore, les hommes et les choses.*

Il arrive aussi que l'omission de la forme plurielle du pronom s'accompagne de l'omission de la marque du pluriel sur le verbe.

Ernest : *Deux suédoises sont autour d'une table. Elle regarde avec envie une carte d'Italie...*

Pourtant, dans un certain nombre de cas, aucun facteur syntaxique apparent n'explique l'omission du nombre dans le SV.

4.4 Pluriel des verbes irréguliers fréquents: *sont, ont, vont et font*

En français, seuls quatre verbes ont un paradigme irrégulier. Ceux-ci sont très fréquents, notamment dans la fonction de verbes auxiliaires (à l'exception de *faire*). Ces verbes ne marquent pas le nombre avec le morphème *-nt* mais avec une forme indépendante : *sont, ont, vont et font*. Pour les raisons mentionnées ci-dessus, irrégularité par rapport au principe général et haute fréquence, ces formes sont probablement gardées de façon individuelle en mémoire (voir Bybee, 1985 ; Ellis, 2002 et aussi Pinker, 1998 ; Clahsen, 1999, Hedbor, 2003 et ce volume, pour les formes irrégulières et fréquentes). De plus, ces formes diffèrent également des formes en *-er* puisque le nombre y est audible. Toutes ces facteurs parlent en faveur d'une acquisition et une production précoce de ces formes en français L2. Du même fait, l'on pourrait également s'attendre à une production plus correcte de ces verbes à la 3pl que des verbes réguliers. Le tableau 3 ci-dessous confirme ces hypothèses.

Tableau 3. *Les verbes irréguliers être, avoir, aller et faire à la 3pl : contextes, occurrences, types et pourcentages corrects versus occurrences et pourcentages incorrects (en gris)*

Stade	Contextes *ont	Occurrences correctes	% corr.	O.c./appr.	Inf .	% inf.	3sg.	% 3sg.
1	21	7	33 %	0,7	6	29 %	8	38 %
2	142	94	66 %	3,2	16	11 %	32	23 %
3	203	171	84 %	5,9	11	5 %	21	10 %
4	95	93	98 %	5,5	-	-	2	2 %
Contr.	69	69	100 %	3,1	-	-	-	-

Légende : *O.c/appr.* = occurrences correctes par apprenant en moyenne, *Inf.* = infinitif à la place d'un verbe conjugué à la 3pl, *3sg.* = la 3^{ème} personne du singulier remplace la 3^{ème} personne du pluriel.

Au stade 1, il n'y a que deux apprenants qui emploient ces verbes à la 3pl. Dans l'ensemble cela correspond à 33 % des occurrences. Les formes en question sont *ont* (5), *vont* et *font*. La façon dont ces verbes sont employés témoigne d'un emploi plutôt figé de ces unités. Le texte d'Annette illustre le fait que *ont* sert de « passe-partout » :

Deux filles imaginaire. Elles ont pour L'Italie sur la voiture. Elles ont sur la rue. Elles arrive a l'Hotel. Elles fâr les clés. Le soleil skiner sur la fenêtre. Il y a trop chaud sur la playa.[...] La fille et le garçon dîner a la restaurant. Amour est sur l'air. Les garçons ont aussi. Enfin

ils ont tillbaka pour la Suede.

Au stade 2, il y a une hausse considérable d'occurrences pour ce groupe de verbes. Sur un total de 29 apprenants, 24 les emploient dans leurs textes. Il est question d'une haute fréquence par rapport aux verbes réguliers en –nt : 3,2 *sont/ont/vont/font* par apprenant comparé à 1,8 verbes réguliers en –nt par apprenant. De même, l'emploi de ces verbes irréguliers et fréquents à la 3pl est de plus en plus correct : 66 %. Il semble même que ce groupe de verbes joue un rôle de modèle pour l'accord en nombre et que le [õ]-final de ces verbes soit contagieux. L'on peut noter de nombreux exemples des morphèmes –ont ou –ons à la place de –nt pour la 3pl, par exemple *ils allont, ils boiyons un verre du vin, ils avont trop beaucoup valises, elles font le soleil*. L'apprenante Bibbi, par exemple, semble marquer le pluriel de la 3pl de façon plutôt conséquente (seulement une fois *elles parlent*) avec les morphèmes –ont/-ons : *elles mangont, elles parlons et riront, ils habitons, elles allons, les fills font les vert valise et y mettons a voiture de Jeanne*.

Il existe deux explications possibles à ce phénomène qui est caractéristique du stade 2. Vu que les formes *sont/ont/vont/font* dominent à ce stade, il est possible d'imaginer que les apprenants suremployent un morphème en [õ] comme marque de la 3pl, tout en hésitant sur sa transcription. L'autre possibilité, en quelque sorte moins systématique, est que ce stade est caractérisé par une confusion totale de la morphologie du pluriel et que les apprenants essaient de distinguer les personnes au pluriel des personnes au singulier en employant divers morphèmes possibles. Il est à remarquer que cette dernière explication devrait également impliquer quelques *nous parlent* parmi les apprenants à ce stade, ce qui n'est pas le cas dans notre étude transversale.

Au stade 3, les formes *sont/ont/vont/font* pour la 3pl sont encore plus fréquentes et le taux de réussite de l'accord en nombre est de 84 %. Le nombre élevé d'occurrences s'explique en partie par le fait que les textes sont de plus en plus longs à ce stade. Chez tous les apprenants qui s'expriment au présent (26/29), l'on retrouve plusieurs occurrences de ces formes. Les apprenants donnent parfois l'impression de s'appuyer sur ces verbes qui ont l'avantage d'exprimer un sens assez général. Un extrait de Carla illustre ce point :

...Elles boient café. Elles sont heureux. Marta et Marie ont une voiture verte. Elles ont beaucoup de valises, et la voiture est trop petite. Elles vont en Italie dans leurs petite voiture. Le temps est beau. Elles sont dans la route...

Parmi les apprenants au stade 3, il ne reste que quelques rares exemples de formes non-normatives en –ont/-ons dans un contexte –nt. Les deux catégories de verbes, les *sont/ont/vont/font* d'un côté et les réguliers en –nt de l'autre, semblent plutôt bien séparées.

Au stade 4, ce groupe de verbes est un peu moins fréquent qu'au stade 3 (5,5 par apprenant en moyenne). Ce phénomène est dû au fait que de nombreux apprenants écrivent leur texte (ou une partie du texte) au passé. De même, les apprenants ne dépendent pas autant de cette catégorie de verbes puisque leur vocabulaire est plus riche et qu'ils s'expriment en employant des verbes plus nuancés du premier groupe. Pour conclure, l'accord en nombre est extrêmement bien maîtrisé pour ce groupe de verbes : 98 %.

Quant au groupe contrôle, les scripteurs français emploient moins souvent les verbes *sont/ont/vont/font* que les apprenants aux stades 3 et 4 (3,1/scripteur). Ils écrivent parfois au passé simple ou au passé composé et, bien évidemment, leur vocabulaire est si riche qu'ils s'expriment

souvent avec des verbes plus nuancés : *aller = se précipiter, débarquer, embarquer, accompagner, rentrer, rouler, se balader, etc.* (cf. tableau 1). En ce qui concerne les apprenants suédois, il semble que ce groupe de verbes soit très important dans la production de texte : ces formes de la 3pl apparaissent déjà au stade 1 et elles sont très fréquentes, voire dominantes, à partir du stade 2.

4.5 Pluriel des verbes avec changement de radical à la 3pl

Un autre groupe de verbes marque le nombre non seulement par l'ajout du suffixe *-nt* à l'écrit mais aussi par l'ajout d'une consonne audible précédant ce suffixe (qui peut correspondre à deux consonnes à l'écrit), par exemple *il boit – ils boivent, il veut – ils veulent, il finit – ils finissent, etc.* Les verbes en question proviennent de la deuxième conjugaison (*-ir*), de la troisième conjugaison (*-re*) ou du groupe des verbes irréguliers. Ces verbes, dans leur forme correcte, ne sont pas très fréquents chez les apprenants de français L2. Par contre, ils apparaissent parfois sous d'autres formes : non-conjugués, 3sg ou avec le suffixe *-nt* mais sans changement de radical (voir la section suivante). Les études de Bartning et Schlyter (2004), Bartning (à paraître) et Schlyter et Bartning (ce volume) sur le français L2 oral, montrent que ces formes de la 3pl sont parmi les dernières à être acquises et qu'elles font leur apparition comme cas isolés au stade 3 et 4 alors qu'elles ne sont productives qu'au stade 5 (avancé moyen). Dans notre corpus écrit, nous observons que ces verbes apparaissent justement comme cas isolés au stade 1 et 2 alors qu'ils sont plus nombreux et fréquents à partir du stade 3. En ce qui concerne ces verbes à la 3pl, le tableau 4 ci-dessous montre que le groupe contrôle diffère des apprenants suédois quant à la diversité de leur emploi (mots types) et quant au taux de réussite de l'accord en nombre (% occ.cor.).

Tableau 4. *Les verbes avec changement de radical à la 3pl : nombre marqué versus nombre non-marqué : occurrences, types et pourcentages*

Stade	Cont.	Occ. -nt + rad. cor.	% occ. cor.	Types cor.	O.c./ appr.	Inf .	% inf.	3sg.	% 3sg	-nt + radical incor.	% -nt + rad. incor.	% nomb. tot.
1	7	1	14 %	1	0,1	3	43 %	3	43 %	-	-	14 %
2	33	3	9 %	3	0,1	4	12 %	8	24 %	18	55 %	64 %
3	86	39	45 %	7	1,4	5	6 %	13	15 %	29	34 %	79 %
4	47	41	87 %	11	2,4	-	-	-	-	6	13 %	100 %
Cont.	66	63	95 %	17	2,9	-	-	3	5 %	-	-	95 %

Légende : *Blanc* = nombre marqué, *Gris* = nombre non-marqué, **Cont.** = nombre de contextes à la 3pl, **Occ. -nt + rad. cor.** = nombre d'occurrences avec radical et suffixe corrects, **% occ. cor.** = pourcentage d'occurrences avec radical et suffixe corrects, **Types cor.** = nombre de types de verbes avec radical et suffixe corrects, **O.c./appr.** = occurrences correctes par apprenant en moyenne, **Inf.** = infinitif à la place du verbe conjugué à la 3pl, **3sg.** = 3^{ème} personne du singulier remplace la 3^{ème} personne du pluriel, **-nt + radical incor.** = verbe accordé en *-nt* mais avec un radical incorrect, **% nomb. tot.** = pourcentage total où le nombre est marqué avec *-nt*.

La partie gauche du tableau 4 montre que les apprenants de français L2 ont des difficultés avec ce groupe de verbes. Ces difficultés ne proviennent pourtant pas de l'accord en nombre, ce qui est indiqué dans la colonne à l'extrémité droite du tableau, mais du changement de radical à la 3pl. Les problèmes du radical sont prononcés, surtout aux stades intermédiaires (55 % et 34 % aux stades 2 et 3), ce qui est confirmé par des études sur l'oral (cf. Schlyter et Bartning, ce volume).

L'analyse de la morphologie de l'écrit ajoute ainsi une dimension à l'étude de l'acquisition de ce groupe de verbes. À l'oral, on ne perçoit que le problème du radical alors qu'à l'écrit on observe également que l'accord en nombre y est plutôt bien maîtrisé chez les apprenants de français L2.

Au stade 1, ce groupe de verbes est très rare et, le plus souvent, la bonne forme est remplacée par un infinitif ou par une forme à la 3sg. Au stade 2, les verbes avec changement de radical sont employés plus souvent, quoique seulement 9 % de façon normative (à comparer avec les 54 % corrects des verbes réguliers en –er et les 66 % corrects des verbes *sont/ont/vont/font* au stade 2). A ce niveau, il est intéressant d'observer qu'une fois sur deux (55 %), les apprenants marquent le pluriel de façon non-normative. L'accord en nombre se manifeste par l'ajout du suffixe –nt à une racine selon la règle des verbes réguliers mais sans changement de radical (voir la section suivante). Bien que ce groupe de verbes soit de mieux en mieux maîtrisé au stade 3 (45 % corrects) et au stade 4 (87 % corrects), la tendance de marquer le nombre sans changement de radical à la 3pl persiste jusqu'à un stade avancé d'acquisition (stade 3 : 34 % et stade 4 : 13 %). Il est à observer que ce marquage du nombre semble se faire selon le modèle des verbes réguliers.

Le manque d'input oral joue certainement un rôle important ici puisqu'il est impossible de maîtriser ce groupe de verbes à la 3pl à l'aide d'une règle. Chez les francophones, il est probable que ces formes irrégulières à la 3pl sont justement des unités gardées individuellement en mémoire. Or, comme elles sont beaucoup moins fréquentes que les verbes *sont/ont/vont/font*, elles sont moins bien maîtrisées en production écrite par les apprenants suédois. Largy et Fayol (2001) font remarquer que pour ce groupe de verbes, les clés orales aident les scripteurs francophones à maîtriser l'accord en nombre silencieux. Ils font moins d'erreurs d'omission du morphème –nt pour les verbes avec changement de radical puisqu'à l'oral, il y a un indice qui les aide à se rappeler du nombre : *ils prend - ils prennent* versus *il parle – ils parlent*. Ce phénomène est visible dans notre corpus écrit puisque, dans le groupe contrôle, la grande majorité des erreurs d'omission du –nt se trouvent parmi les verbes du premier groupe (-er) qui manque de clé orale indiquant le nombre (cf. tableau 3). Pourtant, il est à remarquer que les apprenants suédois au stade 4 sont encore plus corrects que les scripteurs francophones dans le sens qu'ils ne remplacent pas une seule fois une forme à la 3pl par une forme à la 3sg. Leur problème se situe toujours au niveau du changement de radical (voir aussi Hedbor, ce volume).

4.6 Surgénéralisation des morphèmes du nombre

Il existe deux formes de surgénéralisations du nombre qui peuvent apparaître dans le syntagme verbal. D'abord, le –s du nombre nominal peut être transféré au SV, *les garçons italiens parles*, comme on l'a vu chez les enfants francophones (Totereau, Thevenin et Fayol 1997 ; Totereau et al. 1998). Dans ce cas, il est question d'un transfert interphrastique d'un morphème du nombre. Ensuite, il peut être question d'un emploi non-normatif du nombre verbal de la 3pl, –nt, et dans ce cas nous avons affaire à un suremploi de la morphologie régulière (intra-phrastique). Comme dans notre analyse des surgénéralisations dans le cadre du syntagme nominal (Ågren, 2005ms), nous prendrons en compte ce phénomène dans le syntagme verbal.

Le –s comme marqueur du nombre dans le SV

Il est intéressant de constater que l'emploi du –s comme marqueur du nombre verbal apparaît chez les apprenants de français L2, mais dans une moindre mesure, semble-t-il, que chez les enfants

francophones (Totereau, 1998)⁷. La frontière entre SN et SV paraît prononcée chez les apprenants et la différence des morphèmes du nombre, -s et -nt, des deux syntagmes ne pose pas réellement problème. Plus pertinent encore, comme l'indique le tableau 5 ci-dessous, les quelques surgénéralisations qui existent apparaissent surtout au stade 2 et 3, tout comme le suremploi du -s dans le SN (cf. Ågren, 2005ms). Les exemples des stades 2 et 3 ci-dessous illustrent le point de la surgénéralisation du -s⁸ :

Stade 2 : Bengt: On plus elles aines la bier!

Berit: Ils apelles Marco est Eric

Birgit: ...les garçons de nous alles avec nous.

Béa: 1. Karin et Lena checks in dans un très beau hotel...

2. Elles prends un grand chambre avec un très beau vu.

Brynolf: ...elles prends une clés pour la endroit.

Stade 3 : Carina: Elles portes des lunettes de soleil...

Conny: Ils mets leur bagage...

Carl: ... avec deux garçons qui s'appelles Claudio et Cleudo

Claes: Ils appelles Giovanni et Ricardo.

Coralia: 1. Elles retournes a l'hôtel et...

2. Le jour après les garçons qui s'appelles Alejandro et Diego apportent a Maria et Veronique pour faire une sightseeing a Milano!

Il mérite d'être souligné qu'il existe une forme dans le paradigme verbal se terminant en -s, à savoir la 2sg : *tu t'appelles*. Cette forme apparaît par exemple dans les dialogues des manuels de français L2 en Suède. Or, vu que la série d'images n'invite pas à l'emploi de contextes à la 2sg, nous interprétons les exemples ci-dessus comme des formes contaminées du -s du syntagme nominal plutôt que comme des formes à la 2sg.

L'on peut noter que dans deux cas (Carl et Coralia 2), les verbes en question sont localisés dans des propositions relatives syntaxiquement éloignées du sujet de la phrase par une frontière phrastique. Il est pertinent, dans le cas de Coralia, de noter qu'elle emploie le -s pour marquer le pluriel dans la subordonnée alors qu'elle conjugue le verbe dans la principale selon la norme avec -nt. Il n'est pas impossible que la charge cognitive plus lourde impliquée par la subordonnée entraîne cette erreur morphologique (Pienemann, 1998). Malheureusement, nous avons peu d'exemples de ce genre dans le corpus.

La surgénéralisation du modèle des verbes réguliers en -er : racine plus -nt

Pour la grande majorité des verbes (1^{er}, 2^{ème}, 3^{ème} groupes et de nombreux verbes irréguliers), le -nt marque la 3pl, quoique parfois accompagné d'un changement de radical. Il a déjà été constaté que les apprenants suédois de français adoptent vite le principe du -nt comme marque du pluriel verbal à l'écrit (Hedbor, 2003). Selon Hedbor, la production des apprenants à l'écrit de la 3pl semble suivre « the one-to-one principle » (Anderson, 1984) suivant lequel l'interlangue se

⁷ Il faut toutefois être conscient du fait que les méthodes utilisées dans les œuvres mentionnées dans cette section sont très différentes des nôtres, un fait qui rend les résultats difficiles à comparer au niveau quantitatif.

⁸ Les exemples ci-dessous ne correspondent pas à la totalité des formes de ce genre dans l'étude : stade 1 : 1 occurrence, stade 2 : 9 occurrences, stade 3 : 10 occurrences. Pour le stade 4 et le groupe contrôle, il n'y a pas de formes en -s pour la 3pl verbale.

construit selon un système où le sens sous-jacent est exprimé par une forme unique et invariable. De ce point de vue, l'acquisition du nombre verbal ressemble à la règle du –s qui, dans la majorité des cas et de façon bi-univoque, marque le nombre dans le SN (si **nom** et **pluriel**, alors –s, si **verbe** et **3pl**, alors –nt). Souvent les apprenants ajoutent le suffixe –nt à une racine erronée, construite soit sur la forme de la 3sg (*ils font*, *ils conduisent*), soit sur l'infinitif (*ils boient*, *elles conduisent*) ou encore à une forme qui parvient d'une partie de l'infinitif (*elles veulent*, *elles reçoivent*). Il semble clair que de nombreux apprenants sont conscients du fait qu'il faut ajouter le suffixe –nt au verbe si le sujet de la phrase est à la 3pl alors qu'ils hésitent au niveau du radical. La règle donne donc lieu à des formes où le nombre est effectivement marqué mais qui sont malgré cela incorrectes selon la norme. Notre interprétation de ces formes est qu'il est question d'une procédure combinatoire, et non pas d'une mémorisation d'unités individuelles, qui intervient au moment où les apprenants sont très sensibles à la règle (pour une discussion plus approfondie, voir Hedbor, 2003 et Hedbor, ce volume). Comme le montre le tableau 5, cet emploi non-normatif de la règle des verbes réguliers en –er sur des verbes irréguliers est bien plus fréquent que la surrégularisation du –s dans le syntagme verbal et il persiste également plus longtemps chez les apprenants suédois.

Tableau 5. La surgénéralisation du modèle des verbes réguliers en –er sur des verbes irréguliers: racine plus –nt (occurrences)

Stade	-s	Verbes avec changement de radical				Autres verbes irr.	Part./Inf.
		3sg/1sg	infin.	partie	autres		
		p.ex veutent	p. ex. vouloirent t	p.ex. voulent	p.ex boissent	p.ex allent/faient	p.ex. Elles sommes travaillent Elles veulent mettent
1	1	-	-	-	-	-	2
2	9	4	2	9	4	3	1
3	10	8	3	14	4	8	1
4	-	4	1	3	-	4	-
Contr.	-	-	-	-	-	-	2

Légende: En gris = nombre de surgénéralisations du –s comme marqueur du pluriel verbal, **1sg/3sg** = le suffixe est ajouté à la forme de la 3sg (*veut*) ou parfois à la 1sg (*veux*), **infin** = le suffixe est ajouté à un infinitif, **partie infin.** = le suffixe est ajouté à une forme qui est une partie de l'infinitif (*vouloir*, *recevoir*), **autres** = le suffixe est ajouté à une forme influencée par un autre verbe, p.ex *finir* > *ils finissent* qui est transféré à *boire* > *boissent*. **Autres verbes irréguliers** = surtout *aller* mais aussi *faire*, **part./inf.** = le suffixe est ajouté au participe passé ou à l'infinitif précédé d'un auxiliaire.

Au stade 1, l'apprenante Anna emploie deux fois le –nt sur un participe, comme dans l'exemple suivant : *Sur une semaine elles sommes travaill^{nt} en Italie*. Toutes les autres formes avec le morphème –nt sont correctes à ce stade précoce.

C'est surtout aux stades 2 et 3 que l'emploi de la règle –nt laisse de nombreuses traces dans les textes des apprenants. Le plus souvent, ce sont les verbes avec changement de radical à la 3pl qui sont concernés. L'ajout du suffixe –nt à la forme 3sg ou au radical de l'infinitif sont les exemples les plus fréquents, ce qui est illustré dans les exemples suivants : *elles veulent*, *les hommes boient*, *les filles dirent*, *ils prennent*, *elles boient*, *deux garçons voient*, *ils veulent*, *elles*

conduirent, etc.⁹ Parfois, l'on observe le verbe irrégulier *aller* conjugué à la 3pl selon le modèle d'un verbe régulier en –er : *ils allent*. Il est possible de faire une parallèle au SN, où nous avons observé des formes comme *les bateaus*, que les apprenants n'ont jamais vues dans l'input mais qui correspondent à la logique de leur interlangue à ce moment donné de l'acquisition.

Au stade 4, le nombre de formes non-normatives en –nt diminue. Nous avons déjà constaté que les apprenants à ce niveau maîtrisent de mieux en mieux les verbes avec un changement de radical. Toutefois, il existe toujours des traces d'un suremploi de la règle des verbes réguliers en –er, comme dans les exemples *elles allent*, *elles boirent* et *ils conduient*. Il semble donc que le changement de radical soit un phénomène qui retarde la maîtrise de ce groupe de verbes du français L2 écrit.

Chez le groupe contrôle, ces formes n'existent pas au présent. Nous avons déjà constaté que les scripteurs français maîtrisent bien les changements de radical à la 3pl. Pourtant, quelques traces d'un suremploi de –nt demeurent dans une forme de participe passé et une autre à l'infinitif : *les femmes se sont apperçueut* et *...quand elles veulent mettrent*. Comme il a déjà été mentionné, des études sur le français écrit L1 ont montré que les scripteurs francophones sont sensibles aux clés orales pour l'accord en nombre à l'écrit (Largy et Fayol, 2001).

4.7 Le rôle de la sémantique

En nous appuyant sur des études de l'acquisition du français écrit L1 (Totereau, 1998 ; Fayol 2003 ; voir aussi Largy, ce volume), nous avons donné une place importante à la sémantique pour rendre compte du marquage précoce du nombre sur le nom/pronom par rapport à l'adjectif et au verbe en acquisition du français écrit. Le caractère sémantiquement motivé d'une marque sur le nom et le caractère formel de l'accord en nombre du verbe pourraient en partie expliquer la différence entre ces constituants en ce qui concerne le marquage du nombre. Pourtant, l'impact de la sémantique dans la production écrite des apprenants s'observe également dans le SV. Dans un certain nombre de cas, les apprenants accordent un verbe au pluriel bien que le sujet de la phrase soit au singulier. Dans la majorité de ces contextes, la forme formelle du sujet (singulier) ne correspond pas à son contenu sémantique (pluriel). Ces noms sont appelés des « formes ambiguës » d'après Roiné (2000). Le sujet est caractérisé par un contenu sémantique de plusieurs personnes ou d'items, comme par exemple *la famille*, *tout le monde*, *le bagage*, *le couple*, etc. Ce phénomène indique que la sémantique se montre décisive également dans l'accord du nombre dans le SV : du moment où il y a désaccord entre forme et sens, le sens l'emporte sur la forme. Ce phénomène a été étudié par Roiné (2000) pour le français L1 écrit et elle conclut que l'ambiguïté sémantique (l'équipe – le joueur, l'armée – le soldat) joue un rôle en acquisition. Les enfants français font significativement plus d'erreurs d'accord en nombre lorsqu'ils ont affaire à des sujets dits ambigus. Les exemples ci-dessous, tirés de notre corpus, illustrent ce phénomène chez les apprenants de français L2 :

- 2) Bonita : ...toute la groupe vont au "tony's sightseeing".
 Björn : Une famille suédoise discutent qui ils vont.

⁹ Ces exemples mettent d'ailleurs en évidence le fait que, aux yeux des apprenants suédophones, le suffixe de la 3pl est bien –ent. Cf. note 1 à la page 1.

3) Cecilia: Enfin tout le monde vont en Suède...

Carina: ...et toute le monde vivent une bon vie heureuse.

4) Daniel: La famille qui consistent de duex femmes, une fille et sa mère, garent la voiture près de la maison.

Dominique: Dans le soir la famille sont allées d'une bar.

Dans notre corpus, aux stades 2, 3 et 4, l'on perçoit une sensibilité chez les apprenants suédois pour la sémantique du nombre. Dans un nombre non-négligeable de cas, ils marquent le nombre au verbe lorsque le sujet est ambigu. Ce phénomène persiste à un niveau avancé d'acquisition. Toutefois, il n'est visible que chez un seul scripteur du groupe contrôle. Il est évident que dans une structure si complexe que l'exemple de Daniel, la proximité des éléments pluriels influence l'accord verbal qui devrait être dirigé par *la famille*. Or, pour terminer, ces exemples d'accords verbaux au pluriel sont intéressants dans la mesure où ils montrent la sensibilité des apprenants en français L2 pour le nombre sémantique.

5 Conclusion

Dans l'analyse des données, nous pouvons observer que l'acquisition et la mise en œuvre de la flexion du nombre verbal à la 3pl est un phénomène graduel dans l'acquisition du français L2 écrit. En début d'acquisition, les tendances de non-marquage dominant alors qu'à un stade avancé, les apprenants conjuguent les verbes en nombre dans près de tous les contextes. Dans une phase intermédiaire, aux stades 2 et 3, les apprenants marquent de plus en plus souvent le nombre, tout en hésitant sur la forme morphologique du suffixe en question et en suremployant la règle de –nt des verbes réguliers en –er. Pour ce qui est du marquage du nombre nominal, nous avons observé que l'accord en nombre du verbe est maîtrisé bien plus tardivement. Or, au stade 4, les apprenants suédois sont arrivés au même taux de réussite que le groupe contrôle français dans l'expression du nombre verbal au présent, bien qu'il y ait toujours des différences entre les deux groupes quant au marquage du nombre au passé et quant à la distribution des différents groupes de verbes. Nous sommes de l'avis que la différence entre l'accord nominal et verbal en nombre dépend largement du facteur sémantique très important dans l'accord en nombre. D'ailleurs, d'autres données du corpus, plus précisément l'accord en nombre des verbes avec un sujet ambigu, montrent la sensibilité des apprenants pour la sémantique du nombre au moment où il y a désaccord entre forme syntaxique et sens sémantique du sujet de la phrase.

Nous sommes de l'avis que l'emploi du nombre verbal des apprenants laisse entrevoir deux mécanismes différents en action dans l'acquisition de la morphologie du nombre (cf. Pinker, 1998). Les verbes fréquents sont des unités mémorisées dans leur ensemble (racine+suffixe) et ces verbes dominant en début d'acquisition. Ils sont notamment employés de façon plus correcte que les verbes en –er (voir Hedbor ainsi que Schlyter & Bartning, ce volume). À côté de ces unités régit un mécanisme combinatoire qui est très important aux stades intermédiaires. Les apprenants appliquent une règle qui dit que : **si** verbe et 3pl, **alors** –nt. Cette technique conduit à davantage de formes correctes parmi les verbes réguliers ainsi qu'à un nombre important de formes non-normatives parmi les verbes irréguliers moins fréquents.

Les verbes avec un changement de radical à la 3pl posent problème aux apprenants. Ils les

accordent en nombre selon le modèle des verbes réguliers en –er, en omettant le changement de radical et en se basant sur la forme de l’infinitif ou sur la forme de la 3sg. Le groupe contrôle n’a de difficultés ni avec le changement de radical ni avec le suffixe –nt pour ce groupe de verbes. Il est probable qu’ils profitent des clés orales du nombre qui leur facilitent l’accord silencieux. Dans l’input oral restreint des apprenants suédois, ces verbes sont probablement trop peu fréquents pour permettre un tel profit et un emploi correct de ces formes irrégulières. Par contre, les apprenants suédois au stade 4 maîtrisent mieux que les scripteurs francophones l’accord en nombre des verbes du premier groupe où il n’y a pas de clés orales. L’explication la plus plausible est qu’il s’agit là d’un effet de l’importance de l’input écrit pour les apprenants en L2 alors que l’input oral domine pour les scripteurs francophones.

«

Nos données suggèrent que, dans leur volonté de marquer le nombre verbal, les apprenants recourent à des généralisations de divers genres. D’abord, il est possible d’interpréter nos résultats au stade 2, où les verbes *sont/ont/vont/font* dominant par rapport aux autres groupes de verbes à la 3pl, comme une influence de ces verbes sur l’accord en nombre dans le sens où le [õ]-final est transféré à d’autres verbes pour marquer le nombre tout en laissant entrevoir une hésitation sur la mode de transcription (-ont, -ons). Deuxièmement et de façon plus durable, nous avons repéré un suremploi de –nt pour des verbes avec changement de radical (ils prennent) ou avec forme indépendante (ils font). Ces formes ne sont probablement pas suffisamment fréquentes pour assurer une mémorisation d’unités individuelles correcte. Finalement, le –s du nombre nominal est par contre assez rarement transféré au SV. La limite entre les deux syntagmes semble prononcée. Il en existe quelques exemples, surtout au stade 2 et 3, là où il y a également une tendance de suremploi du –s dans le SN. Ces phénomènes montrent que l’application de règles joue un rôle important, comme complément aux unités individuelles, dans l’interlangue écrite des apprenants, surtout à un stade intermédiaire où ils semblent très sensibles aux régularités morphologiques. Les règles du marquage du nombre dans le SV (-nt) aussi bien que dans le SN (-s) sont simples et relativement faciles à mettre en œuvre, ce qui d’après de nombreux chercheurs (cf. Bybee, 1985 ; Clahsen, 1999) plaide en faveur d’un processus combinatoire.

6 Références bibliographiques

- Anderson, J. R. (1984) *The architecture of cognition*. Cambridge MA: Harvard University Press
- Bardovi-Harlig, K. (2000) « Tense and Aspect in Second Language Acquisition : Form, Meaning and Use ». *Language Learning*, Vol 50 : 1.
- Bartning, I. (à paraître) « Late morphosyntactic and discourse features in advanced/very advanced L2 French a view towards the end state ». In Habermatzl, S, (éds.) *Processes and Outcomes. Explaining Achievement in Language Learning*. Berlin: Mouton de Gruyter.
- Bartning, I. et Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies*, 14(3) : 281-299.
- Bybee, J. L. (1985) *Morphology – A study of the relation between meaning and form*. Amsterdam/Philadelphia : John Benjamins Publishing Company.
- Clahsen, H. (1999) « Lexical entries and rules of language : A multi-disciplinary study of German inflection ». *Behavioral and brain sciences* 22 : 991-1060.
- Ellis, N. (2002) « Frequency effects in language processing, A review with implications for theories of implicit and explicit language acquisition ». *Studies in Second Language*

Acquisition 24 : 143-188.

- Ellis, R. (1994) *The study of second language acquisition*. Oxford: Oxford University Press.
- Fayol, M. (2003) « L'acquisition/apprentissage de la morphologie du nombre. Bilan et perspectives ». In *Perspectives : Rééducation Orthographique* – No 213 – mars 2003
- Fayol, M & Bourdin, B (1994) « Is written language production more difficult than oral language production. A working memory approach ». *International journal of psychology*, 29(5) : 591-620.
- Fayol, M. & Largy, P. (1992) « Une approche fonctionnelle de l'orthographe grammaticale », *Langue Française* 95 : 70-80.
- Hedbor, C. (2003), *Les règles d'abord ou la production directe du français écrit ? Étude longitudinale sur l'utilisation ou non des connaissances explicites par cinq élèves de différentes attitudes cognitives*. Thèse de licentiat. Département d'études romanes, l'Université de Lund, Suède
- Jaffré, J-P. & David, J. (1999) « Le nombre : Essai d'analyse génétique ». *Langue Française* 124 : 7-23.
- Klein, W. & Perdue, C. (1992) *Utterance Structure. Developing grammars again*. [Studies in bilingualism 5]. Amsterdam/Philadelphia : John Benjamins.
- Klein, W. & Perdue, C. (1997) « The Basic Variety (or : Couldn't natural languages be much simpler?) ». *Second Language Research* 13(4) : 301-347.
- Largy, P. (2002), *Apprentissage et mise en œuvre de la morphologie flexionnelle du nombre*. Thèse d'habilitation, Université de Rouen, U.F.R de Psychologie, Sociologie, Sciences de l'éducation.
- Largy, P. et Fayol, M. (2001) « Oral cues improve subject-verb agreement in written French ». *International Journal of Psychology*, 36(2) : 121-132
- Largy, P. & Cousin, M-P (2004) « Apprendre implicitement le -s du pluriel ». In Piolat, A. (éds.) *Écriture – Approches en sciences cognitives*. Langues et Écritures. Publications de l'Université de Provence.
- Marcus G. F. et al. (1992) *Overregularisation in Language Acquisition*. Chicago : Monographs of the Society for Research in Child Development, serial number 228, Vol. 57, No. 4. Chicago.
- Pienemann (1998) *Language processing and second language development – Processability Theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Pinker, S. (1998) « Words and Rules », *Lingua* 106 : 219-242.
- Roine, C. (2000) *Étude de l'influence du facteur sémantique sur la performance orthographique d'enfants du CE2 à la 3^{ème} : Le cas de l'accord sujet-verbe*. Mémoire de maîtrise, Psychologie du développement, l'Université de Rouen, France
- Totereau, C. (1998) *Un exemple d'apprentissage du système orthographique : la morphologie écrite du nombre en français*. Thèse de doctorat, U.F.R des Sciences Humaines, Université de Bourgogne, France
- Totereau, C. Thevenin, M-G. & Fayol, M. (1997) « Acquisition de la morphologie du nombre à l'écrit en français ». In Rieben, L., Fayol, M. & Perfetti, C. A. (éds.) *Des orthographes et leur acquisition*, pp. 147-167. Paris : Delacroix & Niestlé.
- Totereau, C. et al. (1998), "Overgeneralisations of number inflexions in the learning of written French : The case of noun and verb". *British Journal of Developmental Psychology*, 16 : 447-464.
- Ågren, M. (2005ms) « Le marquage morphologique du nombre dans la phrase nominale. Une

étude sur l'acquisition du français L2 écrit ». Institut d'études romanes, Université de Lund, Suède. Disponible sur : <http://www.rom.lu.se/durs/usif.htm>

Le traitement implicite et explicite de l'inflexion de la 3^e personne du pluriel du verbe chez les apprenants guidés du français LE

Cecilia Hedbor

Université de Lund
Institut d'études romanes
cecilia.hedbor@rom.lu.se

1 Introduction

Pour pouvoir parler, mais aussi écrire, il faut des connaissances implicites. Ces connaissances permettent un traitement automatique de la langue et facilitent une concentration sur ce que l'on dit, plutôt que sur la façon dont on doit le dire. Toutefois, l'enseignement d'une LE guidée privilégie largement les connaissances explicites. Pour en savoir plus des connaissances explicites et implicites utilisées en production écrite par les apprenants guidés en français LE, nous avons conduit une étude longitudinale sur l'acquisition des connaissances implicites et explicites chez cinq apprenants suédois.

La recherche neurolinguistique en acquisition des langues considère que les connaissances explicites et implicites existent en parallèle et dans des lieux différents du cerveau, voir Paradis (1985, 1994, 1997, 2000). Ces connaissances se développent aussi en parallèle, sans qu'il s'agisse d'un développement parallèle synchronisé, chez les apprenants. Nous avons entre autres pu constater des différences individuelles dans ce développement (Hedbor, 2003, 2005). Certains apprenants utilisent plutôt des connaissances implicites, quand ils produisent du français écrit. Ces connaissances semblent donc se développer avant les connaissances explicites. De cette manière, ils produisent avec beaucoup de fluidité et rapidité. D'autres ont recours à des connaissances explicites en hésitant et réfléchissant, et, de cette manière, les connaissances implicites ont du retard sur les connaissances explicites. La production de ces derniers est ainsi plus lente et saccadée.

Ces différences individuelles mènent à un développement différencié en ce qui concerne l'acquisition de la langue en question. Ainsi, nous avons pu constater des différences liées à une production plutôt implicite ou explicite, entre autres en ce qui concerne l'acquisition de la finitude, de l'accord sujet-verbe et du passé, voir Hedbor (2003, 2005). Ces différences dans l'acquisition de certains phénomènes grammaticaux semblent également liées à la simplicité ou la complexité de la règle déclarative décrivant le traitement du phénomène en question.

Dans cet article, nous allons présenter le traitement d'un de ces phénomènes, géré par une règle assez simple, i.e. l'accord sujet-verbe à la 3^e personne du pluriel. Une règle déclarative qui, d'après le raisonnement pour l'anglais, de Pica (1985) et Krashen (1981, 1985), serait plus facile à

apprendre que d'autres. Il s'agit donc d'un phénomène grammatical qui se prête bien à un traitement et à une mise en œuvre explicites.

Le marquage, *-ent*, du pluriel sur le verbe est intéressant car, pour un grand nombre de verbes, il est inaudible à l'oral, mais extrêmement régulier à l'écrit. Par conséquent, d'après nos données et celles d'Ågren, ce volume, le pluriel semble être marqué assez conséquemment à un stade plus précoce à l'écrit qu'à l'oral, voir aussi Bartning et Schlyter (2002, 2004), Schlyter (2003a et b), ainsi que les articles de Bartning et de Schlyter ce volume. Pour approfondir la question de l'acquisition de l'accord sujet – verbe de la 3^e personne du pluriel à l'écrit, nous nous proposons de présenter son développement explicite ou implicite chez les cinq apprenants de notre étude de cas longitudinale.

2 Corpus et méthode

Le corpus de l'étude longitudinale est composé d'enregistrements de la production écrite libre de cinq apprenants de français langue étrangère du lycée en Suède. Il s'agit de quatre filles et un garçon. Le groupe d'apprenants est assez homogène, et les critères des différences individuelles habituellement retenus dans la littérature sur l'apprentissage des langues étrangères ne jouent pas un grand rôle sur notre corpus (i.e., âge, motivation, attitude, niveau scolaire, intelligence¹, voir par exemple Larsen-Freeman & Long, 1991 ; Kristiansen, 1992 ; Lightbown & Spada, 1993, Ellis, 1997 ? ; Lambert, 1994 ; Cyr, 1996). Ces cinq participants ont été suivis et enregistrés pendant deux ans et demi, du printemps en Seconde jusqu'au printemps en Terminale, i.e., cinq semestres au total. Pendant ce temps, il y a eu six périodes d'enregistrements, chaque période comportant la production de deux textes. A l'exception des textes narratifs (à partir d'images) pour lesquels le temps de production n'était pas contraint, les participants ont eu un temps de production limité à 20 minutes.

Pour enregistrer leur production écrite, qui s'est fait sur ordinateur, nous avons utilisé le logiciel ScriptLog, développé par Strömqvist et Malmsten (1998). Ce logiciel, qui enregistre et mesure tout ce qui se passe sur le clavier lors de la production écrite, chaque frappe, pause, correction, changement etc., permet de suivre la production écrite en temps réel. Toujours dans le but de mieux connaître la nature des connaissances explicites mobilisées en production écrite, l'enregistrement de ScriptLog a été associé à la méthode des protocoles verbaux concomitants. Les participants, ayant eu la consigne de penser à haute voix, ont aussi été filmés lors des sessions de production.

Les épreuves de production écrite consistaient soit à évoquer des souvenirs personnels, soit à résumer une histoire vue en vidéo ou entendue, soit à raconter une histoire à partir d'images. Le pluriel a été élicité dans trois textes lors des trois dernières périodes d'enregistrement, période 4-6. L'élicitation du pluriel s'est fait à l'aide des narrations à partir d'images.

¹ Le critère *sexe* joue naturellement un rôle, comme il y a un garçon et quatre filles dans le corpus.

3 Le contrôle sur l'output et la mise en œuvre du pluriel

Opposant la production écrite en LM et en LE, Barbier (2004) souligne que la production en LE est plus coûteuse en temps et que les processus de formulation sont prioritaires. Les traitements lexicaux et morpho-syntaxiques sont particulièrement importants dans la formulation. Dans son modèle sur la production écrite en LE, basé sur le modèle de Levelt (1989), Zimmerman (2000) propose une beaucoup plus grande activité de formulation avec plusieurs tentatives et rejets avant la transcription en LE. Ce travail impose un fort coût cognitif (notamment parce qu'il mobilise le recours à des connaissances explicites). Ce coût cognitif se manifeste au plan temporel par des pauses dans l'activité de production écrite². Cette activité contrôlée et coûteuse cognitivement peut se dérouler juste avant la production écrite, par le biais d'un contrôle interne (Levelt, 1989) ou d'un monitoring (Krashen, 1981, 1985, 1994).

Dans notre étude, il s'est avéré que les apprenants planifient plus ou moins ce qu'ils vont écrire. Une apprenante, CHR, produit du texte directement en français, sans réfléchir et planifier à l'avance. Sa production de texte semble se faire simultanément à son édition sur le clavier. Les formulations et reformulations, signalées par Zimmerman, n'ont donc pas souvent lieu. Comme CHR ne planifie pas ce qu'elle va écrire, elle n'anticipe pas non plus les problèmes à venir. Elle ne prévoit pas un accord au pluriel, par exemple. Le contrôle de l'output se fait surtout dans le contrôle interne, sans qu'elle n'en dise rien à haute voix.

Les apprenantes SOP et EMM, ressemblent à CHR en ce qu'elles planifient peu ce qu'elles vont écrire. Mais elles forment de temps en temps en suédois avant d'entamer la transcription. On observe ainsi la trace d'un contrôle interne dans leurs protocoles verbaux.

Planifiant presque tout à l'avance en passant par le suédois, les apprenants MAR et OSC présentent un profil opposé à celui de CHR. Ils anticipent aussi les problèmes morpho-syntaxiques, comme l'accord du pluriel. Pourtant, dans la production écrite de ces deux apprenants il y a aussi des occurrences de pauses dues au contrôle interne, avec ou sans traces dans les protocoles verbaux.

Pour cette raison il nous a fallu définir la durée d'une pause indiquant une activité de contrôle, i.e. un recours aux connaissances explicites.

3.1 Comment appréhender le contrôle interne (monitoring) ?

Comme nous utilisons la méthode du protocole verbal, les pauses dues à la phase prolongée de formulation ne nous posent pas de problèmes, car nous pouvons suivre d'assez près ce qui se passe dans le protocole verbal. Par contre, les pauses souvent muettes dues au contrôle interne, le monitoring, sont plus difficiles à discerner. Les pauses muettes sont importantes dans notre étude car nous postulons qu'elles indiquent le recours aux connaissances explicites.

Dans cet article, nous nous intéressons aux pauses ayant à faire avec la mise en œuvre de l'accord sujet – verbe de la 3^e personne du pluriel. Dans une étude sur la production de la morphologie

² Il peut aussi se traduire en erreurs dans la production, voir la recherche de Largy et son article dans cette publication.

flexionnelle en LM, Largy et Fayol (2001) et Largy (2002) ont montré l'existence d'un site de pause juste avant la flexion du verbe. Les participants de l'étude ont fait des pauses significatives avant de transcrire graphiquement le *-nt* du pluriel ou non. En nous basant sur ce qu'ont trouvé Largy et Fayol concernant l'emplacement de la pause devant l'inclinaison du verbe, ainsi que Weingarten, Nottbusch et Will (200 ?), qui a observé des pauses significatives entre les morphèmes, nous avons déterminé une durée individuelle de pause intra-mot muette pour chaque apprenant. Comme nous considérons que la production de verbe en suédois LM se fait implicitement³, cette durée est basée sur le débit de production et la longueur des pauses intra-verbe quand les sujets écrivent en LM. La durée des pauses considérées indiquer une activité explicite varie entre 0,6 s pour les plus rapides sur le clavier et 1,0 s pour les plus lents.

Une fois ces durées individuelles de pauses établies, nous avons procédé à une comparaison statistique entre les apprenants. Les paramètres comparés sont la production de verbes en suédois, la production des verbes réguliers, en *-er* et irréguliers en français. Deux groupes d'apprenants se sont distingués ; un groupe ayant une production plus rapide et donc plus implicite en français, comprenant les apprenants CHR, EMM et SOP, et un groupe ayant une production plus lente et donc plus explicite en français, comprenant les apprenants MAR et OSC. Une analyse de variance entre les deux groupes d'apprenants et la production en suédois et en français montre une interaction significative aussi bien pour les verbes réguliers, $p=0,028$, que pour les verbes irréguliers⁴, $p=0,010$. Une grande différence entre les deux groupes d'apprenants est donc affirmée.

3.2 Les particularités de l'accord à la 3^e personne du pluriel en français

La morphologie verbale en français est silencieuse à l'oral, sauf pour les 1^{re} et 2^e personnes du pluriel. Au contraire de la morphologie, aussi silencieuse du singulier⁵, celle de la 3^e personne du pluriel est très systématique et régulière à l'écrit. À quatre exceptions près, i.e. *sont*, *ont*, *vont*, *font*, on ajoute tout simplement *-ent* (ou *-nt*) au radical pour marquer le pluriel de la 3^e personne. Il s'agit donc d'une règle facile à gérer cognitivement.

Ainsi, Largy (2001) et Largy, Chanquoy et Dédéyan (2004), dans des études relatives à la révision d'accords pluriels, montrent que les enfants gèrent mieux le pluriel en révision qu'en production. Les enfants marquent donc mieux le pluriel quand ils utilisent plus spécifiquement des connaissances explicites. En acquisition de la LM écrite, le traitement du pluriel est apparemment lié aux connaissances explicites avant d'être automatisé. Qu'en est-il en LE ?

4 La mise en œuvre de l'accord au singulier vs l'accord au pluriel chez les apprenants du français LE.

³ Le suédois est une langue où la personne grammaticale n'est pas marquée avec le verbe. Il n'y a pas de différence morphologique sur le verbe entre les différentes personnes dans la conjugaison des verbes.

⁴ Les verbes fréquents formant le pluriel en *-ont* sont exclus de ce groupe.

⁵ Au singulier, les inflexions /suffixes à ajouter au radical varient d'après le groupe du verbe, i.e. *-e*, *-es*, *-e* dans le premier groupe, *-s*, *-s*, *-t* ou *-s*, *-s*, $\emptyset$ ou bien *-x*, *-x*, *-t* dans les autres groupes de verbes.

Avant d'étudier les connaissances implicites ou explicites impliquées dans la mise en œuvre de l'accord sujet-verbe au pluriel, il convient d'étudier celles mises en œuvre dans le cas du singulier chez les apprenants.

4.1 La mise en œuvre de l'accord au singulier

En ce qui concerne la mise en œuvre implicite ou explicite de l'accord sujet-verbe au singulier, nous avons pu constater des différences inter-individuelles, voir Hedbor (2003, 2005). Les apprenants ayant une production plus rapide et implicite en français contrôlent moins l'output (CHR, EMM et SOP). Ce groupe d'apprenants produit plus de formes erronées au singulier que le groupe ayant une production plus explicite (MAR et OSC), voir le tableau 1 :

Tableau 1. *Formes correctes et erronées des verbes réguliers, -er, et des verbes irréguliers au singulier*

	Formes correctes	Formes erronées	Σ	% erronées
Verbes réguliers				
CHR, EMM, SOP	199	19	218	8,7
MAR, OSC	52	1	53	1,9
Verbes irréguliers				
CHR, EMM, SOP	85	60	145	41,4
MAR, OSC	39	10	49	20,4

Légende : Σ = Total d'occurrences correctes et erronées, % erronées = le pour cent de formes erronées du total. Les verbes fréquents *être*, *avoir*, *aller* et *faire* sont exclus de ce tableau. Seules les formes finales, après d'éventuelles corrections, sont comptabilisées.

Le tableau 1 montre qu'il y a une différence entre les deux groupes d'apprenants aussi bien quand il s'agit de verbes réguliers qu'irréguliers au singulier. Cette différence est pourtant plus grande dans le cas des verbes irréguliers, où le contrôle sur l'output de MAR et OSC⁶, leur fait faire beaucoup moins d'erreurs comparé à CHR, EMM et SOP. Il est encore plus intéressant de regarder les chiffres de la production erronée lorsqu'on ne prend en considération que la production faite d'une manière explicite :

⁶ Ces apprenants, ainsi qu'EMM à une occasion, se répètent aussi parfois le paradigme au singulier du verbe en question pour être sûrs de trouver la bonne forme.

Tableau 2. *Formes correctes et erronées des verbes réguliers –er, et verbes irréguliers au singulier produites explicitement*

	Formes correctes	Formes erronées	Σ	% erronées
Verbes réguliers				
CHR, EMM, SOP	40	10	50	20,0
MAR, OSC	33	1	34	2,9
Verbes irréguliers				
CHR, EMM, SOP	32	27	59	45,7
MAR, OSC	28	8	36	22,2

Légende : Voir le tableau 1.

En comparant avec le tableau 1, il apparaît que les connaissances explicites ne favorisent pas tellement la production correcte. Il semble même que pour le groupe, CHR, EMM et SOP, la production explicite est plus erronée que la production totale.

Il existe aussi une différence qualitative entre les formes erronées des verbes réguliers et irréguliers. Dans le cas des verbes réguliers, les formes erronées au singulier sont surtout des formes non-finies, i.e. des infinitifs ou des participes passé. Pour les verbes irréguliers, les formes erronées consistent surtout en un suffixe muet erroné, ou en une forme erronée ressemblant à un traitement régulier d'un verbe irrégulier.

Formes non-finies :

- (1) [CHR 6] on (...) se **reposer*
- (2) [EMM 3] le garçon **acheté*
- (3) [SOP 3] il **poser*

Suffixes muets erronés :

- (4) [CHR 6] le serviteur **viens* et **vois*
- (5) [EMM 4] je **peut*
- (6) [SOP 3] elle **vois*
- (7) [MAR 4] son mari **veux* danser
- (8) [OSC 3] le loup **viens*

Traitement régulier d'un verbe irrégulier :

- (9) [CHR 5] je la **mette* > je la **mettes*⁷
- (10) [EMM 2] elle **dorme*
- (11) [SOP 4] le disco **dorme* > **dormé*⁸
- (12) [MAR 3] Le petit chaperon rouge **prende*
- (13) [OSC 3] elle **entende* > **entendre* > *entend*⁹

⁷ La flèche veut dire que CHR change « je mette » à « je mettes »

⁸ À notre avis, le traitement de SOP du verbe *dormir* montre qu'elle le conçoit comme un verbe régulier –er.

Chez les apprenants du groupe plus implicite, CHR, EMM et SOP, les suffixes muets des verbes irréguliers au singulier sont aussi assez aléatoirement marqués dans une même période d'enregistrement :

- (14) [CHR 3] elle **vois* les fleurs, le garçon *voit* les ballons
- (15) [CHR 3] il **dors*, il *dort*
- (16) [CHR 3] il *comprend*, Le petit chaperon rouge **prends*
- (17) [CHR 6] je **peut* me changer, quand je *peux*
- (18) [EMM 3] tu **veut*, elle *veut*
- (19) [EMM 4] tu **peut*, tu *veux*

L'accord sujet-verbe au singulier semble donc poser certains problèmes aux apprenants. En ce qui concerne les verbes réguliers, *-er*, la forme non-finie d'infinitif et du participe passé est parfois utilisée dans la production des apprenants aux stades assez avancés, voir l'exemple (1)¹⁰. En ce qui concerne les verbes irréguliers, les apprenants produisent surtout des formes finies, mais les suffixes muets sont erronés et parfois mis aléatoirement dans le même texte. Les apprenants ayant une production plus implicite en particulier utilisent peu le contrôle conscient pour trouver le suffixe correct. Les apprenants avec une production plus explicite ont recours à la stratégie de se répéter le paradigme du verbe au présent et trouvent ainsi le suffixe correct.

La question est maintenant de savoir si la mise en œuvre du pluriel suit le même développement. Dans ce cas, le marquage du pluriel serait aussi un peu aléatoire chez une apprenante comme CHR. Ce qui va à l'encontre d'un développement aléatoire du pluriel est sa régularité totale. La seule règle à mémoriser et à mettre en œuvre est celle d'ajouter *-ent* à la racine. Le groupe auquel le verbe appartient ne joue aucun rôle dans ce traitement, comme c'est également le cas pour les suffixes muets du singulier.

3.2 La mise en œuvre de l'accord de la 3^e personne au pluriel

Dans les tableaux 3 et 4, nous présentons le développement implicite et explicite de l'accord sujet-verbe au singulier et au pluriel¹¹. Dans ces tableaux est présentée la production de toutes les occurrences de verbes, sauf les verbes fréquents qui marquent le pluriel avec *-ont*. Les apprenants sont placés en deux groupes d'après leur mode de production de texte : CHR, EMM et SOP produisent d'une manière plus implicite, avec peu de pauses et de commentaires. MAR et OSC produisent d'une manière explicite, avec beaucoup de planification, pauses et commentaires.

⁹ Dans le protocole verbal, OSC dit d'abord qu'*entendre* est un verbe régulier *-er*, ensuite il a l'idée de le traiter comme le verbe *attendre*.

¹⁰ Pour une présentation de l'utilisation des formes avec un /e/ final prononcé dans un contexte fini chez les apprenants suédois, voir l'article de Thomas dans cette publication.

¹¹ Le pluriel est explicitement élicité à partir du 4^e enregistrement.

Tableau 3. *Production implicite et explicite de tous les verbes formant le pluriel –ent à la 3e personne du singulier*

Apprenant	Implicite	Explicite	Σ	% expl / tot
CHR	105	45	150	30,0
EMM	74	37	111	33,3
SOP	35	17	52	32,7
Total	214	99	313	31,6
MAR	21	35	56	62,5
OSC	17	62	79	78,5
Total	38	97	135	71,9

Tableau 4. *Production implicite et explicite de tous les verbes formant le pluriel -ent à la 3e personne du pluriel*

Apprenant	Implicite	Explicite	Σ	% expl / tot
CHR	9	19	28	67,9
EMM	14	23	37	62,1
SOP	6	9	15	60,0
Total	29	51	80	63,8
MAR	5	21	26	80,8
OSC	6	17	23	73,9
Total	11	38	49	77,6

Légende des tableaux 3 et 4 : Σ = somme des verbes produits implicitement et explicitement, % expl / tot = Le pourcentage des verbes produits explicitement sur le total des verbes produits par l'apprenant. Dans ces tableaux il est question des contextes des 3^e personnes du singulier et du pluriel, donc les formes erronées sont incluses.

Malgré le nombre d'occurrences assez bas en ce qui concerne la production du pluriel chez certains apprenants, quelques observations sont à faire. Des tableaux 3 et 4 ressort que la production dans le contexte pluriel est plus explicite chez tous les apprenants, sauf OSC, qui a déjà au singulier une production très explicite. Il semble que le marquage du pluriel se fait en appliquant une règle consciemment, la règle d'ajouter le suffixe, *-ent*. Nous pouvons formuler l'hypothèse :

- En français langue étrangère, tout comme en français LM, l'accord sujet-verbe de la 3^e personne du pluriel se fait d'abord explicitement, i.e. il s'agit d'une application consciente d'une règle apprise.

Une telle hypothèse renvoie aussi aux théories de Krashen (1981, 1985, 1994) et de Pica (1985), qui proposent que seules les règles faciles à gérer consciemment peuvent être apprises d'une manière explicite et ensuite être appliquées explicitement dans la production libre de la langue. La règle en question est d'ajouter *-ent*, ou *-nt*, à la racine pour marquer le pluriel. Cette règle extrêmement simple est valable pour les verbes réguliers et irréguliers, mais elle n'est pas valable pour les verbes fréquents, qui eux ont une forme propre complètement différente de la racine pour la 3^e personne du pluriel, i.e. *sont*, *ont*, *vont*, *font*. Dans ce cas, l'algorithme racine + *-ent*, n'est plus valable et la mise en œuvre du pluriel doit se faire d'une autre manière. Pour cette raison, nous avons pris en considération les pauses faites avant le verbe pour définir une production

explicite¹², voir les tableaux 5 et 6 :

Tableau 5. Production implicite et explicite des verbes fréquents à la 3^e personne du singulier

Apprenant	Implicite	Explicite	Σ	% expl / tot
CHR	82	17	99	17,2
EMM	33	9	42	21,4
SOP	31	9	40	22,5
Total	146	35	181	19,3
MAR	34	13	47	27,7
OSC	22	22	44	50,0
Total	56	35	91	38,5

Tableau 6. Production implicite et explicite des verbes fréquents à la 3^e personne du pluriel

Apprenant	Implicite	Explicite	Σ	% expl / tot
CHR	11	2	13	15,4
EMM	5	5	10	50,0
SOP	3	1	4	25,0
Total	19	8	27	29,6
MAR	6	8	14	42,9
OSC	0	4	4	100,0
Total	6	12	18	66,7

Légende : Voir le tableau 4.

Cet aperçu de la production au singulier et au pluriel des verbes fréquents, qui forment le pluriel en *-ont*, semble en partie contredire notre première hypothèse. L'apprenante CHR ne traite pas la 3^e personne du pluriel plus explicitement que la 3^e personne du singulier. Par conséquent nous formulons une deuxième hypothèse :

- Les apprenants avec une production et un développement plus implicites du français LE, acquièrent l'accord sujet-verbe de la 3^e personne du pluriel de deux manières : explicite pour les verbes réguliers et irréguliers, et implicite pour les verbes fréquents.

Afin d'évaluer ces hypothèses, nous allons dans cette partie présenter la mise en œuvre quantitative et qualitative, les stratégies de production et le recours aux connaissances implicites et explicites. Nous allons comparer les verbes réguliers, irréguliers et fréquents, en commençant par présenter la mise en œuvre quantitative de la 3^e personne du pluriel :

¹² Pour définir la durée de pause significative d'une production explicite, nous nous reposons sur les analyses faites par Stömqvist et Ahlsén (1999) et déterminons la durée d'une pause indiquant une activité explicite à > 2,0s.

Tableau 7. *La production implicite et explicite et correcte et erronée de la 3e personne du pluriel*

	Rég Impl +	Rég Impl-	Rég Expl +	Rég Expl -	Irrég Impl+	Irrég Impl-	Irrég Expl+	Irrég Expl-	Fréq Impl+	Fréq Impl-	Fréq Expl+	Fréq Expl-
CHR 1	0	4	0	0	0	2	0	0	0	1	0	0
CHR 2	0	0	2	1	0	0	1	0	0	1	0	0
CHR 3	0	0	0	0	0	0	0	0	3	1	0	0
CHR 4	2	0	1	0	0	0	1	0	2	1	1	0
CHR 5	0	0	1	0	0	0	0	0	0	0	0	0
CHR 6	0	0	2	0	1	0	9	1	4	0	1	0
Σ	2	4	6	1	1	2	11	1	9	4	2	0
EMM 1	0	0	0	0	0	0	0	0	0	1	0	0
EMM 2	0	1	1	0	0	0	2	0	0	0	0	0
EMM 3	0	0	0	0	0	0	0	0	0	0	0	0
EMM 4	3	0	2	0	1	0	2	0	1	3	3	0
EMM 5	1	0	4	0	0	0	3	3	0	0	2	0
EMM 6	4	2	5	0	2	0	1	0	0	0	0	0
Σ	8	3	12	0	3	0	8	3	1	4	5	0
MAR 1	0	0	0	0	0	0	0	0	0	0	0	0
MAR 2	0	0	0	0	0	0	0	0	0	0	0	0
MAR 3	0	0	0	0	0	0	0	0	0	0	0	0
MAR 4	0	1	4	0	0	0	2	0	0	0	1	0
MAR 5	1	1	3	0	1	0	7	1	4	1	3	0
MAR 6	0	0	2	0	1	0	2	0	0	1	4	0
Σ	1	2	9	0	2	0	11	1	4	2	8	0
OSC 1	0	0	1	0	0	0	0	0	0	0	0	0
OSC 2	0	0	1	0	0	1	0	0	0	0	0	0
OSC 3	0	0	1	0	0	0	0	0	0	0	0	0
OSC 4	2	0	1	0	0	0	0	0	0	0	3	0
OSC 5	0	0	0	0	0	2	8	0	0	0	0	0
OSC 6	1	0	1	1	0	0	3	0	0	0	1	0
Σ	3	0	5	1	0	3	11	0	0	0	4	0

Légende : Rég = verbes réguliers *-er*, Irrég = verbes irréguliers, Fréq = *être, avoir, aller et faire*, Impl = implicite, Expl = explicite, + = pluriel marqué¹³, - = pluriel non marqué. SOP n'est pas incluse dans le tableau, car elle ne marque jamais le pluriel. Dans ce tableau, toutes les occurrences sont incluses, aussi celles qui ont ensuite été corrigées.

Le tableau 7 permet de faire plusieurs observations. Commençons par les verbes réguliers et irréguliers. Chez les trois apprenants qui créent des contextes du pluriel tôt dans l'étude (CHR, EMM et OSC), le pluriel commence aussi à être marqué assez tôt. CHR et EMM le font à partir du deuxième enregistrement et OSC dès le premier. Nous voyons aussi que tous les trois le marquent surtout explicitement quand ils le marquent. Ainsi les formes non-marquées du premier

¹³ Sont incluses aussi les formes erronées marquant le pluriel ; *-ont*, et *ons*.

enregistrement de CHR et EMM sont produites implicitement. Nous pouvons constater qu'une fois que le pluriel commence à être marqué il l'est d'une manière conséquente. Cette impression est renforcée quand nous regardons de près ce qui se passe avec les formes non-marquées dans les enregistrements 4-6. À quatre cas près, dans EMM 6, MAR 4 et OSC 5, ces formes non-marquées sont corrigées en formes du pluriel marquées. Les contextes de trois de ces quatre cas du pluriel non-marqué prêtent à confusion. Dans MAR 4, l'apprenante a d'abord traité le sujet, *les gents*, comme du singulier, *le gent*, en accordant les deux verbes qui suivent avec ce sujet au singulier. Quand elle revient dans le texte pour changer le sujet au pluriel, *les gents*, elle oublie d'accorder le deuxième verbe au pluriel. Pour les deux cas dans OSC 5, la forme verbale se termine avec *par – ent* au singulier : *ils vient* et *ils revient*.

De plus, la production explicite semble garantir une réussite dans l'accord au pluriel. Si nous ne considérons que le produit final, il n'y a pas de pluriel non-marqué explicite. Toutes les formes explicites non-marquées, dans le tableau 7, sont par la suite corrigées. **Une production plus contrôlée, semble favoriser le marquage du pluriel.** Les formes erronées non-corrigées (de EMM 6, MAR 4 et OSC 5), par la suite sont toutes produites implicitement, donc d'une manière moins consciente et moins contrôlée.

Un nouvel argument en faveur de l'application explicite de la règle est fourni par le *logfile* de la production écrite. L'emplacement de la pause significative est souvent juste devant le suffixe marquant le pluriel, voir les exemples (20-23), où le trait vertical désigne l'emplacement de la pause :

- (20) [CHR 6] ils visitlent, ils sortlent
- (21) [EMM 2] les garçons cherchlENT
- (22) [MAR 4] les gents danselNT, une homme et une femme regardlENT
- (23) [OSC 4] les deux personnes regardlent

Dans quelques-uns de ces exemples, il y a une autre indication de l'activité consciente (explicite) de l'accord du pluriel. Les apprenants prononcent assez souvent le suffixe muet en l'écrivant. Cette activité est ici marquée à l'aide des lettres majuscules.

Enfin, le tableau 7 montre que l'accord au pluriel des verbes réguliers et irréguliers semble se faire surtout explicitement tout au long de l'étude. Pour l'accord sujet – verbe à la 3^e personne du pluriel, on ne peut trouver d'indices d'une automatisation ou un développement de connaissances même chez les apprenants ayant un traitement plus implicite du verbe (CHR et EMM).

La mise en œuvre du pluriel des verbes fréquents semble différer de celle des verbes réguliers et irréguliers. Chez tous les apprenants du tableau 7, à l'exception d'OSC, la production des verbes fréquents au pluriel devient de plus en plus implicite vers la fin de l'étude. De plus, CHR, qui a une production assez implicite, traite les verbes fréquents au pluriel d'une manière assez implicite dès le début de l'étude. Dans sa production, nous pouvons suivre comment le pluriel commence à être marqué et constater que ce marquage se fait dès le début d'une manière implicite. CHR 2 et CHR 3 permettent de constater une différence entre les verbes réguliers et irréguliers et les verbes fréquents. Le pluriel non-marqué des deux premiers groupes des verbes est corrigé et ensuite marqué, tandis que CHR ne corrige pas le pluriel non-marqué des verbes fréquents dans CHR 2 et

3. Dans CHR 4 par contre elle le fait. En ce qui concerne le pluriel produit explicitement dans CHR 6, il s'agit de l'application de l'algorithme *racine + -ent*. Elle écrit *ils font*.

L'analyse quantitative du tableau 7 donne des arguments en faveur des hypothèses formulées. Nous allons maintenant examiner la validité des hypothèses en approfondissant la mise en œuvre qualitative du pluriel. De quelle manière les apprenants appliquent-ils la règle simple algorithmique ?

L'emplacement des pauses dans les exemples (20-23) montre un traitement régulier *racine + ent*. Il est maintenant intéressant de voir jusqu'où va cette application régulière. Lardy et Fayol (2001) ont montré que pour l'accord au pluriel en français LM, la morphophonologie joue un rôle important quand il s'agit d'éviter des erreurs de l'accord. Ainsi les verbes qui changent la racine pour le pluriel et qui rendent audible la différence entre le singulier et le pluriel à la 3^e personne, donnent lieu de moins d'erreurs d'accord en LM. La morphophonologie importante à l'oral, donnerait également des indications importantes à l'écrit.

D'après la recherche de Bartning et Schlyter (2002, 2004) et Schlyter (2003 a et b), le pluriel des verbes irréguliers, donc avec une différence morphophonologique entre la 3^e personne du singulier et du pluriel, n'est marqué conséquemment à l'oral qu'à des stades très élevés, i.e. après des années dans le pays chez les apprenants naturels et après des études universitaires chez les apprenants guidés. Les clés orales du pluriel ne sont apparemment pas assez pertinentes pour que les apprenants acquièrent ce pluriel aussi vite que le pluriel des verbes fréquents, *-ont*.

La question est maintenant de savoir si ces indications orales sont prises en considération ou utilisées en français LE écrit. Parallèlement aux formes correctes de certains verbes irréguliers¹⁴, il y a des formes avec une racine incorrecte mais avec un suffixe correct du pluriel. Tout comme Ågren, ce volume, nous pouvons voir une application de la règle *racine + -ent* aussi pour les verbes irréguliers, qui normalement changent de racine à la 3^e personne du pluriel :

(24) [CHR] ils *s'assistent*¹⁵, ils *ditent*, ils *font*

(25) [EMM] ils *inconnuent*, ils *parent*¹⁶

(26) [MAR] ils *ditent*, ils *doivent*, ils *font*, ils *faisent*, ils *pouvent*, ils *prennent*

Comme la production de ces formes est surtout explicite, il semble s'agir d'une application consciente de l'idée de Bybee (1991), une forme de base (dans la plupart des cas la forme de base est la 3^e personne du singulier), à laquelle on ajoute la morphologie du nombre et du temps. Cette régularisation du pluriel est aussi un bon exemple du *one-to-one principle* (Anderson, 1984, cité dans Ellis 1997), d'après lequel une seule forme ou construction invariable exprime le sens sous-jacent. L'application de la règle est parfois si forte que les apprenants ne se rendent pas compte de ce qu'ils prononcent la forme correcte, mais écrivent la forme produite d'après la règle simple :

¹⁴ Surtout EMM, qui a aussi une production un peu plus implicite des verbes irréguliers au pluriel, semble trouver les formes des racines spécifiques au pluriel, ex. ils *veulent* et ils *peuvent*.

¹⁵ Les exemples ils *s'assistent* et ils *inconnuent* sont inclus pour montrer l'extrême régularité dans le traitement, même si les apprenants partent d'un adjectif comme racine au lieu du verbe.

¹⁶ Nous avons inclus ils *parent* dans cette énumération d'exemples car EMM semble d'après le protocole verbal partir de la forme orale *il /parl/*.

(27) [MAR 4] l'homme et la femme ne peuvent regarder la télé

« *ne peut* alors c'est troisième pluriel c'est *pouv/ã/ pouv peuvent (...)* *pou pouv pouv/ã/ ne peuvent* »¹⁷

(28) [MAR 5] ils prenDent > prennaient > prennent > prennent maintenant une mouche

« ils sortent *ils prenD eh un une mouche ils prenD prenD ehm ils prennent* peut-être deux N *NAIT ils prennaient* non c'est une autre temps *prennent NENT ils prennent* oui peut-être c'est non on prend *prenD ENT ils prenDent une mouche* »

Les exemples (27-28) montrent une insensibilité aux clés orales de la part de MAR. Cet apprenant, dont la production écrite en français LE est très explicite et qui se fie beaucoup à ces connaissances déclaratives, donne la priorité à l'application consciente de la règle *racine + ent* sur les clés orales qu'il produit lui-même. La stratégie de se répéter le paradigme du verbe appris par cœur à laquelle les apprenants EMM, MAR et OSC ont recours pour trouver le bon suffixe au singulier, n'est pas utilisée pour le pluriel. Ici la stratégie semble surtout être de prononcer le suffixe à haute voix, comme nous venons de le voir dans les exemples (21-22) et (27-28).

Pour les verbes réguliers et irréguliers, on constate un traitement assez explicite de l'accord à la 3^e personne du pluriel chez les quatre apprenants qui marquent le pluriel dans l'étude. Cette production explicite semble également donner la priorité à un traitement régularisé de *racine + -ent*, aussi bien pour les verbes réguliers que pour les verbes irréguliers. Ce n'est que pour les verbes très fréquents, surtout *être* et *avoir*, que ce traitement ne suit pas cette règle simple et est par conséquent aussi un peu plus implicite. Notre première hypothèse est donc confirmée aussi bien par l'analyse quantitative que qualitative. La deuxième hypothèse est confirmée surtout chez l'apprenante CHR. Chez les autres apprenants, on peut pourtant suivre un léger développement vers une production plus implicite en ce qui concerne les verbes fréquents.

5 Conclusion

Nous venons de voir que la mise en œuvre de l'accord sujet – verbe à la 3^e personne du pluriel diffère de celle du singulier de plusieurs manières. Pour commencer, cet accord, souvent muet, se fait surtout d'une manière consciente chez les apprenants guidés du français LE de cette étude. Surtout en ce qui concerne le groupe des apprenants ayant une production plus implicite, le pourcentage de la production explicite pour le pluriel est beaucoup plus élevé que pour le singulier. Ce niveau élevé de la production explicite indique que les apprenants ont consciemment recours aux règles, quand ils traitent du pluriel.

Deuxièmement, les traitements différents du singulier et du pluriel ont des implications qualitatives. Tandis que les suffixes du singulier des verbes irréguliers continuent à être erronés assez longtemps dans l'étude, nous avons pu constater que le pluriel est marqué d'une manière conséquente, une fois qu'il a commencé à être marqué, aussi bien en ce qui concerne les verbes réguliers, *-er*, que les verbes irréguliers. La conscience de ce qu'il faut ajouter les suffixe *-ent*

¹⁷ Entre citation, nous avons mis la transcription du protocole verbal pour la séquence en question. Ce qui est écrit avec des caractères normaux est dit en suédois et ce qui est écrit en italique est dit en français.

pour marquer la 3^e personne du pluriel sur le verbe est donc très grande chez les apprenants.

Troisièmement, il semble que cette conscience de la règle qu'il faut ajouter *-ent* à la racine du verbe donne une stratégie qui prime sur les autres stratégies observées pour la production de l'accord au singulier. Même les apprenants qui ont souvent recours à se répéter le paradigme du verbe pour trouver le bon suffixe au singulier, ne le font pas au pluriel. Il arrive donc que les verbes qui changent de radical au pluriel ne sont pas détectés et que la règle simple *racine + -ent* est appliquée aussi pour ces verbes. Tout comme l'a observé Bybee (1991) la forme de la 3^e personne du singulier fonctionne comme une forme de base sur laquelle *-ent* est ajouté. Pour cette raison les apprenants produisent des formes erronées mais avec le pluriel marqué du genre *ils peuvent, ils font* etc. Même quand les apprenants se donnent des clés orales en utilisant la forme correcte en réfléchissant à haute voix, ils restent fidèles à l'application de la règle sans s'en rendre compte.

Quatrièmement, le traitement des verbes fréquents, surtout *être* et *avoir* devient plus implicite et semble aussi exclu de l'application de la règle discutée ci-dessus. Tandis que les verbes réguliers et irréguliers sont produits d'une manière surtout explicite jusqu'à la fin de l'étude chez les apprenants, une des apprenants marque le pluriel de ces verbes d'une manière implicite quand elle commence à le faire. Elle continue aussi à produire les formes du pluriel de ces verbes d'une manière implicite tout au long de l'étude. Chez les autres apprenants on peut constater un développement plus prononcé vers une production implicite pour les verbes fréquents comparé aux autres groupes de verbes.

Finalement, cette production explicite du pluriel semble confirmer les théories d'un traitement explicite des règles simples de Pica (1985) et Krashen (1981, 1985, 1994). Au contraire des règles qui gèrent les paradigmes des verbes au singulier, la règle très simple d'ajouter *-ent* pour accorder le verbe à la 3^e personne du pluriel se prête apparemment bien à un recours réussi aux connaissances explicites.

6 Références bibliographiques

- Barbier, M-L. (2004) « Écrire en langue seconde, quelles spécificités ? ». In Piolat, A. (éd). *Écriture : Approches en sciences cognitives*. Aix-en-Provence, Publications de l'Université de Provence.
- Bartning, I. & Schlyter, S. (2002) « Itinéraires et stades acquisitionnels en français LE. Essai de synthèse ». Université de Stockholm et Université de Lund. Paper présenté à EUROSLA 12, Bâle, sept 2002.
- Bartning, I. et Schlyter, S. (2004) « Itinéraires acquisitionnels et stades de développement en français L2 ». *Journal of French Language Studies*, 14:3:281-299.
- Bybee, J. L. (1991) « Natural morphology : the organization of paradigms and language acquisition ». In Huebner, T. & Ferguson, C.A (éds.) *Crosscurrents in Second Language Acquisition and Linguistic Theories*, pp. 67-91. Amsterdam/Philadelphia : John Benjamins.
- Cyr, P. (1996) *Les stratégies d'apprentissage*. Anjou (Québec): Les éditions CEC
- Ellis, R. (1997) *SLA Research and Language Teaching*. Oxford : Oxford University Press.
- Hedbor, C. (2003) *Les règles d'abord ou la production directe du français écrit? Étude*

- longitudinale sur l'utilisation ou non des connaissances explicites par cinq élèves de différentes attitudes cognitives. PERLES n° 18. Institut d'Études Romanes de Lund, Université de Lund.*
- Hedbor, C. (2005) « What impact does your learning attitude have on the progression of your learning of French? », *Forskning och undervisning i främmande språk*. Rapport du workshop les 10-11 juin 2004, Växjö, Växjö universitet.
- Krashen, S. D. (1981) [1988] *Second language acquisition and second language learning*. New York : Prentice Hall.
- Krashen, S. D. (1985) [1992] *The input hypothesis : Issues and implications*. (nouv. éd) New York : Laredo Publishing.
- Krashen, S. D. (1994) « The input hypothesis and its rivals ». In Ellis, N. C. (éd.) *Implicit and explicit learning of languages*, pp. 45-78. London : Academic Press.
- Kristiansen, I. (1992) *Foreign language learning and nonlearning*. Dissertation. Department of Education : University of Helsinki : Research bulletin 82 : 41-79.
- Lambert, M. (1994) « Les profils d'apprenants comme mode de description et d'explication à la variabilité des apprentissages en langue étrangère ». *AILE* 4 : 81-108.
- Largy, P. (2001) « La révision des accords nominal et verbal chez l'enfant ». *L'année psychologique* 101 : 221-245.
- Largy, P. (2002) *Apprentissage et mise en œuvre de la morphologie flexionnelle du nombre*. Thèse d'habilitation. Université de Rouen. UFR de Psychologie, Sociologie, Sciences de l'Éducation.
- Largy, P., Chanquoy, L. & Dédeyan, A. (2004) « Orthographic revision : The case of subject-verb agreement in French ». In Allal, L. et al. (éds) *Revision : Cognitive and Instructional Processes*, pp. 39-62. Boston : Kluwer Academic Publishers.
- Largy, P. & Fayol, M. (2001) « Oral cues improve subject-verb agreement in written French ». *International Journal of Psychology* 36(2) : 121-132.
- Larsen-Freeman, D & Long, M.H. (1991) *An Introduction to Second Language Acquisition Research*. London : Longman.
- Lightbown, P. M. & Spada, N. (1993) *How Languages are Learned*. Oxford : Oxford University Press.
- Paradis, M. (1985) « On the Representation of Two Languages in One Brain ». *Language Sciences* 7(1) : 1-39.
- Paradis, M. (1994) « Neurolinguistic aspects of implicit and explicit memory : Implications for bilingualism and SLA » In Ellis, N.C (éd.) *Implicit and explicit learning of languages*, pp. 393-420. London : Academic Press.
- Paradis, M. (1997) « The Cognitive Neuropsychology of Bilingualism ». In de Groot, A. M. B. & Kroll, J. F. (éds.) *Tutorials in bilingualism : Psycholinguistic Perspectives*, pp. 331-354. New Jersey : Lawrence Erlbaum.
- Paradis, M. (2000) « Neurolinguistics of bilingualism and the teaching of languages ». Paper présenté à International Symposium of Language Awareness, Odense 27-28 avril.
- Pica, T. (1985) « Linguistic simplicity and learnability: Implications for language syllabus design », In Hyltenstam, K. et Pienemann, M. (éds.) *Modelling and Assessing Second Language Acquisition*. Clevedon : Multilingual Matters.
- Schlyter, S. (2003a) « Development of verb morphology and finiteness in children and adults acquiring French ». In Dimroth, C. & Starren, M. (éds.) *Information structure, linguistic structure, and the dynamics of learner language*, pp. 15-44. Amsterdam/Philadelphia : John

Benjamins.

Schlyter, S. (2003b) « Stades développementaux en français chez des apprenants suédophones: Exemples du corpus Lund ». Disponible sur www.rom.lu.se/durs/archive.htm

Strömqvist, S., Ahlsén, E. (1999) *The process of writing : A progress report*. Gothenburg. Gothenburg Papers in Theoretical Linguistics : 83.

Weingarten, R., Nottbusch, G., Will, U. (2004) « Morphemes, syllables and graphemes in written word production ». In Pechman, Th. & Habel, Ch. (éds.) *Multidisciplinary Approaches to Language Production. Berlin*. New York: Mouton de Gruyter.

Zimmerman, R. 2000. « Writing : subprocesses, a model of formulating and empirical findings » : *Learning and Instruction*, 10 : 73-79.

AU REVOIR!

