



La prédication

Martin Haiden

► To cite this version:

Martin Haiden. La prédication. Linguistique. Université François Rabelais - Tours, 2011. tel-00624002

HAL Id: tel-00624002

<https://theses.hal.science/tel-00624002>

Submitted on 15 Sep 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ FRANÇOIS - RABELAIS DE TOURS



Année Universitaire : 2010/2011

HABILITATION A DIRIGER DES RECHERCHES

Tome 1

Discipline : Sciences du langage

présentée et soutenue publiquement

par :

Martin HAIDEN

Le 20 juin 2011

JURY :

(Par ordre alphabétique)

Mme Cornelia	HAMANN	Professeur	Université d'Oldenburg (Allemagne)
M. Philippe	PREVOST	Professeur des Universités	Université de Tours
M. Alain	ROUVERET	Professeur des Universités	Université Paris 7
M. Ur	SHLONSKY	Professeur	Université de Genève (Suisse)
Mme Laurie	TULLER	Professeur des Universités	Université de Tours
M. Edwin	WILLIAMS	Professeur	Université de Princeton (Etats-Unis)

Table des matières

CURRUCULUM VITAE.....	iii
 MEMOIRE : LA PREDICATION	
0. INTRODUCTION	1
1. PERCEPTION	3
1.1 INTRODUCTION.....	4
1.2 LES ROLES THEMATIQUES	6
1.3 CONTRE LA DECOMPOSITION LEXICALE DES PREDICATS	9
1.3.1 <i>Données expérimentales montrant que les composants linguistiques et cognitifs sont dissociés.....</i>	<i>11</i>
1.3.2 <i>Acquisition et induction syntaxique</i>	<i>15</i>
1.3.3 <i>verser, remplir, contenir ; connaissances, perceptions et peur : les paradoxes de la structure lexicale.....</i>	<i>18</i>
1.4 DECOMPOSER LES ROLES ET NON LES EVENEMENTS : LES FONDEMENTS COGNITIFS	26
1.4.1 <i>La perception de la causalité (ToBy).....</i>	<i>26</i>
1.4.2 <i>Action rationnelle (ToMM)</i>	<i>32</i>
1.5 CONCLUSION : LE SENS DES TRAITS THEMATIQUES.....	36
2. CONCATENATION	41
2.1 LE MECANISME GENERATIF : PRESENTATION DU MODELE.....	41
2.1.1 <i>Modèle</i>	<i>42</i>
2.1.2 <i>Exposants morphologiques</i>	<i>50</i>
2.1.3 <i>Ordre linéaire et gabarits</i>	<i>53</i>
2.2 BASE EMPIRIQUE : LA REPRESENTATION SYNTAXIQUE DES PREPOSITIONS FAIBLES DU BERBERE DE CHEMINI (KABYLE)	56
2.2.1 <i>La distribution des prépositions faibles.....</i>	<i>57</i>
2.2.2 <i>Les particules</i>	<i>62</i>
2.2.3 <i>Complémenteurs et prépositions faibles</i>	<i>68</i>
2.2.4 <i>Les clitiques.....</i>	<i>71</i>
2.2.5 <i>Sur la déficience phonologique des prépositions faibles.....</i>	<i>77</i>
2.2.6 <i>Support gabaritique et projection syntaxique.....</i>	<i>84</i>
2.3 CONCLUSION	93

3. ACQUISITION	95
3.1 INTRODUCTION.....	95
3.2 MATERIEL ET METHODES.....	100
3.2.1 Population	100
3.2.2 Matériel expérimental.....	101
3.2.3 Transcription et codage	102
3.2.4 Variables testées	105
3.2.5 Traitement statistique	109
3.3 RESULTATS.....	110
3.3.1 Statut grammatical du clitique élicité : la variable fonction	110
3.3.2 Statut grammatical - résumé.....	124
3.3.3 Différences entre les groupes	130
3.3.4 Différences entre populations – résumé	154
3.3.5 La structure de la population L2	160
3.4 CONCLUSION	183
4. REFERENCES BIBLIOGRAPHIQUES	187

Curriculum Vitae

HAIDEN, Martin

Né le 10 janvier 1969 à Steyr (Autriche)

Nationalité : européenne (autrichienne)

Situation actuelle :

Grade : Maître de conférences

Etablissement : Université François Rabelais de Tours, UFR Lettres et langues,
département Sciences du langage

Equipe de recherche : UMR_S 930 INSERM et ERL 3106 CNRS "Imagerie et cerveau"
(resp. D. Guilloteau), Equipe 1 "Autisme et troubles du développement :
psychopathologie, physiopathologie et thérapeutique" (resp. C. Barthélémy).

Section CNU : 7

Thèse : "Theta Theory"

soutenue le 13 Septembre 2004 à l'Université de Tilburg (Pays-Bas).

directeurs de recherches : H. van Riemsdijk (Tilburg) et M. Prinzhorn (Vienne)

Adresse administrative :

Université François Rabelais

Département SDL

3, rue des Tanneurs

B.P. 4103

37041 Tours Cedex 1

Tél : 02.47.36.66.84

Courriel : martin.haiden@univ-tours.fr

Page web : <http://www.univ-tours.fr/haiden>

Parcours professionnel

- 01/09/2006- Maître de Conférences à l'**Université de Tours**, UFR Lettres et langues. Titulaire depuis le 1 septembre 2007.
- 2004-6
- Post-doctorant **CNRS** à l'**UMR 8163 STL** (dir. F. Blaise, anciennement **UMR 8528 Silex**, dir. Ph. Miller), CNRS & Université Lille 3 (2005/2006).
 - Chargé de cours en linguistique à l'UFR de Lettres Modernes, Université Lille 3 (2005/2006).
 - Post-doctorant à l'**UMR 8528 Silex** (dir. Ph. Miller), CNRS & Université Lille 3 (2004/2005)
- 2003-4
- Enseignant-chercheur (*Lecturer*) en linguistique allemande à **Trinity College Dublin** (Irlande), département d'études germaniques.
 - Equipe de recherche : **CCLS** (Center for Computing and Language studies), **Trinity College Dublin** (resp. C. Vogel).
- 1999-2003
- Allocataire de recherche (*Österreich Kooperation*, allocation du Ministère de la recherche et du Ministère des affaires étrangères autrichien).
 - Moniteur d'allemand à **Aston University**, Birmingham (Royaume-Uni), faculté des langues vivantes.
 - Chargé de cours en linguistique à l'**Université de Wolverhampton** (Royaume-Uni ; en cumul avec le poste à Aston University), département de langues vivantes (2^{ème} semestre 2000/2001).
 - Equipe de recherche : **ISSLS** (Institut for the Study of Language and Society), Aston Univ., Birmingham (Resp. : Ch. Schäffner).

Publications

- Haiden, M. (à paraître). "The content of semantic roles." In : M. Everaert, M. Marelj et T. Siloni (éds.) *The Theta System*. Oxford : Oxford University Press.
- Bendjaballah, S., Haiden, M. (à paraître). "The representational anomalies of floating markers." In : Th. Biberauer et I. Roberts (éds.) *Principles of Linearization*. Berlin/New York : Mouton de Gruyter.
- Bendjaballah, S., Haiden, M. (à paraître). "La représentation syntaxique des prépositions faibles du berbère de Chemini (kabyle)." In : J. Tseng (éd.) *Les prépositions faibles*. (Collection langues et syntaxe). Paris : Hermès.
- Prévost, P., L. Tuller, M. Scheidnes, S. Ferré, Haiden, M. (à paraître). "Computational Complexity Effects in the Acquisition of Wh-Questions in Child L2 French." *Proceedings of Romance Turn 3, Southampton 2008*.
- Haiden, M. 2009. "On Bare Prosodic Structure and the Spell-out of Features." In: K. Grohmann (éd.) *Explorations of Phase Theory: Interpretation at the Interfaces*, pp 67-94. (Interface Explorations). Berlin/New York : Mouton de Gruyter.
- Bendjaballah, S. & Haiden, M. 2008. "A Typology of Emptiness in Templates." In: J. Hartmann, V. Hegedüs et H van Riemsdijk (éds.) *The Sounds of Silence*, pp 23-59. (North Holland Linguistic Series: Linguistic Variations, vol. 63). Amsterdam : Elsevier.
- Haiden, M. 2008. "Linearization in Bare Prosodic Structure." *Poznań Studies in Contemporary Linguistics* 44(3), pp. 401-421.
- Haiden, M. 2006. "Stefan Müller: Complex predicates". *Journal of Comparative Germanic Linguistics* 9/2, pp. 143-160.
- Bendjaballah, S., Haiden, M. 2005. "Präpositionen und Komplementierer im Chemini Berber". *Folia Linguistica* 39/3-4, pp. 319-362.
- Haiden, M. 2005. *Theta Theory*. (Studies in Generative Grammar vol. 78). Berlin, New York : Mouton de Gruyter.
- Haiden, M. 2007. "Sous-spécification et projection argumentale dans le *Theta System*". *Recherches Linguistiques de Vincennes* 36, pp. 153-188.
- Haiden, M. 2005. "Verb-Particle Constructions". In : H. van Riemsdijk et M. Everaert (éds), *Blackwell Companion to Syntax*, Vol. 5, pp. 342-373. Oxford : Blackwell.
- Haiden, M. 2005. "Concatenation and Interpretation". In : Broekhuis, H., N. Corver, R. Huybregts, U. Kleinhenz et J. Koster (éds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, pp. 157-165. Berlin, New York : Mouton de Gruyter.

- Bendjaballah, S. & Haiden, M. 2005. "The Grammar of Prepositions in Berber (Taqbaylit)". *Lingbuzz* 000261 (<http://ling.auf.net>).
- Bendjaballah, S., Haiden, M. 2003. "Templatic Architecture". *Recherches Linguistiques de Vincennes* 32, pp. 157-168.
- Haiden, M. 2003. "Variation in the German Present Participle". In : Zybatow, L. (éd), *Europa der Sprachen: Sprachkompetenz - Mehrsprachigkeit - Translation. Akten des 35. Linguistischen Kolloquiums in Innsbruck 2000. Teil II : Sprache und Kognition*. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang.
- Bendjaballah, S., Haiden, M. 2002. "Templatic Architecture". In : Agbayani, P. Koskinen et V. Samiian (éds), *Proceedings of the Western Conference in Linguistics (WECOL) 2002*. Fresno : California State University.
- Bendjaballah, S., Haiden, M. 2002. "Meaningful Vowels". In : P. Bottari (éd), *Proceedings of the XXVIII Incontro di Grammatica Generativa*. Lecce : Congedo Editore.
- Bendjaballah, S., Haiden, M. 2001. "Templatic Inflection in German". *UPenn Working Papers in Linguistics* 8/1, pp. 29-41.
- Haiden, M. 2001. "The German Present Participle". In : P. Albizu et B. Fernandez (éds), *On Case and Agreement*, pp. 159-213. Leioa : University of the Basque Country Press Service.
- Haiden, M. 1997. "Verbal Inflection and the Structure of IP in German". *Groninger Arbeiten zur Germanistischen Linguistik* 41, pp. 77-106.
- Haiden, M. 1996. "The Aspect of Short Scrambling". *Wiener Linguistische Gazette* 57-59, pp. 121-146.
- Haiden, M. 1993. "Split Topicalization". *Wiener Linguistische Gazette* 48-50, pp. 31-60.

Communications (conférences invitées, colloques, séminaires)

- Haiden, M. 2010. "Blind reference." Workshop CLEAR (Child Language & Eye tracking : Analyses and Rationale), Université de Groningen, Pays-Bas, 10/09/2010.
- Tuller, L., Prévost, P., Haiden, M., Ferré, S., Scheidnes, M. & Zebib, R. 2010. "Computational Complexity Effects in the Acquisition of French by Child L2 Learners Compared to Children with SLI" Meeting of COST action IS0804, Newcastle, Royaume-Uni. 26-28/05/2010.
- Haiden, M., Ferré, S., Prévost, P., Scheidnes, M., Tuller, L. 2010. "Internal factors in the acquisition of French accusative clitics: The role of morphology." DGfS, Berlin, Allemagne, 23-26/02/2010. (Présentation acceptée, mais annulée pour cause de grève du trafic aérien.)
- Haiden, M., Ferré, S., Prévost, P., Tuller, L. 2009. "French accusative clitics: What's the problem? Refining a measure of formal linguistic complexity." Poster, IRIA, Tours, 15-17/04/2009.
- Haiden, M., Ferré, S., Prévost, P., Scheidnes, M., Tuller, L. 2009. "Production and Comprehension of Wh-questions in the Acquisition of French: Comparing L2 Children and Children with SLI." DGfS, Osnabrück, Allemagne, 04/03/2009.
- Haiden, M. 2009. "French Accusative Clitics." Workshop on Eye-tracking and (A)typical Language Acquisition, Tours, 14/12/2009.
- Haiden, M., Prévost, P., Scheidnes, M., Tuller, L. 2009. "SLI in French : What's complex and why." Meeting of COST action IS0804, Amsterdam, Pays-Bas, 16-18/11/2009
- Scheidnes, M., Ferré, S., Haiden, M., Prévost, P., Tuller, L. 2008. "Acquisition of Wh-Questions in French: L2 Children and L1 Children with SLI." SLRF, Hawaï, Etats-Unis, 17-19/10/2008.
- Scheidnes, M., Ferré, S., Haiden, M., Prévost, P., Tuller, L. 2008. "Acquisition of Wh-Questions in French: L2 Children and L1 Children with SLI." RT 3, Southampton, Royaume-Uni, 18-20/09/2008.
- Ferré, S., Haiden, M., Prévost, P., Tuller, L. 2008. "L1 influence or language impairment? A twin case study." The Nature and Development of L2 French., Southampton, Royaume-Uni, 7-8/07/2008.
- Haiden, M. 2008. "Les prépositions légères en berbère de Chemini (kabyle)." Séminaire de linguistique, Université d'Orléans, 07/03/2008.

- Haiden, M. 2008. "On Templates, Linear Order, and the Acquisition of Parameter Settings." GLOW 31, Newcastle, Royaume-Uni, 29/03/2008
- Bendjaballah, S., Haiden, M. 2008. "Templatic Effects in a Concatenative System: German Strong Verbs." Séminaire du CASTL, Université de Tromsø, Norvège, 25/11/2008.
- Bendjaballah, S., Haiden, M. 2008. "Conceptual issues" Séminaire du CASTL, Université de Tromsø, Norvège, 27/11/2008.
- Haiden, M. 2008. "Événements atomiques, rôles décomposés." Journée d'études *Causes, agents, instruments*, Université Lille 3, 06/06/2008.
- Haiden, M. 2007. "Linearization in Bare Prosodic Structure." PLM workshop: The Syntax-Phonology Interface and beyond, Gniezno, Pologne, 13/09/2007.
- Haiden, M. 2007. "Les réciproques : morphosyntaxe et lexique." Séminaire sur les dépendances distributives : pluralité nominale et verbale, Fédération Typologie et Universaux Linguistiques – FR2559, Paris, 07/05/2007.
- Haiden, M. & P. Cabredo Hofherr. 2007. "Syntaxe et typologie des réciproques." Séminaire sur les dépendances distributives : pluralité nominale et verbale, Fédération Typologie et Universaux Linguistiques – FR2559, Paris, 15/01/2007.
- Haiden, M. 2006. "La représentation syntaxique des prépositions faibles du berbère de Chemini." Séminaire de l'équipe *Langues & Représentations*, Université de Tours, 14/12/2006.
- Haiden, M. 2005. "Sous-spécification et projection argumentale dans le *Theta System*." Séminaire de l'UMR 8528, Université Lille 3, 18/11/2005.
- Bendjaballah, S., Haiden, M. 2005. "Prépositions berbères, gabarits et interface phonologie-syntaxe." Réunions parisiennes de phonologie, E.H.E.S.S., Paris, 17/06/2005.
- Haiden, M. 2005. "Adjectival Participles in Morphology and Syntax." Séminaire de l'UMR 8528, Université Lille 3, 18/02/2005.
- Bendjaballah, S., Haiden, M. 2005. "Les prépositions légères du berbère." Journée d'études sur les prépositions, Fédération Typologie et Universaux Linguistiques – FR2559, Université Paris 7, 22/01/2005.
- Bendjaballah, S., Haiden, M. 2005. "A Typology of Emptiness in Templates." Sounds of Silence, Tilburg, Pays-Bas, 20/10/2005.
- Bendjaballah, S., Haiden, M. 2005. "Berber Prepositions, Templates and the PF Interface." GLOW 28, Genève, Suisse, 31/03/2005.

- Haiden, M. 2005. "Modalität und Argumentstruktur im Infinitiv unter *haben* und *sein*." Séminaire du *Fachbereich Sprachwissenschaft*, Universität de Konstanz, Allemagne, 19/05/2005.
- Bendjaballah, S., Haiden, M. 2005. "Leichte Präpositionen im Chemini Berber: Phonologie und Syntax." Séminaire du *Sonderforschungsbereich 471*, Universität de Konstanz, Allemagne, 17/05/2005.
- Bendjaballah, S., Haiden, M. 2004. "Berber Light Prepositions, and the Phonology-Syntax Interface." Séminaire du *Grammatical Models Group*, Université de Tilburg, Pays-Bas, 25/11/2004.
- Haiden, M. 2003. "The Concepts Interface: Thematic Roles in the Mind and in the Lexicon." Séminaire de l'UMR 8528, Université Lille 3, 17/01/2003.
- Haiden, M. 2002. "Die Thetarollen des deutschen Verbs." Österreichische Linguistentagung, Innsbruck, Autriche, 06/12/2002.
- Bendjaballah, S., Haiden, M. 2002. "Templatic Architecture." Western Conference in Linguistics (WECOL), Vancouver, Canada, 02/11/2002.
- Bendjaballah, S., Haiden, M. 2002. "Templatic Architecture and German Strong Verbs." 4èmes Journées internationales du GDR 1954 - Phonologies, Grenoble, 07/06/2002.
- Bendjaballah, S., Haiden, M. 2002. "Templatic Inflection in German: On the Status of Apophony at the Syntax-Phonology Interface." Incontro di Grammatica Generativa 28, Lecce, Italie, 28/02/2002.
- Bendjaballah, S., Haiden, M. 2002. "Meaningful Vowels: On the Necessity of Apophonic Morphology in German." International Conference on the Architecture of Grammar, Hyderabad, Inde, 15-17/01/2002.
- Haiden, M. 2001. "Theory." Celebrating the Link, Birmingham, Royaume-Uni, 17/11/2001.
- Bendjaballah, S., Haiden, M. 2001. "Templatic Inflection in German." 25th Penn Linguistics Colloquium, Philadelphie, Etats-Unis, 04/03/2001.
- Haiden, M. 2000. "The Complementizer /d/ in German." 35. Linguistisches Kolloquium, Innsbruck, Autriche, 29/09/2000.
- Haiden, M. 1999. "The Present Participle and Reduced Relative Pronouns in German." Séminaire de syntaxe, Université du Pays Basque, Vitoria, Espagne, 08/01/1999.
- Haiden, M. 1997. "On Some Strange Properties of Verb-Particle Constructions." International Syntax Round Table, Ybbs/Donau, Autriche, 20/07/1997.

- Haiden, M. 1997. "Die Wahrheit über die Partikelverben." Generative Grammatik im Süden (GGS), Wien, Autriche, 08/05/1997.
- Haiden, M. 1997. "Separable Prefixed Verbs, Phrase Structure and Movement Triggers in German." 30th Poznan Linguistic Meeting, Poznan, Pologne, 02/05/1997.
- Haiden, M. 1996. "Aspect, Case and Conceptual Necessity." Girona International Summer School in Linguistics, Girona, Espagne, 21/07/1996.
- Haiden, M. 1996. "The Aspect of Short Scrambling." Séminaire de syntaxe, Deusto University, Bilbao, Espagne, 24/05/1996.
- Haiden, M. 1995. "Moved or in situ: Some Properties of Low Objects." Séminaire de l'Utrecht Institute of Linguistics OTS, Université d'Utrecht, Pays-Bas, 22/11/1995.
- Haiden, M. 1995. "Distribution and Interpretation of NPs in German." Séminaire de syntaxe, Deusto University, Bilbao, Espagne, 16/06/1995.
- Haiden, M. 1994. "Scramblingstypen und existentieller Abschluß." Generative Grammatik im Süden (GGS), Jena, Allemagne, 27/05/1994.
- Haiden, M., Prinzhorn, M. 1994. "Case-Acquisition and Early Clause Structure." International Morphology Workshop, Visegrad, Hongrie, Janvier 1994.
- Haiden, M. 1993. "VP-fronting and NP-split: Evidence from Upper Austrian." Österreichische Linguistentagung, Salzburg, Autriche, Octobre 1993.
- Haiden, M. 1993. "Split Topicalization." Terzia Giornata di Grammatica Tedesca, Venise, Italie.

Projets financés (depuis la thèse)

- 2009-2013 Action **COST** IS0804 :
"Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment", porteur : S. Armon-Lotem (Bar-Ilan Univ)
- *Substitute member* du Management Committee.
- 2008-2011 **ANR** BLAN 08-3-311192
"Complexité et compensation en acquisition atypique du langage : une approche comparative", porteur : L. Tuller (Univ. de Tours)
- co-responsable (avec S. Ferré) de la *Tâche 2 : Caractérisation de la notion de complexité linguistique et de sa compensation*.

Enseignement

Université de Tours (depuis septembre 2006)

Master Cognition et développement

- *Introduction aux disciplines de la cognition, Linguistique*. M2 Cognition et développement (2006-2008).

Master de linguistique

- *Linguistique française*. M2 Linguistique (2006-2009).
- *Phonologie et morphologie*. M1 Linguistique (depuis 2006).
- *Méthodologie de la recherche*. M1 Linguistique (depuis 2006).

Ecole d'orthophonie

- *Morphologie*. 1^{ère} année, École d'orthophonie et L1 SDL (depuis 2007).
- *Sémantique*. 1^{ère} année, École d'orthophonie et L1 SDL (depuis 2007).

UFR Lettres et langues

- *Morphologie et sémantique*. L3, SDL (2006-2010).
- *Histoire de la linguistique*. L3, SDL (2006-2010) ; L2, SDL (depuis 2009).
- *Structure de la langue*. L3, Lettres (2006-2008)
- *Du son à la phrase*. L2, Lettres (2006-2009)
- *Introduction à la linguistique*. L1, Lettres (2006-2009)

UE Libre

- *Initiation à l'analyse linguistique* (2006-2010)
- *Panorama des sciences du langage* (2008/9, 2010/11)

Université d'été EGG (*Eastern Generative Grammar summer school*, Juillet - Août 2006)

Olomouc, République Tchèque.

- *Interface phonologie-syntaxe*, 2h/jour, 2 semaines (niveau avancé)
- *Interface lexique-grammaire*, 2h/jour, 2 semaines (niveau avancé)

Université Lille 3, UFR de Lettres Modernes (2005 – 2006)

- *Syntaxe et sémantique de l'interrogation. L2, Lettres.*

Trinity College Dublin, département d'études germaniques (2003 – 2004)

Enseignement de linguistique :

- *Acquisition de L2, théories actuelles (Master)*
- *Dialectologie de l'allemand (4^e année)*
- *Introduction à la phonologie de l'allemand (2^e et 3^e années)*
- *Histoire de la langue allemande 2 (2^e année)*
- *Histoire de la langue allemande 1 (1^{ère} et 2^e années)*
- *Introduction à la grammaire de l'allemand (1^{ère} année)*

Enseignement de langue :

- *Langue allemande 2 (2^e année)*
- *Langue allemande 1 (1^{ère} année)*

Aston University, Birmingham, faculté des langues vivantes (1999-2003)

Enseignement de linguistique :

- *Introduction à la grammaire comparée (1^{ère} année)*
- *Histoire de la linguistique en Allemagne (4^e année)*

Enseignement de langue, culture et histoire :

- *Langue et identité (Master)*
- *L'opérette : société et culture "fin de siècle" à Vienne (2^e année)*
- *Histoire de l'Allemagne (1^{ère} année)*
- *Introduction au concept d'identité autrichienne (1^{ère} année)*
- *Langue allemande (niveaux débutants à 4^e année)*
- *Traduction anglais-allemand (2^e année)*

Univ. de Wolverhampton, département de langues vivantes (Janvier - Mai 2001)

Enseignement de linguistique :

- *Introduction à la grammaire générative* (2^e année)
- *Introduction à la théorie du gouvernement et du liage* (2^e année)

Encadrement et jurys

Université de Tours (depuis septembre 2006)

Dupuy, Lucile. 2008/2009. "Acquisition des clitiques objet en développement atypique : l'hypothèse de la complexité morphologique." Mémoire présenté en vue de l'obtention de la capacité d'orthophoniste. Soutenu le 30 septembre 2009.

Guimard-Brunault, M. 2009/2010. "Compréhension des pronoms clitiques en langue française : étude pilote en suivi du regard chez l'adulte sain." M2R Cognition et Neurosciences. (co-direction avec J. Martineau). Soutenu le 17 juin 2010.

Léger, Elodie. En cours 2010/2011. "La compréhension des pronoms clitiques en temps réel : étude en eye tracking chez l'adulte et l'enfant sain." M2R Cognition et Neurosciences. (co-direction avec P. Prévost)

Jurys

Brunie, N. "La complexité phonologique dans le langage oral spontané d'enfants avec dysphasie" Mémoire présenté en vue de l'obtention de la capacité d'orthophoniste. soutenu le 22 octobre 2009.

Berthelot, P. "Traitement de la complexité phonologique chez des enfants avec dyslexie âgés de 7 à 10 ans." Mémoire présenté en vue de l'obtention de la capacité d'orthophoniste. soutenu le 9 octobre 2008.

Trinity College Dublin, département d'études germaniques (2003 – 2004)

Supervision de mémoires de recherche

- *acquisition de L2* au niveau Master (MA in Germanic Studies, 5^{ème} année)
- *linguistique allemande* (BA in Germanic Studies, 4^è année)

Aston University, Birmingham, faculté des langues vivantes (1999-2003)

Supervision de mémoires de recherche en études germaniques (BA in German Studies, 3^è et 4^è années ; MA in German Studies, 5^{ème} année).

Activités d'administration

Université de Tours (depuis septembre 2006)

Liées à l'enseignement

- Membre des jurys d'examens, SDL.
- Membre de la commission pédagogique, SDL (2006-2010).
- Coordinateur ERASMUS, SDL.

Liées au recrutement et à la gestion

- Membre du comité de sélection du poste n° 1763, 7^{ème} section, Université Paris 7, mai 2010.
- Codirection du département SDL (2009/2010)
- Responsable des finances, SDL (2006-2009, et à partir de juillet 2010).
- Responsable du matériel informatique (achat, fonctionnement), site web et mobilier, SDL.
- Participation à la mise en valeur du département et de l'équipe de recherche et aux opérations de recrutement (fête de la science, salon du lycéen etc).

Trinity College Dublin, département d'études germaniques (2003 – 2004)

Liées à l'enseignement :

- Coordinateur du cursus "Traitement automatique, Linguistique et Langues étrangères".

Liées au recrutement et à la gestion :

- Représentant du département d'allemand dans le programme d'entrée à Trinity.
- Participation à la mise en valeur du département et aux opérations de recrutement (journées portes ouvertes, accueil d'étudiants etc).

Aston University, Birmingham, faculté des langues vivantes (1999-2003)

Liées à l'enseignement :

- Conception et évaluation des cursus, adaptation aux directives ministérielles.
- Examineur accrédité du *Österreichisches Sprachdiplom Deutsch* (certificat de langue allemande).
- Tuteur personnel d'étudiants.
- Visites d'étudiants lors de leur année de stage en Allemagne.

Liées au recrutement et à la gestion :

- Représentant du département d'allemand au *Technology Strategy Group*.
- Responsable de l'achat et de l'utilisation des logiciels d'enseignement de l'allemand.
- Administration de l'*Austrian Language Certificate* (certificat de langue allemande).
- Participation à la mise en valeur du département et aux opérations de recrutement (journées portes ouvertes, accueil d'étudiants etc).
- Conseiller du personnel et évaluation des collègues.

La prédication

Martin HAIDEN

2011

0. Introduction

Mes travaux de recherche de ces dernières années s'articulent autour de trois axes, tous liés à un aspect de la combinaison des unités dans les expressions linguistiques :

- (i) la sémantique lexicale et son expression dans les structures syntaxiques et morphologiques
- (ii) la morphologie non concaténative et la théorie de la grammaire
- (iii) l'acquisition des pronoms clitiques du français

Ce mémoire comporte ainsi trois chapitres, qui développent certains aspects de chacun de ces axes respectivement. Dans le premier chapitre, intitulé "Perception", je me concentre sur la relation entre les modules cognitifs centraux et les connaissances lexicales. Ce chapitre constitue l'aboutissement de mes recherches sur les connaissances lexicales et leur expression dans les structures prédicat-argument en syntaxe et en morphologie. Ce travail a été initié lors de ma thèse (soutenue en 2004) et développé depuis dans divers travaux ((Haiden, 2005a, c, 2006, 2007, à paraître)). Je pars de l'hypothèse, défendue dans (Haiden, 2005a), que les structures prédicat-argument en grammaire sont déterminées par des connaissances lexicales, plus précisément dans le sens défini par le *Theta System* (Haiden, 2005a, Reinhart, 2000). J'argumente contre la décomposition du sens lexical des verbes en structures arborescentes et propose des critères indépendants et fiables permettant d'identifier les rôles thématiques.

Le deuxième chapitre, "Concaténation", présente l'aspect formel de mes recherches sur la structure argumentale. J'ai proposé dans (Haiden, 2005a) un modèle de projection syntaxique permettant de formaliser l'insertion des rôles thématiques en syntaxe, et je l'ai ensuite appliqué à des phénomènes morphologiques non concaténatifs ((Bendjaballah et Haiden, 2002, 2003, 2005a, b, c, 2008, to appear, Haiden, 2005b, 2008b, 2009)). J'ai développé cet axe de recherche au cours de mon stage post doctoral au sein de l'UMR 8163 – STL, CNRS & Univ. Lille 3 (anciennement UMR 8528 – SILEX, dir. Ph. Miller). Dans le cadre de ce mémoire, je commence par présenter le modèle que je propose. J'argumente, en particulier, que le système générant les relations hiérarchiques est dissocié de celui régissant l'ordre linéaire, ce dernier étant encodé dans le lexique sous forme de gabarits. Dans un second temps, j'illustre le modèle par

une étude de cas, le comportement des prépositions légères du kabyle de Chemini (berbère, afroasiatique).

Le dernier chapitre, intitulé "Acquisition", développe mes travaux sur l'acquisition de la morphosyntaxe, sujet dominant de ma recherche ces deux dernières années, depuis mon intégration à l'U930 INSERM "Imagerie et cerveau" en janvier 2009. J'ai élaboré ce travail dans le cadre du projet ANR CCAAL BLAN 08-3-311192 "Complexité et compensation en acquisition atypique du langage" (porteur : L. Tuller, Univ. de Tours) et de l'action COST IS0804 "Language Impairment in a multilingual society. Linguistic Patterns and the Road to assessment." (porteur : S. Armon-Lotem, Bar-Ilan Univ.) Je pars d'un constat embarrassant : les travaux sur l'acquisition des clitiques objets du français mentionnent des taux de production très variés. Je développe une méthodologie permettant de comparer ces résultats contradictoires, fondée sur un nouvel échantillonnage des données, et sur de nouvelles mesures. En particulier, je propose de comparer les trajectoires de développement et les contrastes entre variables indépendantes, plutôt que les moyennes. Je présente les premiers résultats de cette recherche en cours et argue que les compétences structurales et leurs réalisations morphophonologiques doivent être dissociées. Je mets en évidence que l'épel morphophonologique ralentit l'acquisition des clitiques, et ce, quels que soient les traits grammaticaux qu'ils expriment. Le facteur morphophonologique est donc peu utile au diagnostic des troubles du développement. Je propose un critère permettant de contrôler ce facteur dans l'évaluation de la performance des apprenants du français en tant que langue seconde.

1. Perception

Les contributions de mes travaux dans le domaine s'articulent autour des points ci-dessous. Dans ce premier chapitre, j'examine la relation entre les modules cognitifs centraux et les connaissances lexicales, i.e. je présente essentiellement le point (b).

(a) Description des entrées lexicales verbales et de leurs rôles thématiques (Haiden, 2005a)

(b) Développement d'un système fournissant une base cognitive à la perception des rôles thématiques (Haiden, 2005a, à paraître)

Je justifie les traits de rôle de (Reinhart, 2000) sur la base de faits cognitifs et présente divers arguments montrant que ces traits encodent les concepts explicatifs centraux de la théorie modulaire de la perception des événements. Un des aspects importants de mes travaux porte sur le rôle de l'intentionnalité dans l'interprétation causale des événements de changement d'état : cet aspect est en effet ignoré en sémantique lexicale. Me fondant sur diverses études importantes dans le domaine de la perception, j'établis des critères falsifiables permettant d'identifier les rôles thématiques d'items lexicaux spécifiques.

(c) Développement d'une théorie syntaxique compatible avec la conception que les rôles thématiques sont des prédicats de rôles (Haiden, 2005a)

J'argumente que les rôles thématiques sont des têtes syntaxiques projetant dans le domaine thématique de la phrase.

(d) Application à certains problèmes de morphosyntaxe et de sémantique de l'allemand et de l'anglais (Haiden, 2005a, c, 2006, 2007).

Dans mon livre (Haiden, 2005a), je décris un large ensemble de verbes allemands, propose une théorie de la projection syntaxique des rôles thématiques sous la forme de prédicats et mène une discussion sur le rôle des prédicats thématiques en morphologie, syntaxe et sémantique. En particulier, j'examine les arguments implicites des participes présent et passé de l'allemand, leurs interprétations affirmatives et modales, les compléments modaux des verbes *haben* "avoir" et *sein* "être", le passif événementiel et adjectival, la sélection de l'auxiliaire dans les temps composés, et enfin les verbes à particule. Certains de ces sujets sont également développés dans divers articles : (Haiden, 2005a, 2006) examine la syntaxe des prédicats complexes constitués d'un verbe et d'une

particule prépositionnelle ; (Haiden, 2005a, 2006) étudie l'influence des rôles thématiques sur l'interprétation modale des constructions infinitivales et (Haiden, 2007) précise, sur la base des verbes non causatifs de l'allemand et des verbes d'émission de l'anglais, quelles sont les conditions thématiques qui permettent la réalisation d'un argument externe.

1.1 Introduction

La sémantique lexicale est un domaine riche en observations mais relativement pauvre en discussion théorique. La grande majorité des publications (voir (Levin et Rappaport Hovav, 2005) pour une revue de la littérature) part de l'hypothèse qu'une description fine des actions, événements et états décrits par les verbes conduit nécessairement à une bonne compréhension de la sémantique verbale. Ces travaux utilisent comme outil de description la décomposition lexicale sous diverses formes (Jackendoff, 1990, Levin et Rappaport Hovav, 1995, Pustejovsky, 1995). Ce choix n'est cependant pas anodin : appliqué dans le cadre de la grammaire formelle, il implique des modifications profondes des notions de compositionnalité et de vérité, notions fondamentales en sémantique. Ces conséquences sont évoquées dans le préface de (Jackendoff, 1983), mais demeurent souvent contestées (voir par exemple la défense de la décomposition lexicale dans le manuel de (Stechow, 2007)).

Mon travail part de l'hypothèse alternative selon laquelle la dénotation est atomique (Fodor et Lepore, 1998) : 270). En particulier, j'admets que les verbes dénotent un événement ou un état de choses de manière indivisible. Il est important de noter que cette hypothèse n'exclut pas une décomposition de la dénotation au niveau ontologique, ni la référence individuelle à des participants à des événements ou états de choses. Cette hypothèse a un avantage immédiat : elle est compatible avec une théorie syntaxique simple et non éliminativiste de la sémantique lexicale, celle des rôles thématiques proposée dans (Chomsky, 1981, Reinhart, 2000).

Sur la base de mes travaux empiriques, je reviens, dans ce chapitre, sur cette question centrale en sémantique lexicale : est-il vraiment indispensable de décomposer la dénotation verbale au niveau lexical ? Dans un article récent (Haiden, à paraître), je passe en revue les arguments empiriques et conceptuels en faveur de la décomposition lexicale et argumente que ni la description linguistique, ni la trajectoire de l'acquisition du

langage, ni la perception des événements n'imposent de décomposer la dénotation verbale dans le lexique. J'adopte l'hypothèse selon laquelle ce sont les rôles thématiques qui doivent être décomposés. Je me fonde sur l'analyse des rôles thématiques en traits proposée par (Haiden, 2005a, Reinhart, 2000), et montre que ces traits ont des fondements cognitifs. En m'appuyant sur des recherches récentes sur le développement cognitif des jeunes enfants, j'argumente que les traits thématiques dénotent les concepts centraux de la théorie modulaire de la perception des événements (Hirschfeld et Gelman, 1994, Leslie, 1994, Leslie et al., 2004), à savoir la causalité et l'intentionnalité. En particulier, et ce point constitue l'un des aspects importants et novateurs de mon travail, je mets en évidence le rôle joué par l'intentionnalité dans la perception des événements de changement d'état. En effet, les arguments affectés par un changement d'état sont connus en linguistique comme étant des exemples prototypiques d'arguments internes, et cela dans des cadres théoriques très variés (Borer, 1994, 1998, Reinhart, 2000). On considère en outre que les changements d'état sont des exemples prototypiques d'événements causaux (Levin et Rappaport Hovav, 2005). Or, jusqu'à très récemment, les travaux menés dans le domaine de la psychologie n'ont trouvé aucune preuve que ce type d'événement est perçu comme causal (voir (Michotte, 1945) sur les "changements qualitatifs", et les travaux qui ont suivis). Seul ((Muentener, 2009, Muentener et Carey, 2010)) a mis en évidence, dans une série d'expériences avec des nourrissons, que les changements d'état peuvent être perçus comme causaux, une condition nécessaire étant que l'événement soit déclenché par un agent intentionnel. J'applique ces résultats à la définition des rôles thématiques.

Ce premier chapitre est organisé comme suit. La section 1.2 introduit les rôles thématiques et les problèmes principaux qu'ils posent. La section 1.3 présente brièvement la stratégie couramment adoptée face à ces problèmes, à savoir la décomposition lexicale. J'argumente que cette approche ne tient aucune de ses promesses concernant la cognition, l'acquisition du langage et la description linguistique de la structure argumentale. La section 1.4 introduit une théorie modulaire de la perception des événements, et en tire des critères pour l'identification des rôles thématiques. La section 1.5 applique ces critères aux traits thématiques du *Theta system*.

1.2 Les rôles thématiques

(Davidson, 1967) propose que l'action décrite par un verbe d'action soit représentée sous la forme d'un terme singulier dissocié du sens du verbe. Le sens du verbe est conçu comme une description, comparable aux descriptions de la dénotation nominale de Russell (Russell, 1905). Par exemple, pour une proposition comme (1a), Davidson propose la forme logique en (1b) ((18) dans l'original). Dans cette représentation, la description verbale est affirmée d'une variable spatio-temporelle x qui se trouve dans la portée d'un quantifieur existentiel.

- (1) a. I flew my spaceship to the Morningstar.
"J'ai piloté mon vaisseau spatial jusqu'à l'étoile du matin."
b. $\exists x$ (flew (I, my spaceship, x) & to (the Morning Star, x))

Selon Davidson, les arguments reçoivent leurs rôles du prédicat verbal. A la suite d'une remarque due à Castañeda (voir (Castañeda, 1967)), (Parsons, 1990), chapitre 5.7) propose de dissocier les rôles argumentaux du sens du verbe. L'exemple de Parsons (ibid : 94) est reproduit en (2a), et sa forme logique en (2b). Cette représentation, une conjonction de prédicats, permet de prédire de façon triviale l'observation que (2a) implique toutes les propositions en (3).

- (2) a. Brutus stabbed Cesar with a knife.
"Brutus poignarda César avec un couteau."
b. $(\exists e)$ (stabbing (e) & Agent (e , Brutus) & Theme (e , Cesar) & with (e , knife))
- (3) a. Brutus stabbed Cesar.
"Brutus poignarda César."
b. Brutus stabbed.
"Brutus effectua l'action de poignarder."
c. Cesar was stabbed.
"César fut poignardé."

Dans le cadre de la théorie néo-davidsonienne de la structure argumentale, on utilise généralement, à la suite de Parsons, des prédicats de rôles tels que *Agent* et *Theme* en

(2b). Cette théorie encode les observations de Davidson sur la singularité de la dénotation verbale (un verbe décrit un événement, et un seul) et les implications des structures prédicat-argument. Par ailleurs, elle se prête bien au transfert d'information du lexique à la forme logique via la syntaxe (tel que le requiert le modèle génératif) : les items lexicaux verbaux sont insérés en syntaxe munis d'une grille thématique (*theta-grid* ; (Chomsky, 1981)) sans que la syntaxe ait besoin de traiter de l'information strictement sémantique.

L'approche néo-davidsonienne est confrontée à divers problèmes empiriques dont les plus importants sont (i) le caractère arbitraire de toute liste de rôles et (ii) l'opacité des prédicats de rôles, qui rend difficile la formulation de généralisations transversales, ainsi que l'identification de classes naturelles de rôles (cf. (Levin et Rappaport Hovav, 2005) pour une liste plus exhaustive de problèmes).

Concernant le premier de ces problèmes, (Levin et Rappaport Hovav, 2005) observent qu'il est difficile d'identifier un ensemble unique et universel de rôles et elles argumentent qu'une décomposition des rôles en traits ne permettrait pas de résoudre ce problème. Elles citent ainsi (Ostler, 1980) qui propose 8 traits, générant 256 rôles possibles, dont seuls 48 sont utilisés. (Levin et Rappaport Hovav, 2005) argumentent qu'il est donc nécessaire d'avoir une théorie contraignant la combinaison des traits, mais elles doutent de la possibilité de construire une telle théorie : "*Considering the care that Ostler takes in motivating this set of features and the wide range of phenomena he attempts to deal with, it is doubtful that other attempts at defining a comprehensive feature system will meet with much more success.*" (Levin et Rappaport Hovav, 2005) : 46 ("Etant donné le soin avec lequel Ostler motive cet ensemble de traits et la variété des phénomènes dont il tente de rendre compte, il est peu probable que d'autres tentatives de définition d'un système complet de traits rencontrent beaucoup plus de succès.", traduction MH).

(Dowty, 1991) discute en détail le second problème. Selon Dowty, une classification fine des rôles thématiques réussirait bien à décrire des distinctions fines, mais elle manquerait les généralisations larges, tandis qu'une classification large ne permettrait pas de capter des distinctions fines (p. 554). Son dilemme peut être illustré par le rôle "cause immédiate" de (Levin et Rappaport Hovav, 1995). La cause immédiate est un rôle très large incluant les agents, les causes et les émetteurs, mais néanmoins pertinent en tant que tel pour la réalisation grammaticale des rôles : la cause immédiate doit être réalisée comme argument externe. Elle est donc un bon candidat au statut de

prédicat de rôles. Cependant, il serait alors impossible de formuler des distinctions plus fines entre agents, causes et émetteurs, distinctions qui sont également pertinentes en grammaire. Par exemple, les émetteurs permettent l'inversion locative (4b), tandis que les agents ne la permettent pas (4a) (exemples tirés de (Levin et Rappaport Hovav, 1995) : 255). Si l'on utilise une liste de rôles comme celle de Parsons, il est impossible d'exprimer ces deux généralisations sans contradiction ou redondance.

- (4) a. *In the factory worked young boys at the age of 7.
dans l'usine travaillaient jeunes garçons à l'âge de 7 ans
- b. And in their wake rumbled trucks to haul off the remains.
"Et à leur suite, des camions sont passés bruyamment pour dégager les restes"

On peut envisager deux directions pour résoudre ces problèmes.

La première, celle poursuivie ici, rejette la position de (Levin et Rappaport Hovav, 2005) selon laquelle une analyse des rôles en traits est vouée à l'échec. Dans le *Theta system*, les rôles thématiques sont modélisés sous forme de deux traits binaires formant des faisceaux (*clusters*). Ces traits pouvant être sous-spécifiés, le système génère un total de neuf faisceaux (4 faisceaux entièrement spécifiés, 4 faisceaux sous-spécifiés et un faisceau par défaut, dont les deux traits sont sous-spécifiés). Cet ensemble est suffisamment restreint pour être universellement attesté. Par ailleurs, il est possible de faire référence à chacun des deux traits séparément, ce qui permet de formuler des généralisations transversales.

La majorité des travaux en sémantique lexicale adopte une approche radicalement différente : les entrées lexicales des verbes sont représentées sous forme de structures arborescentes projetées par un ensemble réduit de prédicats de base. Dans ce cadre, les rôles thématiques sont déterminés par la position des arguments dans la structure lexicale.

Je développerai plusieurs arguments contre cette seconde approche. L'hypothèse de la décomposition lexicale ne peut pas être motivée par référence à la réalisation syntaxique des rôles. En effet, la structure argumentale syntaxique requiert deux échelles orthogonales de proéminence lexico-sémantique, à savoir la causalité et l'intentionnalité (section 1.3.3). Or, deux échelles orthogonales ne se laissent pas facilement représenter

sous la forme d'une arborescence globale. L'avantage de la décomposition lexicale devrait donc être ailleurs, hors du domaine syntaxique propre. Je discuterai et rejetterai deux avantages potentiels de la décomposition lexicale : l'hypothèse d'une structure conceptuelle unique (section 1.3.1) et l'hypothèse de la conformité aux processus d'acquisition du langage (section 1.3.2). (Jackendoff, 1983) propose qu'une structure conceptuelle unique imposerait sa forme au lexique. Or, diverses expériences sur la perception et la mémorisation des événements montrent que les représentations cognitives et linguistiques sont parfois dissociées. Cette hypothèse n'est donc pas tenable. Ensuite, l'hypothèse selon laquelle les structures lexicales faciliteraient l'acquisition du langage en induisant la structure syntaxique (*semantic bootstrapping*, (Pinker, 1984)) ne tient pas non plus. Il a en effet été montré, dans des études en acquisition (voir (Fisher et al., 2010) pour une revue récente de la littérature), que les connaissances syntaxiques *précèdent* la sémantique verbale. En fait, les configurations syntaxiques sont utilisées par les apprenants pour induire la sémantique verbale (*syntactic bootstrapping*).

J'en conclus que l'hypothèse des structures arborescentes lexicales n'aide ni à la compréhension de la sémantique lexicale, ni à son acquisition, ni à la compréhension de sa place en cognition. Je rejetterai donc cette hypothèse. Par conséquent, les rôles thématiques restent le meilleur outil pour représenter l'information lexicale concernant la structure argumentale.

1.3 Contre la décomposition lexicale des prédicats

A la suite de Jackendoff et de Talmy (Jackendoff, 1972, 1983, Talmy, 1975, 1976, 1983), il est souvent admis que la dénotation des verbes correspond à des sections d'une structure conceptuelle universelle.

- (5) The Conceptual Structure Hypothesis (Jackendoff, 1983 : 17)
 "There is a *single* level of mental representation, *conceptual structure*, at which linguistic, sensory, and motor information are compatible."

L'hypothèse de la structure conceptuelle (HSC)

"Il existe un niveau de représentation mentale unique, la *structure conceptuelle*, auquel les informations linguistiques, sensorielles et motrices sont compatibles."
(trad. MH)

L'hypothèse de Jackendoff est intéressante, car elle propose une théorie explicite de l'interface entre les composants majeurs de l'esprit. Il s'agit là d'un avantage considérable, qui dépasse des considérations de portée plus réduite, par exemple, purement linguistiques – à la condition cependant que l'hypothèse s'avère correcte...

Selon l'hypothèse de la structure conceptuelle (HSC), la représentation lexicale d'un verbe est une structure syntaxique comportant des positions indexées réservées aux arguments. Pour un verbe comme *to butter* 'beurrer', Jackendoff propose la représentation en (6c) ci-dessous. Lorsque nous affirmons que Nicky a beurré une tartine, nous ne désignons pas une seule et unique action (ce que dirait Davidson), mais nous décrivons une séquence d'événements (en principe indépendants) : une action "déclencheur", un mouvement directionnel et un état de localisation (cf. la paraphrase de (6a) en (6b)).

- (6) a. Nicky buttered the toast. (Jackendoff 1983 : 184-5)
 "Nicky a beurré la tartine."
- b. Nicky put butter on the toast.
 "Nicky a mis du beurre sur la tartine."
- c. *butter* 'beurrer' :
 [Event CAUSE ([Thing x], [Event GO ([Thing BUTTER], [Path TO ([Place ON
 ([Thing y])])])])]]]

Les implications de l'HSC sont très fortes. Dans la préface de son livre (p. x), Jackendoff affirme que "la conséquence la plus fondamentale de l'approche adoptée ici est peut-être que les notions standard de vérité et référence ne jouent aucun rôle significatif dans la sémantique des langues naturelles" ("*perhaps the most fundamental consequence of the approach adopted here is that the standard notions of truth and reference play no significant role in natural language semantics.*", trad. MH). Cependant, cela ne me

semble pas être la conséquence la plus radicale de l'HSC. Celle-ci est encore plus surprenante en elle-même : il existerait un seul et unique niveau de représentation rassemblant de l'information linguistique, motrice et perceptive. Il s'avère que cette hypothèse est problématique à la fois du point de vue cognitif et du point de vue linguistique.

Tout d'abord, l'HSC prédit que les asymétries de représentation des événements entre les langues seraient corrélées à des asymétries cognitives. Or, des données expérimentales (Gennari et al., 2002, Papafragou et al., 2008) montrent que ceci n'est vrai que dans les tâches mobilisant des capacités linguistiques. Il est donc nécessaire de pouvoir dissocier, au moins en partie, la capacité linguistique et la capacité cognitive. Ensuite, si elle existe, la structure conceptuelle unique devrait être accessible lors de l'acquisition du lexique. Les dénотations verbales devraient donc être stables à un stade précoce de l'acquisition du langage, et la syntaxe devrait être initialisée par la sémantique (*semantic bootstrapping*). Or, la recherche en acquisition montre le contraire : les paramètres syntaxiques sont fixés dès l'âge auquel on peut les observer (Wexler, 1998) tandis que les dénотations verbales ne se stabilisent que tardivement (Fisher et al., 2010). Je développe successivement ces deux points dans la suite.

1.3.1 Données expérimentales montrant que les composants linguistiques et cognitifs sont dissociés

L'HSC est une hypothèse empirique, qui fait des prédictions. Entre autres, elle prédit que la représentation linguistique des événements reflète la structure conceptuelle sous-jacente. Si les langues se distinguent systématiquement en ce qui concerne la représentation des événements, alors ces différences devraient révéler des faits importants sur l'organisation des mécanismes cognitifs en général. La représentation du déplacement directionnel est un exemple bien documenté (Aske, 1989, Jackendoff, 1990, Talmy, 1975), qui permet de tester les prédictions de l'HSC.

Dans certaines langues, dites "langues à allure" (*manner languages*, par exemple, l'anglais, l'allemand, le russe et le chinois mandarin), les verbes ont tendance à décrire l'allure du déplacement (*marcher* vs *courir* vs *sauter*), le but du mouvement étant indiqué par des PP directionnels et des particules. Dans d'autres langues, les "langues à chemin" (*path languages*, par exemple l'hébreu, le japonais, le grec, le turc et les langues

romanes), la directionnalité du mouvement est généralement indiquée par le verbe (*entrer* vs *sortir*). Les exemples suivants (tirés de (Papafragou et al., 2008), leurs ex. (1-3)) illustrent cette observation typologique. L'exemple anglais (7) exprime un déplacement achevé à l'aide d'un verbe d'allure (*to fly* 'voler') et d'un PP directionnel (*to its nest* 'vers son nid'). L'exemple grec correspondant (8) ne peut pas avoir cette interprétation. Le PP *sti folia tu* 'vers son nid' ne peut avoir qu'une interprétation de localisation, pas de but. Si un locuteur du grec veut décrire un mouvement achevé, il doit utiliser un verbe directionnel (*bike* 'est entré', 9a). Or, dans ce contexte, l'ajout de modifieurs de manière (entre parenthèses) réduit l'acceptabilité de la proposition. La façon correcte d'exprimer la manière est celle en (9a) sans le modifieur d'allure. La généralisation typologique est donc que les "langues à allure" expriment à la fois l'allure et la directionnalité du mouvement, tandis que les "langues à chemin" ont tendance à n'exprimer que le second de ces éléments.

(7) anglais :

The bird	flew	to its nest.
"L'oiseau	a volé	vers son nid"

(8) grec moderne :

To puli	petakse	sti folia tu.
"L'oiseau	a volé	dans son nid"

(9) grec moderne :

(a)	To puli	bike	sti folia tu	(petontas).
	"L'oiseau	est entré	dans son nid	(en volant)"
(b)	To puli	petakse	ke bike	sti folia tu.
	"L'oiseau	a volé	et est entré	dans son nid"

L'existence même d'une variation typologique de ce type est problématique pour l'HSC : en effet, une structure conceptuelle *unique* devrait s'imposer de manière universelle. On peut envisager deux façons d'adapter l'HSC afin de rendre compte des faits : (i) admettre de la variation entre les langues en cognition, (ii) admettre un minimum d'indépendance

entre cognition et langage. Appelons la première de ces options *l'option relativiste*, et la seconde *l'option modulariste*.

En tout état de cause, une partie de la structure conceptuelle doit être acquise. (Jackendoff, 1990, 1996) argumente que les "langues à chemin" instancient la configuration sous-jacente.

Si nous adoptons *l'option relativiste* (i), nous prédisons que les locuteurs des "langues à chemin" sont moins sensibles à l'information concernant l'allure du mouvement en général, car ce type d'information est moins saillant dans leurs représentations cognitives. Si nous adoptons *l'option modulariste* (ii), nous nous attendons à ce que les locuteurs des "langues à allure" soient moins sensibles à l'information concernant l'allure du mouvement dans les tâches non linguistiques, car ce type d'information ne fait pas partie de la structure conceptuelle universelle.

Afin d'évaluer le lien entre cognition et langage, (Gennari et al., 2002) ont fait l'expérience suivante. Ils ont présenté des vidéos représentant des déplacements directionnels (par exemple, un homme tire une planche vers l'intérieur d'une pièce) à des locuteurs de l'anglais et de l'espagnol, en distinguant trois conditions expérimentales, que je détaille immédiatement dans la suite : (i) dénomination en premier (*naming first*), (ii) mémorisation libre (*free encoding*) et (iii) ombre (*shadow*). Dans la condition "dénomination en premier", les sujets étaient priés de décrire la scène observée par une seule phrase. Cette condition favorise la construction d'une représentation linguistique de l'événement observé. Dans la condition de "mémorisation libre", les participants regardaient les scènes en silence, en sachant qu'ils devaient les mémoriser pour une partie ultérieure de l'expérience. Enfin, dans la condition "ombre", les participants devaient répéter des syllabes dénuées de sens pendant qu'ils regardaient les scènes. Il s'agissait par là d'empêcher la construction de représentations linguistiques.

Après avoir vu les scènes, les participants devaient effectuer des tests de reconnaissance et de jugements de similitude. Durant ces tests, ils étaient confrontés à nouveau aux scènes originales, et, en outre, à deux types de variantes : l'une concernant l'allure du mouvement (par exemple "entrer dans la pièce en marchant" vs "en galopant") et l'autre la directionnalité du mouvement (par exemple "entrer" vs "sortir de la pièce"). Dans le test de reconnaissance, les sujets devaient dire s'ils avaient déjà vu la scène. Dans le test de jugement de similitude, ils devaient regarder les trois scènes (l'original et ses 2 variantes), puis déterminer laquelle des 2 variantes était la plus proche de

l'original. Après ces tests, les sujets des conditions (ii) et (iii) ont été à nouveau confrontés aux scènes originales et ils ont dû les décrire en une phrase.

Gennari et al ont mis au jour des différences entre les langues dans le test de jugement de similitude. Les locuteurs de l'espagnol considèrent plus souvent que les locuteurs anglophones que les scènes ayant la même allure (mais pas le même chemin) sont similaires. Ce résultat est conforme à ce que l'on attend : il confirme la différence typologique entre l'anglais et l'espagnol mentionnée précédemment. Cependant, ce contraste n'existe que pour la condition linguistique (*naming first*). Dans les conditions non linguistiques, on n'obtient pas de différence entre les locuteurs des deux langues. Les auteurs en concluent que les propriétés des langues n'exercent un effet sur l'activation des représentations cognitives que si des ressources linguistiques sont mobilisées dans la création de ces représentations. Si une représentation cognitive est créée sans que le langage intervienne, alors le type de langue parlée par le sujet n'a aucune influence sur les propriétés de cette représentation cognitive.

Ce résultat est confirmé par une étude de suivi du regard menée par (Papafragou et al., 2008). Ces auteurs ont testé des locuteurs de l'anglais et du grec. Afin d'identifier clairement des zones d'intérêt, ils se sont servis d'un mouvement impliquant un instrument (ski, planche à voile, etc.), soit dirigé vers un objet cible (mouvement directionnel), soit sans objet cible (mouvement non directionnel). Dans leur expérience, la dernière image de la vidéo reste visible sur l'écran quelques secondes après que l'événement soit achevé. Tout comme Gennari et al, ce n'est que dans la condition linguistique que Papafragou et al ont mis en évidence un effet de langue. Pendant le déroulement de la scène, le centre d'intérêt du regard (l'allure du déplacement ou l'objet cible) varie en fonction de la langue maternelle du sujet : les locuteurs grecs ont tendance à regarder d'abord l'objet cible, puis l'instrument du déplacement, tandis que les anglophones font l'inverse. Dans la tâche non linguistique, tous les sujets explorent de la même façon les scènes en déroulement.

Papafragou et al mettent au jour que la langue maternelle a une influence sur la cognition dans les phénomènes de mémoire aussi. Après l'arrêt de la vidéo en effet, les anglophones se concentrent sur l'objet cible et négligent l'allure du mouvement, tandis que les locuteurs du grec regardent tout autant la cible que l'allure du mouvement. Les auteurs attribuent ces différences à l'intervention de ressources linguistiques lors de la mémorisation des scènes.

Ces deux expériences montrent que, même si elles interagissent à plusieurs niveaux, les représentations cognitives et linguistiques doivent parfois être dissociées. Or, une structure conceptuelle unique est incapable de modéliser des composants indépendants. La perception des événements ne peut donc pas motiver une structure cognitive unique, qui s'imposerait de manière automatique au lexique. Dans la section suivante, j'argumente que l'acquisition du lexique ne fournit pas non plus d'argument en faveur de structures lexicales.

1.3.2 Acquisition et induction syntaxique

Si le sens des entrées lexicales est généré par des structures universelles, on s'attend à ce que ces structures soient activées lors de l'acquisition du langage. En particulier, on prédit que les enfants apprennent le sens et le cadre syntaxique des verbes à l'aide de la structure lexicale (*semantic bootstrapping*, (Pinker, 1984)). Ainsi, on prédit que les enfants établissent des généralisations syntaxiques au fur et à mesure de leur développement lexical, durant une longue période de leur développement (Tomasello, 1992). En fait, diverses études empiriques, sur lesquelles je reviens immédiatement ci-dessous, établissent le contraire : les paramètres syntaxiques sont fixés correctement dès l'âge auquel ils peuvent être observés (Wexler, 1998, 2002), et la sémantique verbale reste flexible pendant une longue période (pour certaines classes de verbes, elle le reste jusqu'à l'adolescence, voire l'âge adulte). En outre, ce sont les configurations syntaxiques (en particulier le nombre de DP dans une phrase), et non une structure conceptuelle sous-jacente, qui initialisent la sémantique verbale (*syntactic bootstrapping*, (Gleitman, 1990, Landau et Gleitman, 1985, Naigles, 1990)). Les enfants sont très tôt capables de déduire une interprétation sémantique à partir de la distribution des DP, et ce, sans la moindre indication concernant le type de l'événement (pour une revue récente de la littérature, voir (Fisher et al., 2010)). Ce phénomène est en accord avec l'hypothèse selon laquelle la seule structure hiérarchique pertinente pour un verbe est la syntaxe. Si une structure conceptuelle différente de la syntaxe existe, alors elle n'est pas pertinente pour l'acquisition du langage.

A partir des années 1980, divers travaux portant sur l'acquisition des verbes ont montré que, dans certaines conditions, les enfants sont prêts à accepter et à interpréter des verbes employés dans des phrases agrammaticales ((Gleitman, 1990, Landau et

Gleitman, 1985, Naigles, 1988) ; pour un aperçu récent, cf. (Fisher et al., 2010)). Par exemple, ils acceptent la phrase **The elephant comes the giraffe* ("L'éléphant arrive la girafe") et interprètent *l'éléphant* comme agent et *la girafe* comme thème. Pour un locuteur adulte, cette interprétation correspondrait à la phrase *The elephant brings the giraffe* ("L'éléphant apporte la girafe"). En d'autres termes, les enfants possèdent des connaissances sur le sens typique des constructions transitives et intransitives et ils appliquent ces connaissances lorsqu'ils rencontrent de nouveaux verbes, ou bien des verbes connus, employés dans un nouveau contexte. Ce comportement est dit "conforme au cadre" (*frame compliant*).

Les adultes ne se comportent pas de la même manière. Lorsqu'ils sont confrontés à une phrase agrammaticale comme **The elephant comes the giraffe* ("L'éléphant arrive la girafe"), ils "réparent" soit la syntaxe (par exemple en répétant *The elephant comes to the giraffe* "L'éléphant arrive vers la girafe"), soit le sens (par exemple en interprétant que l'éléphant vient seul). Ceci indique que les adultes connaissent la distribution syntaxique des verbes et qu'ils refusent qu'un verbe soit employé en dehors de son contexte habituel. Ce comportement est dit "conforme au verbe" (*verb compliant*).

Le comportement des enfants devient graduellement "conforme au verbe", en fonction du type de la construction, des classes de verbes, voire des items lexicaux individuels. (Naigles et al., 1992) ont étudié 5 groupes de locuteurs natifs de l'anglais, âgés de 5 à 13 ans. Les sujets devaient mimer (*act out*) des phrases impliquant 4 constructions différentes : transitive + PP, transitive simple, intransitive + PP, intransitive simple. Pour chacune de ces constructions, on leur donnait des phrases grammaticales et agrammaticales. Les auteurs ont constaté que le taux de conformité au verbe augmente avec l'âge pour toutes les constructions, sauf la plus complexe (transitive + PP). Notons que même les adultes présentent un degré élevé de conformité au cadre pour cette construction (80% environ). En fait, chaque construction se développe à un rythme différent. Pour la construction intransitive simple, la conformité au verbe est atteinte à l'âge de 5 ans ; pour la construction transitive simple, elle est atteinte à l'âge de 10 ans ; pour la construction intransitive + PP, elle ne l'est qu'à l'âge de 12 ans. Ceci montre que le passage de la conformité au cadre à la conformité au verbe n'est pas un processus général du développement cognitif, mais un processus spécifique au développement linguistique, qui varie en fonction des structures grammaticales, voire des items lexicaux individuels.

A la suite de ces travaux, divers auteurs ont établi que la structure syntaxique est l'information la plus importante pour l'acquisition de la sémantique verbale, et cela même dans des langues où, contrairement à l'anglais, le nombre de syntagmes nominaux exprimés dans une phrase n'est pas un indice fiable de la structure argumentale. Je mentionne deux exemples : le mandarin et le kannada (langue dravidienne, parlée dans le sud-ouest de l'Inde). En mandarin, les arguments exprimant de l'information connue peuvent être omis. Le nombre de syntagmes nominaux exprimés dans une phrase n'est donc pas un indice fiable de la transitivité du verbe. Cependant, (Lee et Naigles, 2008) ont montré que, à l'âge de 2-3 ans, les enfants apprenant le mandarin utilisent le nombre de syntagmes nominaux pour interpréter des verbes inconnus. En kannada, la morphologie est un indice bien plus fiable de la structure argumentale que le nombre d'arguments. Par exemple, il existe un marqueur de causatif unique et stable. Si ce morphème est affixé à une base verbale, l'interprétation du verbe ne peut être que causative. Néanmoins, (Lidz et al., 2003) établissent que, lorsqu'ils apprennent le sens de nouveaux verbes, les enfants d'âge moyen 3;6 ainsi que les adultes se fient au nombre de syntagmes nominaux, et ignorent la morphologie.

Le rôle de la structure syntaxique dans l'acquisition de la sémantique verbale est donc clair dans des langues typologiquement différentes. De plus, ce rôle est indépendant des propriétés individuelles des événements décrits par le verbe. Ce résultat a été établi dans une expérience de regard préférentiel (*preferential looking*) impliquant des enfants anglophones de 28 mois (Yuan et Fisher, 2009). Les enfants ont été confrontés à des vidéos représentant des personnages en train de dialoguer. Dans ces dialogues, les personnages employaient des verbes inventés dans des constructions soit transitives, soit intransitives. Ni les vidéos, ni les dialogues ne donnaient d'indice concernant le type d'événement décrit par les verbes fictifs. Le seul indice était donc la distribution syntaxique du verbe, indiquée par le nombre de DP dans la phrase. Lors de la phase de test, les enfants ont regardé des paires de vidéos représentant des scènes impliquant un et deux participants respectivement. Leur tâche consistait à identifier la scène illustrant l'événement représenté par le verbe fictif, énoncé sans arguments (par exemple *Find blinking* ! - "Trouve *blinking* !"). Les auteurs ont montré que la distribution syntaxique entendue par les enfants dans les dialogues et la direction de leur regard étaient corrélés de façon significative. Si le verbe avait été employé dans des constructions transitives, ils préféraient les scènes à deux participants ; si le verbe avait été employé dans des constructions intransitives, ils préféraient les scènes à un seul

participant. (Yuan et Fisher, 2009) concluent que les enfants sont capables d'utiliser de l'information abstraite, en particulier le nombre de DP dans la phrase, lorsqu'ils construisent les représentations lexicales de nouveaux verbes.

En résumé, les données de l'acquisition montrent (i) que les paramètres clés de la syntaxe sont acquis précocement, tandis que la sémantique verbale continue à se développer pendant une longue période et (ii) que l'acquisition de la sémantique verbale est initialisée par la structure syntaxique. L'acquisition de la sémantique verbale dépend bien de connaissances structurales, mais le lieu de ces connaissances est la *syntaxe* de la phrase. Les faits tirés de la recherche en acquisition ne justifient donc pas l'existence de structures lexicales arborescentes.

1.3.3 verser, remplir, contenir ; connaissances, perceptions et peur : les paradoxes de la structure lexicale

Les structures arborescentes de la sémantique lexicale sont censées décrire des propriétés lexicales pertinentes pour la projection syntaxiques des items lexicaux. Ces structures devraient donc être capables de prédire la réalisation syntaxique des rôles thématiques. Mais, ce n'est pas le cas.

Tout d'abord, nous constatons que certains prédicats de base dénotant des états ont des réalisations syntaxiques contradictoires. Ce problème peut être résolu si l'on autorise les structures lexicales à être réanalysées en fonction d'une échelle de force (Croft, 1991, 1998, Talmy, 1988, 2000) Cependant, l'échelle de force ne peut analyser correctement les prédicats psychologiques (Haiden, 2005a). Si le lexique doit prédire la structure argumentale syntaxique, alors il doit encoder (au moins)¹ deux échelles orthogonales (force/causalité et implication psychologique), et non une seule échelle. Or, les structures arborescentes admises en sémantique lexicale ne sont pas adaptées à la représentation de deux échelles orthogonales. Les structures arborescentes ne sont donc pas adaptées à la modélisation de la sémantique verbale.

Prenons comme point de départ un exemple tiré de (Gropen et al., 1991, Pinker, 1989) : quelqu'un effectue l'action de transférer de l'eau dans un verre. Si l'on veut décrire cette action, deux verbes viennent immédiatement à l'esprit : *verser* ('to pour' en

¹ Il est probable que deux échelles ne sont pas suffisantes. Les propriétés aspectuelles, ignorées dans ce chapitre, ont en effet également une influence sur la projection argumentale : voir (Borer, 2005, Verkuyl, 1993) pour une discussion et des références.

anglais) et *remplir* ('to fill' en anglais). En (10), ces deux verbes décrivent le même événement, cependant, leur structure argumentale est différente. Le verbe *to pour* 'verser' en (10a) prend comme objet un liquide (*water* 'eau'), le verbe *to fill* 'remplir' en (10b) un contenant (*the glass* 'le verre').

- (10) a. He poured water into the glass.
 "Il a versé de l'eau dans le verre."
- b. He filled the glass with water.
 "Il a rempli d'eau le verre."

Pinker, Gropen et al. argumentent que les deux verbes *to pour* 'verser' et *to fill* 'remplir' ont des structures argumentales différentes car ils ont des propriétés sémantiques différentes : ces verbes décrivent l'action selon deux perspectives différentes. Si l'on veut présenter le liquide comme un objet affecté, on utilise *to pour* 'verser', si c'est le contenant que l'on veut présenter comme objet affecté, on utilise *to fill* 'remplir'.

On peut représenter cette différence en décomposant le sens des verbes en question. La partie de leur représentation sémantique décrivant l'activité serait différente (*agent manipule liquide* vs *agent manipule contenant*), tandis que celle décrivant l'état résultant serait identique (*liquide est localisé dans contenant*). Les structures en (11) tentent de représenter fidèlement les structures informelles de (Gropen et al., 1991) en utilisant les prédicats de base CAUSE, DO, LOC. Le prédicat de base CAUSE représente une relation de causalité entre deux propositions, le premier argument étant la cause, le second la conséquence. Le prédicat de base DO représente les actions, avec comme premier argument l'agent et comme second argument l'objet affecté par l'action. Le prédicat LOC représente la localisation d'un objet (premier argument) dans un lieu (second argument).²

- (11) a. *pour* 'verser' : CAUSE (DO(z,x), ... LOC(x,y))
 b. *fill* 'remplir' : CAUSE (DO(z,y), ... LOC(x,y))

² Je laisse ici de côté la possibilité d'avoir des prédicats aspectuels portant sur le second argument de CAUSE tels que *BECOME*.

(11a) signifie donc qu'un agent *z* agit sur *x*, de façon à ce que *x* soit localisé dans *y* ; (11b) signifie que l'agent *z* agit sur *y*, de façon à ce que l'objet *x* soit localisé dans *y*. Si l'on admet avec Jackendoff (1990) que le prédicat d'action (la ligne des actions, *action tier*) a la priorité dans l'application des structures lexicales à la syntaxe, on peut considérer que (11) explique le comportement syntaxique des deux verbes en question.

La théorie de Gropen et al est intéressante, car elle s'étend naturellement à un autre emploi du verbe *fill* 'remplir', celui en (12a). En (12a), le verbe *fill* 'remplir' décrit une localisation. Ce verbe peut être comparé au verbe *contain* 'contenir' (en 11c). Les deux verbes décrivent le même état, mais ont une structure argumentale différente : le liquide est sujet avec *fill* 'remplir' en (12a), mais objet avec *contain* 'contenir' (en 12c). Le contenant est objet avec *fill* 'remplir', mais sujet avec *contain* 'contenir'.

- (12) a. Water fills the glass.
"De l'eau remplit le verre."
- b. *fill* 'remplir' : LOC (*x*,*y*)
- c. The glass contains water.
"Le verre contient de l'eau."
- d. *contain* 'contenir' : LOC (*x*,*y*)

Selon les conventions précédentes (le premier argument de LOC est un objet localisé, le second est un lieu), (12b) décrit bien le fait que de l'eau est dans le verre. (12d), en revanche, pose problème : par parité des formes, (12c) devrait être interprétée comme une affirmation du fait qu'un verre est dans l'eau. Cette interprétation est évidemment absurde.

Afin de résoudre ce problème, on pourrait affaiblir l'hypothèse de la décomposition lexico-sémantique à la manière de (Croft, 1991, 1998, Talmy, 1988, 2000) ; voir aussi (Jackendoff, 1990). Ces auteurs proposent que les verbes décrivent des relations asymétriques de "force" entre les arguments, où la notion de "force" correspond à une représentation naïve de l'énergie mécanique (voir la section 1.4.1 ci-dessous). Chaque prédicat prend une liste ordonnée d'arguments, et la position de chaque argument dans cette liste représente sa position dans une échelle de force. Dans

ce cadre, la réalisation syntaxique n'est pas déterminée par les relations hiérarchiques de la structure lexicale, mais par l'échelle de force. Cette hypothèse est tout à fait compatible avec la décomposition lexicale des prédicats, mais elle la rend redondante : les généralisations portant sur les rôles sémantiques et leur réalisation syntaxique sont désormais énoncées à l'aide de l'échelle de force. La structure lexico-sémantique devrait donc être motivée par des données psychologiques ou tirées de l'acquisition. Encore une fois, nous sommes renvoyés aux problèmes détaillés précédemment.

Ceci noté, voyons maintenant de quelle façon l'approche en termes d'échelle de force peut expliquer le comportement des verbes *verser*, *remplir* et *contenir*. Je commencerai par les emplois statifs, i.e. ceux en (12). Le verbe *contenir* décrit la localisation du liquide, et dans cette description, l'attention est portée sur le contenant. Celui-ci est présenté comme un support enfermant le liquide, il est donc situé plus haut sur l'échelle de force, et, de ce fait, est réalisé comme sujet de la phrase (12c). Pour le verbe *remplir*, la perspective est inversée. Ce verbe dirige l'attention sur le liquide, présenté comme la source d'une propriété du contenant. Le verbe *remplir* liste donc le liquide plus haut que le contenant sur l'échelle de force. Ainsi, c'est le liquide qui devient sujet de la phrase (12a). Examinons maintenant les emplois agentifs (10). Pour *remplir*, la position relative entre le liquide et le contenant reste telle qu'elle était dans l'emploi statif. L'agent est ajouté plus haut dans l'échelle de force. L'élément le plus bas (le contenant) est réalisé comme complément direct, et l'élément le plus haut (l'agent) est réalisé comme sujet (10b). Le verbe *verser* peut être considéré comme la contrepartie agentive du verbe d'état *contenir*, avec l'agent situé en haut, le liquide en bas et le contenant au milieu de l'échelle de force. Dans cette configuration, l'agent est réalisé comme sujet et le liquide comme complément direct (10a).

(Croft, 1991, 1998) propose que les événements non causaux soient conceptualisés à l'aide de l'échelle de force. Cette échelle serait notamment en jeu dans la conceptualisation et la réalisation des prédicats psychologiques. Ces prédicats présentent en fait le même paradoxe structural que les prédicats statifs de localisation. Prenons comme exemple les verbes allemands *ärgern* 'énervé' et *erklären* 'expliquer' en (13-14).

- (13) a. Er erklärt mir die Lösung.
 Il explique moi_{DAT} la_{ACC} solution
 "Il m'explique la solution."

- b. Er ärgert mich (mit seiner Arroganz).
 Il énerve moi_{ACC} (avec son arrogance)
 "Il m'énerve (par son arrogance)."

A première vue, cette paire peut être traitée de la même manière que la paire de Gropen et al. Le verbe *erklären* 'expliquer' affirme qu'un agent manipule des idées (qui sont à l'accusatif ; et qui sont représentées par y en (14)) pour que l'expérimenteur (au datif ; représenté par x en (14)) les comprenne. La représentation (14), adaptée de (Rapp, 1997) : 62f, rend compte de cette description. (Rapp, 1997) emploie le prédicat de base PSYCH, qui affirme qu'un individu (le premier argument) est dans un état mental (le second argument). Le verbe *ärgern* 'énervé' décrit l'expérimenteur comme étant affecté par une propriété de l'agent : l'agent fait une action qui énerve l'expérimenteur. Par application du raisonnement de Gropen et al., la représentation de ce verbe est celle donnée en (15).³

(14) *erklären* 'expliquer': CAUSE (DO(z, y), ... PSYCH(x,y))

(15) *ärgern* 'énervé': CAUSE (DO(z, x), ... PSYCH(x,y))

Dans les verbes décrivant des états psychologiques, on s'attend, par définition, à ce que le prédicat de base PSYCH surgisse seul. Prenons par exemple le verbe *kennen* 'savoir/connaitre'. Dans les constructions impliquant ce verbe, l'expérimenteur est réalisé sujet et la représentation mentale complément direct (16a). Ceci correspond à ce que (Rapp, 1997) : 58-60, ainsi que tous les auteurs utilisant le prédicat de base PSYCH, prédisent. Cependant, la configuration inverse, où la représentation mentale est réalisée en position de sujet et l'expérimenteur en position d'objet, est également bien attestée. C'est le cas par exemple de la version stative du verbe *ärgern* 'énervé' en (16b).

³ (Rapp, 1997) ne discute pas le cas du verbe *ärgern* 'énervé', mais celui de verbes comparables tels que *enttäuschen* 'décevoir'. De tels verbes sont appelés *Wirkungsverben* (verbes de causation indirecte). Rapp propose (p. 75-78) que ces verbes aient la représentation lexicale suivante : CAUS (y, ... , PSYCH (x,y)). Dans cette représentation, le premier argument de CAUS est un individu, et non une proposition. Cette option peut probablement être exploitée pour l'implémentation des effets de la hiérarchie de force dans le cadre théorique des prédicats de base. Cependant, les détails d'une telle théorie ne sont pas pertinents ici, puisque cette théorie ne propose pas de solution au problème plus sérieux discuté dans le texte, à savoir le fait que l'intentionnalité est un facteur indépendant déterminant la prééminence syntaxique.

- (16) a. Sie kennt die Lösung.
Elle connaît la solution
- b. Seine Arroganz ärgert mich.
Son arrogance énerve moi
"Son arrogance m'énervé."

La décomposition des prédicats verbaux en prédicats de base ne dérive donc pas, sans hypothèses auxiliaires, les classes naturelles pertinentes pour la réalisation syntaxique. Ceci constitue un problème grave pour cette approche, mais ce n'est pas encore le problème le plus grave.

Tentons de résoudre le problème soulevé en (16) par l'échelle de force. Nous devrions alors paraphraser *kennen* 'savoir/connaître' par une version très littéraire de *comprendre*, qui signifierait que le sujet exerce de la force envers le contenu de la connaissance. Avec le verbe statif *ärgern* 'énervé', la configuration devrait être inversée : le stimulus s'impose à l'expérimenteur, entraînant la configuration sujet-complément désirée.

Cette analyse implique des paraphrases très métaphoriques et nous allons voir immédiatement qu'elle échoue, même si l'on est prêt à accepter de telles paraphrases. Le vrai problème n'est pas la nature métaphorique de l'échelle de force, mais le fait qu'elle n'est pas adaptée à la description des relations de prééminence dans le domaine des attitudes mentales (voir la notion d'anomalisme du mental chez (Davidson, 1970) et la discussion dans la section 1.4). Prenons l'exemple du verbe anglais *fear* 'craindre'. La crainte est rarement désirée ou volontairement produite par l'expérimenteur : elle s'impose d'elle-même à l'expérimenteur (si ce n'est pas le cas, c'est une *attitude*, et non une *crainte*). Pourtant, c'est bien l'expérimenteur qui demeure sujet, et le contenu de l'émotion (munie de force) doit être réalisé en position d'objet :

- (17) *John fears the devil.*
John craint le diable

De nombreux verbes, qui lient le physique et le mental, suivent ce modèle, en particulier dans le domaine des perceptions. Les verbes de perception ont diverses variantes : *watching* et *looking at* 'regarder' sont des activités, mais *hearing* 'entendre' et *feeling*

‘ressentir’ le sont moins ; si quelqu'un surprend une conversation secrète malgré lui (*to overhear a secret conversation*), il ne le fait certainement pas exprès. Il en est de même de *feeling pain* ‘ressentir une douleur’ ou *suffering a heart attack* ‘faire une crise cardiaque’ : de telles expressions décrivent des événements subis. On devrait alors s'attendre à ce que l'expérienceur soit situé plus bas que l'autre argument sur l'échelle de force. L'expérienceur devrait donc être réalisé comme objet direct. Ce n'est pas le cas.

On pourrait envisager de sauver l'échelle de force en analysant ces verbes problématiques comme des inaccusatifs à deux places. Cependant, cette option ne fonctionne pas pour deux raisons principales (cf. chapitre 2 de (Haiden, 2005a) pour une discussion et des données supplémentaires) : (i) une analyse en termes d'inaccusatifs à 2 places n'est pas applicable de manière universelle et, même si elle l'était, (ii) certaines configurations de liage par un quantifieur sont contradictoires avec l'échelle de force.

Premier point : prenons l'exemple du verbe allemand *erleiden* ‘subir, supporter’ (18a). Ce verbe se comporte systématiquement comme un verbe transitif normal au regard des tests d'inaccusativité standards tels que l'interprétation du participe et le liage par un quantifieur. Quand le participe de *erleiden* ‘subir/supporter’ est employé comme adjectif, le N modifié doit être l'argument interne du participe (18b), qui correspond à l'accusatif de la proposition finie (18a). L'accusatif est donc bien un argument interne.

- (18) a. Der Patient erlitt einen Herzinfarkt.
le patient subit une attaque-cardiaque
- b. der erlittene Herzinfarkt
la subi attaque-cardiaque
- c. *der erlittene Patient
le subi patient (visé : "le patient affecté")

Second point : le test du liage par un quantifieur en (19) montre que l'expérienceur (au nominatif) c-commande l'accusatif (19a). L'inverse est impossible, même si l'accusatif, en position de topique, précède le nominatif (19b). L'expérienceur doit être structurellement plus haut que l'autre argument. Ceci va à l'encontre de l'hypothèse de

l'échelle de force, même si les verbes en question sont analysés comme des inaccusatifs à deux places.

- (19) a. Jeder_i Trader erlitt sein_i Waterloo.
chaque_{NOM} trader subit son_{ACC} Waterloo
"Chaque trader a subi sa propre défaite personnelle."
- b. *Jedes_i Waterloo erlitt sein_i Trader.
chaque_{ACC} Waterloo subit son_{NOM} trader
visé : "Pour chaque défaite, il existe un trader qui l'a subie."

J'en déduis que, pour certains verbes (ceux qui réalisent l'expérienceur dans une position syntaxique supérieure à celle occupée par la perception/l'expérience qui s'impose à l'expérienceur), l'échelle de force ne peut décrire correctement les schémas de réalisation des rôles. Ces verbes suggèrent qu'il existe une seconde échelle, indépendante de l'échelle de force, qui hiérarchise les arguments en fonction de leur implication mentale dans l'événement ((Haiden, 2005a)). La pertinence de deux échelles de proéminence est difficilement compatible avec le schéma arborescent adopté dans le cadre de la décomposition lexicale des événements : les arbres représentent en effet une hiérarchie unique, globale.

Pour conclure cette section, rappelons le problème qui nous avait conduits à rejeter en première analyse les prédicats de rôles thématiques, et qui avait motivé la décomposition lexicale du sens des verbes en structures arborescentes projetées par des prédicats de base. Le problème était que les rôles atomiques ne permettent pas de formuler des classes naturelles de rôles et qu'ils ne sont pas adaptés à la formulation de généralisations sur leur réalisation en syntaxe. Clairement, la décomposition lexicale des événements a les mêmes problèmes. Mais il y a plus grave : l'existence de non pas une, mais deux échelles de proéminence (causation/force et intentionalité/implication mentale) est incompatible avec des structures lexicales arborescentes. La décomposition lexicale des événements ne pourrait ainsi être motivée que par des faits externes. Malheureusement, cette tentative a également échoué. Tout d'abord, l'HSC (qui imposerait une structure conceptuelle hypothétique au lexique) a dû être rejetée. Ensuite, nous avons vu que l'acquisition du sens des verbes ne requiert aucune

connaissance structurale autre que celle apportée par la syntaxe. J'en conclus que la décomposition lexicale du sens des verbes en structures arborescentes projetées par des prédicats de base ne tient pas ses promesses. Je n'adopte donc pas cette approche. Les rôles thématiques sont de nouveau au centre du jeu.

1.4 Décomposer les rôles et non les événements : les fondements cognitifs

Supposons donc, contre (Levin et Rappaport Hovav, 2005) : 46), qu'une liste de rôles thématiques analysés en traits peut être adéquate empiriquement. (Reinhart, 2000) propose deux traits, *c* et *m*, indiquant la causation et l'implication mentale respectivement. Comme nous l'avons vu dans la section précédente, ces traits correspondent aux propriétés des descriptions d'événement pertinentes pour la structure syntaxique. J'adopterai cette proposition et mon objectif dans cette section sera d'identifier des indices perceptuels indépendamment motivés en cognition qui permettent de définir les rôles thématiques et les traits qui les composent. Me fondant sur l'hypothèse modulariste de l'esprit ((Fodor, 1983)) et son application par (Hirschfeld et Gelman, 1994)) aux processus centraux (à savoir à la cognition proprement dite), j'explore les deux mécanismes cognitifs qui gèrent les propriétés pertinentes des traits thématiques : le *Theory of Body Mechanism* (ToBy) et le *Theory of Mind Mechanism* (ToMM) ((Leslie, 1994)).

1.4.1 La perception de la causalité (ToBy)

Dès 1945, Albert (Michotte, 1945)) a proposé, sur la base de certaines propriétés des événements de déplacement, une théorie détaillée de la perception de la causalité. L'ouvrage de Michotte est important, car il est le premier à proposer une théorie modulaire de la perception de la causalité.

Notons tout d'abord que la causation est une construction mentale qui met en relation différents événements. L'adjacence temporelle et/ou la covariance statistique ne permettent pas à elles seules d'établir un lien causal entre deux événements. Ainsi, comme (Saxe et Carey, 2006)) nous le rappellent, même si le jour et la nuit se succèdent

inexorablement, nous ne voudrions pas dire qu'ils entretiennent une relation de causalité.

Dans son étude des événements de lancement et d'entraînement, Michotte postule l'existence d'un analyseur d'input perceptuel spécifiquement dédié au traitement des impressions causales. Selon Michotte, si deux événements de déplacement satisfont certaines conditions, cet analyseur produit automatiquement et obligatoirement une interprétation causale. En particulier, "il a montré que le lancement est perçu quand et seulement quand les paramètres des deux mouvements sont cohérents avec ceux d'un seul mouvement transféré d'un premier objet vers un second objet ayant des caractéristiques perceptuelles différentes" (trad. MH de ((Saxe et Carey, 2006) : 146) : *"he showed that launching is perceived when and only when the two motions have parameters consistent with a single motion transferred from one object to a second, perceptually distinct one"*). En d'autres termes, la perception d'un seul événement de déplacement unique équivaut à une interprétation causale. Dès qu'un objet "cause" n'entre pas en contact avec l'objet qu'il est supposé mettre en mouvement, et plus généralement, dès qu'il existe des indices suggérant que l'événement implique deux déplacements indépendants, aucun lien de cause à effet ne sera établi entre ces déplacements.

De nombreuses études sur les interprétations causales des nourrissons et des enfants se fondent sur ce résultat. En général, ces expériences comportent une première partie dite d'*habituation*, au cours de laquelle les sujets sont habitués à un événement causal partiellement masqué. Cette première partie est généralement construite selon le protocole classique de (Ball, 1973)), illustré ci-dessous en figure 1. Les sujets doivent regarder un objet B partiellement masqué par un écran (le cercle le plus à droite dans la première illustration ci-dessous). Un objet A (le cercle de gauche) fait son apparition, se déplace vers B, puis disparaît derrière l'écran. Au moment où A aurait dû entrer en contact avec B, B commence à se déplacer selon la trajectoire attendue sur la base du mouvement de A (seconde illustration ci-dessous). Lorsque les sujets sont habitués, *i.e.* lorsque leur intérêt pour l'événement diminue, ils effectuent le test. Lors du test, ils sont confrontés aux mêmes événements, mais cette fois-ci sans écran. Le test comporte deux conditions, illustrées sur la seconde ligne de la figure 1 : dans la première condition (condition avec contact), l'objet A touche effectivement B et B commence à se déplacer tout seul. Dans la seconde condition (condition sans contact), A s'arrête avant de toucher B et B commence à se déplacer tout seul.

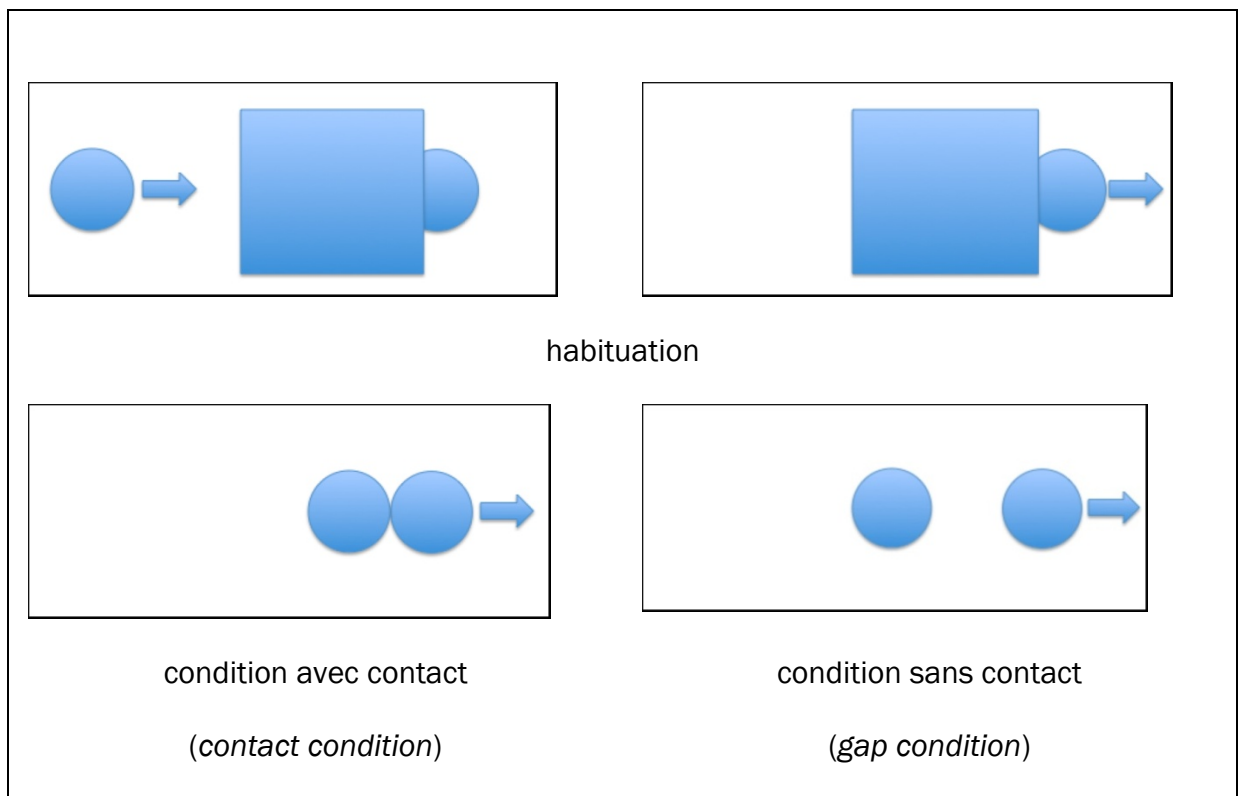


Figure 1. Protocole de Ball (1973)

Considérons à la suite de Michotte que le contact est une condition nécessaire à la perception de la causalité. Faisons en outre l'hypothèse que, en l'absence de source externe de propulsion, on s'attend par défaut à ce que les objets stationnaires restent stationnaires. Alors, on jugera qu'un sujet accorde une interprétation causale aux événements masqués s'il manifeste de la surprise dans la condition sans contact. En utilisant des variantes de ce paradigme expérimental, ((Spelke et al., 1995)) montrent que les enfants attribuent une interprétation causale aux événements de lancement masqués dès l'âge de 6 mois (la surprise étant mesurée par une augmentation du temps de regard).

(i-iv) résumant nos hypothèses jusqu'à présent :

i.) Domaine :

Le domaine principal de la perception de la causalité est celui du déplacement dans l'espace.

ii.) Inertie :

En l'absence de source de propulsion, les objets stationnaires restent stationnaires et les objets en déplacement ne changent pas de trajectoire.

iii.) Unicité de l'événement :

Une interprétation causale ne peut émerger que si les mouvements de l'objet lanceur et de l'objet lancé peuvent être perçus comme un déplacement unique, continu.

iv.) Contact :

Le contact est une condition nécessaire à la perception de la causalité.

Michotte a étudié les objets inanimés, et il a montré que ni la taille, ni le poids ou la solidité perçus, ni encore la vitesse des objets impliqués n'ont une influence qualitative sur l'émergence de la perception d'une causalité. Cependant, des travaux plus récents ont établi que la présence d'indices d'intentionnalité ainsi que certaines propriétés dispositionnelles des objets inanimés peuvent modifier la perception de la causalité.

(Leslie et Keeble, 1987)) ont ainsi montré que les prédictions sur le comportement d'un objet dans des interactions causales nouvelles varient selon que cet objet a été catégorisé comme cause ou bien comme affecté. Leur expérience était construite de la façon suivante. Ils ont habitué deux groupes de nourrissons de 6 mois à des événements de lancement entièrement visibles. Dans le premier groupe, un objet A touche un objet B, qui est alors mis en mouvement (*direct launching*). Dans le second groupe, l'objet B ne se met en mouvement qu'après un laps de temps (*delayed launching*). Dans la phase de test, l'ordre temporel des événements est inversé : l'objet B touche A, et A commence à se déplacer. En d'autres termes, les rôles sont inversés : l'objet connu comme étant la cause devient l'objet affecté, et l'objet connu comme étant affecté devient cause. Leslie & Keele ont montré que cet inversement de rôles ne surprend que les nourrissons ayant été habitués au lancement direct (*direct launching*). Ceci confirme que la réaction immédiate est nécessaire à la perception causale. Par ailleurs, cette expérience montre que les nourrissons distinguent les rôles agentifs et réceptifs dans les interactions causales. Une fois qu'ils ont catégorisé des participants à l'événement comme agents et récepteurs, ils sont surpris de voir ces rôles inversés dans le même type d'événement.

Les expériences menées par (Kotovskiy et Baillargeon, 2000)) confirment ces résultats. Dans une expérience impliquant des nourrissons de 6 mois, ils montrent que l'interprétation causale d'événements de lancement cachés dépend crucialement de la catégorisation préalable de l'objet lancé. Si cet objet est classifié comme inerte, les enfants s'attendent à ce que l'objet lancé se déplace après contact et demeure inerte en l'absence de contact. En revanche, si des indices suggèrent que l'objet lancé peut se déplacer par lui-même, cette attente n'existe pas.

Des études ultérieures montrent que des groupes de nourrissons âgés de 7 et 10 mois sont capables d'attribuer des rôles dispositionnels stables aux objets. Lorsqu'ils sont confrontés au mouvement spontané d'un objet qu'ils avaient précédemment catégorisé comme inerte, ils infèrent la présence d'agents cachés, et préfèrent ne pas modifier les rôles qu'ils avaient précédemment établis ((Saxe et al., 2007)).

Nous pouvons donc ajouter la généralisation suivante à nos observations précédentes :

v.) Rôle dispositionnel :

Soient deux objets préalablement classifiés comme actif et inerte respectivement. L'objet actif est prédit garder cette propriété dans les événements à venir, et l'objet inerte est prédit rester inerte dans les événements à venir.

(Leslie, 1984)) a montré que certains indicateurs d'intentionnalité influencent eux aussi la production d'interprétations causales. Dans son étude importante des événements de ramassage par une main (*manual pick-up event*),⁴ il utilise des vidéos montrant soit une main, soit un objet inanimé ramassant une poupée. Dans la condition de contact (*picking up*), l'agent (i.e. la main ou l'objet inanimé) ramasse la poupée. Dans la condition à distance (*reaching and following*), l'agent s'approche de la poupée, mais ne la touche pas, puis se retire, et la poupée le suit, à distance constante. Après habituation, les conditions sont inversées. Les nourrissons qui avaient été habitués à la condition à distance sont confrontés aux événements de contact et vice versa. La population de Leslie (d'âge moyen = 28 semaines, i.e., environ 6 mois et demi) effectue une distinction claire entre les mains humaines et les objets inanimés dans l'interprétation causale des événements. Si l'agent est une main, tous les nourrissons, i.e. aussi bien ceux qui avaient

⁴ I.e., d'événement d'entraînement dans la terminologie de Michotte.

été habitués à la condition de contact que ceux qui avaient été habitués à la condition à distance, sont sensibles à l'inversion de la condition (contact/à distance). Si l'agent est un objet inanimé, ils ne manifestent aucune surprise face à l'inversion de la condition, et ce, dans les deux cas : ils s'attendent à ce que la poupée suive l'objet inanimé, avec ou sans contact.

Ce n'est que très récemment que les propriétés dispositionnelles des participants à un événement ont été mises en relation avec la perception des événements de changement d'état ((Muentener, 2009, Muentener et Carey, 2010)). Le changement d'état est un cas intéressant car (Michotte, 1945)) n'avait trouvé aucune preuve que ce type d'événement déclenche une perception causale (cf. chapitres XV-XVI). Or, en sémantique lexicale, le changement d'état est presque toujours considéré comme étant l'exemple prototypique des événements causaux. Ceci est tout particulièrement vrai du *Theta System* de (Reinhart, 2000) : la causalité y est en effet représentée par le trait *c* dénotant le concept "cause d'un changement" (*cause change*).

((Muentener, 2009, Muentener et Carey, 2010)) montrent qu'un changement d'état peut effectivement déclencher une perception causale, à la condition que l'agent présente des indices d'intentionnalité. Muentener a étudié des nourrissons de 8 mois, en suivant une variante du paradigme de Ball (cf. figure 1 plus haut). Lors de l'habituation, une boîte est partiellement cachée derrière un écran. Un objet fait son apparition, s'approche de la boîte et disparaît derrière l'écran. Au moment où l'on attend l'impact de l'objet sur la boîte, cette dernière subit un changement d'état : soit elle s'effondre, soit elle change de couleur, soit elle commence à émettre de la musique. Lors de la phase de test, les nourrissons voient les événements sans l'écran occultant. Dans la condition à distance, le changement d'état a lieu sans qu'il y ait contact entre l'objet (agent) et la boîte. Dans la condition de contact, le changement d'état a lieu suite à l'impact. Muentener a mis en évidence que, lorsque l'objet qui va percuter la boîte est un petit train, les nourrissons ne font pas de différence entre les 2 conditions (contact et à distance). Ceci confirme l'observation de Michotte : les événements de changement d'état ne déclenchent pas la perception d'une causalité. Cependant, et ceci va à l'encontre de ce que Michotte proposait, Muentener a mis en évidence que, lorsque l'objet qui va percuter la boîte est une main (humaine), il existe une différence entre les 2 conditions. Dans ce cas, les nourrissons sont surpris par les événements à distance. Ceci suggère que les événements de changement d'état ne peuvent être classifiés comme causaux que si le changement est déclenché par un agent intentionnel : "A la différence

des expériences antérieures, qui ont montré que les représentations des traits dispositionnels des agents ont un effet sur *la façon* dont les événements de mouvement causaux sont construits, cette expérience suggère que les représentations qu'ont les nourrissons des traits dispositionnels de l'agent peuvent déterminer *si* un événement est perçu comme causal ou non." (soulignement et trad. MH, "*Unlike previous experiments, which have demonstrated the effects of representations of the dispositional features of agents on how a causal motion event is construed, this experiment suggests that infants' representations of the dispositional features of the agent influenced whether or not the event was represented as causal.*" (Muentener et Carey, 2010) : 14)

vi.) Changement d'état :

Les changements d'état ne sont interprétés de manière causale que si le déclencheur du changement est un agent intentionnel.

1.4.2 Action rationnelle (ToMM)

A peu près à l'époque où Michotte travaillait sur la perception de la causalité, (Heider et Simmel, 1944)) étudiaient la production d'explications anthropomorphiques et intentionnelles pour des événements sans aucun indice d'implication d'un être animé. Leurs expériences utilisent des figures géométriques : un carré stationnaire et des triangles et cercles en mouvement. Dans certaines conditions, un angle du carré s'ouvre, comme s'il s'agissait de la porte d'une maison. Les cercles et les triangles se déplacent de façon autonome, sans impulsion externe visible. Dans ce type de mouvement, les sujets de (Heider et Simmel, 1944)) ont identifié les cercles et les triangles comme étant des agents intentionnels, et ils ont expliqué leur comportement en employant des termes réservés à l'interaction humaine, par exemple les deux triangles "se battent", le grand triangle "enferme" le cercle dans une maison, etc. Cette expérience montre donc que les adultes peuvent identifier une action comme étant intentionnelle exclusivement sur la base d'indices comportementaux. L'étude de ces indices et des conditions de prédiction d'une action intentionnelle constitue aujourd'hui encore un domaine de recherche important.

Des travaux récents ont pu établir que les agents intentionnels sont identifiés sur la base d'indices figuratifs et comportementaux (pour une revue de la littérature, cf. (Biro

et Leslie, 2007)). L'œil humain est sans doute l'indice figuratif le plus clair ((Baron-Cohen, 1995)), mais les mains indiquent également une intentionalité (cf. la sous-section précédente). Les traits comportementaux indiquant une intentionalité sont, entre autres, les suivants : le mouvement autonome, auto-initié ((Baron-Cohen, 1994, Premack, 1990)), la réaction à distance ((Leslie, 1994, Rochat et al., 1997, Schlottman et Surian, 1999)), la poursuite d'un but, la satisfaction lorsque le but est atteint ((Montgomery et Montgomery, 1999)), enfin, la capacité à adapter son action et à la rendre efficace pour la poursuite d'un objectif ((Gergely et al., 1995)).

vii.) Indices d'agentivité :

- a. indices figuratifs : les yeux, la main
- b. indices comportementaux : déplacement autonome, réaction à distance, poursuite d'un but, satisfaction, efficacité

On sait, depuis (Heider et Simmel, 1944)), que les adultes préfèrent interpréter et expliquer les actions intentionnelles à l'aide de concepts d'états mentaux tels que les croyances et les désirs : A a fait B, car il croit que B va lui donner C, qu'il désire. Cependant, les concepts d'états mentaux ne sont accessibles aux enfants qu'à partir d'un âge assez tardif : le raisonnement sur les croyances et les désirs peut être observé à partir de l'âge de 4-5 ans seulement. C'est à cet âge en effet que les enfants réussissent le test classique des "fausses croyances" ((Wimmer et Perner, 1983)). Ce test est structuré de la façon suivante. L'agent A met un jouet dans une boîte, puis disparaît. L'agent B entre, prend le jouet dans la boîte, le met dans une autre boîte et disparaît. A entre à nouveau. L'enfant doit alors montrer la boîte dans laquelle il pense que A va chercher le jouet. Avant 4 ans, les enfants indiquent la boîte où se trouve le jouet, et non celle où A l'avait mis (et où il devrait le chercher).

viii.) Action intentionnelle :

Les adultes (et les enfants à partir de 4-5 ans) expliquent les actions intentionnelles à l'aide de concepts mentaux (croyances, désirs, etc.)

La raison pour laquelle les enfants de moins de 4-5 ans ne réussissent pas le test des fausses croyances est objet de débats. (Ruffman et al., 2001) par exemple ont argumenté que, si un enfant connaît la vérité, alors il la dira par défaut (*true belief*

default). Les enfants les plus jeunes ne peuvent inhiber cette réaction. (Leslie et al., 2004), par ailleurs, suggèrent que "si des tests non verbaux sont employés, les enfants, même de moins de 3 ans, résolvent des tâches de fausses croyances" (trad. MH, "*children even younger than 3 years solve false-belief tasks when non-verbal measures are used*" (Leslie et al., 2004) : 531). Il est néanmoins difficile de repérer chez les jeunes enfants le raisonnement sur les états mentaux d'autrui.

Le paradigme expérimental de (Gergely et al., 1995) s'avère ici pertinent puisqu'il permet d'étudier un trait central de l'intentionnalité, l'action efficace et rationnelle, et cela chez les nourrissons de 6 à 12 mois. Ce protocole est illustré ci-dessous en figure 2.

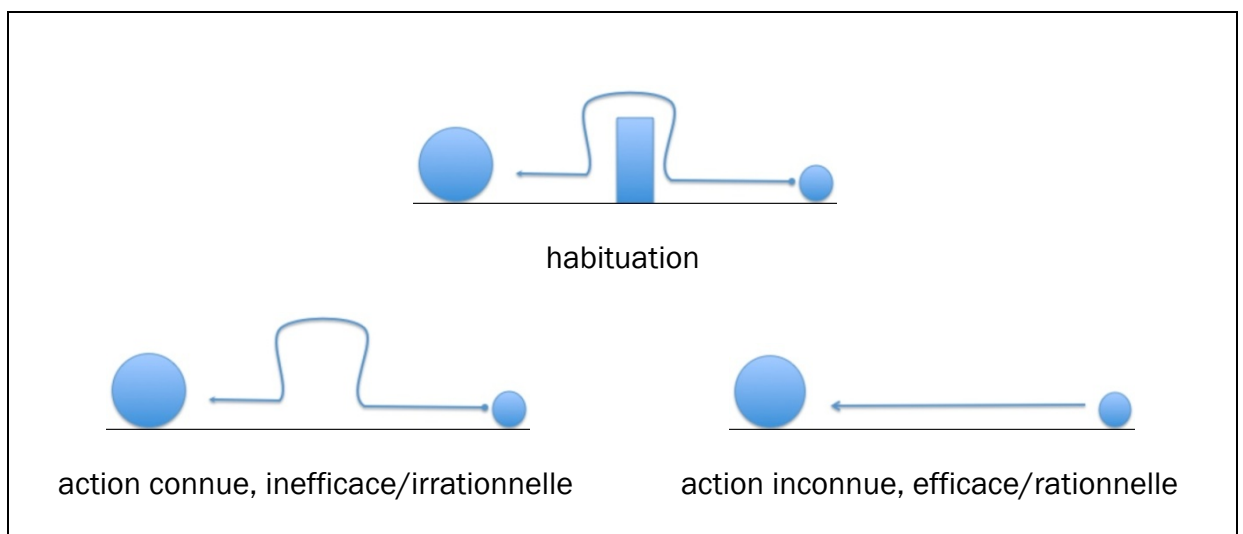


Figure 2. Protocole de Gergely et al. (1995)

Se fondant sur les travaux de (Gibson et al., 1978), qui montrent que la transformation autonome de la surface d'un objet est un indice d'agentivité, (Gergely et al., 1995) présentent aux enfants des balles qui se rétrécissent et s'élargissent en réponse l'une à l'autre. Ils élicitent ainsi le jugement que les balles sont animées. Dans la phase d'habitude, les deux balles sont séparées par une barrière, la petite balle saute par dessus la barrière pour rejoindre la grande, et les deux balles, "contentes", restent ensemble. Dans la phase de test, les enfants regardent le même événement, mais sans barrière. Un premier groupe d'enfants regarde exactement le même mouvement que celui observé lors de l'habitude : la petite balle fait un saut au cours de sa trajectoire, et fait donc un détour dans son déplacement vers la grande balle. Cet événement est familier, mais irrationnel dans le sens où la trajectoire pour arriver au but est plus longue que nécessaire. Un autre groupe d'enfants observe la petite balle s'approcher de la

grande selon un mouvement rectiligne. Ce mouvement est nouveau, mais rationnel. Gergely et al. observent que les enfants sont surpris à la vue de l'événement familier mais irrationnel. Ils en concluent que même les enfants très jeunes (âgés de 12 mois) sont capables de produire des explications intentionnelles des événements dirigés vers un but. ("*by the end of the first year infants are indeed capable of taking the intentional stance ((Dennett, 1987)) in interpreting the goal-directed behavior of rational agents.*" ((Gergely et al., 1995) : 184). Il est important de rappeler qu'à cet âge, la rationalité n'est pas attribuée à un agent ou à l'esprit d'autrui, mais à l'action elle-même ("*when interpreting behavior as goal-directed, "rationality is attributed as a property of the action, and not of the agent (or the agent's mind).*" (Gergely et Csibra, 1997) : 233, cf. aussi (Csibra, 2008)). Il est fait abstraction des concepts d'états mentaux et c'est ainsi que la tâche devient suffisamment facile pour les nourrissons.

ix.) Action rationnelle :

La rationalité [peut être] imputée à l'action et non à l'agent (ou à l'esprit de l'agent). (trad. MH, "*[R]ationality [can be] attributed as a property of the action, and not of the agent (or the agent's mind).*" ((Gergely et Csibra, 1997) : 233)

Des études ultérieures vont dans le même sens et établissent que les enfants de 6 à 8 mois peuvent interpréter les actions rationnelles, et qu'ils le font même pour des actions anatomiquement impossibles. (Southgate et al., 2008)) par exemple, adaptant le paradigme de (Gergely et al., 1995)), font intervenir un bras humain séparé d'un ballon par trois barrières. Dans la phase d'habituation, les barrières sont placées de façon à ce qu'il suffise de déplacer une seule barrière pour attraper le ballon. Dans la phase de test, les barrières sont placées de manière différente : cette fois-ci, le bras doit déplacer deux barrières pour pouvoir saisir le ballon de façon anatomiquement plausible. Cette condition est dite condition *inefficace*. Dans la condition *efficace*, le bras ne déplace aucune barrière, mais se tord comme un serpent, et parvient ainsi à attraper le ballon directement. (Southgate et al., 2008)) ont mis en évidence que les nourrissons de 6 à 8 mois sont surpris à la vue de l'action inefficace, anatomiquement plausible, mais pas à la vue de l'action efficace, qui est cependant anatomiquement impossible. Ils en concluent que les enfants ne font pas intervenir exclusivement leur expérience antérieure dans la perception des actions efficaces, dirigées vers un but. Les enfants utilisent un mécanisme cognitif consacré spécifiquement à l'interprétation des actions efficaces,

rationnelles. Ce mécanisme fournirait la base des interprétations mentalistes des actions dirigées vers un but chez les sujets plus âgés ((Gergely et Csibra, 1997) : 232).

1.5 Conclusion : le sens des traits thématiques

Pour conclure, j'examine dans cette section le contenu des traits thématiques et la façon dont ils encodent l'information fournie par les modules de la perception. Je présente des listes d'indices permettant d'identifier les traits thématiques.

On pourrait envisager que l'initialisation syntaxique (*syntactic bootstrapping*) permette d'identifier les traits de rôles. En fait, ce n'est pas le cas : l'initialisation syntaxique produit bien des asymétries d'interprétation grossières entre le rôle interne et le rôle externe (initialisant par là l'acquisition des verbes), mais ce mécanisme n'est d'aucune utilité pour l'identification de distinctions plus fines, en particulier de celles concernant les sous-classes du rôle interne et du rôle externe. En outre, les individus ne demeurent pas conformes au cadre (*frame compliant*) tout au long de leur vie. Le mécanisme d'initialisation doit donc être complété par un système de connaissances qui impose (graduellement) la conformité au verbe (*verb compliance*). Je propose que ce système de connaissances repose sur la perception des événements dénotés par les verbes en cours de développement.

En distinguant deux traits *c* et *m*, Reinhart rend compte du fait linguistique que la réalisation des arguments dépend à la fois de la proéminence causale et de la proéminence intentionnelle (cf. section 1.3.3 plus haut). Ce modèle est en outre compatible avec la théorie modulaire de la perception, selon laquelle deux domaines de perception des événements doivent être distingués, l'un étant causal/mécanique, l'autre intentionnel/rationnel (cf. section 1.4). Dans (Haiden, 2005a, 2007)), j'ai donc proposé de fonder les traits *c* et *m* dans chacun de ces deux systèmes cognitifs, celui responsable de la perception des événements causaux/mécaniques, et celui responsable de la perception des événements intentionnels/rationnels respectivement. Ceci nous permet de faire référence à des mécanismes indépendamment motivés pour définir les traits de rôles thématiques.⁵

⁵ On pourrait songer à utiliser des critères morphosyntaxiques pour définir les rôles thématiques. Cependant, dans le cadre du *Theta System*, une telle définition serait circulaire : le *Theta System* génère la structure argumentale syntaxique sur la base de l'information thématique. En outre, je doute qu'il soit

Dans la suite de ce chapitre, je propose des listes d'indices perceptuels permettant de reconnaître chacun des traits $[/+m]$, $[-m]$, $[-c]$ et $[/c]$. Le chiffre romain figurant entre parenthèses renvoie aux principes et observations de la section précédente. Précisons préalablement que les différents indices proposés ci-dessous ne constituent pas des conditions nécessaires à la postulation d'un trait et de sa valeur. En ce sens, cette liste est similaire à la liste des "*proto-entailments*" de (Dowty, 1991)). Si un argument présente systématiquement, pour un trait donné, des indices contradictoires, alors on en déduira que ce trait est sous-spécifié.

Commençons par le trait *m*. Un participant à l'événement est codé comme $[/+m]$ s'il présente des traits d'intentionnalité pertinents pour la description de l'événement dénoté par le verbe. Cela inclut les indices figuratifs d'intentionnalité bien connus donnés en (20b), mais également des indices comportementaux tels que l'autonomie (20a), la réaction à distance (20c), et l'action efficace orientée vers un but (20d-e). Un facteur supplémentaire, ignoré dans la littérature, intervient : le déclenchement d'un changement d'état (20f). Enfin, si un verbe attribue un état mental à un argument, il encode cette propriété par le trait $[/+m]$ (20g).

(20) Signaux de la présence de $[/+m]$ chez l'argument A du verbe V

- a. A se déplace de façon auto-initié, autonome, pertinent pour l'événement (vii)
- b. A présente des traits humains pertinents pour l'événement, par exemple, il utilise ses mains, des organes sensoriels (vii)
- c. A produit une réaction à distance pertinente pour l'événement (iv, vii)
- d. A poursuit de façon continue et efficace un objectif pertinent pour l'événement (vii-x)
- e. A manifeste de la satisfaction lorsqu'il a atteint un but pertinent pour l'événement (vii)
- f. A déclenche le changement d'état d'un autre participant pertinent pour l'événement (vi)
- g. A est dans un état mental pertinent pour l'événement (viii)

possible d'établir une définition des rôles thématiques empiriquement adéquate sur la base de la morphosyntaxe.

Le trait $[-m]$ peut désigner des contenus mentaux tels que les croyances et les désirs (21e), mais il existe d'autres cas, moins évidents. Si un objet situé à distance déclenche une réaction (21a), il est interprété comme $[-m]$. C'est le cas par exemple des sources et des cibles non atteintes (par exemple dans *aiming at X* 'viser à'). Il en est de même des objets poursuivis avec persistance et efficacité par un agent (21b) et des objets, qui, lorsqu'ils sont atteints, déclenchent chez un agent une satisfaction pertinente pour l'événement (21c). Enfin, ainsi que noté précédemment, les objets décrits comme étant affectés par un changement d'état causal sont $[-m]$ (21d).

(21) Signaux de la présence de $[-m]$ chez l'argument A du verbe V⁶

- a. A déclenche une réaction à distance chez un individu intentionnel (cibles non atteintes, sources, etc.) (iv, vii)
- b. A est poursuivi avec persistance et efficacité (vii, ix)
- c. Si A est atteint, cela déclenche une satisfaction chez un agent (vii, ix)
- d. A subit un changement d'état (vi, ix)
- e. A est le contenu d'une représentation mentale pertinente pour l'événement (viii)

Le trait *c* encode en grande partie les relations causales dans le domaine du mouvement, et je propose à ce sujet une adaptation conceptuelle importante de (Reinhart, 2000)). Dans le travail de Reinhart, *c* est un raccourci pour *cause d'un changement* (*cause change*). Or, nous avons observé que le changement d'état ne représente en fait qu'un cas marginal de la perception de la causalité, il dépend en fait d'une intentionnalité pertinente pour l'événement. Par conséquent, *c* signifie *cause d'un changement* si et seulement si le trait *m* est présent sous une forme ou sous une autre dans l'entrée lexicale du verbe en question (22d).

Si l'intentionnalité est encodée sur l'argument déclencheur ($[/+m]$), il en résulte un faisceau entièrement spécifié $[+c, +m]$. Ce faisceau décrit les agents volitionnels. Si l'intentionnalité est décrite de la perspective d'un objet affecté par une intention, ou bien

⁶ On interprète parfois $[-m]$ comme "ne fournit pas d'activité mentale pertinente" (Reinhart, 2003). Je ne vois pas comment une telle condition pourrait être remplie. Prenons par exemple le verbe anglais *to eat* 'manger' dans la phrase *A shark ate my friend Paul* "Un requin a mangé mon ami Paul." Dans cet événement, l'objet, *my friend Paul* 'mon ami Paul', produira sans aucun doute un état mental, causé par l'événement. Le verbe *to eat* 'manger' ne fait certainement pas référence à un tel état mental, mais je ne pense pas que son entrée lexicale doive l'exclure explicitement. En effet, si tel était le cas, la proposition *A shark ate my friend Paul* serait fausse si *Paul* devait ressentir de la douleur au cours de l'événement...

d'une cible d'une intention, il en résulte le faisceau [-c, -m]. Ce faisceau correspond typiquement au *thème*. C'est la spécification [/m] dans le faisceau du thème qui autorise le verbe à sélectionner un argument cause ([+c]) (22d) et à spécifier [-c] dans le thème lui-même (23c). On peut étendre ce raisonnement aux changements d'états dans le domaine mental (22e). Dans ce cas, l'intentionnalité nécessaire à la perception de la causalité dans les événements de changement d'état est encodée sur l'expérienceur, sous la forme du faisceau [-c, +m]. C'est la spécification [/+m] dans le faisceau de l'expérienceur qui autorise le verbe à sélectionner un argument cause ([+c]) (23e) et à spécifier [-c] dans l'expérienceur lui-même (23c).

En l'absence de contribution mentale pertinente, le trait c correspond par défaut à une interprétation de la causation à la Michotte, *i.e.* à un mouvement dans l'espace. Les objets décrits comme se déplaçant indépendamment, comme touchant, lançant ou poussant d'autres objets (22a-c) sont encodés sous la forme [/+c]. Les objets décrits comme étant atteints/touchés (*i.e.* les buts), poussés ou tirés par d'autres objets (23a, b) sont encodés sous la forme [-c].

(22) Signaux de la présence de [/+c] chez l'argument A du verbe V

- a. A se déplace indépendamment pertinent pour l'événement (i-ii, v)
- b. A entre en contact avec un autre participant à l'événement (iii-iv)
- c. A déclenche un mouvement chez un autre participant à l'événement (ii-iii, v)
- d. A déclenche un changement d'état chez un autre participant à l'événement et une intervention intentionnelle peut être envisagée (vi)
- e. A déclenche un changement d'état mental pertinent pour l'événement chez un autre participant à l'événement (vi-viii)

(23) Signaux de la présence de [-c] chez l'argument A du verbe V

- a. A est touché ou atteint par un autre participant à l'événement (i-v)
- b. Le mouvement de A est décrit comme dépendant de l'activité d'un autre participant à l'événement (ii-iii, v)
- c. A subi un changement d'état pertinent pour l'événement et une intervention intentionnelle peut être envisagée (vi)

Le lecteur pourra trouver dans (Haiden, 2005a)) une analyse et une discussion d'un large ensemble de verbes allemands employés dans diverses constructions, dans (Haiden,

2007)) une analyse des verbes d'émission de l'anglais et de l'allemand, et dans (Haiden, 2005a, 2006) une discussion explicite du lien entre modalité et rôles thématiques. Dans le cadre de ce chapitre, je voudrais simplement brièvement illustrer le raisonnement par le verbe anglais *to break* 'casser, briser' (24-25). Les rôles thématiques proposés par (Reinhart, 2003)) pour ce verbe figurent en (26).

(24) Max broke the window.

Max a-cassé la fenêtre

(25) The storm broke the window.

l'orage a-cassé la fenêtre

(26) *break* 'casser, briser' [+c] [-c, -m]

Le verbe *to break* 'casser, briser' décrit un changement d'état (une désintégration). En outre, il décrit ce changement d'état comme ayant une cause. En application de (vi), ce verbe devrait donc encoder l'intentionnalité sous une forme ou sous une autre. Cependant, le verbe *to break* 'casser, briser' peut tout à fait avoir une cause inanimée comme sujet (25). Dans ce cas, on ne peut pas considérer que le sujet encode l'intentionnalité. En fait, on a l'intuition que (25) n'attribue pas d'intentionnalité du tout. Ainsi, si les résultats présentés dans la section 1.4 ont quelque conséquence que ce soit sur l'interprétation linguistique des changements d'état, nous sommes obligés de conclure soit que *the storm* 'l'orage' en (25) n'est causalement pas pertinent pour l'état de la fenêtre, soit que, malgré les apparences, le verbe *to break* 'casser, briser' encode de l'intentionnalité. La première de ces options n'est pas tenable : elle ne rend pas compte du sens du verbe, qui décrit un changement d'état produit par une cause. La seconde de ces options consiste à poser que l'intentionnalité est présente de manière indirecte grâce au trait [/-m] sur le faisceau de l'argument interne. On peut alors se demander de quelle façon le trait [/-m] est interprété dans une proposition qui ne fait intervenir aucun agent intentionnel (comme c'est le cas en (25)). Je propose que l'état mental soit dans ce cas une propriété de l'observateur : l'observateur représente l'objet de la perspective d'un "changement qualitatif" (pour reprendre le terme de Michotte), auquel il attribue une structure causale. Des cas d'intentionnalité externe sont d'ailleurs attestés dans le contexte de la modalité déontique (Haiden, 2010).

2. Concaténation

Ce chapitre est consacré à la théorie linguistique et à sa base empirique. Je propose une théorie générale de la projection des unités lexicales et je l'applique à certains problèmes de morphologie non concaténative. J'argumente que les structures non concaténatives sont entièrement compositionnelles.

Ma recherche dans ce domaine s'organise autour des trois points suivants. Dans ce chapitre, je me concentrerai sur les points (a) et (c).

- (a) Développement d'un modèle de la grammaire capable de traiter des marqueurs non concaténatifs de façon compositionnelle : (Bendjaballah et Haiden, 2003, submitted, Haiden, 2005b, 2008a, 2009).
- (b) Application du modèle à la projection syntaxique des rôles thématiques : (Haiden, 2005a).
- (c) Application du modèle à la morphologie d'une langue non concaténative : (Bendjaballah et Haiden, 2005b, c, 2008, to appear).

Ce chapitre est constitué de deux grandes parties : la première est consacrée à la présentation du modèle théorique, la seconde à une étude de cas, l'analyse des prépositions faibles du kabyle de Chemini.

Les résultats sont fondés en grande partie sur une série de travaux menés en collaboration avec Sabrina Bendjaballah ((Bendjaballah et Haiden, 2003, 2005b, c, 2008, submitted, to appear)). Le lecteur souhaitant y faire référence est prié de citer également ces travaux.

2.1 Le mécanisme génératif : présentation du modèle

Dans les systèmes morphologiques non concaténatifs, les traits/catégories morphosyntaxiques ne sont pas toujours exprimés par des entrées lexicales discrètes, mais parfois par des *alternances* phonologiques. Ce type de morphologie est problématique pour les modèles dits *item-based* car l'identification et la segmentation des marqueurs n'est pas évidente. Je propose une théorie qui permet de réduire des structures apparemment non concaténatives, impliquant des exposants "exotiques", à des objets entièrement compositionnels. Cette théorie utilise (i) un algorithme

mathématique très simple pour dériver les opérations pertinentes du système computationnel (*concaténation* et *interprétation*), et (ii) les gabarits, définis comme des règles d'ordonnement d'unités temporelles, pour cibler les positions internes actives dans la dérivation, *i.e.* les marqueurs non concaténatifs.

Cette première partie est organisée de la façon suivante. La section 2.1.1 détaille le modèle, la section 2.1.2 présente les différents types d'exposants morphologiques dérivés dans ce cadre, la section 2.1.3 définit le rôle des gabarits à l'interface syntaxe-phonologie, et présente brièvement un exemple illustrant la façon dont le modèle traite la morphologie non concaténative.

2.1.1 Modèle

Toute théorie traitant des exposants morphologiques doit être explicite quant à ses hypothèses sur les deux composants pertinents, la morphosyntaxe et la phonologie. Suivant (Hauser et al., 2002), le modèle adopte l'hypothèse d'une composante de "syntaxe étroite" (*narrow syntax*), définie comme un module computationnel extrêmement contraint. L'output de la syntaxe étroite est relié aux interfaces pertinentes, à savoir l'interface sensori-motrice et l'interface conceptuo-intentionnelle, par d'autres modules.

Tout modèle des représentations linguistiques doit minimalement distinguer 3 types d'opérations. La première prend deux objets A et B et les assemble. Cette opération sera appelée *assemblage*. La seconde établit une relation hiérarchique entre A et B, ce sera l'opération *hiérarchie*. Enfin, la troisième assigne une interprétation à A et B, appelons-la *interface*. Il est évidemment souhaitable que ces opérations soient formalisées à l'aide d'un appareillage technique minimal.

(Chomsky, 1995) propose de rassembler *assemblage* et *hiérarchie* en une seule opération, *merge*, qui doit être distinguée d'une seconde opération que j'appellerai *interface* (et qui est parfois appelée TRANSFER). En conséquence, le pouvoir de la syntaxe étroite est considérablement réduit : la dérivation syntaxique ne crée pas d'information nouvelle, mais ne fait qu'arranger les items. Cette dernière condition est connue sous le terme de "condition d'uniformité" (*uniformity condition*). En revanche, la condition d'uniformité ne s'applique pas à la dérivation des représentations à l'interface.

Dans le modèle de (Chomsky, 1995), *merge* combine deux opérations (*assemblage* et *hiérarchie*), et cette caractéristique rend la structure de la syntaxe étroite finalement assez complexe. Je propose, dans le modèle développé ici, de regrouper différemment les 3 opérations *assemblage*, *hiérarchie* et *interface* : l'opération d'*assemblage* sera définie indépendamment, et ce seront *hiérarchie* et *interface* qui seront regroupées. Ceci nous permet de modéliser l'ensemble des traits formels muni de l'opération d'*assemblage* comme un monoïde, la structure algébrique la plus simple possible. En particulier, l'opération d'*assemblage* sera définie comme une loi de composition interne (*l.c.i.*), i.e. de telle façon que son résultat fasse partie de l'ensemble dans lequel les deux variables de départ prennent leurs valeurs. Dans ce modèle, la syntaxe étroite est donc encore plus "étroite" qu'elle ne l'est dans le cadre de Chomsky, ce qui est souhaitable, d'un point de vue minimaliste. Les opérations *hiérarchie* et *interface* seront regroupées dans une même opération, appelée *tête*. *tête* assigne automatiquement une interprétation aux éléments dépendants. L'opération *tête* n'est pas une *l.c.i.* : l'ensemble de départ est celui des traits formels et l'ensemble d'arrivée est celui des traits interprétables. L'architecture générale du modèle est schématisée et comparée à celle de (Chomsky, 1995) dans la figure 3 ci-dessous.

	<i>assemblage</i>	<i>hiérarchie</i>	<i>interface</i>
(Chomsky, 1995)	<i>merge</i> (syntaxe étroite)		<i>transfer</i>
ce chapitre	<i>assemblage</i> (syntaxe étroite)	<i>tête</i>	

Figure 3. Architecture générale

2.1.1.1. *Assemblage*

Supposons, à la suite de (Chomsky, 1995) et des travaux qui l'ont suivi, qu'il est nécessaire de distinguer deux types de traits syntaxiques : les *traits formels* et les *traits interprétables*. Les traits formels déterminent la dérivation dans le sens où ils entretiennent des relations de "vérification" (*checking*), de "recherche" (*search*), et d'"accord" (*agree*). Les traits interprétables reçoivent une interprétation à l'interface pertinente. Les traits formels sont effacés au cours de la dérivation, et une représentation d'interface n'est bien formée que si elle ne comporte que des traits interprétables. Supposons en outre que l'ensemble des traits formels inclue au moins les

étiquettes catégorielles N, V, T, etc. Informellement, on peut considérer que les traits formels servent de points d'ancrage aux traits interprétables que sont les descriptions d'événement, les rôles thématiques, les opérateurs, et les variables argumentales.

Dans le domaine de la phonologie, il est établi, depuis le développement du modèle autosegmental ((Goldsmith, 1979, Leben, 1978) et les travaux qui ont suivi), que les représentations doivent distinguer deux niveaux : le niveau squelettal et le niveau segmental. Le premier de ces niveaux héberge les points d'ancrage squelettaux, et le second, les traits segmentaux (segments, éléments, matrices de traits etc selon le cadre choisi). Je considérerai que les positions squelettiques sont les *traits formels* et les segments les *traits interprétables* de la composante phonologique. Les traits interprétables sont associés aux traits formels selon les principes de la phonologie autosegmentale (un à un, de gauche à droite). Plus précisément, le cadre adopté sera la phonologie du gouvernement ((Kaye et al., 1985)) et le modèle "CV" ((Lowenstamm, 1996) et les travaux qui ont suivi). Dans ce modèle, le niveau squelettal consiste en une stricte alternance de positions C (consonantiques) et V (vocaliques), i.e. d'unités CV. L'unité CV est l'unité minimale de temps phonologique. Ainsi, les unités CV seront les traits formels de la composante phonologique, et l'ensemble des traits interprétables comporte les tons et les éléments. On considèrera qu'une dérivation phonologique est bien formée à l'interface sensori-motrice si elle ne comporte que des traits interprétables.

La figure 4 ci-dessous résume nos hypothèses sur les types de traits manipulés par le modèle :

	SYNTAXE	PHONOLOGIE
<i>traits formels</i>	étiquettes catégorielles	unités CV
<i>traits interprétables</i>	descriptions d'événements...	éléments...

Figure 4. Traits formels et traits interprétables

Aussi bien en morphosyntaxe qu'en phonologie, les traits apparaissent en combinaison avec d'autres traits : les terminaux syntaxiques se combinent pour former des syntagmes et des propositions, et les unités CV se combinent pour former des unités prosodiques telles que les pieds et les syntagmes prosodiques.

En syntaxe, la présence de frontières entre les constituants n'est pas encodée sous la forme d'un objet spécifique, et ceci est admis par tous. En revanche, en phonologie, les frontières sont marquées sous la forme d'un diacritique (#, cf. (Chomsky

et Halle, 1968)).⁷ Pourtant, rien ne s'oppose à ce qu'en phonologie comme en syntaxe, les frontières soient définies indirectement, par l'absence d'*assemblage* à un certain niveau de représentation. C'est cette hypothèse qui sera poursuivie ici. Soit le symbole $\circ$, signifiant qu'un assemblage a été établi. L'opération d'*assemblage* en phonologie est illustrée en (27). (27a) et (27b) représentent chacun un constituant, (27c) représente deux constituants consécutifs séparés par une frontière : le premier est composé de deux unités CV, et le second d'une seule unité CV. Les représentations plus habituelles de ces structures figurent dans la partie droite de l'illustration.

L'opération d'*assemblage* dérive les frontières, mais cette stratégie a une implication généralement rejetée : les constituants phonologiques sont le résultat d'une concaténation. J'accepte cette implication.

(27) Assemblage et frontière en phonologie

- | | | |
|----|--------------------------|-----------|
| a. | CV $\circ$ CV | #CVCV# |
| b. | CV $\circ$ CV $\circ$ CV | #CVCVCV# |
| c. | CV $\circ$ CVCV | #CVCV#CV# |
- où C = position consonantique, V = position vocalique

Passons maintenant aux structures définies sur la base des opérations et des variables que nous venons de présenter. Dans le modèle que je propose, les traits formels de la syntaxe et de la phonologie sont des variables ; l'ensemble des traits formels syntaxiques (noté désormais FF_{syntaxe}) et l'ensemble des traits formels phonologiques (noté désormais FF_{phono}), munis de l'opération *assemblage*, sont définis comme des monoïdes commutatifs (28). Un monoïde est une structure algébrique (29) munie d'une loi de composition interne, associative, admettant un élément neutre (30).

(28) (FF , *assemblage*) est un monoïde commutatif ssi

- (FF , *assemblage*) est un monoïde
- *assemblage* est commutative : $\forall x, y \in FF : x \circ y = y \circ x$

(29) (FF , *assemblage*) est une structure algébrique :

- FF est un ensemble
- *assemblage* est une opération binaire

⁷ Le statut de ce diacritique est objet de débats (cf. (Scheer, 2010) pour une revue et une discussion).

- (30) $(FF, assemblage)$ est un monoïde ssi
- c'est une structure algébrique
 - *assemblage* est une loi de composition interne associative, munie d'un élément neutre
 - *assemblage* est une loi de composition interne (l.c.i.) ssi :

$$\forall x, y \in FF : x \circ y \in FF$$
 - *assemblage* est associative dans FF ssi :

$$\forall x, y, z \in FF : (x \circ y) \circ z = x \circ (y \circ z)$$
 - élément neutre : $\exists e \in FF : \forall x \in FF : x \circ e = x = e \circ x$

Nos hypothèses jusqu'à présent sont résumées dans la figure 5.

	SYNTAXE	PHONOLOGIE
ensemble de variables :	FF _{syntax}	FF _{phono}
opération binaire :	assemblage	
(propriétés : l.c.i., associative, commutative, élément neutre)		

Figure 5. Une structure de monoïde commutatif en syntaxe et en phonologie

2.1.1.2. Hiérarchie et interface

La seconde opération du modèle, *tête*, désigne une variable (ou un *assemblage* de variables) comme tête de la configuration et élimine la variable non tête de la dérivation. En d'autres termes, la variable désignée comme tête reste active dans la dérivation, et la variable complément est éliminée de la dérivation active. *tête* prend des traits formels comme input, et génère des traits interprétables comme output ; *tête* n'est donc pas une loi de composition interne pour l'ensemble des traits formels. Ce mécanisme effectue ce que la partie "étiquetage" (*label*) de l'opération *merge* effectue chez Chomsky (l'output de *merge* hérite sa catégorie de la tête), mais il produit également des représentations interprétables aux interfaces. Ceci correspond au contenu empirique de l'opération de *transfer* de Chomsky.

Si l'opération d'*assemblage* est modélisée comme un produit, et *tête* comme une somme (31, 32), alors toutes les conditions précédemment détaillées sont remplies (Bendjaballah et Haiden, 2003, Haiden, 2005b). *Assemblage* a les propriétés (28-30). La

(31) assemblage
assemblage: $(x, y) \rightarrow xy$

(34)

CV		
CV	CV	
CV	CV	CV
pa	ta	t

2.1.1.3. Application PF-syntaxe

Jusqu'à présent, nous avons défini la dérivation des représentations d'interface en syntaxe et en phonologie. La syntaxe livre des représentations à l'interface conceptuo-intentionnelle, la phonologie livre des représentations à l'interface sensori-motrice. Un individu équipé de ces opérations sera capable de raisonner (en utilisant sa syntaxe) et de produire des sons bien formés, mais sera incapable d'exprimer ses raisonnements en utilisant ses sons. Afin d'y parvenir il/elle doit établir une relation entre les sons et les sens. Cette relation est définie comme une application injective de FF_{phono} vers FF_{syntaxe} (35).

(35) Application PF-syntaxe

$$f : FF_{\text{phono}} \rightarrow FF_{\text{syntax}}$$

f est une application injective

Le domaine de départ et le domaine d'arrivée de f sont limités aux traits formels. Ceci a deux conséquences majeures. Tout d'abord, l'application est restreinte aux objets actifs dans la dérivation, ce qui réduit le coût computationnel. Ensuite, un élément phonologique donné ne peut être l'exposant d'un trait syntaxique qu'à la condition d'être associé à une unité CV. Ce point sera crucial pour la segmentation des marqueurs non concaténatifs.

L'application f est définie de FF_{phono} vers FF_{syntaxe} , car la syntaxe est la structure la plus riche des deux ensembles (FF_{phono} et FF_{syntaxe}). Il existe en effet des items lexicaux qui ont une représentation syntaxique mais pas de forme phonologique (par exemple pro). En revanche, il n'existe pas d'item lexical qui ait une forme phonologique, mais aucune représentation syntaxique (les explétifs, qui sont susceptibles de ne pas avoir de sens, doivent avoir des traits formels). En définissant f comme une application (36), nous

garantissons que tous les éléments de l'ensemble de départ (la phonologie) reçoivent une valeur unique dans l'ensemble d'arrivée (la syntaxe).

(36) Application ((Lipschutz et Lipson, 2001) : 171, version française : 121)

"Soient A et B deux ensembles quelconques. Supposons qu'à chaque $a \in A$ on associe un élément unique de B ; l'ensemble de ces correspondances est appelé une fonction ou application de A dans B . "

La propriété d'injectivité indiquée en (35) est nécessaire en raison d'une caractéristique importante du lexique connue sous le terme d'unicité (*uniqueness* (Naigles et al., 1992)) ou de contraste lexical (*Lexical Contrast*, (Bloom, 2000) : 65ff) : il ne peut y avoir deux items lexicaux ayant la même valeur syntactico-sémantique.

(37) Injection ((Lipschutz et Lipson, 2001) : 173, version française : 123)

"Une application $f: A \rightarrow B$ est dite injective si des éléments distincts de A ont des images distinctes ; c'est-à-dire :

si $a \neq a'$ implique $f(a) \neq f(a')$

ou ce qui est équivalent

si $f(a) \neq f(a')$ implique $a \neq a'$ "

Dans le modèle proposé, l'application phonologie-syntaxe n'est pas définie comme une relation dérivationnelle transportant, telles quelles, les entrées lexicales de la phonologie vers la syntaxe. L'application phonologie-syntaxe doit plutôt être conçue comme une contrainte générale imposée par le lexique : elle garantit l'intégrité des items lexicaux dans les dérivations phonologiques et syntaxiques. Ainsi, la directionalité de l'application a un statut très différent de celui qu'elle a dans le cadre minimaliste standard, dérivationnel.

PHONOLOGIE		SYNTAXE
FF _{phono}	$\rightarrow$ f	FF _{syntaxe}
propriétés : application injective, définie dans le lexique sur les items lexicaux		

Figure 6. Application PF – syntaxe

2.1.1.4. Et l'ordre linéaire ?

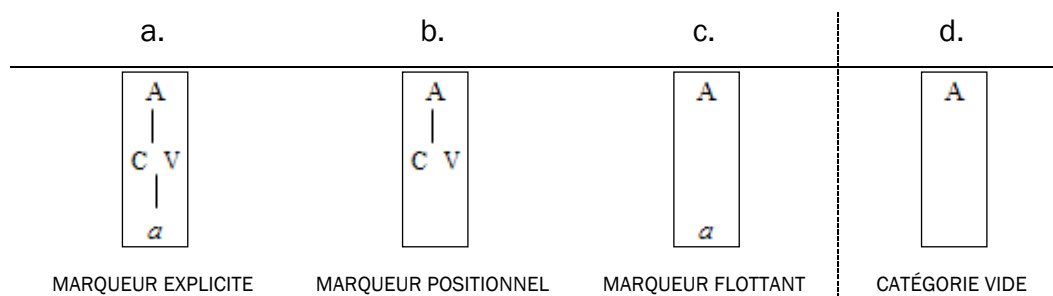
Nous avons défini l'opération d'*assemblage* comme un produit, il s'agit donc d'une opération commutative, qui en tant que telle, n'impose pas d'ordre linéaire. L'ordre linéaire doit donc être stipulé par ailleurs. Cette propriété du modèle est souhaitable pour plusieurs raisons. Tout d'abord, l'ordre linéaire est paramétrique, et non universel. De ce fait, il ne doit pas être encodé dans le noyau universel du système dérivationnel. Ensuite, les systèmes morphologiques non concaténatifs nous montrent que les items lexicaux peuvent être linéairement discontinus. Puisque ces systèmes sont évidemment acquis par leurs locuteurs, les configurations non concaténatives ne doivent pas réclamer un effort dérivationnel différent de celui requis par les systèmes concaténatifs ; elles doivent découler des mécanismes du système dérivationnel. Je propose que l'ordre linéaire, tel que nous l'observons dans les expressions explicites en surface, est déterminé par les gabarits. Les gabarits sont des règles d'ordonnancement d'unités CV (Haiden, 2008b, 2009).

2.1.2 Exposants morphologiques

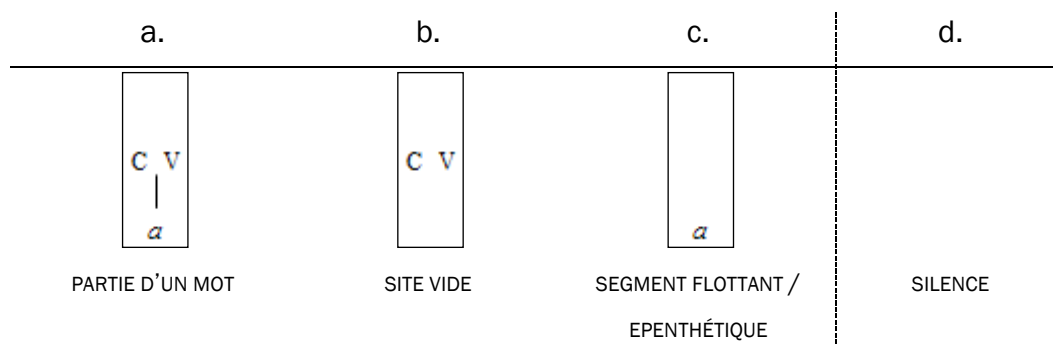
Examinons maintenant tour à tour les possibilités logiques de représentation dérivées par le modèle. En principe, $2^4 = 16$ combinaisons de traits interprétables et de traits formels en phonologie et en syntaxe sont dérivées. Les traits syntaxiques interprétables seront cependant exclus de la discussion, car ces traits soulèvent des questions concernant la représentation syntaxique d'objets sémantiques non réalisés tels que les variables d'événement, questions qui dépassent le cadre de ce chapitre. Cela nous laisse les $2^3 = 8$ configurations listées ci-dessous en (38)-(39). Pour faciliter la visibilité des représentations, l'application entre PF et syntaxe, définie en (35), est représentée par une ligne verticale reliant les unités CV (en bas) aux terminaux syntaxiques (en haut). Cette notation est un simple raccourci, elle ne signifie en aucun cas que les terminaux syntaxiques dominent les traits phonologiques.

Le modèle établit une distinction entre deux classes principales de configurations : les objets ayant une valeur morphosyntaxique par convention lexicale, i.e. les exposants morphologiques (38), et les objets sans valeur morphosyntaxique (39).

(38) exposants



(39) objets sub-lexicaux



Parmi les types d'exposants listés en (38), notons d'emblée que la configuration (38d) ne nous concernera plus dans la suite : elle représente une catégorie syntaxique dénuée d'expression phonologique.

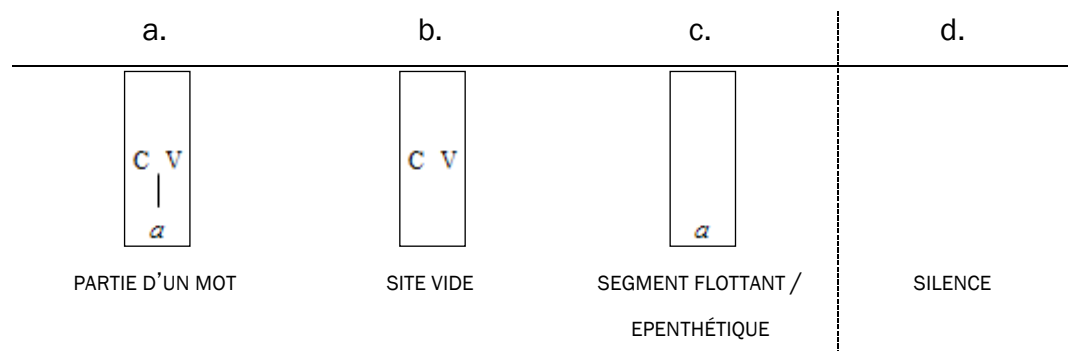
(38a) est un exposant entièrement spécifié, muni à la fois de matériel segmental et de matériel squelettal. C'est le type d'exposant que nous nous attendons à trouver dans les systèmes concaténatifs.

(38b) est un exposant sous-spécifié. Il exprime un trait syntaxique A, en fournissant une unité CV, mais il n'a pas de traits segmentaux (interprétables). Ce type d'exposant n'aura pas d'épel audible, à moins que l'unité CV qu'il introduit soit identifiée par du matériel segmental adjacent. (38b) est la représentation des morphèmes impliqués dans les phénomènes de propagation consonantique, vocalique, ou bien de reduplication CV à valeur morphologique. On peut citer divers exemples, tels que le marqueur d'"imperfectif" en kabyle (propagation consonantique, berbère, afroasiatique, cf. (Bendjaballah, 2001))), le marqueur d'"incomplétif" en muskoguee (propagation vocalique, creek, cf. (Haas, 1940, Ségéral, 2005))), ou le marqueur "d'intensité/distributif" en lamang (reduplication CV, tchadique, afroasiatique, cf. (Newman, 1990) : 67).

(38c) est la représentation lexicale d'un marqueur flottant. Cette configuration est déficiente, puisqu'elle n'a pas de trait formel phonologique. Par conséquent, la valeur segmentale de ce marqueur ne peut pas être réalisée telle quelle : je rappelle que l'opération *assemblage* ne "voit" pas les segments (cf. section 2.1.1.1 plus haut). Le marqueur (38c) ne peut être réalisé qu'à la condition de recevoir le support d'un gabarit externe qui puisse l'héberger. En particulier, ce gabarit externe doit remplir une condition importante, qui découle de la condition d'unicité (cf. section 2.1.1.3 plus haut) : il ne peut pas avoir une valeur syntaxique différente de celle du marqueur flottant. L'existence de marqueurs flottants a été établie au départ sur la base de certaines langues à tons (Goldsmith, 1979, Leben, 1978). En bédja par exemple (couchitique, afroasiatique), le marqueur possessif de 1ère personne est un ton bas, et celui de 3ème singulier est un ton haut (Appleyard, 1991) : 7). Ensuite, cette phénoménologie a été étendue aux éléments. Des processus de palatalisation et de labialisation à valeur morphosyntaxique (impliquant les éléments flottants I et U respectivement) ont par exemple été identifiés dans diverses langues. On peut ainsi citer la 2nde personne féminine du singulier dans les formes de "non perfectif" en chaha (éthio-sémitique, afroasiatique), qui est marquée par une palatalisation, *i.e.* par un élément I flottant (cf. (McCarthy, 1983): 179, (Rose, 1997), et (Lowenstamm, 2000) pour une discussion).

Les configurations en (39), répétées en (40) ci-dessous pour plus de commodité, n'encodent pas l'appariement d'un son et d'un sens. Ces configurations ne peuvent donc pas représenter des exposants morphologiques, puisque sons et sens sont reliés par une application : par définition, tout élément de FF_{phono} doit être associé à un élément de FF_{syntaxe} . Cependant, ces configurations jouent un rôle crucial dans les gabarits, en facilitant la réalisation d'autres exposants. Nous allons donc les examiner une par une. Notons que (40d), représentation du silence, ne nous concernera plus dans la suite.

(40) objets sub-lexicaux



(40a) est une position squelettale, munie d'un segment, mais sans valeur morphosyntaxique. Puisque cet objet possède à la fois du matériel segmental et du matériel squelettal, il est phonétiquement audible. Cependant, il ne fournit aucun trait morphosyntaxique propre. Cette configuration apparaît de façon triviale dans les items lexicaux de longueur supérieure à une unité CV.

(40b) est un site gabaritique vide, i.e. un objet comportant un trait formel phonologique (une unité CV), et n'ayant ni trait interprétable phonologique, ni valeur syntaxique. Cette configuration sert typiquement de support aux marqueurs flottants (de type (38c)). Nous en verrons un exemple dans la seconde partie de chapitre consacrée aux prépositions légères du berbère kabyle de Chemini (Bendjaballah et Haiden, to appear). Comme nous examinerons ce cas très soigneusement, je ne donne pas plus de détails ici.

(40c) est un segment flottant sans valeur morphosyntaxique. Cette configuration représente un segment épenthétique : les segments épenthétiques en effet ne créent pas de structure dans le sens où ils n'ont pas de traits formels.⁸ Les segments épenthétiques font surface lorsque des conditions phonotactiques nécessitent l'épel d'une position CV. Le cas le plus clair est celui des alternances voyelle-zéro (par exemple en arabe classique, cf. (Kaye et al., 1985, 1990)) : si une position V vide est proprement gouvernée, elle reste silencieuse, si elle n'est pas proprement gouvernée, elle est épelée par une voyelle neutre.

2.1.3 Ordre linéaire et gabarits

Le formalisme proposé ici, de par sa conception, n'impose pas d'ordre linéaire. La source de la linéarisation doit donc être ailleurs. Cette question est particulièrement cruciale dans les systèmes non concaténatifs, dans lesquels les exposants morphologiques peuvent être linéairement discontinus. Je propose que l'ordre linéaire soit imposé par des règles ordonnant les unités CV (Haiden, 2008b, 2009). De telles règles sont plus connues sous le nom de gabarits. Dans cette section, je précise la façon dont les gabarits interagissent avec le formalisme présenté précédemment, et mets au jour le mécanisme de dérivation des objets morphologiques non concaténatifs.

⁸ Pour une hypothèse similaire, cependant pour des raisons différentes, cf. (Scheer and Ségéral, 2005).

Prenons un exemple simple. En berbère kabyle, l'imperfectif des verbes triconsonantiques de type $C_1C_2\alpha C_3$ est formé par gémination de la seconde consonne radicale ((Bendjaballah, 1999)). Par exemple, l'imperfectif du verbe $x\ddot{o}m$ 'travailler' est $x\ddot{o}\ddot{o}m$ 'travailler habituellement'. Cette formation est de type non concaténatif à deux égards. Tout d'abord, le marqueur d'imperfectif n'est pas situé à la périphérie de la base, mais en son milieu. Ensuite, la valeur segmentale du marqueur, [dd] dans notre exemple, dépend de celle des segments de la base (ici, l'identité de la seconde consonne radicale). Nous allons voir que, malgré les apparences, cette forme est entièrement compositionnelle.

Supposons que l'on puisse isoler en berbère kabyle un morphème "racine", appariant une séquence ordonnée de segments, $x\ddot{o}m$ dans notre exemple, et un sens, *travailler* (cf. (Prunet et al., 2000), (Idrissi, 2000), entres autres, pour des arguments en faveur de la racine). Puisque les racines apparaissent dans des radicaux ayant des catégories variées, nous avons deux options. Soit les racines n'ont pas de catégorie syntaxique (Marantz, 1997), soit elles ont une catégorie maximale sous-spécifiée, compatible avec toutes les catégories de classe majeure (appelons-la $\sqrt{}$). Je poursuivrai la seconde de ces hypothèses. Je supposerai également que les racines n'ont pas de traits phonologiques formels : elles n'ont pas de support squelettal. Ainsi, les racines sont des objets de type (38c). (41) illustre l'entrée lexicale de la racine $x\ddot{o}m$.

(41) racine $x\ddot{o}m$

$\sqrt{}$

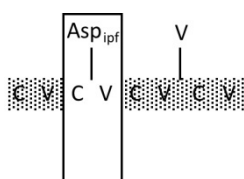
$x \ \ddot{o} \ m$

$\sqrt{}$ étant une catégorie sous-spécifiée, les racines peuvent être hébergées par des marqueurs positionnels (type 38b) de toute catégorie (lexicale), mais elles ne peuvent pas être hébergées par le gabarit d'une catégorie fonctionnelle.

L'imperfectif est formé par ajout d'une unité CV qui n'a pas de contenu segmental indépendant, *i.e.* d'un marqueur de type (38b). L'imperfectif est marqué sur une position spécifique dans le gabarit, la seconde position. Je propose donc que le marqueur

d'imperfectif fasse partie d'un gabarit consistant en une unité CV initiale (objet de type 39c), suivie du marqueur positionnel d'imperfectif (de type 38b), suivi de deux sites vides (objets de type 39c). Les première, troisième et quatrième unités CV, ensemble, forment un exposant de type (38b). Ces positions, grisées en (42), hébergent la base verbale. Comme les gabarits établissent une relation entre les terminaux syntaxiques et les positions CV (i.e. une application entre phonologie et syntaxe), ils doivent être listés dans le lexique (cf. section 2.1.1.3 plus haut). Supposons que tel est le cas.

(42) gabarit d'imperfectif



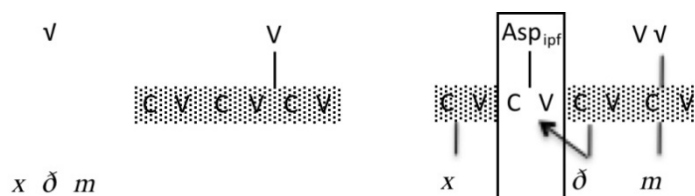
En tant qu'items lexicaux, les gabarits peuvent imposer un ordre linéaire, mais ils n'ont ni structure syntaxique, ni structure phonologique intrinsèque. Les assemblages et structures sont générés au cours de la dérivation dans la composante syntaxique et dans la composante phonologique. Voyons maintenant de quelle façon.

La racine entre dans une "matrice lexicale" (*lexical array*) avec le gabarit d'imperfectif. En syntaxe, V(erbe) est assemblé à $\sqrt{\quad}$. En phonologie, les positions CV appliquées à V(erbe) sont jointes par assemblage (43a). Or, la racine est un marqueur flottant (41), elle ne peut donc être réalisée que si elle reçoit le support de sites gabaritiques. V(erbe) et $\sqrt{\quad}$ étant assemblées, et ces deux catégories n'étant pas distinctes, le contenu segmental de la racine peut être associé au support squelettal du verbe, en respectant les propriétés de l'application phonologie-syntaxe définies en section 2.1.1.3.

Nous avons maintenant deux options. Soit les unités CV du verbe sont épelées, soit elles sont à nouveau soumises à l'opération d'*assemblage*. Si l'on suit la première de ces options, le radical verbal est scellé tel quel, et seuls des marqueurs externes (i.e. concaténatifs) peuvent lui être affixés. Le gabarit (42) nous force à exclure cette option. Si l'on suit maintenant la seconde option, les unités CV du radical restent actives dans la dérivation. L'opération *assemblage* étant commutative, les objets nouvellement affixés peuvent être placés n'importe où dans la séquence active dans la dérivation. Puisque le gabarit (42) exige que le marqueur d'imperfectif soit réalisé en seconde position, il sera linéarisé dans cette position.

Enfin, la configuration obtenue est interprétée par l'opération tête. Les catégories $\sqrt{\text{ }}$ et V(erbe) étant indistinctes, V(erbe) hérite sa sémantique de la racine, et les segments radicaux seront interprétés sur les positions CV verbales. En syntaxe, tête élimine V(erbe) $\sqrt{\text{ }}$, et définit par là Asp_{ipf} comme tête. En phonologie, les segments radicaux sont interprétés sur les positions CV verbales, selon les règles standard de la phonologie autosegmentale, un par un, de la gauche vers la droite. Puisque seules des positions CV verbales sont interprétées à ce stade, la position aspectuelle qui a été intercalée n'interfère pas dans l'association des segments radicaux aux positions squelettales. Une fois que les unités CV verbales ont été interprétées, leurs valeurs segmentales peuvent se propager sur le site infixal lexicalement vide (43b).

- (43) a. *assemblage* de $\sqrt{\text{ }}$ et V b. *identification* de la position infixale



2.2 Base empirique : La représentation syntaxique des prépositions faibles du berbère de Chemini (kabyle)

Dans la seconde partie de ce chapitre, j'illustre ce modèle par une étude de cas portant sur les prépositions faibles du berbère de Chemini (kabyle).⁹ Le berbère possède un nombre restreint de morphèmes à valeur relationnelle ou spatio-temporelle préfixés sur le nom. Ces morphèmes sont intéressants dans la perspective de la théorie que je viens d'énoncer, car ce sont des marqueurs flottants, *i.e.* de type (38c), et leur comportement morphosyntaxique permet de motiver l'existence de positions gabaritiques vides, *i.e.* d'objets de type (39b/40b).

Cette partie est structurée de la manière suivante. Dans la section 2.2.1, je présente les données de départ. Dans les sections 2.2.2, 2.2.3 et 2.2.4, je mets en

⁹ Cette partie reprend un manuscrit accepté pour publication en 2006 dans un volume qui n'a pas encore vu le jour (*Prépositions et postpositions – Approches typologiques et formelles*, J. Tseng (éd.), Hermès Lavoisier).

évidence que les prépositions faibles ne peuvent être analysées ni comme des marqueurs casuels, ni comme des clitiques pronominaux. Dans les sections 2.2.5 et 2.2.6, j'examine précisément la représentation phonologique des prépositions faibles et leur comportement à l'interface PF-syntaxe. Enfin, j'étudie le comportement des prépositions faibles dans le contexte des dépendances A' à longue distance.

2.2.1 La distribution des prépositions faibles

Les prépositions faibles en kabyle de Chemini sont au nombre de six : *f* 'sur', *s* 'avec (instrumental)', *ð* 'avec (comitatif)', *g* 'dans', *i* 'datif' et *n* 'génitif'. Elles apparaissent préfixées sur le nom, ainsi qu'illustré en (44).^{10,11}

- (44) a. *i-sərs-it^s* *f-θk^wərsit^s*
 3MS-poser.PF-OD:3FS sur-petite chaise.EA
 "Il l'a posée sur la petite chaise."
- b. *jə-ttʃa* *s-θkənʒawθ*
 3MS-manger.PF avec-cuillère.EA
 "Il a mangé avec une cuillère."
- c. *jə-ttʃa* *ð-wəmʃiʃ*
 3MS-manger.PF avec-chat.EA
 "Il a mangé avec le chat."

¹⁰ Les abréviations employées dans cette partie sont les suivantes : MS = masculin singulier, FS = féminin singulier, PF = perfectif, AOR = aoriste, PART = participe, INT = intensif, EA = état d'annexion, EL = état libre, OD = objet direct, OI = objet indirect, DAT = datif, GEN = génitif, PRT = particule, COMP = complémentateur, NEG = négation, DEM = démonstratif, DIR = directionnel, PR.TON = pronom tonique, POSS = possessif.

¹¹ Certaines des séquences présentes dans les exemples sont sujettes à diverses assimilations. Dans un premier temps, afin de maintenir la clarté de la segmentation, ces assimilations ne sont pas notées. Elles seront explicitement abordées en section 2.5.1.

- d. i-qəlβ-it^s g-wəxxam
3MS-chercher.PF-OD:3FS dans-maison.EA
"Il l'a cherchée dans la maison."
- e. jə-fka-jas-θ i-θməttuθ
3MS-donner.PF-OI:3S-OD:3MS à-femme.EA
"Il l'a donné à la femme."
- f. axxam n-θməttuθ
maison.EL de-femme.EA
"La maison de la femme"

Un nom en berbère a deux formes, appelées traditionnellement *état libre* et *état d'annexion*.¹² Les deux formes des noms employés en (44) figurent en (45). Notons que la distinction entre *état libre* et *état d'annexion* n'est pas corrélée à une distinction en termes de définitude : en berbère, un nom est interprété comme défini ou indéfini selon le contexte. Un nom à l'état d'annexion tout comme un nom à l'état libre peut être interprété comme défini.

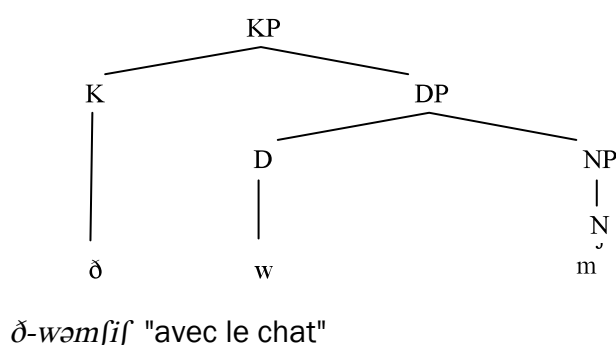
(45)	<u>état libre</u>	<u>état d'annexion</u>	
	θak ^w ərsit ^s	θk ^w ərsit ^s	'petite chaise'
	θakənɯawθ	θkənɯawθ	'cuillère'
	amɕiɕ	wəmɕiɕ	'chat'
	axxam	wəxxam	'maison'
	θaməttuθ	θməttuθ	'femme'

Le nom qui suit une préposition faible apparaît à l'état d'annexion. Cette observation est centrale dans la théorie de (Guerssel, 1987, 1992). Guerssel propose que (a) un nom à

¹² On trouve aussi le terme *état construit* pour qualifier la forme d'état d'annexion du berbère. Le terme *état construit* est cependant très généralement employé pour décrire une notion différente dans la grammaire des langues sémitiques ; pour une discussion des différences entre l'état d'annexion en berbère et l'état construit en sémitique, voir (Bendjaballah and Haiden, to appear). Afin d'éviter toute confusion, j'emploierai donc le terme *état d'annexion* dans la suite.

l'état libre possède un marqueur de cas tandis qu'un nom à l'état d'annexion n'en possède pas, et (b) les prépositions faibles sont elles-mêmes des marqueurs de cas. Guerssel explique alors la distribution des prépositions faibles de façon simple : les noms à l'état libre étant déjà marqués casuellement, les prépositions faibles ne peuvent s'y attacher. Elles s'attachent à un nom non marqué casuellement, un nom à l'état d'annexion (46).

(46) Analyse des P faibles comme marqueurs de cas



Dans la suite, je maintiendrai l'analyse des marqueurs d'état comme marqueurs casuels, mais argumenterai contre l'hypothèse que les prépositions faibles sont également des marqueurs casuels.

Dans le parler kabyle de Chemini, les prépositions faibles s'attachent à la gauche de certaines particules. Dans ce contexte, la préposition faible suit le nom, et elle doit être adjacente à la particule. Le nom ainsi que l'élément *wh* apparaissent à l'état libre, et non à l'état d'annexion. Ceci est illustré en (47a) avec la préposition *s* 'sur': *s* apparaît directement à la gauche des particules *i*, *ara* et *u(r)*. Par contraste, les prépositions faibles ne peuvent pas s'attacher à la particule *a(ə)* (47b).¹³

- (47) a. anta θaqəndurθ s-i-t^s i-ɣra
 quel.F.EL chemise.EL avec-PRT-OD:3FS 3MS-voir.PF
 "Avec quelle chemise/robe l'a-t-il vue?"

¹³ Le statut syntaxique de ces particules sera détaillé en section 2.2.2. Nous verrons que *i*, *ara* et *u(r)* sont des compléments tandis que *a(ə)* est un marqueur de temps.

anta θaqəndurθ s-ara-t^s i-zər
 quel.F.EL chemise.EL avec-PRT-OD:3FS 3MS-voir.AOR
 "Avec quelle chemise/robe la verra-t-il?"

anta θaqəndurθ s-u-t^s i-zr-ara
 quel.F.EL chemise.EL avec-PRT-OD:3FS 3MS-voir.PF-NEG
 "Avec quelle chemise/robe ne l'a-t-il pas vue?"

b. * anta θaqəndurθ s-a-t^s i-zər
 quel.F.EL chemise.EL avec-PRT-OD:3FS 3MS-voir.AOR
 visé : "Avec quelle chemise/robe la verra-t-il?"

Les prépositions de datif *i* et de génitif *n* présentent quelques particularités qui les distinguent des autres prépositions légères. Je présente ici deux de ces caractéristiques. Tout d'abord, lorsqu'elles sont employées à la gauche des particules *i*, *ara* et *u(r)*, les prépositions *i* (datif) et *n* (génitif) ne sont pas réalisées sous leur forme "normale" : c'est un *m* qui fait surface dans ces contextes (48b, 48d, 49b).¹⁴

(48) a. j-uçr-as-θ i-θəqɿjθ
 3MS-voler.PF-OI:3S-OD:3MS DAT-fille.EA
 "Il l'a volé à la fille."

b. anta θaqɿjθ m-i-θ j-uçər
 quel.F.EL fille.EL *m*-PRT-OD:3MS 3MS-voler.PF
 "A quelle fille l'a-t-il volé?"

c. jə-fka-jas-θ i-θəqɿjθ
 3MS-donner.PF-OI:3S-OD:3MS DAT-fille.EA
 "Il l'a donné à la fille."

¹⁴ Le segment *m* peut être analysé soit comme un allomorphe de *i*, soit comme un morphème différent. Dans le cadre de ce chapitre, je ne tranche pas entre ces deux analyses.

- d. anta θaɟʃiʃθ m-i-θ jə-fka
 quel.F.EL fille.EL m-PRT-OD:3MS 3MS-donner.PF
 "A quelle fille l'a-t-il donné?"

- (49) a. jə-nka aɟʒun n-wərgaz-aki
 3MS-tuer.PF chien.EL GEN-homme.EA-DEM
 "Il a tué le chien de cet homme."

- b. anwa argaz m-i gə-nka aɟʒun
 quel.M.EL homme.EL m-PRT 3MS-tuer.PF chien.EL
 "Il a tué le chien de quel homme?"

Ensuite, la préposition *n* (génitif) a pour particularité de pouvoir apparaître comme complément de N dans une construction [P_{faible} N [P_{faible} N]]. Les autres prépositions faibles sont exclues dans ce contexte (Guerssel, 1987).¹⁵

¹⁵ Ces constructions sont rarement discutées dans la littérature, et leur comportement n'est pas encore vraiment compris. Voici une brève présentation de la situation. Commençons par un exemple illustratif du français : *Il parle avec la femme sur la terrasse*. Cette phrase est ambiguë : *sur la terrasse* peut être soit un complément de phrase, soit un complément de nom. Dans le premier cas, la phrase ne peut être vraie que si le sujet se trouve sur la terrasse. Dans le second cas (complément de nom), la phrase peut être vraie si le sujet ne se trouve pas sur la terrasse. Examinons maintenant la phrase correspondante en berbère de Chemini :

- (i) ihəddər t^sməttuθ flatekas
 3MS.parler.INT avec.femme.EA sur.terrasse
 a. "Il parle [avec la femme] [sur la terrasse]."
 b. "Il parle de la terrasse avec la femme."

Cette phrase n'est pas ambiguë de la même manière qu'en français. Le point essentiel est que *f-latekas* (sur la terrasse) ne peut pas être complément de nom. Il est complément de phrase (i.a). La seule interprétation où le sujet peut ne pas être sur la terrasse et la phrase reste vraie est celle en (i.b) : le verbe *həddər* (parler) prend la préposition *f* comme complément (*parler de*), et la phrase signifie alors que le sujet a parlé de la terrasse avec la femme. Pour obtenir l'interprétation où *f-latekas* est complément de nom, il faut employer une structure différente, une relative :

- (ii) ihəddər t^sməttuθ igəllan flatekas
 3MS.parler.INT avec.femme.EA être.PART sur.terrasse
 "Il parle avec la femme qui est sur la terrasse."

Par contraste, la préposition faible *n* apparaît de façon privilégiée comme complément de nom, même dans la structure [P_{faible} N [P_{faible} N]] :

- (iii) iruħəd [ggwəxxam ppwəmʃiʃ]
 3MS.aller.PF.DIR dans.maison.EA GEN.chat.EA
 "Il vient de la maison au chat."

Pour conclure, la distribution des prépositions faibles est résumée dans le tableau (50). Toutes les prépositions faibles s'attachent à la gauche d'un nom à l'état d'annexion. Dans le domaine des particules, on observe deux asymétries. Tout d'abord, les prépositions *f*, *s*, *ǫ* et *g* s'attachent aux particules phrastiques *i*, *ara* et *ur*, tandis que les prépositions *n* (génitif) et *i* (datif) ne s'y attachent jamais sous leur forme "normale". Ensuite, les particules *i*, *ara* et *ur* peuvent être hôtes de préposition faible, tandis que la particule *aǫ* ne l'est jamais.

(50)

	<i>f, s, ǫ, g</i>	<i>i</i>	<i>n</i>
$_{-} N_{EA}$	+	+	+
$_{-} i/ara/ur$	+	+ (<i>m</i>)	+ (<i>m</i>)
$_{-} a(ǫ)$	-	-	-

Dans la suite, nous verrons que ces deux asymétries sont corrélées aux propriétés syntaxiques et phonologiques des éléments impliqués. Plus précisément, nous verrons que la particule *a(ǫ)* n'a pas le même statut syntaxique que les particules *i*, *ara* et *ur*, puis que la préposition *n* (génitif) ne se comporte pas comme les autres prépositions faibles vis-à-vis des phénomènes d'assimilation entre préposition et nom.

2.2.2 Les particules

En berbère de Chemini, les quatre particules suivantes peuvent apparaître en initiale de proposition :

- (51) *a(ǫ)* particule *irrealis*, marque souvent le futur
- i* introduit une proposition *realis*, souvent avec un verbe au perfectif
- ara* introduit une proposition *irrealis*
- ur* introduit une proposition négative

Dans cette section je présente d'abord trois arguments en faveur de l'hypothèse que *i* et *ara* sont des complémenteurs (des arguments similaires sur la particule *ur* peuvent être trouvés dans (Bendjaballah et Haiden, to appear)), tandis que *a(ð)* est un marqueur flexionnel, puis j'argumente que les prépositions faibles ne peuvent pas être analysées comme des marqueurs casuels au sens fort.

2.2.2.1. Argument 1 : les particules dans les propositions interrogatives

L'ordre des mots neutre dans la proposition affirmative en berbère de Chemini est VSO. Dans une proposition principale affirmative dont le verbe est au perfectif, le verbe apparaît en position initiale (52a). Dans une proposition principale *irrealis*, le verbe doit être précédé de la particule *að*, et le verbe est employé à une forme traditionnellement appelée "aoriste" (52b). Les particules *i* et *ara* sont toujours agrammaticales en initiale de proposition principale affirmative neutre (53).

- (52) a. jə-ʒra wərgaz θakθəβθ
 3MS-voir.PF homme.EA livre.EL
 "L'homme a vu le livre."

- b. að i-ʒər wərgaz θakθəβθ
 PRT 3MS-voir.AOR homme.EA livre.EL
 "L'homme verra le livre."

- (53) a. * i jə-ʒra wərgaz θakθəβθ
 PRT 3MS-voir.PF homme.EA livre.EL

- b. * ara i-ʒər wərgaz θakθəβθ
 PRT 3MS-voir.AOR homme.EA livre.EL

Dans une question *wh* dont le verbe est au perfectif, le constituant *wh* apparaît en initiale, suivi de la particule *i* puis du verbe : (54a) correspond à (52a). Dans les

propositions interrogatives *irrealis*, ce n'est plus la particule *að* qui est employée, mais la particule *ara* : (54b) correspond à (52b).

- (54) a. anta θakθəβθ i jə-zra wərgaz
 quel.F livre.EL PRT 3MS-voir.PF homme.EA
 "Quel livre l'homme a-t-il vu ?"
- b. anta θakθəβθ ara i-zər wərgaz
 quel.F livre.EL PRT 3MS-voir.AOR homme.EA
 "Quel livre l'homme verra-t-il ?"

L'interrogation déclenche un mouvement *wh*, qui conduit à la lexicalisation du complémenteur. Dans les interrogatives, les particules *i* et *ara* apparaissent. J'en déduis que *i* et *ara* sont des complémenteurs.

2.2.2.2. Argument 2 : la distribution complémentaire avec le complémenteur *uqβəl* "avant"

Les particules *i* et *ara* sont en distribution complémentaire avec des éléments ayant la fonction de complémenteur. Ce fait constitue un second indice en faveur de l'analyse de *i* et *ara* comme complémenteurs.

Considérons la préposition *uqβəl* 'avant', qui apparaît en tête de syntagme prépositionnel en (55a), dans un emploi adverbial en (55b), et en tête de proposition subordonnée en (55c).

- (55) a. θə-pp^wð-əd uqβəl amfiʃ
 3FS-arriver.PF-DIR avant chat.EL
 "Elle est arrivée avant le chat."

- b. $\theta\alpha$ -pp^wǾ-əd uqβəl
 3FS-arriver.PF-DIR avant
 "Elle est arrivée avant."
- c. $\theta\alpha$ -pp^wǾ-əd uqβəl a n-əttf
 3FS-arriver.PF-DIR avant PRT 1P-manger.AOR
 i. "Elle est arrivée avant que nous ayons mangé."
 ii. * "Elle est arrivée, et nous avons mangé avant cela."

Trois faits indiquent qu'en (55c) *uqβəl* est un complémenteur, et non un adverbe.

Tout d'abord, *uqβəl* détermine la forme du verbe qui le suit : seul le non-réel en *aǾ* + *aoriste* est possible (56). Notons ici que l'incompatibilité de *uqβəl* et *ara* n'est pas sémantique. En effet, *aǾ* et *ara* sont sémantiquement équivalents (ils indiquent le non réel).¹⁶

- (56) a. * $\theta\alpha$ -pp^wǾ-əd uqβəl nə-ttʃa
 3FS-arriver.PF-DIR avant 1P-manger.PF
- b. * $\theta\alpha$ -pp^wǾ-əd uqβəl ara n-əttf
 3FS-arriver.PF-DIR avant PRT 1P-manger.AOR

Ensuite, l'adverbe apparaît typiquement à la droite du verbe qu'il modifie, et non à sa gauche (55b). Or, en (55c), *uqβəl* apparaît à la gauche du verbe *əttf* 'manger'.

Enfin, un argument sémantique établit que *uqβəl* ne modifie pas le verbe qui le suit (en tant qu'adverbe), mais prend comme complément la proposition dont le verbe *əttf* "manger" est la tête. Cet argument a la forme suivante. *uqβəl* ordonne deux temps : le complément de *uqβəl* est postérieur à celui de son argument externe. Ceci est clair en (55a), où *uqβəl* est employé comme préposition : le temps du complément 'l'arrivée du

¹⁶ En revanche, *uqβəl* 'avant' est sémantiquement incompatible avec le complémenteur *i*, qui introduit une proposition *realis*.

chat' est postérieur à celui de l'argument externe, 'son arrivée à elle'. Si *uqβəl* est utilisé comme adverbe sans complément explicite (55b), il localise un temps de référence implicite après l'événement dénoté par le verbe. En (55c), *uqβəl* localise le temps de la subordonnée après le temps de la principale (55c.i). Si *uqβəl* était un adverbe de la subordonnée en (55c), nous nous attendrions à ce qu'il localise le temps de 'manger' avant un temps implicite de référence, et à ce que ce temps de référence puisse être identique au temps de la principale, *elle arrive* (55c.ii). Puisque (55c.ii) n'est pas une interprétation possible de (55c), nous concluons que *uqβəl* est bien un complémenteur en (55c).

Le complémenteur *uqβəl* 'avant' est compatible avec la particule *að* (55c), mais ne l'est pas avec *ara* (56b). J'en déduis que *að* est un marqueur flexionnel aspectuo-temporel tandis que *ara* est un complémenteur.

2.2.2.3. Argument 3 : le phénomène d'anti-accord

Notre classification des particules *i*, *ara* et *a(ð)* est confirmée par un phénomène indépendant connu sous le nom d'anti-accord (Ouhalla, 1993).

En berbère, le verbe porte des marques d'accord avec le sujet en nombre, genre et personne. Ces marques disparaissent si le sujet subit un mouvement local de type A'. Le verbe apparaît alors à la forme appelée "participe", qui n'est marquée ni pour le genre, ni pour le nombre, ni pour la personne. L'emploi du participe est illustré en (57a-b) par une question *wh*, et en (57c-d) par une construction clivée introduite par la particule *ð*.

- (57) a. *mənhu i-θ jə-ttʃa-n*
 qui PRT-OD:3MS manger.PART
 "Qui l'a mangé?"

b. mənhu ara-θ jə-ttʃ-ən
 qui PRT-OD:3MS manger.PART
 "Qui va le manger?"

c. ð nəkk i-θ jə-ttʃa-n
 PRT PR.TON.1S PRT-OD:3MS manger.PART
 "C'est moi qui l'ai mangé."

d. ð nəkk ara-θ jə-ttʃ-ən
 PRT PR.TON.1S PRT-OD:3MS manger.PART
 "C'est moi qui vais le manger."

La simple montée du sujet à l'intérieur de la proposition ne déclenche pas l'anti-accord en kabyle de Chemini : en (58a) le verbe doit s'accorder avec le sujet. Ceci est également vrai lorsque le sujet est réalisé à la gauche de la particule a(ð) (58c).

- (58) a. θaməttuθ-aki θə-ttʃa aɕsum
 femme.EL-DEM 3FS-manger.PF viande.EL
 "Cette femme a mangé la viande."
- b. * θaməttuθ-aki jə-ttʃa-n aɕsum
 femme.EL-DEM manger.PART viande.EL
 "Cette femme a mangé la viande."
- c. θaməttuθ-aki að θə-ttʃ aɕsum
 femme.EL-DEM PRT 3FS-manger.AOR viande.EL
 "Cette femme va manger la viande."
- d. * θaməttuθ-aki að jə-ttʃ-ən aɕsum
 femme.EL-DEM PRT manger.PART viande.EL
 "Cette femme va manger la viande."

En revanche, la montée du sujet à la gauche de *i* et *ara* déclenche l'anti-accord :

- (59) a. θaməʈtuθ-aki i-θ jə-ttʃa-n
 femme.EL-DEM PRT-OD:3MS manger.PART
 "C'est cette femme qui l'a mangé."
- b. θaməʈtuθ-aki ara-θ jə-ttʃ-ən
 femme.EL-DEM PRT-OD:3MS manger.PART
 "C'est cette femme qui va le manger."
- c. * θaməʈtuθ-aki i-θ θə-ttʃa
 femme.EL-DEM PRT-OD:3MS 3FS-manger.PF
- d. * θaməʈtuθ-aki ara-θ θə-ttʃ
 femme.EL-DEM PRT-OD:3MS 3FS-manger.AOR

La distribution du participe constitue un argument supplémentaire en faveur de l'hypothèse que *i* et *ara* sont des complémenteurs tandis que *að* ne l'est pas. Si nous analysons *i* et *ara* comme des complémenteurs, nous prédisons que le mouvement du sujet vers la gauche de ces éléments est un mouvement A'. Le verbe ne doit donc pas s'accorder. Tel est bien le cas (59a-b).¹⁷

2.2.3 Complémenteurs et prépositions faibles

Nous pouvons désormais préciser le tableau (50) de la section 2.2.1 : les prépositions faibles s'attachent aux complémenteurs, mais non aux marqueurs aspectuo-temporels.

¹⁷ Le mouvement A' met en valeur le constituant déplacé. Pour cette raison, les exemples en (59a-b) sont traduits par des clivées.

(60)

	<i>f, s, ḏ, g</i>	<i>i</i>	<i>n</i>
$-N_{EA}$	+	+	+
$-C$	+	+(<i>m</i>)	+(<i>m</i>)
$-T$	-	-	-

Si nous analysons les prépositions faibles comme des marqueurs casuels au sens strict, alors nous devons analyser les complémenteurs auxquels elles s'attachent comme des têtes nominales portant une marque casuelle. Cette hypothèse pourrait se fonder sur des travaux comme (Galand, 1957, Mettouchi, 2005)p.92 ss. qui argumentent que la particule *i* comporte un élément pronominal. Nous nous attendons alors à ce que les séquences P-Comp-Proposition soient attestées dans les mêmes contextes que les séquences P-N. Nous allons voir que tel n'est pas le cas : la séquence P-Comp ne se comporte pas comme un nom marqué casuellement.

Soient les verbes *hḏar* 'parler', *ṣḥāḇ* 'écrire' et *nnak* 'se battre', qui régissent un complément introduit par *f* 'sur' : (61a), (62a) et (63a). L'analyse des prépositions faibles comme marqueurs casuels au sens strict prédit que ces verbes prennent aussi un complément phrastique introduit par *f*. Or tel n'est pas le cas : (61b), (62b) et (63b). Le groupe P-Comp n'est pas un nom portant une marque de cas. Nous en concluons que les prépositions faibles ne sont pas des marqueurs casuels au sens strict.

Notons que lorsqu'un nom, comme *afu* 'chose', est introduit, la construction devient grammaticale (61c). La structure grammaticale en (62c) implique un élément d'origine arabe *bəlli* (Chaker, 1983, pp. 434-435). (63c) est rendue grammaticale par le déplacement d'un constituant *wh*.

- (61) a. *i-hḏr-iji-d* *f-ḥmurḥ-is*
 3MS-parler.PF-OI:1S-DIR sur-pays.EA-POSS:3S
 "Il m'a parlé de son pays."

- b. * i-hǝr-iji-d f-i-d j-usa
 3MS-parler.PF-OI:1S-DIR sur-COMP-DIR 3MS-venir.PF
 visé : "Il m'a parlé de ce qu'il est venu."
- c. i-hǝr-iji-d f-waʃu i-d j-usa
 3MS-parler.PF-OI:1S-DIR sur-chose.EA COMP-DIR 3MS-venir.PF
 "Il m'a parlé de la raison pour laquelle il est venu."
- (62) a. jə-çθβ-iji-d θaβrat^s f-θmurθ-is
 3MS-écrire.PF-OI:1S-DIR lettre sur-pays.EA-POSS:3S
 "Il m'a écrit une lettre sur son pays."
- b. * jə-çθβ-iji-d θaβrat^s f-i-d i-pp^{wəð}
 3MS-écrire.PF-OI:1S-DIR lettre sur-PRT-DIR 3MS-arriver.PF
 visé : "Il m'a écrit une lettre sur le fait qu'il est venu."
- c. jə-çθβ-iji-d θaβrat^s bəlli i-pp^{wəð}
 3MS-écrire.PF-OI:1S-DIR lettre sur PRT-arriver.PF
 "Il m'a écrit une lettre comme quoi il est arrivé."
- (63) a. nnuɁ-ən f-wəxxam
 se battre.PF-3MP sur-maison.EA
 "Ils se sont battus pour la maison."
- b. * nnuɁ-ən f-i jə-pp^{wəð} ð-amənzɔ
 se battre.PF-3MP sur-PRT 3MS-arriver.PF PRT-premier
 visé : "Ils se sont battus pour arriver les premiers."
- c. t^snaɁ-ən anwa ara j-awð-ən ð-amənzɔ
 se battre.INT-3MP quel.M PRT arriver.PART PRT-premier
 "Ils se battent pour arriver les premiers."

2.2.4 Les clitiques

Les prépositions faibles ne sont pas des marqueurs casuels au sens strict. Voyons maintenant si elles peuvent être analysées comme des marqueurs casuels au sens large. (Abney, 1987) propose que les traits grammaticaux du nom soient réalisés syntaxiquement sur une tête D distincte du nom. Dans cette perspective, les prépositions faibles peuvent être analysées comme des éléments de type D, et leur position à la gauche du complémenteur peut être attribuée à un processus de cliticisation de D vers Comp.

Prenons comme exemple la phrase (47a), répétée ci-dessous en (64). Selon l'hypothèse que nous examinons maintenant, la préposition faible a son origine à la gauche du nom, en [t_i]. Suite à un processus de cliticisation, elle est attachée au complémenteur. Cette hypothèse ne rencontre pas les problèmes présentés dans la section précédente.¹⁸

- (64) [[t_i] anta θaqəndurθ] s_i- -i-t^s i-zra
 quel.F.EL chemise.EL avec PRT-OD:3FS 3MS-voir.PF
 "Avec quelle chemise/robe l'a-t-il vue?"

Les prépositions faibles sont des éléments phonologiquement "faibles" - elles consistent en un seul segment - et elles sont réalisées sur un hôte. L'hypothèse qu'il s'agisse de clitiques est donc à première vue tout à fait plausible. Cependant, nous allons voir que la distribution des prépositions faibles est différente de celle des clitiques. Nous en concluons que les prépositions faibles ne peuvent pas être analysées comme des clitiques au sens propre.

En berbère de Chemini, comme dans les autres langues berbères, les clitiques pronominaux et les particules directionnelles s'attachent toujours à la droite de leur hôte (cf. entre autres (Dell et Elmedlaoui, 1989, 1991, Ouhalla, 2005, Shlonsky, 2004). Ils apparaissent dans deux contextes : (a) enclitiques sur le verbe dans une proposition à verbe initial, et (b) enclitiques sur la particule initiale de proposition s'il y en a une (c'est-

¹⁸ Notons que *anta θaqəndurθ* est à l'état libre dans l'exemple (64). Ce fait est problématique pour l'analyse évaluée dans cette section : s 'avec' régit en effet l'état d'annexion. J'ignore ce point pour le moment : mon objectif ici est de présenter clairement les grandes lignes de l'analyse des P faibles comme marqueurs casuels au sens large.

à-dire directement à la droite de $u(r)$, i , ara et $a(\delta)$). Les clitiques objets indirects, objets directs et directionnels apparaissent dans un ordre fixe ; leur distribution est résumée en (65) et quelques exemples sont donnés en (66).

(65) Distribution des clitiques :

- a. $V - CL_{OI} - CL_{OD} - CL_{DIR}$
- b. $\{u(r), i, ara, a(\delta)\} - CL_{OI} - CL_{OD} - CL_{DIR} \quad (...) \quad V$

(66) a. $jə-ʒra-θ$

3MS-voir.PF-OD:3MS

"Il l'a vu."

b. $pp^w i-κ-as-θ-id$

apporter.PF-1S-OI:3S-OD:3MS-DIR

"Je le lui ai apporté."

c. $mənhu \quad i-θ \quad jə-ʒra-n$

qui COMP-OD:3MS voir.PART

"Qui l'a vu?"

d. $a-s-θ-id \quad awi-κ$

PRT-OI:3S-OD:3MS-DIR apporter.AOR-1s

"Je le lui apporterai."

Les prépositions faibles ont une distribution différente (67). En particulier, elles apparaissent à la gauche, et non à la droite, du complémenteur (68).¹⁹

(67) Distribution des P faibles :

- a. $P - N_{EA}$
- b. $P - \{u(r), i, ara\}$

¹⁹ En (68), la préposition faible s 'avec' ne peut pas être considérée comme hôte du groupe de clitiques $i-ts$. En effet, si elle était hôte, elle apparaîtrait sous sa forme forte, donnée plus bas en (69).

- (68) anta θaqəndurθ s-i-t^s i-zra
 quel.F.EL chemise.EL sur-COMP-OD:3FS 3MS-voir.PF
 "Avec quelle chemise/robe l'a-t-il vue?"

Les prépositions faibles peuvent prendre elles-mêmes des pronoms clitiques comme complément. En berbère de Chemini, les prépositions faibles apparaissent sous une forme "forte" lorsqu'elles sont hôtes de clitique. Ces formes sont données en (69), et quelques exemples figurent en (70).²⁰

(69) <u>préposition faible</u>	<u>avec clitique pronominal</u>	
n	(j)in	'génitif'
f	fəll	'sur'
s	jis	'avec (instr.)'
ð	jið	'avec (com.)'
g	ðɣ	'dans'

- (70) a. að i-sərs θaβrat^s f-taβla
 PRT 3MS-posér.PF lettre.EL sur-table
 "Il posera la lettre sur la table."

- b. að i-sərs θaβrat^s fəll-as
 PRT 3MS-posér.PF lettre.EL sur-OI:3S
 "Il posera la lettre dessus."

- c. jə-ttʃa ð-wultma-s
 3MS-manger.PF avec-soeur.EA-OI:3S
 "Il a mangé avec sa soeur."

²⁰ La préposition du datif *i* ne figure pas dans le tableau (69) car il existe un paradigme spécifique de clitiques pronominaux objets indirects.

- d. jə-ttʃa jið-əs
 3MS-manger.PF avec-OI:3S
 "Il a mangé avec elle."

La distribution des groupes P-clitique est variable selon les langues berbères. En chleuh par exemple, les groupes P-clitique sont eux-mêmes des clitiques : ils apparaissent enclitiques au complémenteur, à la droite des autres clitiques (71). Un exemple est donné en (72).

(71) Chleuh (Dell et Elmedlaoui, 1989), p.170)

CL:	datif	objet	déictique	adverbes	groupes prépositionnels
	1	2	3	4	5

(72) Chleuh

ur	a	di-	-s	i-	-ʃtta
neg	ar	avec	3s	3s	manger.IPF

"Il ne mange pas avec elle." (Dell et Elmedlaoui, 1989), p. 173, (22))

La situation en kabyle de Chemini est différente. Un groupe P-clitique se comporte comme un groupe prépositionnel plein, et non comme un clitique. En particulier, un groupe P-clitique n'est jamais attiré par les particules *u(r)*, *i*, *ara* et *a(ð)*. Ce fait est illustré en (73a-c). Les syntagmes prépositionnels *fəll-as* 'sur lui/elle' et *jið-əs* 'avec lui/elle' ne sont pas attirés par les hôtes de clitique *u(r)* et *a(ð)*. Ils apparaissent à la droite du verbe, et en sont séparés par le sujet. Les exemples en (73d-e) illustrent le fait que les groupes P-clitique peuvent apparaître en position initiale.

- (73) a. ur i-qqim-ara wəqʃiʃ fəll-as
 NEG 3MS-être assis.PFNEG-NEG garçon.EA sur-OI:3S
 "Le garçon ne s'est pas /n'était pas assis dessus."

- b. ur i-sətt^s-ara wəqfiʃ jið-əs
 NEG 3MS-manger.INT-NEG garçon.EA avec-01:3s
 "Le garçon ne mange pas avec lui/elle."
- c. a-t^s i-sərs wəqfiʃ fəll-as
 PRT-OD:3FS 3MS-posér.PF garçon.EA sur-01:3s
 "Le garçon la posera dessus."
- d. jið-əs i-d i-pp^{wəð}
 avec-01:3s COMP-DIR 3MS-arriver.PF
 "C'est avec lui/elle qu'il est arrivé."
- e. fəll-as i-t^s i-sərs
 sur-01:3s COMP-OD:3FS 3MS-posér.PF
 "C'est dessus qu'il l'a posée."

Les affixes possessifs constituent un cas particulier de groupe P-clitique : ils peuvent être directement suffixés au nom qu'ils modifient, ou bien apparaître dans diverses positions dans la phrase, par exemple en initiale (74). Notons que (74b) est une proposition avec une copule vide. Je reviendrai sur cette construction plus bas en section 2.2.6.2.

- (74) a. axxam-in-əs (= axxam-is)
 maison.EL-GEN-3s
 'sa maison'
- b. (j)in-əs wəxxam-ənni
 GEN-3s maison.EA-DEM
 "Elle est à lui/elle, cette maison."

(Chaker, 1983) décrit un autre parler kabyle (Aït Iraten, Grande Kabylie). Dans ce parler, les groupes P-clitiques ont la même distribution que dans celui de Chemini. Les exemples en (75a-b) illustrent le fait que les groupes P-clitique peuvent apparaître en initiale, et

l'exemple en (75c) montre que le groupe P-clitique n'est pas attiré par les particules phrastiques.

(75) a. fəll-as i θə-rna
 sur-oi:3s COMP 3FS-ajouter.PF
 "Elle est née juste après lui." (Chaker, 1983), appendice ligne 501)

b. ǝγ-əs i zəðɤ-ən
 dans-oi:3s COMP habiter.PF-3MP
 "Ils y habitaient." (Chaker, 1983), appendice ligne 532)

c. aǝ θ-ərr-əǝ ǝγ-əs irðən
 PRT 2S-mettre.AOR-2S dans-oi:3s blé.EL
 "pour que tu y mettes du blé" (Chaker, 1983), appendice ligne 111)

(Chaker, 1983) note que c'est seulement dans certains cas exceptionnels que les groupes P-clitique sont attirés par des particules phrastiques : *"le [...] phénomène d'attraction existe aussi, dans la langue poétique et dans certaines expressions figées ("scories diachroniques"), pour les syntagmes constitués d'une préposition + pronom affixe personnel."* (p. 140). Il donne l'exemple en (76b) comme variante poétique de la structure "normale" en (76a).

(76) a. aǝ i-ru fəll-i
 PRT 3MS-pleurer.AOR sur-oi:1s
 "Il pleurera sur moi = à mon propos."
 b. aǝ fəll-i i-ru
 PRT sur-oi:1s 3MS-pleurer.AOR
 "Il pleurera sur moi = à mon propos."

Pour résumer, les prépositions faibles et les clitiques du berbère de Chemini n'apparaissent jamais dans le même contexte. Les distributions respectives sont

données dans le tableau (77). J'en conclus que les prépositions faibles ne sont pas des clitiques.

(77) hôte	clitique au sens propre	prép faible (sauf génitif) ²¹
_C	-	+
C_	+	-
_T	-	-
T_	+	-
_V	-	-
V_	+	-
_N	-	+
N_	-	-

Les prépositions faibles du kabyle de Chemini ne peuvent donc pas être analysées comme des marqueurs casuels, ni au sens strict (c'est-à-dire comme traits du nom), ni au sens large (c'est-à-dire comme déterminants). Ce sont des prépositions.

2.2.5 Sur la déficience phonologique des prépositions faibles

Commençons par la question suivante : pourquoi les prépositions faibles doivent-elles être cliticisées ? Dans cette section, nous verrons d'abord que les prépositions faibles font partie du domaine phonologique de leur complément nominal. Ensuite, sur la base de l'analyse des séquences P-Comp, nous mettrons en évidence que ce sont des segments flottants : elles ne possèdent pas de matériel squelettal propre. Ainsi, les prépositions faibles sont des marqueurs flottants, i.e., selon la terminologie présentée dans la première partie de ce chapitre, des exposants de type (38c). Elles ont besoin d'un hôte qui leur fournisse le matériel squelettal nécessaire à leur réalisation phonétique. Le choix de l'hôte est conditionné par des principes syntaxiques, et non phonologiques.

²¹ Rappelons que le génitif est particulier à deux égards : il est enclitique sur le nom dans les constructions possessives et il n'est pas proclitique sur le complémenteur.

2.2.5.1. Prépositions faibles et domaines phonologiques

Nous observons diverses assimilations entre une préposition faible et son complément nominal. Quelques exemples impliquant les éléments *n* (génitif), *g* ‘dans’ et *f* ‘sur’ sont donnés en (78), et illustrés en (79).

- (78) a. *n* (génitif) : (i) /n + θ/ → [tt] *assimilation optionnelle*
 (ii) /n + w/ → [pp^w] *assimilation obligatoire*
 b. *g* 'dans' : /g + j/ → [gg] *assimilation obligatoire*
 c. *f* 'sur' : /f + wV/ → [ffV] *assimilation obligatoire*
 où V = i, a, u

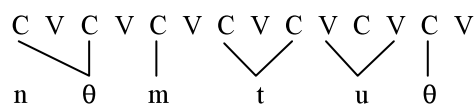
- (79) a. i. axxam n-θəqfiθ → [axxam ttəqfiθ]
 maison.EL de-fille.EA
 "la maison de la fille"
- ii. axxam n-wərgaz → [axxam pp^wərgaz]
 maison.EL de-homme.EA
 "la maison de l'homme"
- b. g-jəxxamən → [ggəxxamən], *[gjəxxamən]
 dans-maisons.EA
 "dans les maisons"
- c. f-waman → [ffaman], *[fwaman]
 sur-eau.EA
 "sur l'eau"

La règle (78a.i) s'applique typiquement à l'intérieur du mot, par exemple entre le radical et le marqueur de féminin /θ _ θ/ (80a), dans le domaine de la flexion verbale (80b), ainsi que dans celui de la cliticisation des pronoms objets (80c).

- (80) a. θ-aqɿiθ-θ → [θaqɿiθ] 'fille'
 F-garçon-F
 θ-uɿɿən-θ → [θuɿɿətt] 'chacal fém.'
 F-chacal-F
- b. xəðm-n → [xəðmən] 'ils ont travaillé'
 travailler.PF-3MP
 xəðm-n-θ → [xəðmətt] 'elles ont travaillé'
 travailler.PF-3MP-F
- c. jið-sən → [jiðsən] 'avec eux'
 avec-3MP
 jið-sən-θ → [jiðsətt] 'avec elles'
 avec-3MP-F

En berbère de Chemini, θ gémine sous la forme [tt] et w sous la forme [pp^w]. Les assimilations entre la préposition n et la consonne initiale du nom qui la suit sont donc à analyser toutes les deux comme une gémination de la consonne initiale du nom.

- (81) a. n + θməttuθ → [ttməttuθ]²²



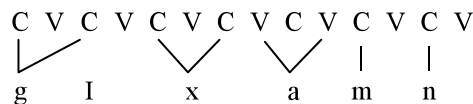
²² Les voyelles périphériques du berbère sont phonologiquement longues (Bendjaballah, 1999, 2005).

b. $n + wərgaz \rightarrow [p^wərgaz]$

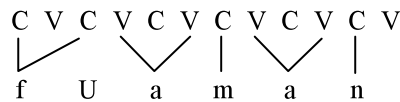


Examinons maintenant les prépositions *g* ‘dans’ et *f* ‘sur’. Celles-ci déclenchent les assimilations (78b) et (78c). Les représentations phonologiques de ces assimilations figurent en (82) : le contenu segmental de la préposition se propage sur la position occupée par le glide initial de nom, qui est désassocié.

(82) a. $g + jəxxamən \rightarrow [ggəxxamən]$



b. $f + waman \rightarrow [ffaman]$



Les assimilations décrites ci-dessus témoignent de la formation d’un domaine unique entre la préposition faible et son complément nominal. A nouveau, la préposition *n* (génitif) a un comportement particulier : elle déclenche une assimilation régressive, tandis que les autres prépositions faibles déclenchent une assimilation progressive.

2.2.5.2. Les prépositions faibles sont des éléments flottants

Les prépositions faibles sont réalisées à l’initiale du domaine phonologique du nom. Ceci implique qu’il y ait une position consonantique à la gauche du nom qui héberge le matériel segmental de la préposition. Deux configurations sont envisageables : soit la préposition faible possède une position consonantique squelettale intrinsèque (83a), i.e. est un marqueur de type (38a), soit la position consonantique à laquelle la préposition faible est associée est apportée par le nom (83b), et la préposition est un marqueur de type (38c).

- (83) a. $[\underset{\substack{| \\ f}}{C} \underset{\substack{| \\ \theta}}{V} [\underset{\substack{| \\ \theta}}{N} \underset{\substack{| \\ f}}{C} V \dots]]$ b. $[\underset{\substack{| \\ f}}{[\underset{\substack{| \\ \theta}}{N} C V C V \dots] }]$

Afin de décider en faveur de l'une ou l'autre de ces options, nous allons examiner les séquences P-Comp. Nous verrons que les prépositions faibles ne possèdent pas de matériel squelettal intrinsèque : ce sont des segments flottants (hypothèse (83b)).

Considérons tout d'abord le comportement de la préposition *ar* 'vers'. *ar* gouverne exclusivement l'état d'annexion (84a). Dans une proposition interrogative, *ar*, comme les prépositions faibles, est attirée à la gauche du complémenteur. Cependant, dans ce contexte, elle n'est pas réalisée sous sa forme normale (84b). Elle fait surface sous une forme faible, *ɛr* (84c).

- (84) a.
- | | | | | | |
|--|--------------|--|------|--|------------|
| | θ-ruħ | | ar | | wasif |
| | 3FS-aller.PF | | vers | | rivière.EA |
- "Elle est allée vers la rivière."
- b. *
- | | | | |
|--------|------------|-----------|--------------|
| anwa | asif | ar-i | θ-ruħ |
| quel.M | rivière.EL | vers-COMP | 3FS-aller.PF |
- c.
- | | | | |
|--------|------------|-----------|--------------|
| anwa | asif | κ-i | θ-ruħ |
| quel.M | rivière.EL | vers-COMP | 3FS-aller.PF |
- "Vers quelle rivière est-elle allée?"

Examinons maintenant le comportement de la préposition *i* (datif). Cette préposition est la seule préposition faible consistant en une voyelle, et non en une consonne. Lorsqu'elle apparaît à la gauche du complément, elle ne fait pas surface sous sa forme normale (85b). A la place, c'est un segment consonantique, *m*, qui apparaît (85c). A nouveau, nous observons que le segment réalisé à la gauche du complément consiste en une seule consonne.

- (85) a. j-uçr-as-θ i-θəqʃiʃθ
 3MS-voler.PF-OI:3S-OD:3MS DAT-fille.EA
 "Il l'a volé à la fille."
- b. * anta θəqʃiʃθ i(j)-i-θ j-uçər
 quel.F fille.EL DAT-COMP-OD:3MS 3MS-voler.PF
- c. anta θəqʃiʃθ m-i-θ j-uçər
 quel.F fille.EL m-COMP-OD:3MS 3MS-voler.PF
 "A quelle fille l'a-t-il volé?"

Les compléments *i* et *ara* sont à voyelle initiale, leur représentation phonologique comporte donc une position consonantique vide initiale (86). Or, à la gauche du complément, les prépositions doivent être réalisées sous la forme d'un seul segment consonantique. La réalisation d'une préposition à la gauche du complément semble donc directement dépendre du matériel squelettal disponible sur le complément.

- (86) a. *i* b. *ara*
-

Examinons maintenant les prépositions faibles consistant en un segment consonantique en (87) et (88).

- (87) a. anwa axxam g-i-t^s ruʒa-n
 quel.M.EL maison.EL dans-COMP-OD:3FS attendre.PF-3MP
 "Dans quelle maison l'ont-ils attendue?"
- b. anta θəqəndurθ s-i-t^s i-ʒra
 quel.F.EL chemise.EL avec-COMP-OD:3FS 3MS-voir.PF
 "Avec quelle chemise l'a-t-il vue?"

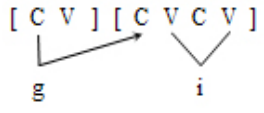
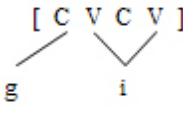
- c. anwa ak^wərsi f-i θə-qqim
 quel.M.EL chaise.EL sur-COMP 3FS-être assis.PF
 "Sur quelle chaise s'est-elle assise?"

- (88) a. anwa axxam g-ara-t^s raʒu-n
 quel.M maison.EL dans-COMP-OD:3FS attendre.AOR-3MP
 "Dans quelle maison vont-ils l'attendre?"

- b. anta θaqəndurθ s-ara-t^s i-zər
 quel.F chemise.EL avec-COMP-OD:3FS 3MS-voir.AOR
 "Avec quelle chemise/robe va-t-il la voir?"

- c. anwa ak^wərsi f-ara θə-qqim
 quel.M chaise.EL sur-COMP 3FS-s'asseoir.AOR
 "Sur quelle chaise va-t-elle s'asseoir/être assise?"

Dans tous les cas, la préposition est réalisée sous la forme d'une consonne simple, elle n'est jamais réalisée géminée. Si la préposition faible possédait une position squelettale consonantique propre, comme en (89a), nous nous attendrions à ce qu'elle gémisse devant le complémenteur.²³ Le fait que la préposition ne gémisse pas indique que le segment de la préposition est lié à une seule position C, la position initiale du complémenteur (89b). Nous en concluons que les prépositions faibles sont des *marqueurs flottants*.

- (89) a. *  b. 
 * [ggɪ] [gɪ]

²³ Les géminées initiales sont autorisées en berbère. L'absence de gémination dans les séquences P_{faible} - Comp ne peut donc pas être attribuée à une interdiction générale des géminées initiales dans la langue.

2.2.6 Support gabaritique et projection syntaxique

La représentation phonologique des prépositions faibles ayant été précisée, examinons maintenant plus en détail leur statut syntaxique. Les prépositions faibles ne se combinent jamais avec des noms à l'état libre : elles régissent exclusivement des noms à l'état d'annexion. L'analyse des P faibles comme marqueurs casuels pouvait expliquer simplement ce fait : le marqueur d'état libre et les prépositions faibles étant des marqueurs casuels, ils sont en distribution complémentaire. Cependant, l'analyse des prépositions faibles comme marqueurs casuels a dû être abandonnée. Il nous faut donc maintenant trouver une nouvelle façon d'expliquer le fait que l'état libre est exclu après une préposition faible.

Pour commencer, (90) illustre quelques exemples représentatifs de noms masculins et féminins à l'état libre et à l'état d'annexion.

(90)	<u>état libre</u>	<u>état d'annexion</u>	
<u>masc.</u>	amʃiʃ	wəmʃiʃ	'chat'
	axxam	wəxxam	'maison'
<u>fém.</u>	θakənʒawθ	θkənʒawθ	'cuillère'
	θaməʔtuθ	θməʔtuθ	'femme'

Les noms masculins à l'état libre sont à voyelle initiale. Toutes les autres formes du paradigme en (90) sont à consonne initiale. D'un point de vue phonologique, la configuration préposition faible + nom masculin à l'état libre est donc bien formée : la position C initiale du nom est disponible pour accueillir le segment flottant de la préposition. Il n'y a aucune raison phonologique d'exclure la séquence *gaxxam*. Cependant (91) est agrammatical. La configuration préposition faible + nom à l'état libre doit donc être exclue pour des raisons morphosyntaxiques, et non phonologiques.

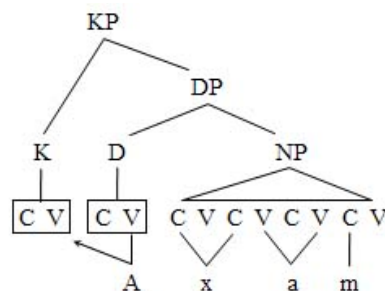
(91)	* C V C V C V C V C V
	g a x a m

Dans la suite, je me fonde sur une propriété cruciale du modèle développé dans la partie précédente : un segment ne peut projeter une tête en syntaxe que s'il est associé au niveau squelettal des représentations phonologiques (section 2.1.1.3). J'argumente que les prépositions faibles ne peuvent projeter la catégorie P que lorsqu'elles sont affixées à un nom à l'état d'annexion. Un nom masculin à l'état libre possède certes une position squelettale C(onsonantique) vide, mais cette position exprime déjà une catégorie syntaxique, K. De ce fait, la projection de P est impossible. Enfin, je reviendrai sur les séquences P-Comp et conclurai que la préposition faible à la gauche du complémenteur n'a pas de réalité syntaxique distincte de Comp : elle exprime un trait syntaxique de Comp (cf. (Emonds, 1985)).

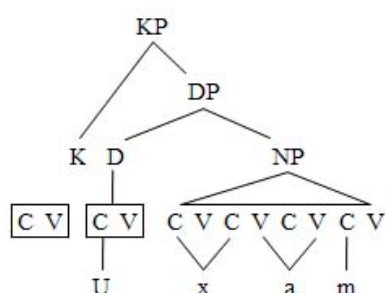
2.2.6.1. *Emploi prénominal : P-K-D-N*

Examinons les représentations du nom *axxam* 'maison' à l'état libre et à l'état d'annexion en (92). A l'état libre (92a), le cas (K) est exprimé phonologiquement par la position CV initiale du gabarit nominal (exposant de type 38b). A l'état d'annexion, le cas est exprimé par une catégorie vide (92b) (exposant de type 38d). La position CV initiale du gabarit nominal reste alors vide (configuration de type 39b).

(92) a. *axxam* 'maison' (EL)

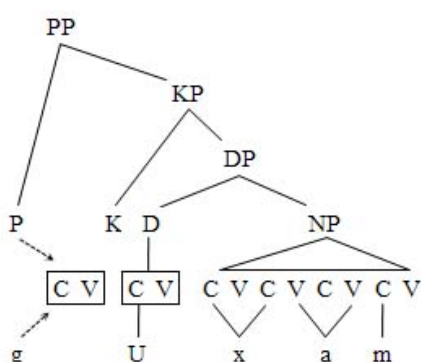


b. *wəxxam* ‘maison’ (EA)



Ainsi, en (92b), la position CV initiale du gabarit n'est identifiée ni segmentalement, ni syntaxiquement. Si nous faisons l'hypothèse que cette position est tout de même présente dans le gabarit du nom (Bendjaballah et Haiden, 2005b), nous prédisons qu'elle peut servir d'hôte à des marqueurs flottants. En effet, les marqueurs flottants possèdent une catégorie syntaxique et du matériel segmental, mais pas de support squelettal (marqueur de type 38c). Dans la section précédente, nous avons vu que les prépositions faibles ont exactement ces propriétés. Je propose donc que la position CV initiale vide d'un nom à l'état d'annexion serve de support gabaritique aux prépositions faibles. Une préposition faible attribue une interprétation à la fois segmentale et syntaxique à cette position (93).

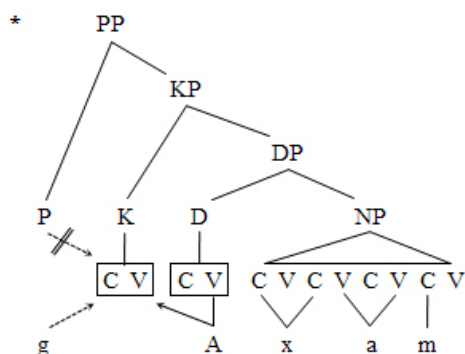
(93) *g-wəxxam* ‘dans la maison’



A l'état libre, la position CV initiale du nom a une interprétation à la fois syntaxique et segmentale. Au niveau segmental, elle est identifiée par la voyelle initiale du nom, au niveau syntaxique, elle est interprétée comme un marqueur de cas K. Cette configuration offre de la place pour le matériel segmental de la préposition faible, la position C(onsonantique) initiale, mais non pour son interprétation syntaxique. L'affixation d'une préposition faible à un nom à l'état libre équivaldrait à l'effacement syntaxique de la

préposition faible (94). Par conséquent, les prépositions faibles ne sont pas affixées aux noms à l'état libre.

(94) * gaxxam

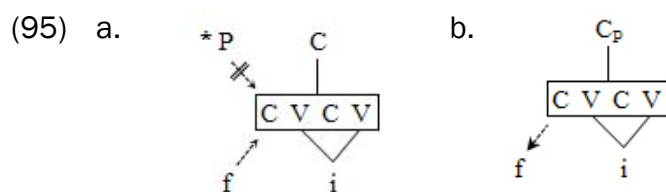


2.2.6.2. Emploi avec Comp

Selon l'analyse détaillée jusqu'à présent, une préposition faible n'est visible en syntaxe que si elle a accès à un support squelettal sans fonction syntaxique intrinsèque. Or, lorsqu'une préposition faible est affixée au complémenteur, elle est associée à la position C(onsonantique) initiale du gabarit du complémenteur (section 2.2.5.2), c'est-à-dire à une position CV qui exprime la catégorie Comp. La préposition faible n'est donc pas visible syntaxiquement, et cette configuration devrait être agrammaticale (95a). Pourquoi ne l'est-elle pas ?

Suivant (Emonds, 1985), je fais l'hypothèse que les P et les Comp ont la même catégorie syntaxique, à savoir P. Une préposition faible peut donc exprimer un trait du complémenteur (95b) tout comme la racine exprime les traits de V (voir la section 2.1.3). La combinaison *g-i* deviendra alors un complémenteur local, comparable sémantiquement au *où* du français dans "Je vois où ça nous mène." par exemple. En d'autres termes, je propose que l'affixation de la préposition faible au complémenteur constitue un cas de morphologie compositionnelle, et non d'incorporation syntaxique. De ce fait, je m'attends à ce que les combinaisons P-Comp présentent une tendance à la lexicalisation, ce qui est effectivement noté par (Chaker, 1983).²⁴ Dans la suite j'examinerai quelques implications syntaxiques de cette analyse.

²⁴ "Par leur stabilité et leur fréquence dans certains parlers, ces complexes sont certainement l'indice - d'une tendance à la constitution d'un paradigme de supports spécifiques,



Examinons tout d'abord en (96) les trois possibilités de former une question dont le constituant *wh* est en relation avec une préposition faible.

- (96) a. **g-wənwa** axxam i-t^s ruʒa-n
dans-quel.M.EA maison.EL COMP-OD:3FS attendre.PF-3MP
"Dans quelle maison l'ont-ils attendue?"
- b. **g-wənwa** axxam **g-i-t^s** ruʒa-n
dans-quel.M.EA maison.EL dans-COMP-OD:3FS attendre.PF-3MP
"C'est dans quelle maison qu'ils l'ont attendue ?"
- c. **anwa** axxam **g-i-t^s** ruʒa-n
quel.M.EL maison.EL dans-COMP-OD:3FS attendre.PF-3MP
"Quelle est la maison où ils l'ont attendue?"

En (96a), la préposition faible est affixée au mot interrogatif *anwa*, qui apparaît donc à l'état d'annexion : *wənwa*. La structure peut être analysée sans problème comme un cas de mouvement *wh* du PP. En (96b), la préposition est redoublée sur le complémenteur. Selon mon analyse, seule la première préposition, celle qui est préfixée au mot interrogatif, est visible syntaxiquement. La deuxième occurrence de l'item lexical *g* 'dans' appartient au paradigme du complémenteur. Cette construction pourrait être analysée en termes de mouvement *wh* du PP, mais une telle analyse ne peut répondre à la question

- d'une tendance à la pronominalisation des éléments supports d'origine prépositionnelle; combiné au nominal *ay/i*, le support prépositionnel acquiert des latitudes fonctionnelles similaires à celles des "relatifs" pronominaux. [...] Du point de vue de la dynamique de la langue, ces formes représentent l'amorce d'une homogénéisation du paradigme." (Chaker, 1983), p.398).

suivante : pourquoi la forme préfixée du complémenteur est-elle choisie dans ce contexte ? Enfin, en (96c), l'item lexical *g* 'dans' n'apparaît qu'une fois, sur le complémenteur. Cette structure ne peut pas être analysée comme un cas de mouvement *wh* du PP car, par hypothèse, le *g* préfixé au complémenteur n'est pas visible syntaxiquement. Le constituant *wh anwa axxam* 'quelle maison' doit donc être un DP indépendant. Et, de fait, il est réalisé à l'état libre. Une nouvelle question se pose alors : comment le DP *anwa axxam* acquiert-il sa fonction grammaticale de complément de lieu du verbe *razu* 'attendre' ?

Voici l'analyse que je propose : toutes les structures impliquant une préposition légère préfixée sur le complémenteur sont des constructions clivées avec une copule vide. Notons d'abord que les constructions à copule vide existent en berbère (97a). En particulier, ces constructions peuvent aussi fonctionner comme interrogatives (97b).

- (97) *jin-əs wəxxam-ənni*
 GEN-OI.3S maison.EA-DÉM
 a. "Cette maison est à lui."
 b. "Est-ce que cette maison est à lui?"

L'analyse des phrases (96b-c) figure en (98a-b) respectivement. OP est un complément de lieu vide du verbe *razu* 'attendre'. Cet opérateur se déplace jusqu'au [Spec, CP] (cf. (Chomsky, 1986)). Dans cette position, il est identifié référentiellement par l'expression *wh gwənwa/anwa axxam*. Cette expression *wh* est elle-même complément de la copule vide (et non du verbe *razu* 'attendre'). Afin de clairement illustrer l'analyse syntaxique proposée, j'emploierai des gloses informelles dans ce qui suit.

- (98) a. [_v *ø*] *g-wənwa axxam* [OP_i *g-i-t^s ruza-n* [_{t_i}]]
 c'est dans quelle maison où-cl ont attendu-cl
 "Dans quelle maison l'ont-ils attendue ?"

- b. [_v Ø] anwa axxam [OP_i g-i-t^s ruʒa-n [t_i]]
 c'est dans quelle maison où-cl ont attendu-cl
 "C'est dans quelle maison qu'ils l'ont attendue ?"

Considérons maintenant les constructions à long mouvement *wh*.²⁵ Outre (99), qui ne comporte pas de complémenteur complexe, et qui constitue donc un cas standard de mouvement cyclique du PP (100), toutes les constructions en (101) sont grammaticales.

- (99) g-wənwa axxam i-s θə-qqar nə-t^sraʒu-t^s
 dans quelle maison que-cl elle-dit nous-attendons-cl
 "Dans quelle maison pense-t-elle que nous l'attendons?"

- (100) [_{CP} [g-wənwa axxam]_i i-s θə-qqar [_{CP} [t_i] [nə-t^sraʒu-t^s [t_i]]]]
 dans-quelle maison que-cl elle-dit nous attendons-cl

- (101) a. anwa axxam gi-s θə-qqar i-t^s nə-t^sraʒu
 quelle maison où-cl elle-dit que-cl nous-attendons
 b. g-wənwa axxam gi-s θə-qqar i-t^s nə-t^sraʒu
 dans-quelle maison où-cl elle-dit que-cl nous-attendons
 c. anwa axxam i-s θə-qqar gi-t^s nə-t^sraʒu
 quelle maison que-cl elle-dit où-cl nous-attendons
 d. g-wənwa axxam i-s θə-qqar gi-t^s nə-t^sraʒu
 dans-quelle maison que-cl elle-dit où-cl nous-attendons

Ainsi, selon l'analyse proposée ici, un complémenteur complexe est l'indice d'une construction clivée à copule vide. Dans les exemples (101a-b), la copule est la tête de la proposition principale (102). La relation entre le constituant *wh* *anwa axxam* 'quelle maison' et la position définissant sa fonction grammaticale (complément de lieu du verbe *raʒu* 'attendre') est établie par un opérateur vide. Cet opérateur se déplace, et atteint la

²⁵ Voir (Ouhalla, 1993) pour d'autres cas de longs mouvements *wh* en berbère.

position de spécifieur du complémenteur complexe *gi*. Dans cette position, l'opérateur est identifié référentiellement par le constituant *wh*. La même analyse s'applique à la structure (101b), où le complément de la copule vide est un PP.

- (102) a. [_v $\emptyset$] anwa axxam [**OP** g-i-s $\theta\acute{a}qqar$ [[t] i-t^s nət^srazu [t]]]
 "Dans quelle maison pense-t-elle que nous l'attendons?"
- b. [_v $\emptyset$] g-wənwa axxam [**OP** g-i-s $\theta\acute{a}qqar$ [[t] i-t^s nət^srazu [t]]]
 "Dans quelle maison pense-t-elle que nous l'attendons?"

La situation en (101c-d) est différente. La construction clivée est enchâssée sous le verbe *penser*. Nous observons alors deux mouvements indépendants : l'opérateur vide, dont l'origine est le complément du verbe *attendre*, monte jusqu'à la gauche du complémenteur complexe, où il est identifié par le constituant *wh*, qui, lui, monte vers le début de la phrase.

- (103) a. [anwa axxam]_i i-s $\theta\acute{a}qqar$ [_v $\emptyset$] [t_i] [**OP**_j g-i-t^s nət^srazu [t_j]]
 "Dans quelle maison pense-t-elle que nous l'attendons?"
- b. [g-wənwa axxam]_i i-s $\theta\acute{a}qqar$ [_v $\emptyset$] [t_i] [**OP**_j g-i-t^s nət^srazu [t_j]]
 "Dans quelle maison pense-t-elle que nous l'attendons?"

Les cinq configurations grammaticales en (100) et (101) sont en contraste avec les deux configurations suivantes, agrammaticales :

- (104) a. * g-wənwa axxam gi-s $\theta\acute{a}$ -qqar g-i-t^s nə-t^srazu
 dans-quelle maison oû-cl elle-dit oû-cl nous-attendons
- b. * anwa axxam gi-s $\theta\acute{a}$ -qqar g-i-t^s nə-t^srazu
 quelle maison oû-cl elle-dit oû-cl nous-attendons

Ces deux exemples comportent deux complémenteurs complexes, et ils sont agrammaticaux. Considérons en (105a) la dérivation de (104a), et le même

raisonnement s'appliquera à (105b). Puisque le complémenteur le plus élevé est un complémenteur complexe, nous devons postuler une clivée à ce niveau de la phrase. Cependant, nous devons aussi postuler une clivée enchâssée sous le verbe *penser*. Par conséquent, (105a) comporte deux chaînes de mouvement, et deux opérateurs indépendants. L'origine de OP_j est le complément de lieu du verbe *attendre*, OP_i est le complément de la copule vide sous *penser*. Ce dernier opérateur vide est identifié par une catégorie pleine, le PP interrogatif. En revanche, le premier opérateur vide (OP_j) ne l'est pas. Par conséquent, la construction est agrammaticale.

- (105) a. *gwənwə axxam [_v∅][OP_i gis θəqqar [_v∅] [t_i] [OP_j git^s nət^srazu [t_j]]]
 b. *anwa axxam [_v∅] [OP_i gis θəqqar [_v∅] [t_i] [OP_j git^s nət^srazu [t_j]]]

En me fondant sur des considérations phonologiques, j'ai argumenté que les prépositions légères affixées sur le complémenteur ne sont pas des têtes syntaxiques. Cette hypothèse m'a contraint à analyser les constructions interrogatives comme des clivées, ce qui m'a alors permis de prédire l'agrammaticalité de certaines constructions à long mouvement *wh*. J'en déduis que l'hypothèse est correcte : les prépositions faibles affixées au complémenteur ne sont pas des têtes syntaxiques.

Pour conclure cette section, je présente un argument syntaxique indépendant en faveur de cette analyse.²⁶ Si nous faisons l'hypothèse que les prépositions légères sur le complémenteur sont des têtes syntaxiques indépendantes, c'est-à-dire des prépositions orphelines, nous devrions analyser l'exemple grammatical (101c) comme en (106).

- (106) [anwa axxam]_i is θəqqar [_{CP} [_{PP} g [t_i]]_j it^s nət^srazu [t_j]]

Selon (106), le PP dont la tête est *g* 'dans' monte vers la position [Spec, CP] intermédiaire, où la préposition est laissée en arrière (*stranded*), tandis que le DP continue son mouvement vers le début de la phrase. Or (Merchant, 2002, Postal, 1972) argumentent que, très généralement, de telles structures ne sont pas grammaticales : une préposition orpheline ne peut pas être laissée dans une position [Spec, CP]

²⁶ Cet argument est fondé sur une remarque due à Uri Shlonsky.

intermédiaire. (101c) devrait être agrammatical, au même titre que l'exemple anglais (107a) tiré de (Merchant, 2002). Une préposition orpheline doit être laissée dans sa position initiale (107b).

- (107) a. * Who₁ do you think [_{CP} [_{PP} for t₁]₂ [_{IP} she bought it t₂]]?
 b. Who do you think [_{CP} [t₁] [_{IP} she bought it [for [t₁]₂]]?
 "Pour qui penses-tu qu'elle l'a acheté? "

L'analyse proposée explique la grammaticalité des exemples de type (101c) en kabyle de Chemini. La préposition à la gauche du complémenteur n'est pas une tête syntaxique, mais un complément lexical du complémenteur. La structure correcte est (60a), et non (64), et nous ne sommes pas obligés de remettre en cause la généralisation de Postal.

2.3 Conclusion

Ce chapitre a été consacré aux aspects descriptifs et formels du langage. Le modèle esquissé dans la première partie propose une solution au long débat concernant le statut des positions infixales en morphologie (Guerssel et Lowenstamm, 1990, Lowenstamm et Kaye, 1986, Lowenstamm, 2003). Le mécanisme proposé peut être brièvement résumé comme suit. Tout d'abord, l'opération d'*assemblage* est définie de façon à autoriser des marqueurs infixaux. Ensuite, les gabarits imposent un ordre linéaire aux marqueurs. Enfin, les positions infixales ne posent plus problème à l'association des segments radicaux car, dans le modèle, les segments ne « cherchent » pas des positions, mais ce sont les positions qui sont épelées par les segments. Les points d'ancrage squelettaux munis de la valeur morpho-syntaxique V(erbe) sont interprétés par les segments de la racine. Les positions infixales sont ainsi immédiatement exclues : leur interprétation doit attendre une étape dérivationnelle ultérieure. Ce modèle a en outre l'avantage de maintenir une stricte modularité de la phonologie et de la syntaxe, en dehors du lexique, là où diverses théories affaiblissent considérablement le principe de modularité. Pour prendre un seul exemple, (Prince et Smolensky, 2004) : 42 et (McCarthy et Prince, 1993) : 102) autorisent des contraintes phonologiques à défaire les contraintes sur le placement des affixes.

La seconde partie illustre ma démarche empirique avec les prépositions légères du kabyle de Chemini. J'ai adopté trois hypothèses fondamentales. Tout d'abord, suivant (Guerssel, 1987; 1992), j'ai admis que le cas est une catégorie grammaticale, et donc une tête syntaxique. Ensuite, j'ai considéré qu'une préposition faible est un seul et même item lexical dans ses différents emplois : la préposition faible employée avec le nom est la même que celle employée avec le complémenteur. Enfin, j'ai adopté l'hypothèse qu'un item lexical a une catégorie grammaticale fixe : une préposition est une préposition dans tous ses emplois. Si l'on accepte ces hypothèses, les prépositions faibles du kabyle de Chemini ne peuvent être analysées comme des marqueurs casuels. En effet, les prépositions faibles à la gauche du complémenteur n'ont pas les propriétés des marqueurs casuels. J'ai analysé les prépositions faibles comme des prépositions, et ai imputé leur nature affixale à une propriété phonologique : les prépositions faibles sont des éléments sans support squelettal intrinsèque. Des considérations sur le gabarit des complémenteurs m'ont conduit à la conclusion que les prépositions faibles affixées au complémenteur ne peuvent pas être visibles en syntaxe : elles réalisent un trait de leur catégorie hôte, le complémenteur. Cette analyse m'a permis d'expliquer les restrictions portant sur le redoublement des prépositions faibles dans les constructions à longue extraction A'. Pour résumer : lorsqu'elle apparaît à la gauche du nom, une préposition faible est visible syntaxiquement ; le gabarit nominal lui fournit en effet une position CV sans valeur syntaxique prédéterminée (i.e. un objet de type (39b)). Lorsqu'elle apparaît à la gauche du complémenteur, la préposition faible n'est qu'un affixe lexical du complémenteur, sans projection syntaxique indépendante ; le complémenteur en effet ne fournit qu'une position consonantique faisant partie de son gabarit propre, mais pas de site gabaritique disponible pour une interprétation syntaxique indépendante.

3. Acquisition

Ce troisième et dernier chapitre est consacré à l'acquisition des pronoms clitiques du français. Il s'agit là d'un travail en cours s'articulant autour de deux grands axes :

- (a) l'étude de la production induite
- (b) l'étude de la compréhension en temps réel.

Ce chapitre présente les résultats empiriques du volet (a).

J'ai présenté ces résultats à plusieurs occasions : colloque annuel de la DGfS à Osnabrück (2009), réunions de l'action COST IS0804 à Amsterdam (2009) et Newcastle (2010), journées d'études sur la méthodologie du suivi du regard à Tours (2009) et à Groningen (CLEAR 2010).

J'ai par ailleurs encadré un mémoire de capacité d'orthophonie (Dupuy, 2009), et co-encadré un mémoire de M2R (Guimard, 2010) sur le sujet ; j'en co-encadre actuellement un autre (Léger, prévu pour 2011) sur l'utilisation du suivi du regard pour l'étude de la compréhension des clitiques objets en temps réel.

3.1 Introduction

Les pronoms clitiques objets du français ont fait l'objet de débats en théorie syntaxique depuis bien avant l'émergence du paradigme génératif. La grammaire de Port-Royal (Arnauld et Lancelot, 1660), par exemple, leur consacre un chapitre, dans lequel l'emploi clitique dit "avant le verbe" est explicitement distingué d'un emploi tonique dit "partout ailleurs". Certaines particularités de l'emploi clitique y sont décrites en détail. Dans le cadre génératif, la classification distributionnelle en pronoms toniques et pronoms clitiques est reprise dans l'ouvrage de référence, (Kayne, 1975), qui définit le domaine empirique ainsi que les questions théoriques au centre du débat. Empiriquement, le sujet des clitiques semble être épuisé : leurs propriétés sont bien connues. Cependant, un accord sur l'analyse théorique reste toujours à trouver : l'analyse de Kayne en termes de déplacement du clitique d'une position argumentale vers sa position préverbale est souvent modifiée, voire rejetée. Selon (Uriagereka, 1995), par exemple, le clitique n'est pas un argument au sens plein du terme, mais le déterminant d'un argument. Quand le

clitique se déplace, il laisse en arrière un syntagme nominal qui peut être prononcé ou non, en fonction d'autres contraintes en vigueur dans la langue. L'analyse d'Uriagereka rend compte du redoublement du clitique (attesté par exemple en espagnol), ainsi que du fait que, dans de nombreuses langues, les articles définis et les clitiques accusatif de troisième personne sont homophones. Ces analyses (de Port-Royal à Kayne et Uriagereka) partagent une hypothèse de base : le clitique est (ou fait partie d') un argument. Cette hypothèse est cependant contestée, et ce dans divers contextes théoriques (Borer, 1984, Jaeggli, 1982, 1986, Miller et Sag 1997, Roberge, 1990, Sportiche, 1996). Ces auteurs, entre autres, proposent que le clitique est un élément flexionnel, indiquant que le complément du verbe est soit absorbé par une règle lexicale (cadre HPSG), soit réalisé par une catégorie vide (cadre principes et paramètres). A ce jour, la question du statut argumental du clitique n'est pas résolue.

Les clitiques objets font également l'objet de débat du point de vue de l'acquisition. A la suite de (Clark, 1985, Hamann et al., 1996, Zesiger et al., in press) entre autres, il a été observé que les clitiques objet de troisième personne sont acquis tardivement, qu'ils constituent une source de difficulté en acquisition atypique saine (White, 1996) et pathologique (Jakubowicz et al., 1998, Tuller et al., 2011), et qu'ils restent difficiles pour les individus dysphasiques au-delà de l'adolescence (Tuller et al., 2007). Ce phénomène a reçu divers types d'explication, et le lecteur pourra utilement consulter (Grüter, 2006) pour une revue des analyses proposées.

Mon objectif, dans ce chapitre, est d'ordre à la fois empirique et méthodologique. Je me concentrerai sur un problème souvent noté, mais rarement traité, dans la littérature : les études sur l'acquisition des clitiques objet donnent des taux de production très variés pour des populations cependant comparables. Ainsi, pour le groupe d'enfants de 3 ans de développement typique (TD3), (Varlokosta et al, in prep) relèvent des taux de production entre 13,2% (Perez-Leroux et al., 2008) et 39,2% (Jakubowicz et Nash, to appear). Pour les TD4, les taux varient entre 41,7% (Perez-Leroux et al., 2008) et 88,1% (Chillier Zesiger et al., 2006). Des variations de cette ampleur sont aussi observées pour les TD5 et TD6. (Varlokosta et al, in prep) argumentent de façon convaincante que des différences dialectales ne peuvent pas à elles seules expliquer ces variations, car la variation reste importante pour la seule variante métropolitaine standard du français. Ces auteurs indiquent également que les taux donnés sont souvent accompagnés d'écarts-type élevés et que les tâches d'élicitation diffèrent. Pour remédier à ce problème, ils proposent une tâche censée être comparable dans les diverses langues à

clitique. Cependant, cette tâche a pour inconvénient de n'induire que des clitiques objet suivant un sujet DP. Or, la plupart des études sur le français portent sur la configuration la plus naturelle en langue parlée, à savoir les *groupes* de clitiques sujet et objet. Ainsi, l'étude de Varlokosta et al. permet de comparer les résultats à travers les langues, mais elle n'est pas directement compatible avec la littérature existante sur le français.

Dans ce chapitre, je propose une autre stratégie. J'admettrai que les taux de production de clitiques sont variables, et je chercherai à identifier des moyens permettant de dégager des généralisations sur la base de données variées. A cette fin, j'exploiterai des données déjà récoltées, et leur appliquerai de nouveaux échantillonnages et de nouvelles mesures. Ma stratégie principale consistera à réduire l'importance des moyennes dans l'interprétation des données, et à donner plus d'importance aux contrastes. Je poserai les questions suivantes, classées de la plus générale à la plus spécifique :

1. Les mesures de production, d'omission et de substitution mesurent-elles un développement unique, ou bien plusieurs développements ? S'il s'agit de plusieurs développements, ces développements sont-ils liés ou indépendants ?
2. Quel est le degré d'influence du regroupement des clitiques sur les taux de production, d'omission et de substitution ?
3. Quel est le statut des erreurs morphologiques ? Le taux de production de clitiques cible, qui par définition exclut les erreurs morphologiques, mesure-t-il le même phénomène que le taux de production, qui, lui, inclut les clitiques morphologiquement erronés ?
4. Quel est le statut des datifs grammaticaux ? Si un sujet emploie une construction avec un clitique datif à la place d'une construction avec clitique accusatif seul (par exemple "Il me pique la jambe." au lieu de "Il me pique."), doit-on comptabiliser cette réponse comme réponse cible ?

Je proposerai des réponses à ces questions en me fondant sur deux types de données : a) des données d'enfants à développement typique entre 4 et 11 ans et b) des enfants anglophones apprenant le français comme L2 dans un contexte d'immigration. Cette dernière population me conduira à poser les deux questions supplémentaires suivantes :

5. Les enfants apprenant le français en tant que L2 présentent-ils la même trajectoire de développement que les enfants à développement typique, ou bien présentent-ils des spécificités liées à l'acquisition L2 ? Si oui, lesquelles ?
6. Est-il possible d'identifier les enfants présentant un risque de trouble de développement linguistique à l'aide des clitiques objet ?

Je partirai du principe que deux mesures (produisant des taux différents) quantifient le même développement si, toutes choses égales par ailleurs, elles font ressortir les mêmes effets par tranche d'âge et si elles distinguent les mêmes tranches d'âge. Ce raisonnement sera appliqué de manière générale, et, tout particulièrement, pour répondre à la question 1.

Je me fonderai sur une opposition bien connue concernant les clitiques accusatif de troisième et de première personne et les réfléchis : les réfléchis émergent plus précocement que les accusatifs de première personne, qui précèdent à leur tour les accusatifs de troisième personne (Tuller et al., 2011). Ainsi, on peut observer des taux de production significativement différents pour ces trois fonctions grammaticales (réf, 1acc et 3acc) au cours du développement, puis ces différences évoluent, avant de disparaître complètement. Si deux mesures données font ressortir des oppositions différentes entre fonctions grammaticales pour une tranche d'âge, ou si ces oppositions évoluent de manière différente, alors elles mesurent des phénomènes différents.

J'aborderai la question 2. à l'aide d'un nouvel échantillonnage des données. Dans la littérature, les mesures de production, d'omission et de substitution par un élément tonique portent sur la totalité des réponses pertinentes. Or, certains sujets produisent des réponses à enchâssement. On trouve, par exemple, "Il veut m'écraser." à la place de "Il m'écrase." Ces réponses sont tout à fait grammaticales et pragmatiquement adaptées à la tâche, mais elles requièrent la séparation des clitiques sujet et objet, tandis qu'une réponse sans enchâssement requiert le regroupement des clitiques sujet et objet. On peut ainsi isoler deux échantillons de données, l'un comportant les réponses avec regroupement des clitiques, l'autre les réponses avec séparation des clitiques. Si les taux de production, d'omission et de substitution se développent de manière différente dans les deux échantillons, alors on aura montré que le regroupement des clitiques a une influence sur la trajectoire de l'acquisition des clitiques.

J'aborderai la question 3. à l'aide de deux nouveaux types de mesure. La première quantifie les clitiques objet cibles, quel que soit le statut du sujet : clitique (cible, erroné

ou omis) ou DP plein. La seconde quantifie les groupes cibles, c'est-à-dire les réponses comportant à la fois un clitique sujet cible et un clitique objet cible. Si ces deux mesures donnent des résultats qui diffèrent de ceux obtenus par la mesure traditionnelle de la production, on pourra en déduire que le choix de la forme morphologique est un facteur indépendant qui influence l'acquisition des clittiques. Par ailleurs, la comparaison de la production de clittiques objet cibles et de la production des groupes de clittiques cibles permettra de renforcer ou bien d'infirmer les observations concernant la question 2.

La question 4. concernant le statut des datifs grammaticaux pose un problème méthodologique très spécifique, mais tout aussi important. Partant du principe que ce sont les clittiques accusatif qui présentent un développement particulier et intéressant, la production d'une réponse avec clitique datif devrait être exclue du taux de production. Cependant, les datifs grammaticaux sont souvent inclus dans ce taux, l'argument étant qu'un clitique est bien produit. Afin d'étudier le statut des datifs grammaticaux, j'encoderai les mesures de production cible sous formes de paires de variables, l'une excluant les datifs, l'autre les incluant. Si l'une de ces variables fait ressortir un développement plus cohérent que l'autre, c'est celle-ci qui doit être retenue comme la plus performante. Sur la base des calculs effectués dans ce chapitre, je conclurai que les datifs grammaticaux doivent être comptabilisés comme réponses cibles.

Enfin, j'examinerai les questions 6. et 7. concernant les spécificités de la population L2 en comparant les résultats spécifiques de ce groupe à ceux du groupe d'enfants à développement L1 typique. Je m'attacherai à déterminer si la trajectoire d'acquisition en L2 dépend a) de l'âge de l'apprenant au moment du test, b) de l'âge de l'apprenant au moment du début de l'acquisition L2, c) de la durée d'exposition au français, d) du niveau de l'apprenant dans sa langue maternelle, ou bien enfin d'une combinaison de ces facteurs. J'observerai l'asymétrie suivante : pour

- les clittiques accusatif de troisième personne, les taux ont tendance à être corrélés à la durée d'exposition et au niveau en langue maternelle
- les clittiques accusatifs de première personne et les réfléchis, les taux ne sont jamais corrélés au niveau en langue maternelle ; en revanche, ils ont tendance à être corrélés à la durée d'exposition et à l'âge chronologique des enfants au moment du test.

Je définirai un indicateur absolu de trouble linguistique, i.e. une mesure ne dépendant que peu ou pas du tout des variables d'âge et de durée d'exposition, et corrélée exclusivement au niveau en langue maternelle. Je montrerai que cette mesure permet d'isoler un groupe d'individus se comportant très différemment des autres.

D'autres mesures seront considérées comme candidats au statut d'indicateur relatif (*i.e.* de mesures corrélées au niveau en langue maternelle et à une variable temporelle), et il s'avère qu'effectivement certaines mesures corrélées au niveau en langue maternelle et à la durée d'exposition sont pertinentes. J'établirai le coefficient de corrélation entre la mesure et la durée d'exposition, et identifierai un seuil critique, variable selon la durée d'exposition, en dessous duquel un individu peut être considéré comme présentant un risque de trouble d'acquisition du langage.

3.2 Matériel et méthodes

3.2.1 Population

La population analysée dans ce chapitre fait partie de celle étudiée dans le cadre du projet ANR CCAAL BLAN 08-3-311192 "*Complexité et compensation en acquisition atypique du langage*" (porteur : L. Tuller, Univ. Tours). Elle est composée de 80 sujets répartis comme suit.

- 50 enfants à développement typique, regroupés en 4 tranches d'âge :
 - TD4, N=14, âge : 4;8 – 4;11, moyenne 4;10
 - TD6, N=12, âge : 6;2 – 6;7, moyenne : 6;4
 - TD8, N=12, âge : 7;10 – 8;7, moyenne : 8;2
 - TD11, N=12, âge : 11;1 – 11;9, moyenne : 11;4

- 30 enfants de l'immigration britannique apprenant le français comme langue seconde, caractérisés par les données suivantes :
 - âge au début de l'exposition au français : 4;6 – 10;10, moyenne 7;1
 - âge au moment du test : 6;4 – 13;1, moyenne : 9;7
 - durée d'exposition au français : 0;9 – 4;5, moyenne : 2;6
 - sexe : 18 f, 12 m

Dans la population L2, l'âge au moment du test est très significativement corrélé à l'âge au début de l'exposition ($p < 0,001$ pour les tests de *Kendall* et de *Spearman*). Il existe une corrélation, cependant faiblement significative ($p < 0,05$ pour les tests de *Kendall* et *Spearman*), entre l'âge au moment du test et la durée d'exposition. Ces corrélations résultent de la façon dont le recrutement a été opéré puisqu'il s'agissait d'exclure toute

corrélation entre l'âge au début de l'exposition et la durée d'exposition. Une telle corrélation aurait en effet pu masquer d'éventuels effets de période critique. Ces informations devront être prises en considération lors de l'interprétation des données.

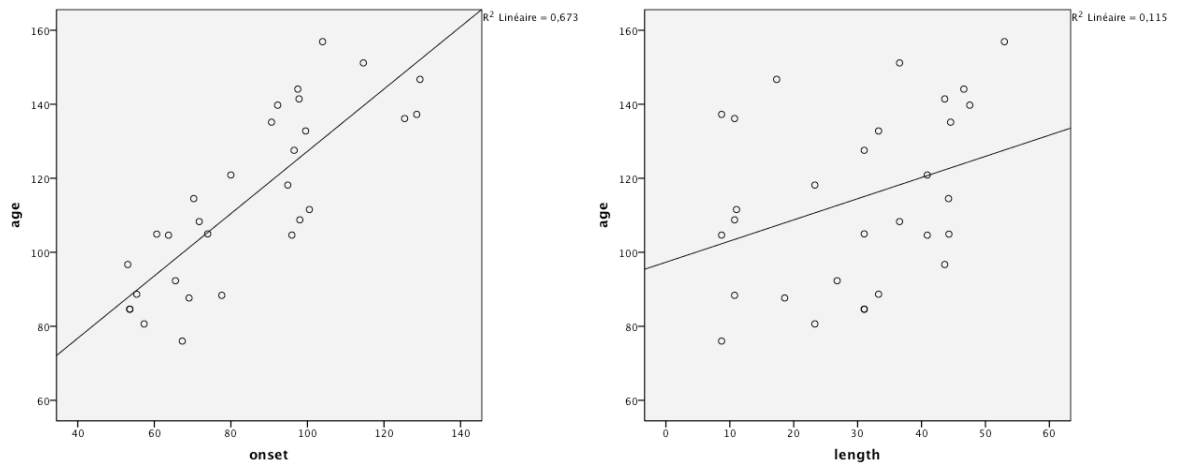


Figure 7. Structure de la population L2

3.2.2 Matériel expérimental

Les données ont été élicitées à l'aide du protocole pour la production de pronoms clitiques (PPPC, (Tuller et al., 2004)). Dans ce protocole, une série d'images est présentée aux sujets, accompagnée d'un stimulus verbal prononcé par l'expérimentateur. Le stimulus verbal introduit les deux arguments qui devront être repris sous la forme de pronoms clitiques dans la réponse du sujet.

Voici un exemple : en présentant l'image 8 ci-dessous, l'expérimentateur dit : "Regarde bien l'image et écoute bien. Que fait Marie avec le fil de téléphone ?" La réponse attendue est : "Elle le coupe."

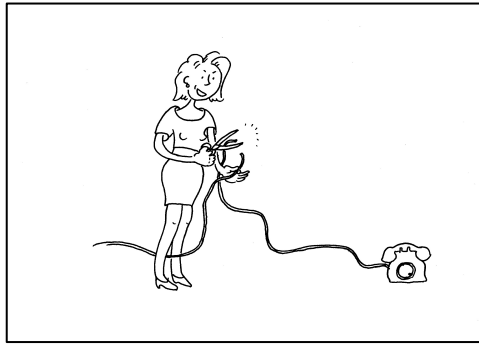


Figure 8. PPC, item 20 (Tuller et al., 2004)

Cette réponse requiert la production d'un groupe de clitiques consistant en un clitique sujet suivi d'un clitique objet adjoint au même verbe. Cependant, d'autres réponses sont également possibles et compatibles avec la pragmatique de la tâche. En particulier, la réponse peut comporter des modificateurs aspectuels ou modaux comme, par exemple, dans "Elle vient de le couper." Ce type de réponse est effectivement produit spontanément par certains sujets et ce fait est particulièrement pertinent, car l'enchâssement requiert la séparation du groupe de clitiques et leur adjonction à deux verbes différents.

Après trois items d'entraînement, les sujets doivent répondre à 32 items de test, structurés de la manière suivante : 16 items élicitent une réponse transitive, 16 élicitent une réponse réfléchie. Dans chacun des 2 groupes, les personnes et les genres sont distribués de manière uniforme, selon le schéma suivant :

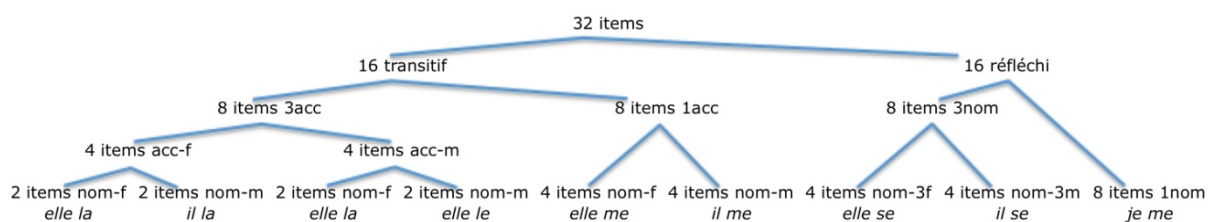


Figure 9. PPC, design

3.2.3 Transcription et codage

Les passations ont été enregistrées et transcrites dans le format CHILDES (MacWhinney, 2000), dans le cadre des projets en cours du groupe langage de l'E1, U930 INSERM (ANR CCAAL BLAN 08-3-311192 "Complexité et compensation en acquisition atypique du

langage", porteur : L. Tuller, Univ. Tours, et action COST IS0804 "*Language Impairment in a multilingual society. Linguistic Patterns and the Road to assessment.*", porteur : S. Armon-Lotem, Bar-Ilan Univ.) J'ai recodé les données élicitées dans une feuille de calcul spécifiquement élaborée pour la présente étude (voir aussi le mémoire de (Dupuy, 2009)).

Dans cette feuille sont saisis : la réponse du sujet et le codage du statut des clitiques sujet et objet. Le codage des clitiques spécifie les informations suivantes :

- pour le sujet : si le clitique
 - o correspond à la cible
 - o comporte une erreur de personne ou de genre
 - o est omis
 - o est remplacé par un élément tonique.
- pour l'objet : si le clitique
 - o correspond à la cible
 - o est remplacé par un datif grammatical
 - o comporte une erreur de personne, de genre, de position ou de cas
 - o est omis
 - o est remplacé par un élément tonique.

Enfin, la feuille précise s'il s'agit d'une réponse à enchâssement requérant la séparation des clitiques et si la réponse est non pertinente ou bien absente.

La feuille extrait ensuite une quarantaine de scores binaires pour chaque item. Ces scores correspondent aux possibilités de combinaison des clitiques sujet et objet cibles et non cibles dans deux échantillons. Le premier échantillon rassemble les réponses à enchâssement requérant la séparation des clitiques et le second échantillon rassemble les réponses racine requérant la production du groupe de clitiques. Ensuite, la feuille calcule les scores absolus et les moyennes par sujet pour les catégories de l'objet de la Figure 9, à savoir les transitifs classifiés par la personne de l'objet (3acc, 1acc), les réfléchis classifiés par la personne (3ref, 1ref), les transitifs et réfléchis regroupés (acc, ref), ainsi que les personnes regroupées (all1st, all3rd) et la moyenne totale (Figure 10). Ces informations sont finalement regroupées de façon à être clairement lisibles sur d'autres onglets (Figure 11).

pop indiv	production	S			O			clit err			prod			omission	DP lex	séparé ?	NC/NR	séparés target S/O
1	Je me lave	clit ok	clit err pers	clit err gen	clit ok	clit ok dat	clit err pers	clit err gen	clit err cas	clit err	prod	omission	DP lex					0
2	Elle se maquille	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
3	Il me lèche	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
4	Il le pèse	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
5	Elle me mord	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
6	Je me coiffe	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
7	Il se cache	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
8	Elle la regarde	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
9	Il la coiffe	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
10	Je me rase	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
11	Elle me lèche	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
12	Elle m'éclabousse	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
13	Il la réveille	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
14	Elle le coupe	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
15	Je me maquille	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
16	Il se pèse	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
17	Il m'écrase	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
18	Je me sèche	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
19	Elle le lave	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
20	Elle se coupe	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
21	Je me regarde	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
22	Il se peigne	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
23	Elle se réveille	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
24	Il me pique	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
25	Il se regarde	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
26	Il me poursuit	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
27	Je me pique	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
28	Il le cache	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
29	Elle se lave	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
30	Je me pèse	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
31	Elle me pique	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
32	Elle la maquille	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0
3 acc		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1 acc		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3 ref		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1 ref		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
acc		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ref		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
all 1st		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
all 3rd		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
all		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3 acc		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
1 acc		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
3 ref		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
1 ref		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
acc		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
ref		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
all 1st		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
all 3rd		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!
all		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	#DIV/0!

Figure 10. Extrait de la grille de codage

[illegible]

3.2.4 Variables testées

J'ai retenu, pour la présente étude, un sous-ensemble des mesures calculées par la feuille de données. Ces mesures sont réparties en 3 groupes, correspondant à 3 échantillons de données différentes.

1. Le premier ensemble regroupe les mesures portant sur la totalité des réponses, à l'exclusion des réponses non pertinentes et des non réponses. Il s'agit des mesures suivantes :

- Le taux de production de clitiqes objet (**O prod trad.**)
Cette mesure inclut les objets cibles, les datifs grammaticaux (produits en remplacement d'un accusatif), les objets clitiqes comportant une erreur morphologique, et les énoncés à clitique en position argumentale canonique.
- Le taux d'omission (**O omission trad.**)
Cette mesure compte les réponses sans objet réalisé.

- Le taux de compléments toniques remplaçant un clitique (**DP trad.**)
Cette mesure rassemble les réponses comportant un complément DP ou un pronom tonique.
- Le taux de production d'environnements requérant la séparation des clitiques sujet et objet (**% sep**).

2. Le second ensemble rassemble les mesures calculées seulement à partir des réponses requérant un regroupement des clitiques, à savoir la totalité des réponses à l'exclusion a) des réponses non pertinentes ou non réponses, et b) des réponses à enchâssement requérant la séparation des clitiques.

- Le taux de production de clitiques objet (**O prod in clust**)
Cette mesure inclut les objets cibles, les datifs grammaticaux (produits en remplacement d'un accusatif), les objets clitiques comportant une erreur morphologique, et les énoncés comportant un clitique en position argumentale canonique.
- Le taux d'omission (**O omission in clust.**)
Cette mesure compte les réponses sans objet réalisé.
- Le taux de compléments toniques remplaçant un clitique (**DP clust.**)
Cette mesure rassemble les réponses comportant un complément DP ou un pronom tonique.
- Le taux de clitiques objet cibles (**O ok in clust**)
Cette mesure rassemble les réponses sans erreur morphologique, ni erreur de placement, ni substitution par un datif grammatical, quel que soit le statut du sujet.
- Le taux de clitiques objet cibles ou datif grammatical (**O ok+dat in clust**)
Cette mesure est la somme de la mesure précédente (**O ok in clust**) et du taux des datifs grammaticaux, quel que soit le statut du sujet.

- Le taux de groupes sujet - objet cibles (**Target SO in clust**)
 Cette mesure compte les réponses avec un clitique sujet cible et un clitique objet cible. Sont exclues : les réponses comportant une erreur morphologique sur l'un des clitics, une erreur de placement du clitique objet, et les réponses dans lesquelles l'objet est remplacé par un datif grammatical.
- Le taux de groupes sujet - objet cibles, groupes sujet - datif inclus (**Target SO+dat in clust**)
 Cette mesure est la somme de la mesure précédente (**Target SO in clust**) et du taux des datifs grammaticaux.

3. Enfin, le dernier ensemble rassemble les mesures calculées sur la base des seules réponses à enchâssement requérant la séparation des clitics sujet et objet.

Notons que ce type de réponse n'a pas été élicité dans l'expérience : certains sujets l'ont produit spontanément. Ces réponses sont très intéressantes d'un point de vue théorique, puisqu'elles correspondent à une stratégie grammaticale choisie spontanément par une partie de la population. Cependant, il faudra garder en mémoire que, d'un point de vue statistique, l'absence de valeurs pour un nombre important de sujets rend les résultats moins certains que ceux portant sur la totalité des sujets.

- Le taux d'omission (**O omission in sep.**)
 Cette mesure compte les réponses sans objet réalisé.
- Le taux de compléments toniques remplaçant un clitique (**DP sep.**)
 Cette mesure rassemble les réponses comportant un complément DP ou un pronom tonique.
- Le taux de clitics objet cibles (**O ok in sep**)
 Cette mesure rassemble les réponses sans erreur morphologique, ni erreur de placement, ni substitution par un datif grammatical, quel que soit le statut du sujet.

- Le taux de clitiques objet cibles ou datif grammatical (**O ok+dat in sep**)
Cette mesure est la somme de la mesure précédente (**O ok in sep**) et du taux des datifs grammaticaux, quel que soit le statut du sujet.

- Le taux de groupes sujet - objet cibles (**Target SO in sep**)
Cette mesure compte les réponses avec un clitique sujet cible et un clitique objet cible. Sont exclues : les réponses comportant une erreur morphologique sur l'un des clitiques, une erreur de placement du clitique objet, et les réponses dans lesquelles l'objet est remplacé par un datif grammatical.

- Le taux de groupes sujet - objet cibles, groupes sujet - datif inclus (**Target SO+dat in sep**)
Cette mesure est la somme de la mesure précédente (**Target SO in sep**) et du taux des datifs grammaticaux.

Les 18 variables ainsi obtenues ont été classées en fonction de deux conditions indépendantes nominales (i.e. non numériques) :

- la population (variable *population*) : L2, TD4, TD6, TD8, TD11

- le statut grammatical de l'objet élicité (variable *fonction*) : 3acc, 1acc et réf, correspondant à des items de test élicitant un clitique accusatif de troisième personne (le/la/l'), un clitique accusatif de première personne (me/m') et un clitique réfléchi (se/s'/me/m') respectivement.

Pour le groupe L2, les variables numériques indépendantes suivantes s'ajoutent : l'âge du sujet au moment du test (variable *age*), l'âge au début de l'exposition au français (variable *onset*), la durée de l'exposition au français au moment du test (variable *length*) et le score de l'enfant au CELF IV (variable *celf*).²⁷

²⁷ Afin de m'assurer de la validité des résultats, j'ai calculé le score au CELF IV sous forme de deux variables indépendantes, le score standardisé sur 100 et l'écart-type. Dans toutes les corrélations calculées, ces deux mesures ont donné le même résultat.

3.2.5 Traitement statistique

Les deux critères suivants ont été centraux dans mon choix des tests statistiques appropriés : la dépendance/indépendance des variables les unes par rapport aux autres et la normalité de la population. En effet, les tests paramétriques ne s'appliquent qu'à des variables indépendantes et à des échantillons tirés de populations normales. Si l'une de ces deux conditions n'est pas remplie, nous devons utiliser des tests non paramétriques. Après application d'un test de normalité à un échantillon (*Kolmogorov-Smirnov*), j'ai constaté que la distribution de chacune des 18 mesures sur l'ensemble de la population, et sur divers groupes de population, n'est pas normale. Les tests paramétriques sont donc exclus.

Je considère que deux variables sont liées si et seulement si la valeur de l'une dépend de la valeur de l'autre. Ainsi, par exemple, les mesures de production (**O prod trad.**) et d'omission (**O omission trad.**) sur un échantillon donné (i.e. sur l'ensemble des items élicitant un clitique 3acc, ou bien celui des items élicitant un clitique 1acc ou bien celui des items élicitant un clitique réf) sont liées, car chaque omission observée fait baisser le taux de production. Plus précisément, la somme des valeurs de la production et de l'omission ne peut jamais dépasser 1 (100%) pour un individu et une catégorie (3acc, 1acc ou réf) donnés. Cette définition implique en particulier que les valeurs prises par les catégories 3acc, 1acc et réf (la variable *fonction*) ne sont pas liées : la variable *fonction* définit des échantillons indépendants.

Parmi les tests non paramétriques disponibles, j'ai choisi :

- pour la comparaison des taux issus de plus de deux échantillons indépendants :

l'ANOVA de *Kruskal-Wallis* (KW). Ce test est une extension à plusieurs (= plus de 2) variables du test U de Mann-Whitney. Il constitue l'alternative non paramétrique de l'ANOVA à un facteur.

- pour les comparaisons deux à deux :

les tests de *Kruskal-Wallis*, de *Mann-Whitney* (MW) et de *Kolmogorov-Smirnov* (KS). Les tests de KS et MW comparent les taux sous la forme de rangs et ils permettent de comparer deux distributions par leurs lieux (moyenne, médiane). Le test de KS permet d'étudier en outre la forme d'une distribution, et de déterminer si deux distributions sont engendrées par la même loi.

J'ai en outre étudié la structure de la population L2 : j'ai calculé les corrélations entre les mesures mentionnées précédemment en section 3.2.4 et quatre variables supplémentaires indépendantes : l'âge chronologique au moment du test, l'âge au moment du début de l'exposition au français, la durée d'exposition au français et le niveau en L1. Ces corrélations ont été établies à l'aide des 2 tests non paramétriques suivants :

- le *rho de Spearman*, variante non paramétrique (calculé sur la base de rangs) du coefficient de corrélation de *Pearson*
- le *tau-b de Kendall*, qui comme *Spearman* est calculé sur la base de rangs. A la différence du *tau-a de Kendall*, le *tau-b* est corrigé pour les ex-aequo.

Je précise enfin que ces tests statistiques ont été effectués à l'aide des logiciels *StatPlus:mac 2009* et *SPSS 19*.

3.3 Résultats

3.3.1 Statut grammatical du clitique élicité : la variable *fonction*

Il est bien connu (Tuller et al., 2011) que le statut grammatical du clitique élicité exerce une influence sur le taux de production observé. Ainsi, à certains âges, le taux de production des réfléchis est plus élevé que celui des accusatifs de première personne, et les accusatifs de troisième personne ont les taux de production les plus faibles. Cependant, cette asymétrie évolue avec l'âge des apprenants et elle disparaît au fur et à mesure de l'évolution vers la production cible de tous les clitiques. L'évolution de cette asymétrie est susceptible de nous renseigner sur la trajectoire de l'acquisition.

J'ai testé l'évolution de la variable *fonction* par des ANOVAs de *Kruskal-Wallis* à un facteur, avec des comparaisons paire par paire. Ce test a été appliqué aux groupes L2, TD4, TD6, TD8, TD11. Dans les graphiques suivants, les histogrammes montrent les moyennes par groupe (L2, TD4/6/8/11), accompagnées d'écarts-type. Les oppositions significatives sont indiquées dans la légende à la droite des graphiques. Le niveau p

figure entre parenthèses, avec les conventions suivantes : (*) indique $p < 0,05$, (**) indique $p < 0,01$ et (***) indique $p < 0,001$.

En ce qui concerne le taux de production de clitiqes calculé sur la base de la totalité des réponses pertinentes (**O prod trad.**, en Figure 12), on observe une différence entre :

- acc (i.e. 3acc et 1acc) et réf pour les TD4
- 3acc et réf pour les TD6
- 3acc et 1acc pour les L2

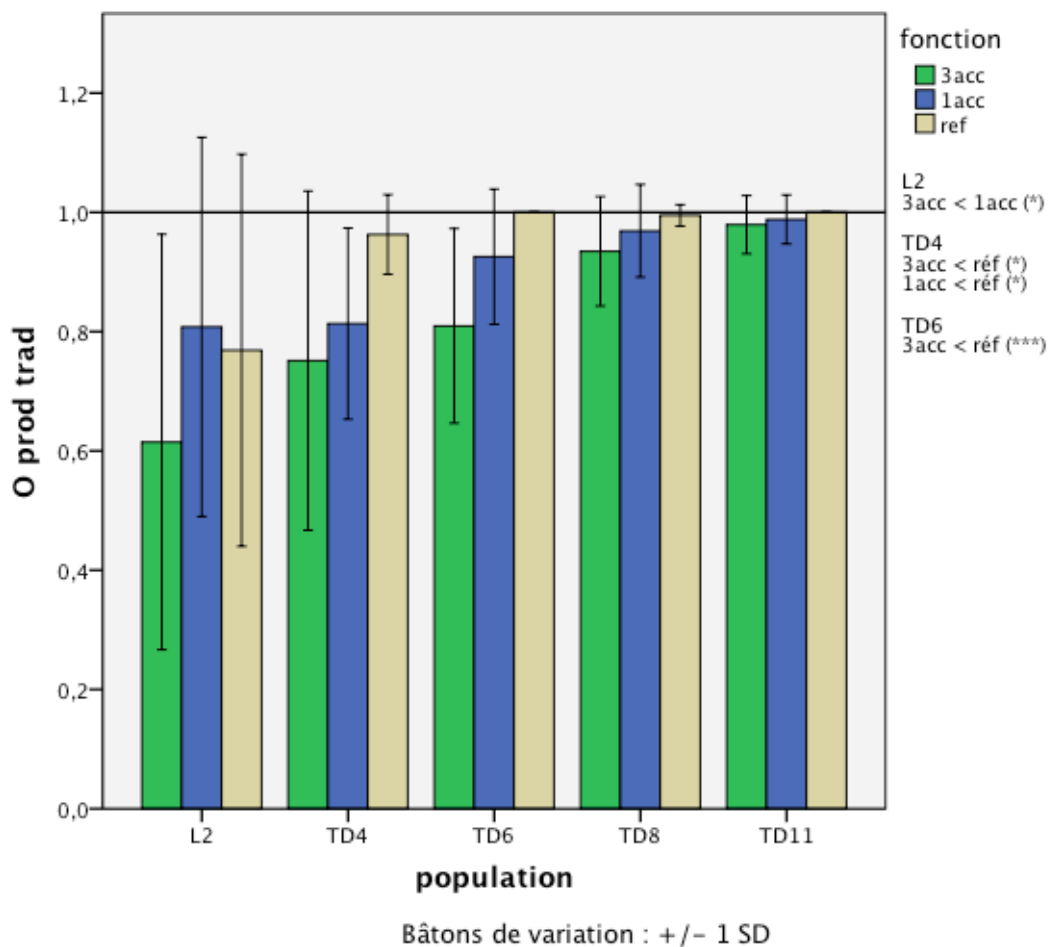


Figure 12. O prod trad. : opposition entre les fonctions

Pour l'omission du clitique objet calculée sur la base de la totalité des réponses pertinentes (**O omission trad.**, en Figure 13), on observe une différence entre 3acc et réf pour les TD6. Aucune autre comparaison n'est significative.

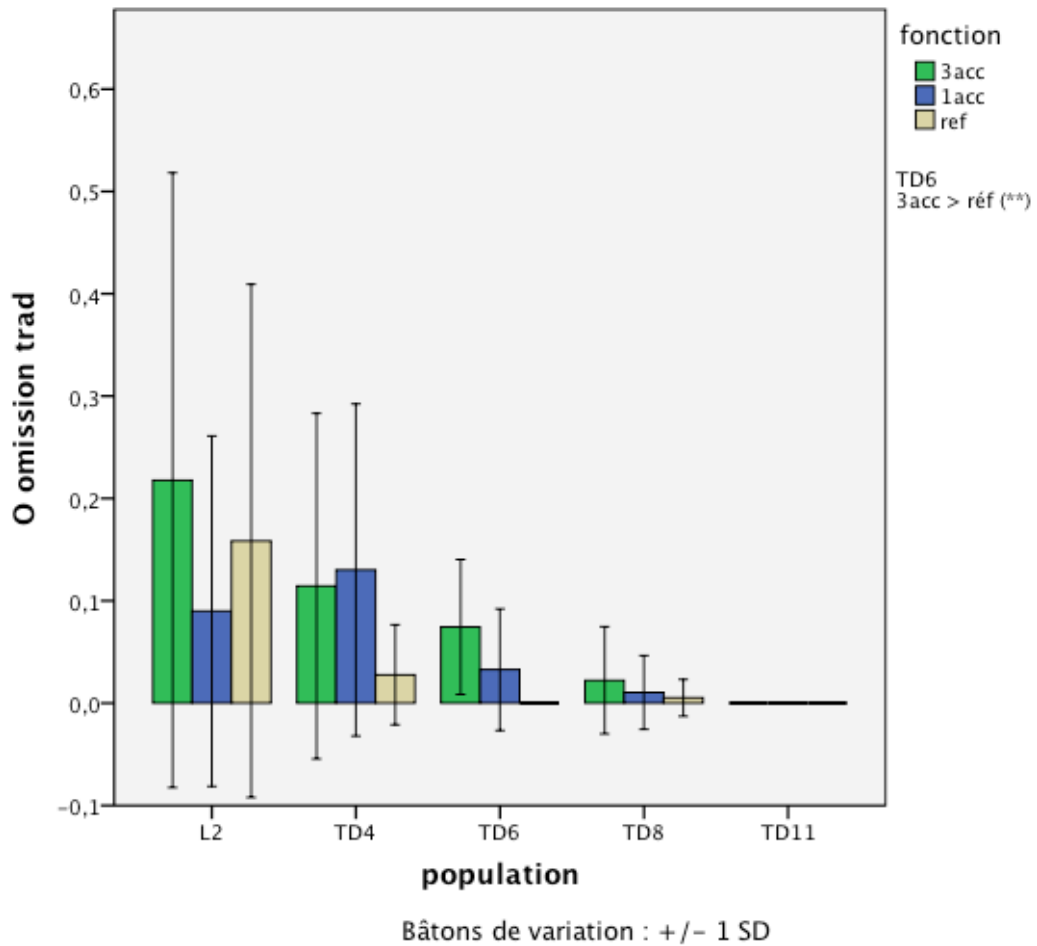


Figure 13. O Omission trad. : opposition entre les fonctions

En ce qui concerne la substitution du clitique objet par un élément tonique calculée sur la base de la totalité des réponses pertinentes (**DP trad.**, en Figure 14), on observe une différence entre 3acc et réf pour les TD4 et les TD6.

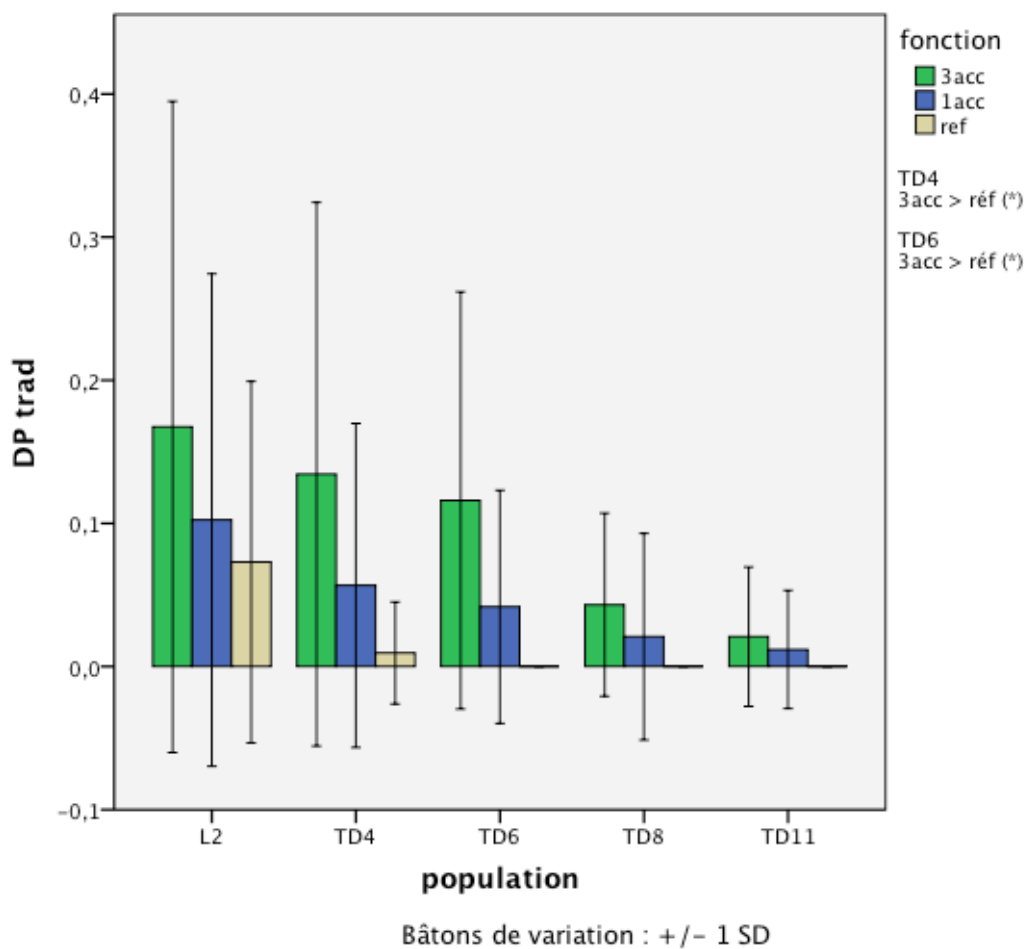


Figure 14. DP trad. : opposition entre les fonctions

Le taux de réponses à enchâssement requérant la séparation des clitiques sujet et objet (% **sep** en Figure 15) ne donne lieu à aucune différence significative entre les fonctions.

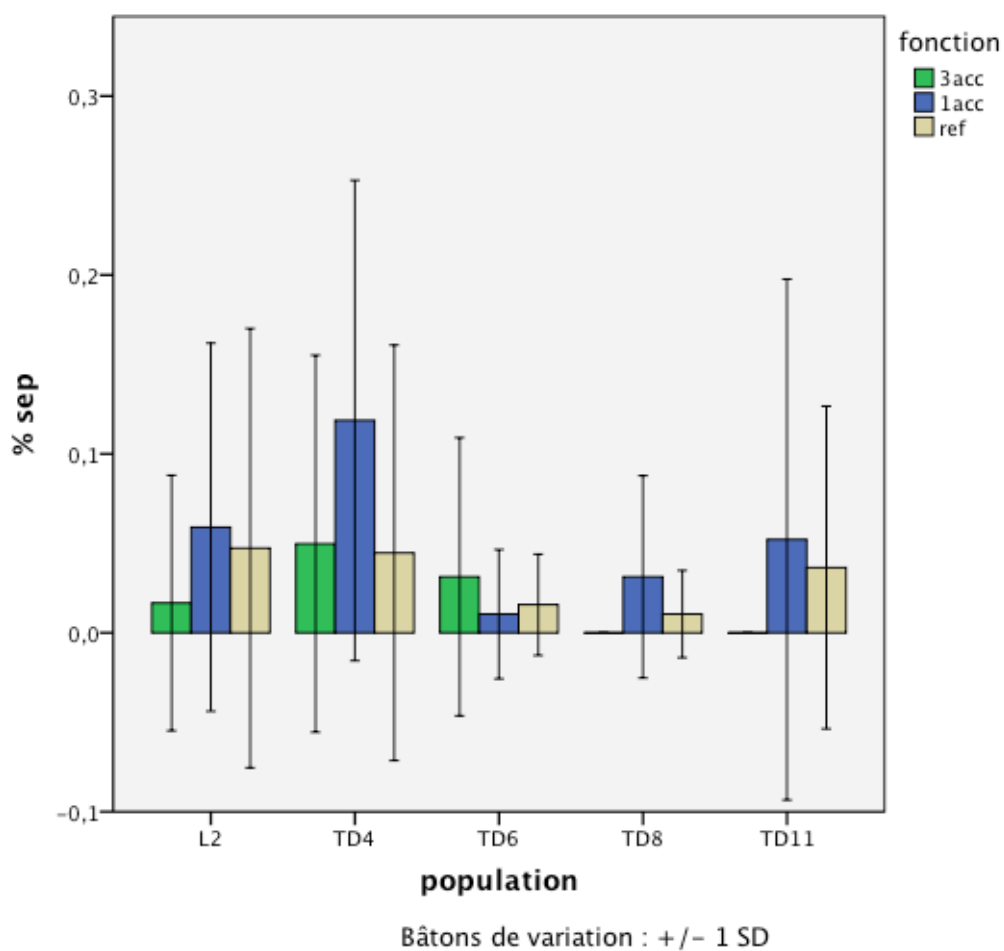


Figure 15. % sep. : opposition entre les fonctions

Pour les objet cible dans des contextes requérant un groupement de clitiques (**O ok in clust** en Figure 16), on observe une différence entre :

- acc (i.e. 3acc et 1acc) et réf pour les TD4
- 3acc et 1acc/réf pour les L2

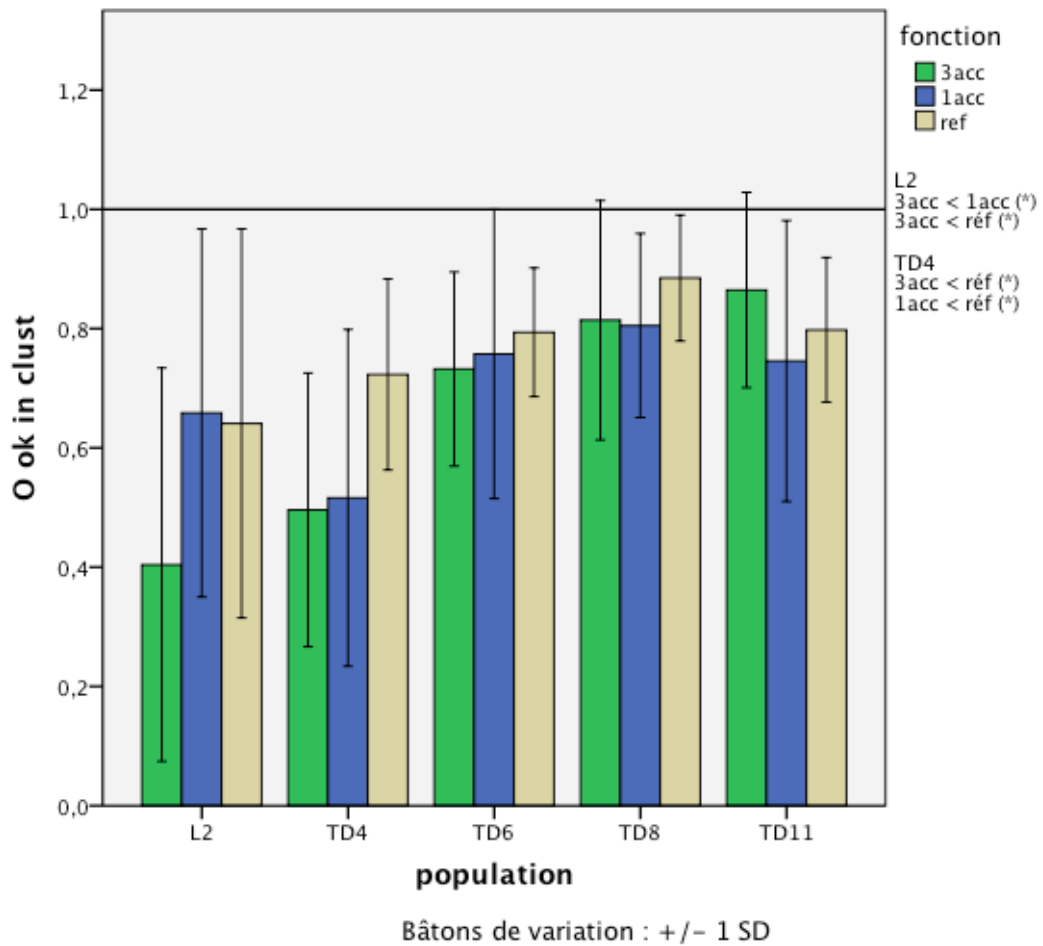


Figure 16. O ok in clust. : opposition entre les fonctions

Si l'on inclut maintenant les réponses avec clitique datif grammatical, on constate que les différences entre les fonctions sont bien plus marquées (Figure 17). On observe alors une différence entre :

- acc (i.e. 3acc et 1acc) et réf pour les TD4
- 3acc et réf pour les TD6/TD8
- 3acc et 1acc/réf pour les L2.

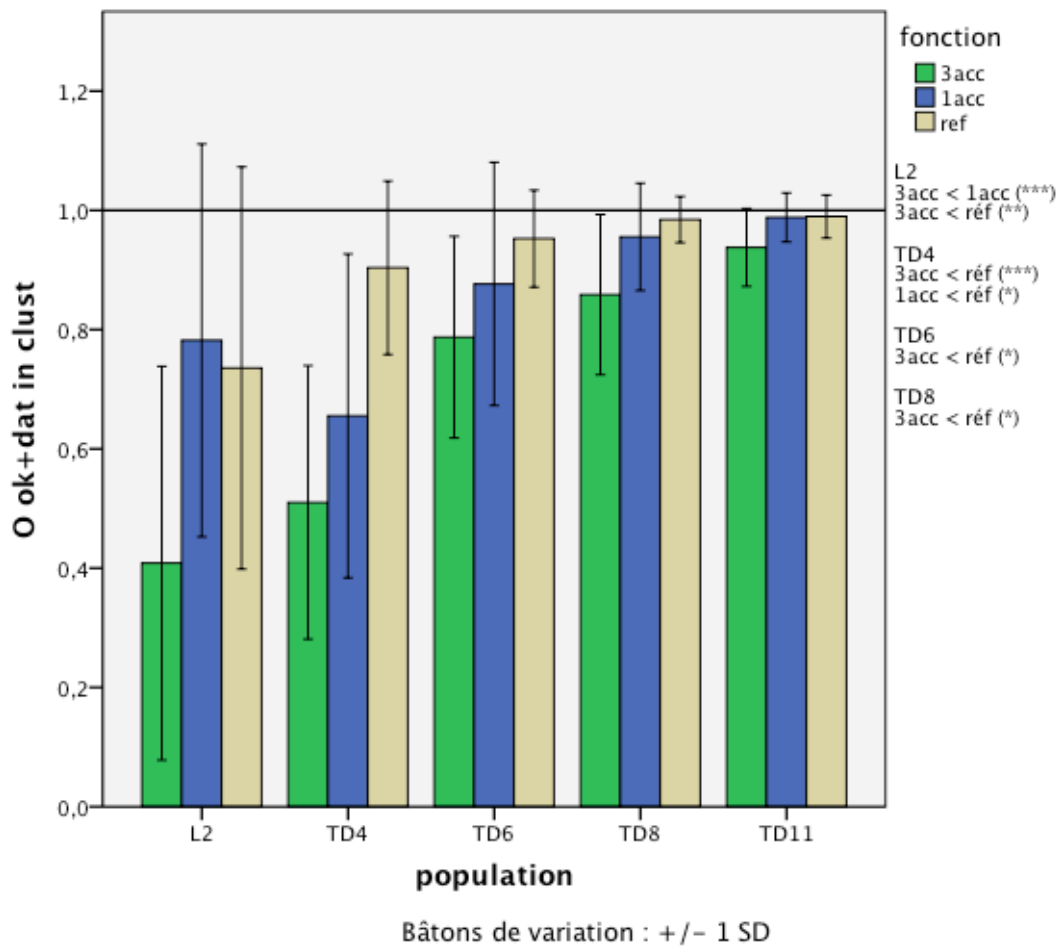


Figure 17. O ok+dat in clust : opposition entre les fonctions

La production de clitiques objet dans les contextes requérant un groupe de clitiques (**O prod in clust**) donne lieu à des distinctions similaires, mais pas identiques, à celles observées pour **O prod trad.** (comparer Figure 12 et Figure 18). On observe une différence entre :

- acc (i.e. 3acc et 1acc) et réf pour les TD4
- 3acc et réf pour les TD6
- 3acc et 1acc pour les L2

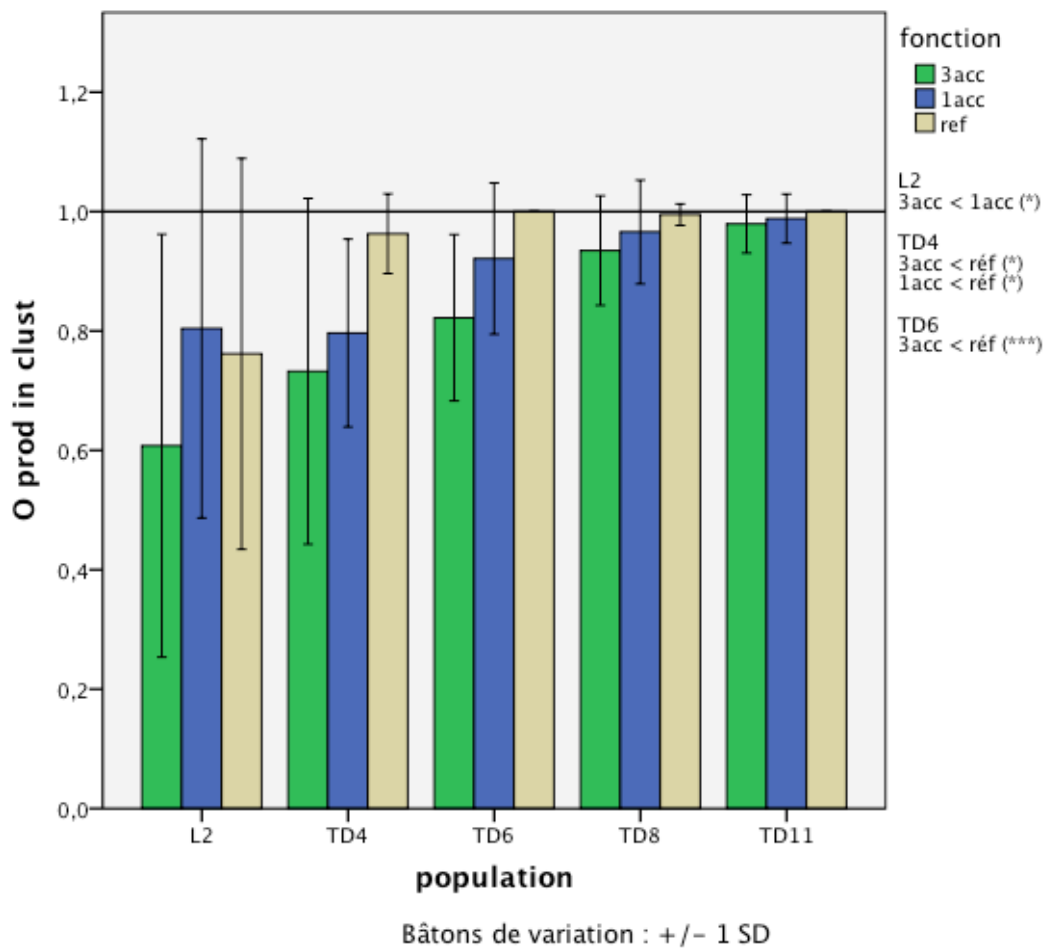


Figure 18. O prod in clust : opposition entre les fonctions

Pour la variable "omission", l'exclusion des contextes à enchâssement ne change rien aux oppositions entre fonctions : **O omission in clust** se comporte exactement comme **O omission trad.**, à savoir : 3acc se distingue de réf pour les TD6 (comparer Figure 13 et Figure 19). Notons l'absence d'opposition entre les fonctions dans le groupe des TD4.

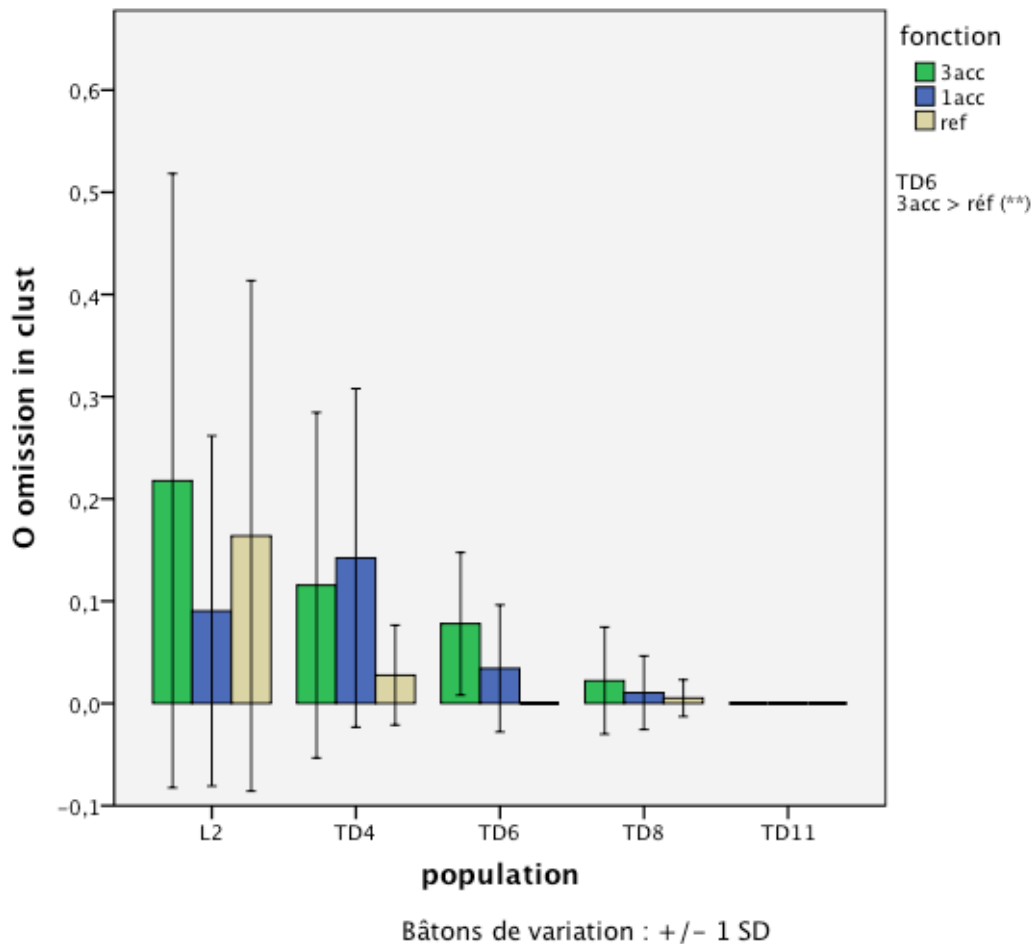


Figure 19. O omission in clust : opposition entre les fonctions

Pour la substitution du clitique objet par un élément tonique, on observe le même phénomène : l'exclusion des contextes à enchâssement ne change pas les oppositions entre fonctions (comparer Figure 14 et Figure 20). Comme pour **DP trad.**, on observe une différence entre 3acc et réf pour les TD4 et TD6.

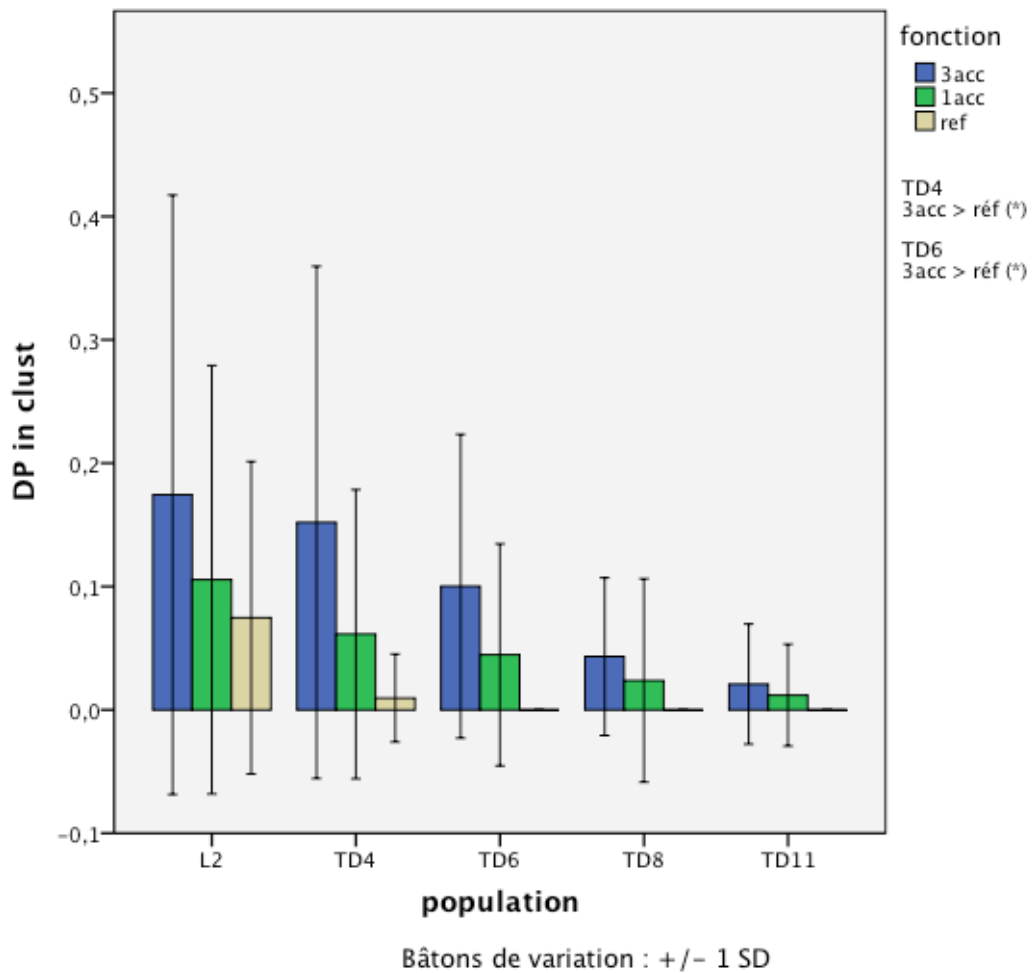


Figure 20. DP in clust : opposition entre les fonctions

Le taux de réponses à sujet et objet cible dans les contextes requérant un groupe de clitics (**Target SO in clust**, en Figure 21) présente une différence entre

- acc (i.e. 1acc et 3acc) et réf pour les TD4 et les L2
- 3acc et 1acc dans le groupe des TD11 : le taux de 3acc est supérieur au taux de 1acc.

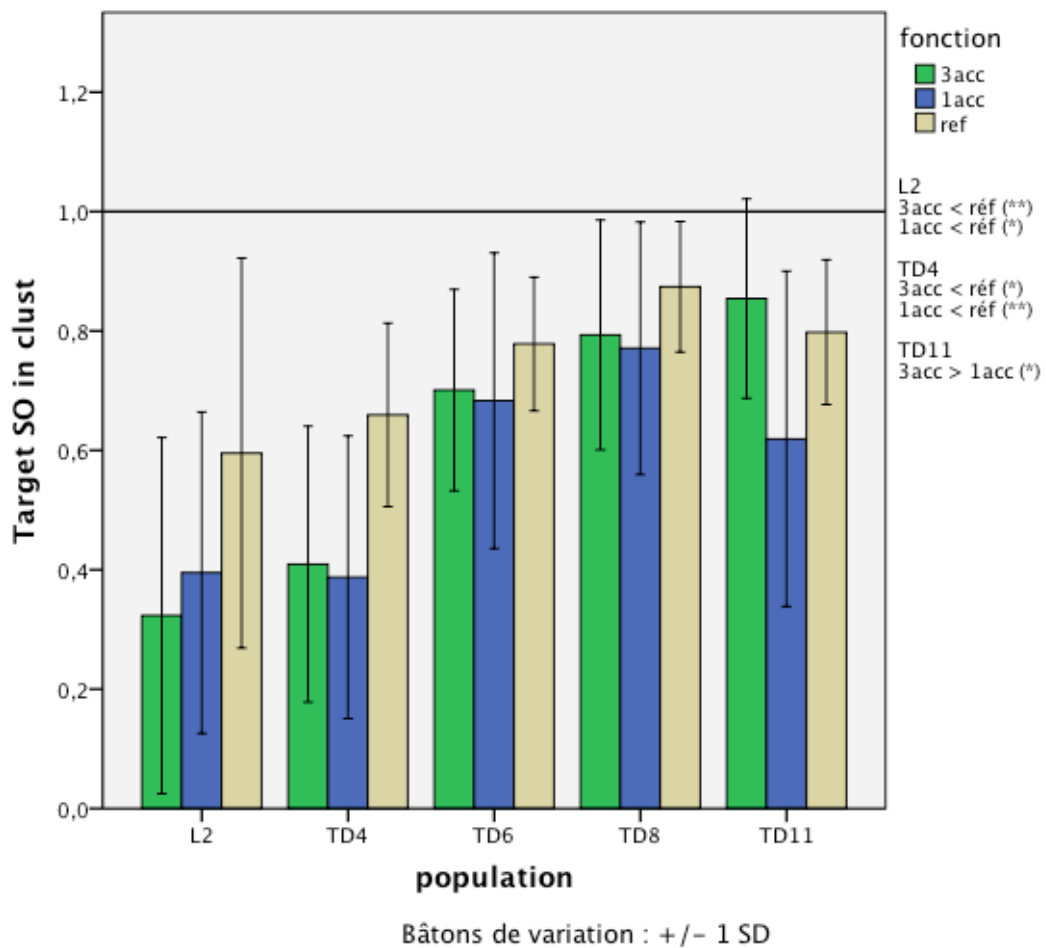


Figure 21. Target SO in clust : opposition entre les fonctions

Si l'on inclut les datifs grammaticaux dans le calcul, les asymétries entre les fonctions deviennent plus marquées et moins surprenantes, sauf pour les TD11 (comparer Figure 21 et Figure 22). On observe une différence entre :

- acc (3acc et 1acc) et réf pour les TD4 et les L2
- 3acc et réf pour les TD6 et les TD8
- 1acc et réf pour les TD11 : nous remarquons à nouveau que le taux de 1acc est bas dans ce groupe.

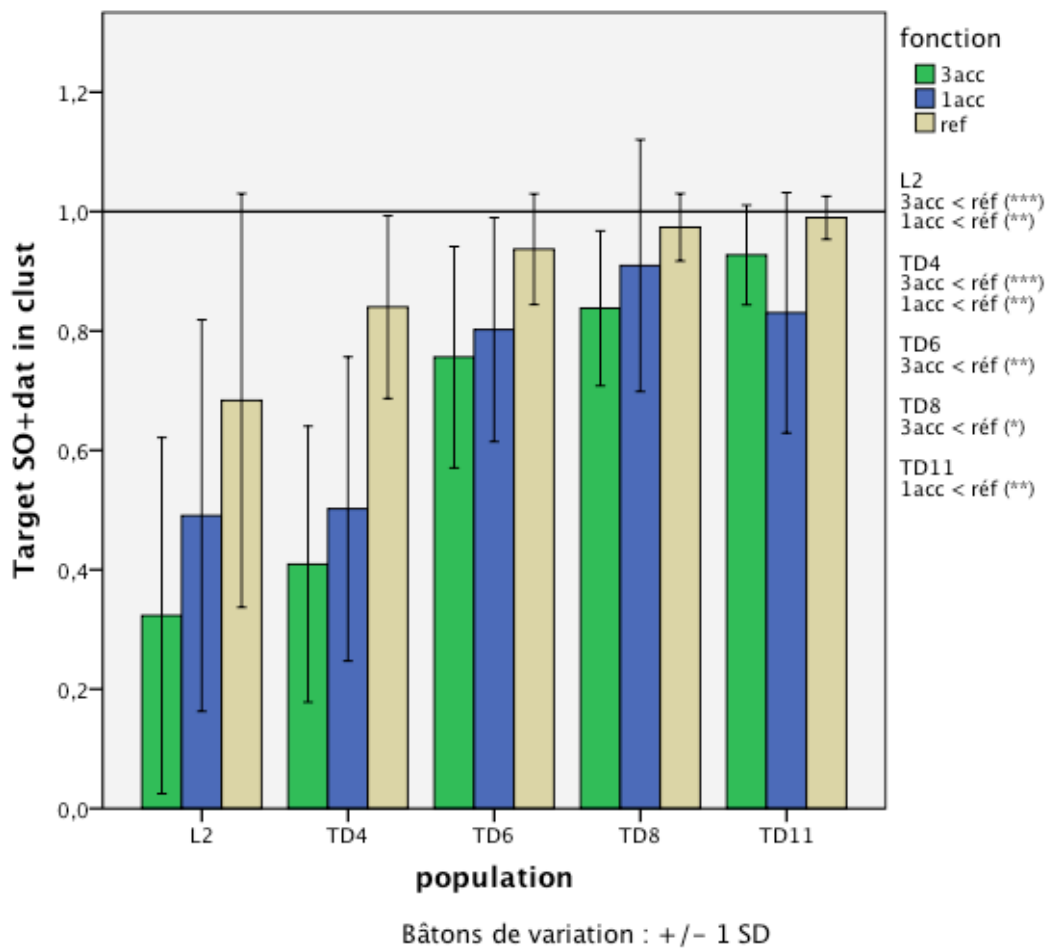


Figure 22. Target SO+dat in clust : opposition entre les fonctions

Pour les variables portant sur les réponses à enchâssement requérant la séparation des clitiques sujet et objet, les chiffres absolus sont faibles. Pour cette raison, il n'est souvent pas possible d'établir de distinction entre les fonctions. On retrouve néanmoins une opposition entre 3acc et 1acc sur la population entière (TD et L2) pour les variables **O ok+dat in sep** (Figure 23), **O prod in sep** (Figure 24) et **DP in sep** (Figure 25). Notons aussi que les taux de réussite sont très élevés dans cette catégorie, même aux âges les plus jeunes. Je reviendrai sur ce fait plus bas.

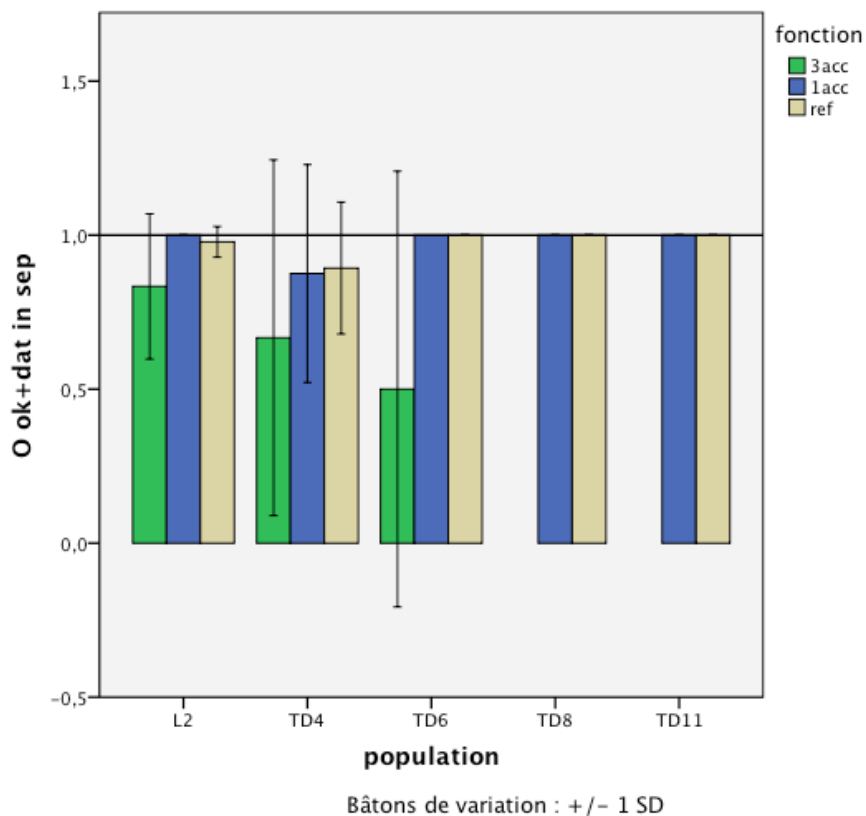


Figure 23. O ok+dat in sep : opposition entre les fonctions

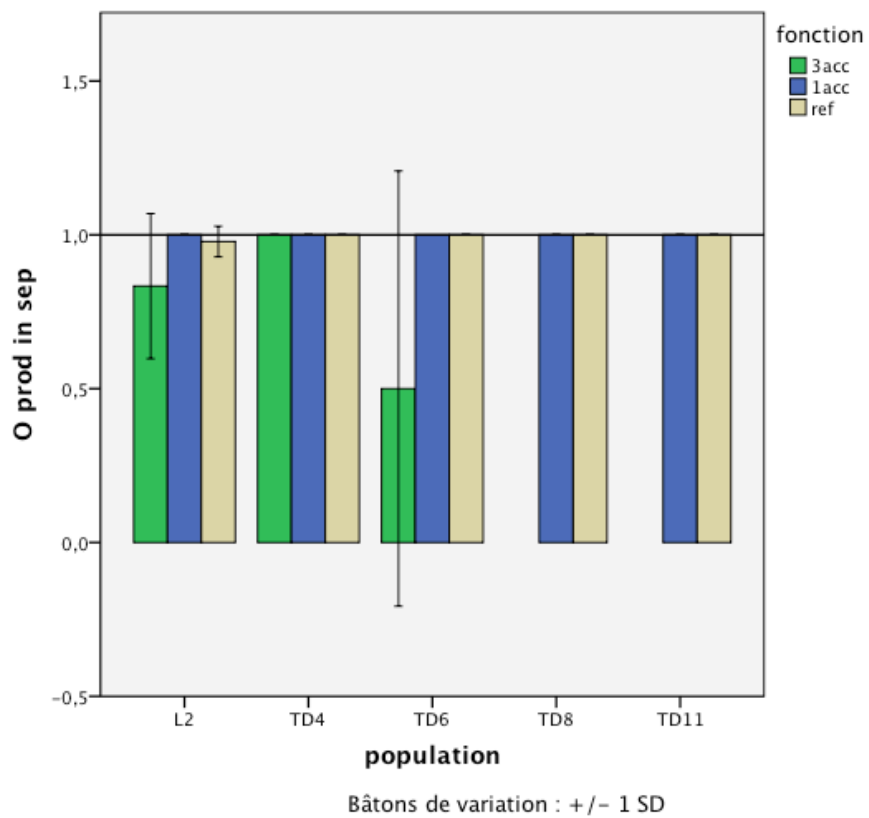


Figure 24. O prod in sep : opposition entre les fonctions

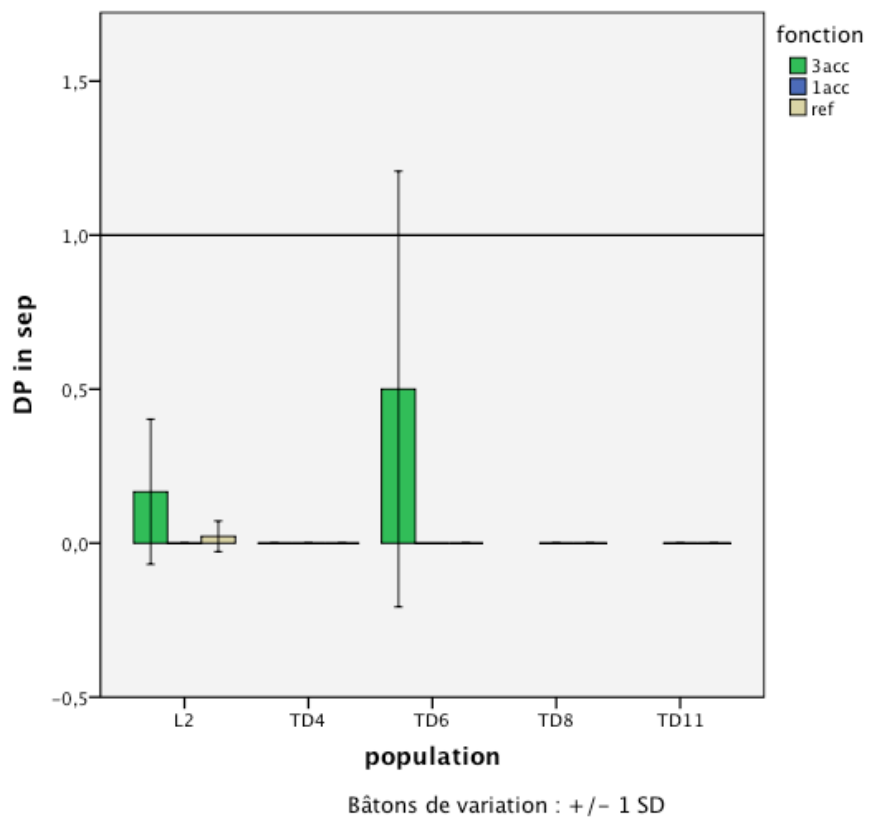


Figure 25. DP in sep : opposition entre les fonctions

3.3.2 Statut grammatical - résumé

Je résume dans le tableau suivant les différences observées entre les fonctions. Le tableau est organisé de la façon suivante : les mesures pertinentes, i.e. celles qui mettent au jour des oppositions entre fonctions, figurent en colonne ; les tranches d'âges figurent en ligne. Chaque case de la première partie du tableau donne ainsi les oppositions significatives entre fonctions par tranche d'âge. La seconde partie du tableau donne les oppositions significatives chez les L2. Les cellules correspondant à une opposition significative donnée sont de la même couleur, et les correspondances entre oppositions et couleurs sont les suivantes :

opposition	trame de fond	opposition	trame de fond
acc vs réf	marron	3acc vs 1acc	bleu
3acc vs réf	rouge	spécificités des TD11	jaune
3acc vs 1acc/réf	orange		

	L2	TD4	TD6	TD8	TD11
O prod trad	3acc vs 1acc	acc vs réf	3acc vs réf		
O prod in clust	3acc vs 1acc	acc vs réf	3acc vs réf		
O omission trad			3acc vs réf		
O omission in clust			3acc vs réf		
DP trad		3acc vs réf	3acc vs réf		
DP in clust		3acc vs réf	3acc vs réf		
O ok in clust	3acc vs 1acc/réf	acc vs réf			
O ok+dat in clust	3acc vs 1acc/réf	acc vs réf	3acc vs réf	3acc vs réf	
Target so in clust	acc vs réf	acc vs réf			3acc vs 1acc
Target so+dat in clust	3acc vs réf	acc vs réf	3acc vs réf	3acc vs réf	1acc vs réf

Tableau 1. Oppositions entre fonctions

Ce tableau met au jour deux phénomènes, que je détaillerai immédiatement ci-dessous. Premièrement, la distinction largement adoptée dans la littérature entre les clitiques 3acc, 1acc et réf se trouve confirmée. Mon étude montre que cette opposition n'est cependant pas rigide mais évolue. Ce résultat doit être mis en relation avec les résultats de (Zesiger et al., in press) : 20s. Dans une étude qui manipule le nombre (sg vs pl), mais non la personne, des clitiques accusatifs, ces auteurs mettent en évidence une influence de l'âge et du type de clitique (accusatif ou réfléchi) sur le taux de production de clitiques grammaticaux et morphologiquement erronés ; cependant, ils ne trouvent pas d'interaction significative entre l'âge et le type de clitique (accusatif ou réfléchi). J'ai trouvé des indices d'une telle interaction, mais j'ai également observé que cette interaction varie selon la mesure. Deuxièmement, l'épel morphophonologique des groupes de clitiques à un effet retardataire. Cet effet retardataire touche les diverses fonctions et les diverses tranches d'âge (y compris les L2) de manière (plus ou moins) uniforme. Je détaille maintenant ces observations, en 7 points.

1. L'état le plus précoce observé dans le cadre de cette étude est caractérisé par l'opposition entre accusatif et réfléchi.

Cette étape est marquée en marron dans le Tableau 1. On l'observe chez les TD4 pour les mesures de production (**O prod trad.**, **P prod in clust**) et de production cible (**O ok+dat in clust**, **Target SO in clust**, **Target SO+dat in clust**). Notons l'absence d'opposition pour la mesure d'omission (**O omission trad.**, **O omission in clust**) et la présence d'une opposition différente (3acc vs réf) pour la mesure de substitution (**DP trad.**, **DP in clust**).

2. A l'étape suivante, 3acc s'oppose toujours à réf, mais 1acc ne s'oppose plus à réf.

Cette étape est caractérisée par l'augmentation du taux des 1acc, qui s'approche de celui des réf. Cette étape est marquée en rouge dans le Tableau 1. Elle est caractéristique des TD6, qui présentent une opposition 3acc vs réf dans les mesures de production (**O prod trad.**, **P prod in clust**) et de production cible (**O ok+dat in clust**, **Target SO+dat in clust**). On trouve également l'opposition 3acc vs réf dans les mesures d'omission (**O omission trad.**, **O omission in clust**).

Pour **O ok+dat in clust** et **Target SO+dat in clust** (i.e. les mesures de la production cible) l'opposition 3acc vs réf s'étend aux TD8. Pour les mesures de substitution **DP trad.** et **DP in clust**, l'opposition 3acc vs réf est visible dès les 4 ans.

3. L'étape la plus avancée de l'acquisition est caractérisée par l'absence d'opposition entre les fonctions.

L'absence d'opposition entre les fonctions est accompagnée de taux de réussite très élevés pour les trois catégories de clitiques. Cette étape est figurée par des cases blanches/vides dans le Tableau 1.

Deux ensembles de cases blanches semblent à première vue surprenants. Tout d'abord, chez les TD4, il n'y a pas d'opposition entre les fonctions pour les mesures d'omission. Ensuite, pour les mesures de production cible excluant les datifs (**O ok in clust**, **Target SO in clust**), les fonctions ne s'opposent plus dès l'âge de 6 ans, alors que les mêmes mesures montrent une opposition 3acc vs réf pour les TD6 et les TD8 dès qu'elles incluent les datifs (**O ok+dat in clust** et **Target SO+dat in clust**).

Pour rendre compte de la première observation, je fais l'hypothèse que les apprenants emploient un clitique zéro sous-spécifié qui est remplacé par les clitiques entièrement spécifiés de la langue adulte au fur et à mesure qu'ils sont acquis (cf. le principe *Elsewhere* (Kiparsky, 1973)). Les TD4 employant un morphème sous-spécifié (le même dans les trois catégories 3acc, 1acc et réf), l'absence d'oppositions est attendue. Les oppositions deviennent visibles au moment où les clitiques réfléchis ne sont (presque) plus omis (chez les TD6) : les réfléchis prononcés se distinguent bien des autres clitiques, y compris du clitique zéro.

Pour rendre compte de la seconde observation, je propose que c'est l'exclusion des réponses avec datif qui masque les oppositions entre fonctions. Si l'on inclut en effet les datifs dans le décompte, le développement suit la trajectoire attendue (TD4 : acc vs réf ; TD6/8 : 3acc vs réf). Ces observations indiquent que les réponses avec datif se comportent comme des réponses cible en ce qui concerne l'état du développement des TD.

4. Nous avons noté à plusieurs reprises une anomalie chez les TD11.

Ce groupe présente des taux de réussite bas pour les clitiques 1acc. Ceci se traduit par l'existence d'une opposition 1acc vs 3acc pour **Target SO in clust** et 1acc vs réf pour **Target SO+dat in clust** (en jaune dans le Tableau 1). Notons aussi que l'opposition 3acc vs 1acc chez les TD11 n'est pas comparable à celle observée chez les L2 (en bleu dans le Tableau 1) : contrairement aux TD11, les L2 produisent les 1acc à un taux plus élevé que les 3acc.

5. Le passage de l'étape 1. (caractérisée par l'opposition acc vs réf) à l'étape 2. (caractérisée par l'opposition 3acc vs réf), puis 3. (caractérisée par l'absence d'opposition) s'effectue de façon différente selon la mesure considérée.

Cette observation montre que les mesures considérées ciblent des connaissances différentes. Plus précisément, le facteur distinguant les mesures est de nature morphologique : (i) plus on mesure la conformité morphologique des clitiques exprimés aux clitiques cibles, plus les oppositions entre fonctions ressemblent à celles typiques d'un état précoce du développement, et (ii) plus on fait abstraction de la réalisation morphologique des clitiques, plus les oppositions entre fonctions ressemblent à celles typiques d'un état de développement avancé.

Afin de préciser ce point, je donne ci-dessous une échelle du poids de la morphologie dans les variables étudiées (du plus élevé au plus bas) :

- (i) **Target SO** : mesure l'expression correcte des 2 clitiques, sujet et objet, et exclut tout clitique comportant une erreur morphologique ; i.e. mesure les groupes de clitiques grammaticaux
- (ii) **O ok** : mesure l'expression correcte du clitique objet seulement ; i.e. mesure les clitiques objet grammaticaux, mais pas les groupes de clitiques grammaticaux
- (iii) **O prod** : mesure la réalisation d'un clitique objet, qu'il soit morphologiquement bien formé ou pas
- (iv) **O omission** : mesure l'absence d'un clitique objet dans les contextes où aucun objet DP plein n'est réalisé. Dans ce contexte, l'absence du clitique est agrammaticale en langue adulte ; je fais l'hypothèse (voir ci-dessus) que les apprenants emploient un clitique zéro sous-spécifié qui est remplacé par les clitiques entièrement spécifiés de la langue adulte au fur et à mesure qu'ils sont acquis. Je considère ainsi que l'omission est affectée indirectement par le facteur morphologie, à savoir par l'apparition des clitiques phonologiquement réalisés (voir la section 3.3.4).
- (v) **DP** : mesure la substitution du clitique par un DP plein. Puisqu'aucun clitique n'est ciblé dans ces énoncés, cette mesure n'est pas affectée par la réalisation morphologique d'un clitique.

A cette hiérarchie s'ajoute un autre facteur, le regroupement des clitiques. Il existe en effet des conditions qui portent sur les groupes de clitiques, qui ne sont pas une simple somme des conditions portant sur chacun des clitiques du groupe. Ce fait génère les asymétries suivantes :

← + importante	Réalisation morphologique	- importante →
échantillon dans lequel les groupes sont obligatoires	échantillon dans lequel les groupes sont possibles	échantillon dans lequel les clitiques sont séparés
O prod in clust	O prod trad.	O prod in sep
O omission in clust	O omission trad.	O omission in sep
DP in clust	DP trad.	DP in sep

Ainsi, chez les TD4, l'opposition acc vs réf est présente pour les mesures de production et de production cible, mais elle est absente pour les mesures d'omission et de substitution. Ensuite, l'opposition 3acc vs réf se retrouve chez les TD6 pour **O ok+dat in clust**, mais chez les TD6+TD8 pour **Target SO+dat in clust**. D'ailleurs, **Target SO(+dat)** est la seule mesure produisant des oppositions entre fonctions chez les TD11. Enfin, je n'ai trouvé aucune opposition entre les fonctions dans l'échantillon dans lequel les clitiques sont séparés.

6. Le groupe L2 ne présente pas le même type d'oppositions que les TD.

En particulier, on observe chez les L2 une opposition 3acc vs 1acc/réf (en orange dans le Tableau 1) qui n'existe jamais chez les TD. Je montrerai plus bas (sections 3.3.3, 3.3.4) que ce comportement des L2 est dû à une performance relativement basse pour la fonction réf dans toutes les mesures.

7. Le facteur morphologique (cf. point 5 plus haut) touche la population L2 de la manière suivante : (i) plus on mesure la conformité morphologique des clitiques exprimés aux clitiques cibles, plus les L2 ressemblent aux TD jeunes, et (ii) plus on fait abstraction de la réalisation morphologique des clitiques, plus les L2 se distinguent des TD jeunes.

Ainsi, pour les mesures d'omission et de substitution, on ne trouve aucune opposition entre les fonctions chez les L2. Il en est de même chez les TD8/11. Si l'on compare le triplet **O prod**, **O ok+dat** et **Target SO+dat**, on observe que la spécificité des L2 (l'opposition 3acc vs 1acc) est la seule opposition observée pour **O prod** ; pour **O**

ok+dat, une opposition 3acc vs réf (typique des TD6/8) s'ajoute à l'opposition 3acc vs 1acc, et pour **Target SO+dat**, 3acc vs réf demeure la seule opposition significative.

3.3.3 Différences entre les groupes

Afin de caractériser les différences entre les groupes de population, j'ai examiné deux questions. Tout d'abord, j'ai déterminé si le facteur "population" est source ou non de variation pour chacune des 18 variables étudiées. J'ai ainsi calculé, et cela dans les trois catégories de fonctions (3acc, 1acc et réf), des ANOVAs de *Kruskal-Wallis* à un facteur, avec la variable "population" comme variable indépendante et les 18 mesures présentées plus haut en section 3.2.4 comme variables dépendantes. J'ai appliqué ce test deux fois, une fois sur la population TD tous âges confondus et une fois sur la totalité de la population (TD et L2). J'ai pu établir les résultats suivants :

- La variable population est source de variation dans les 3 catégories de fonction, sur la totalité de la population (L2+TD) ainsi que la population TD tous âges confondus pour les mesures suivantes : **O prod trad.**, **O prod in clust**, **O omission trad.**, **O omission in clust**, **O ok+dat in clust**, **Target So in clust** et **Target SO+dat in clust**.
- La variable **O ok in clust** établit elle aussi une distinction entre les groupes de population TD pour les trois fonctions. Cependant, une fois le groupe L2 ajouté, i.e. sur la totalité de la population, elle ne varie plus significativement pour la fonction réf.
- **%sep** ne varie que pour les clitiques 1acc chez les TD et sur la totalité de la population.
- Les variables **DP trad.** et **DP clust** ne varient que pour réf, et cela sur la totalité de la population (L2+TD) - sur aucun échantillon plus réduit. Notons en particulier l'absence de variation dans la catégorie 3acc pour les variables de substitution. Cette observation deviendra importante dans la section 3.3.5.3 concernant l'identification des sujets présentant un risque de trouble du développement.

- Pour toutes les mesures portant sur les contextes d'enchâssement (variables dont le nom se termine par **in sep**), on n'observe aucune différence entre les groupes. Cette observation deviendra importante dans la suite. En effet, nous verrons que les propriétés structurales du clitique accusatif hors cluster doivent être en place très précocement. Ceci suggère que l'hypothèse dominante, selon laquelle les clitiques 3acc du français se développent tardivement et graduellement (voir (Grüter, 2006) pour une revue des analyses proposées), doit être adaptée. Le clitique accusatif (et plus tard le clitique 3acc) a effectivement un statut particulier qui n'est pas contesté dans cette étude. Cependant, le caractère prolongé et graduel de son développement est lié en grande partie à l'épel morphophonologique des clitiques sous forme de groupes. J'argumenterai que ce facteur s'applique de la même façon à tous les clitiques, y compris les réfléchis.

Ma seconde question porte plus spécifiquement sur les différences entre les sous-groupes de la population. J'ai effectué des comparaisons par paire pour tous les groupes de population en utilisant les tests de *Mann-Whitney*, de *Kolmogorov-Smirnov* et de *Kruskal-Wallis*. Comme les fonctions ont un comportement différent (mis en évidence dans la section précédente), j'ai calculé les différences entre les groupes séparément pour chaque fonction.

Pour chaque opposition significative, je donnerai le niveau p de significativité dans la légende de l'histogramme. Les conventions adoptées sont les suivantes : les tests sont indiqués par des lettres (m pour *Mann-Whitney*, k pour *Kruskal-Wallis*, s pour *Kolmogorov-Smirnov*), suivies d'un astérisque * pour $p < 0,05$, deux astérisques ** pour $p < 0,01$ ou trois astérisques *** pour $p < 0,001$.

3.3.3.1. Mesures sur la totalité des réponses pertinentes

Considérons tout d'abord la production des clitiques (**O prod trad.**, Figure 26) dans l'échantillon qui regroupe les réponses simples (requérant le regroupement des clitiques) et les réponses à enchâssement (requérant leur séparation).

Pour les clitiques 3acc, on observe une opposition entre

- les TD4 et les TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)
- les TD6 et les TD11 (3 tests)
- les L2 et les TD8/TD11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- les TD4 et les TD6 (*Mann-Whitney* et *Kruskal-Wallis*)/TD8/TD11 (3 tests).
- les L2 et les TD11 (*Kruskal-Wallis*)

Pour les clitiques réf, on observe une opposition entre

- les TD4 et les TD6/TD11 (*Kruskal-Wallis*)
- les L2 et les TD4 (*Mann-Whitney* et *Kruskal-Wallis*), TD6/TD8/TD11 (3 tests)

Pour les 3acc, les enfants jeunes (TD4 et TD6) constituent un groupe s'opposant aux enfants âgés (TD11). Ceci indique que la production des 3acc continue à se développer au delà de 6 ans. Les L2 se comportent comme les TD jeunes.

Pour les 1acc, les TD4 ont un comportement particulier, qui les oppose aux TD plus âgés, ce qui indique que la production des clitiques 1acc ne se développe plus significativement au-delà de l'âge de 6 ans. Les L2 ne se comportent pas comme les TD jeunes (ie, les TD4), ils ne se distinguent que légèrement des TD11.

Pour les réf, la production ne se développe plus significativement chez les TD dès l'âge de 4 ans : les TD4 ne s'opposent que légèrement aux TD6/TD11. En revanche, les L2 se distinguent par un taux de production inférieur même à celui des TD les plus jeunes. Je mettrai ce phénomène en relation (section 3.3.4) avec le paramètre régissant la dérivation des prédicats réfléchis soit en syntaxe, soit dans le lexique (Reinhart et Siloni, 2005).

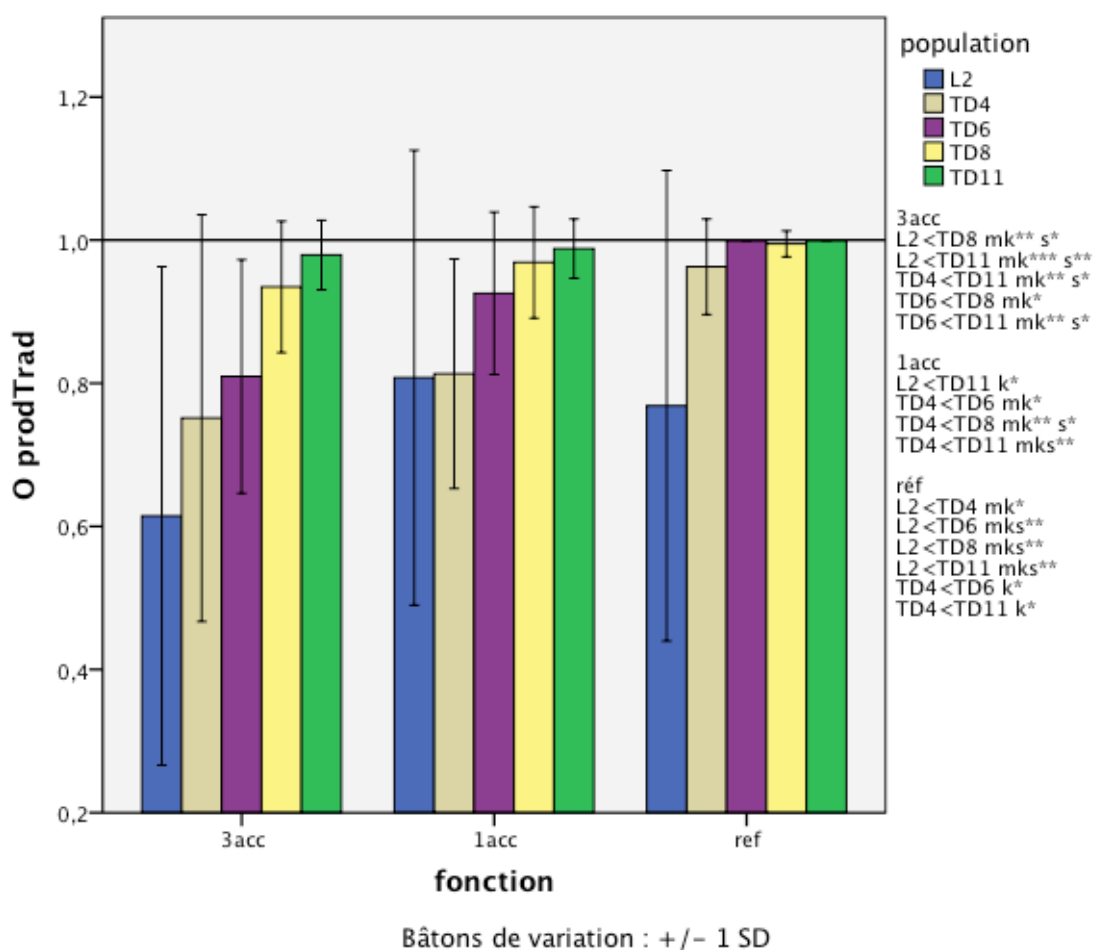


Figure 26. O prod trad : comparaison des groupes

La variable d'omission (**O omission trad.**, Figure 27) établit un schéma général similaire.

Pour les clitiques 3acc, on observe une opposition entre

- les TD4 et les TD11 (*Kruskal-Wallis*)
- les TD6 et les TD11 (3 tests)
- les L2 et les TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- les TD4 et les TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)
- les L2 et les TD11 (*Kruskal-Wallis*)

Pour les clitiques réf, on observe une opposition entre

- les TD4 et les TD6/TD11 (*Kruskal-Wallis*)
- les L2 et les TD6 (*Mann-Whitney* et *Kruskal-Wallis*)/TD6/8/11 (3 tests)

Ces observations confirment le schéma observé pour **O prod trad**. Le développement TD ne s'achève qu'après 6 ans pour les clitiques 3acc, il s'achève avant 6 ans pour les clitiques 1acc et il est largement achevé à 4 ans pour les clitiques réf. Les L2 se comportent comme les TD jeunes (TD4/6) pour les clitiques 3acc, comme les TD6/8 pour les clitiques 1acc et ils présentent un retard notable pour les clitiques réf.

On notera toutefois que les oppositions observées sont moins significatives que celles observées en production (Figure 26) et l'absence d'une opposition entre les TD6 et les TD11 pour 3acc. Ces observations indiquent que la connaissance mesurée par l'omission est plus avancée que celle mesurée par la production. J'ai déjà mentionné plus haut (section 3.3.2, voir aussi les sections 3.3.4, 3.3.5.3) que l'omission résulte de l'emploi d'un clitique zéro sous-spécifié. Cette forme est évidemment moins affectée par le facteur retardataire de l'épel morphologique que les clitiques phonologiquement réalisés.

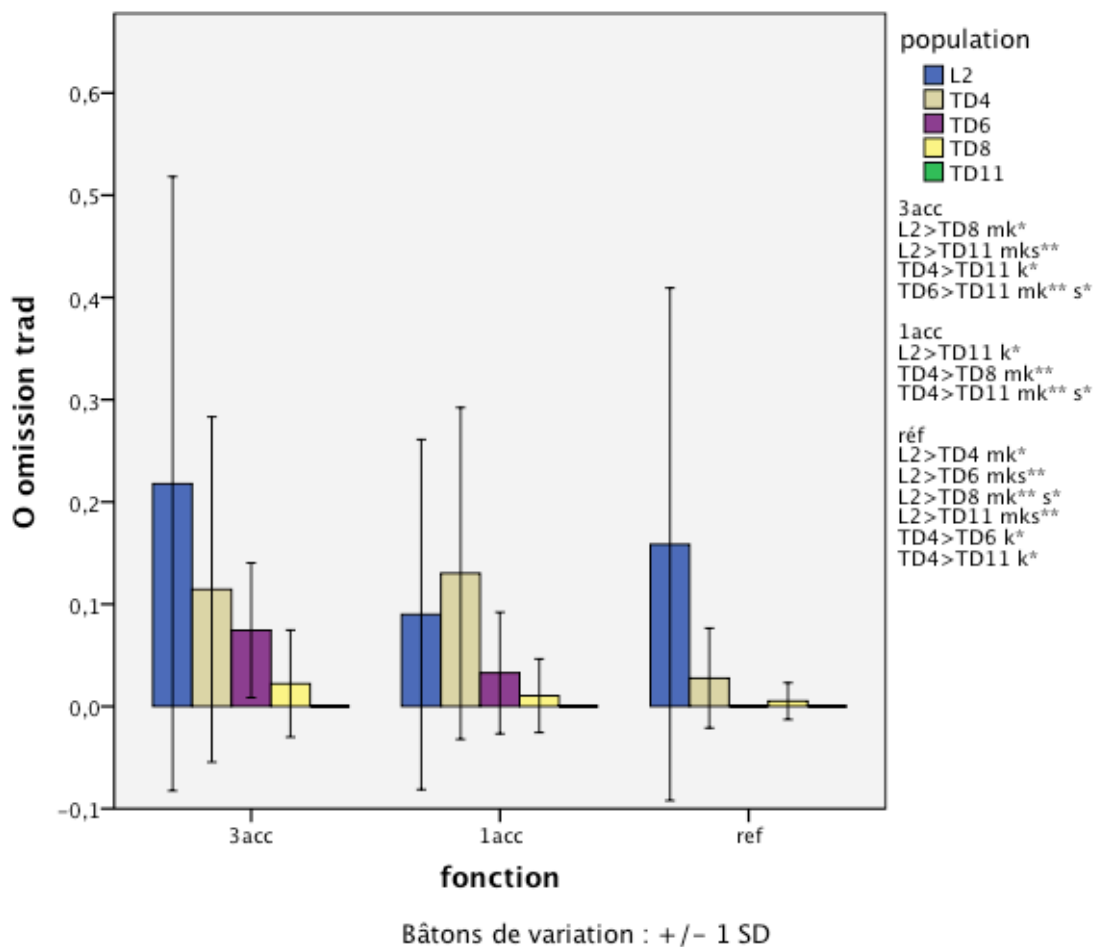


Figure 27. O omission trad : comparaison des groupes

Les oppositions sont encore moins significatives pour la mesure de la substitution (**DP trad.** Figure 28).

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD11 (*Kruskal-Wallis*)
- L2 et TD11 (*Mann-Whitney* et *Kruskal-Wallis*).

Pour les clitiques 1acc, aucune opposition n'est significative.

Pour les clitiques réf, les L2 se distinguent des TD4 (*Mann-Whitney* et *Kruskal-Wallis*)/TD6/TD8/TD11 (3 tests), ce qui confirme le retard des L2 dans cette catégorie.

Même si les moyennes semblent baisser avec l'âge, l'absence presque totale d'oppositions significatives chez les TD indique que la connaissance mesurée par la substitution ne se développe pas (ou plus) dans la tranche d'âge étudiée. Cette stabilité est remarquable. Je proposerai plus bas 3.3.5.3 de considérer la substitution des 3acc comme marqueur de trouble d'acquisition L2.

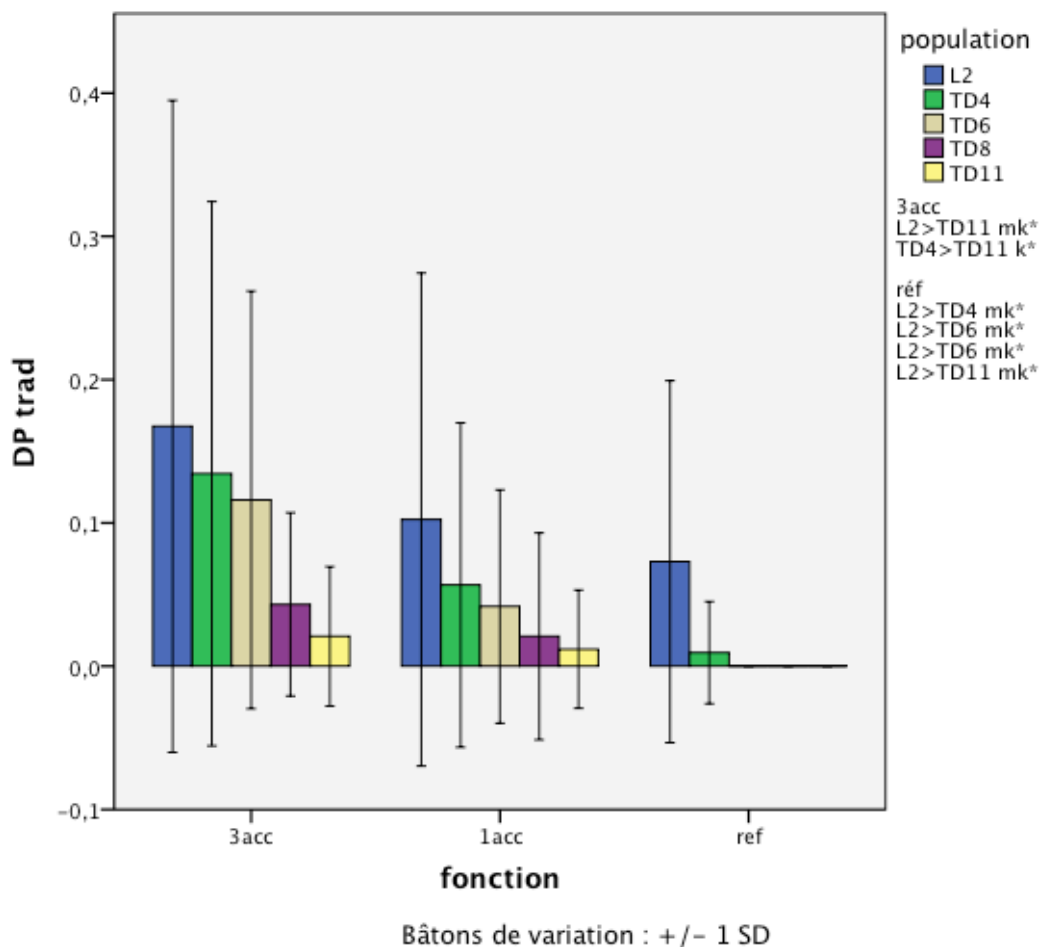


Figure 28. DP trad : comparaison des groupes

La production spontanée d'enchâssements requérant la séparation des groupes de clitiques (**%sep**, Figure 29) est peu fréquente chez les TD6 et TD8 en général. On note cependant une opposition significative en ce qui concerne les clitiques 1acc : plus précisément, entre les TD4 et les TD6 (*Mann-Whitney* et *Kruskal-Wallis*) / TD8 (*Kruskal-Wallis*).

On remarquera que les TD4 ne se distinguent pas significativement des TD11, qui produisent des enchâssements à un taux relativement élevé. Ceci indique que le recours spontané et non élicité à l'enchâssement séparant les clitiques est déterminé par des facteurs autre que l'acquisition des clitiques dans les populations TD et L2 tout venant. Cette conclusion reste néanmoins à vérifier dans les populations avec pathologie.

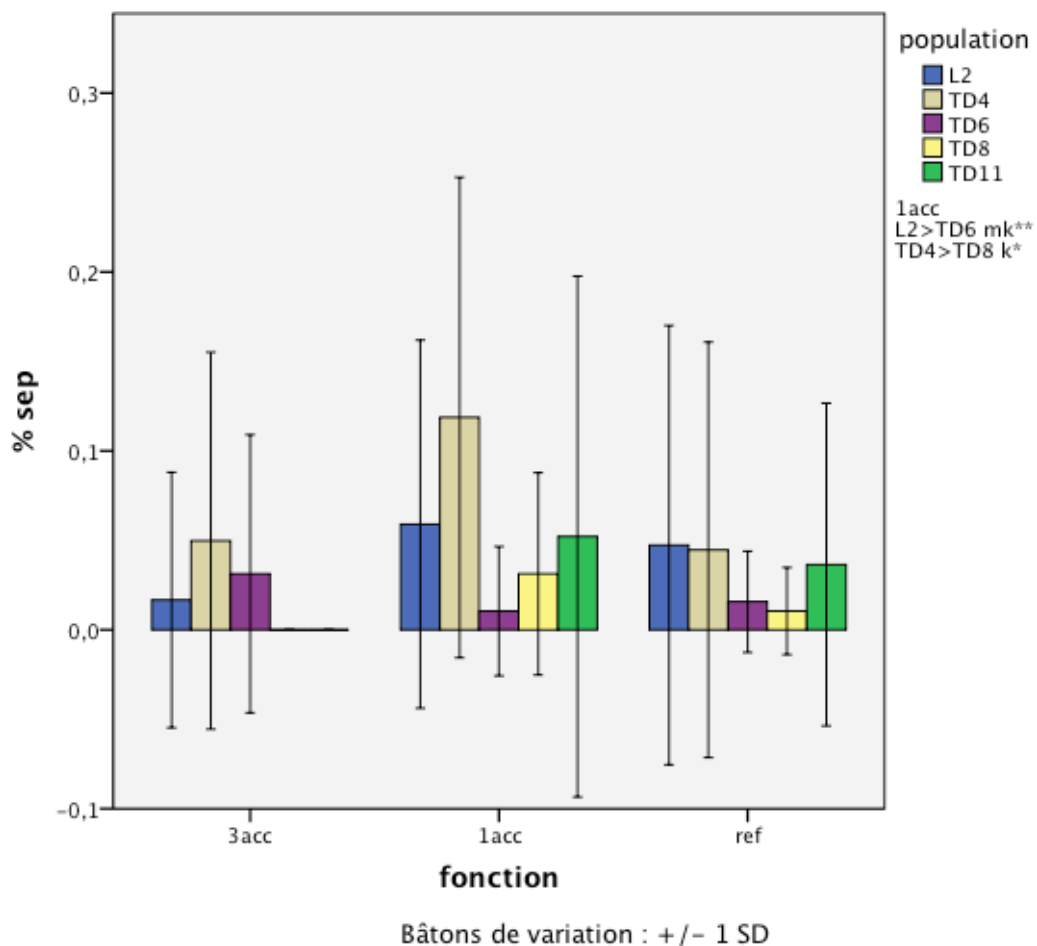


Figure 29. % sep : comparaison des groupes

3.3.3.2. Mesures sur l'échantillon de réponses requérant le regroupement des clitiques

Passons maintenant aux contextes requérant le regroupement des clitiques. La production de clitiques objet (qu'ils soient ou non grammaticaux) dans des contextes requérant un groupe de clitiques (**O prod in clust**, Figure 30) met en évidence des oppositions similaires à celles observées dans **O prod trad.** (Figure 26). Ces oppositions sont plus marquées dans **O prod in clust** dans le sens où les oppositions significatives observées dans **O prod trad.** constituent un sous-ensemble propre de celles observées dans **O prod in clust**.

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD8 (*Kruskal-Wallis*)/TD11 (3 tests)
- TD6 et TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)
- L2 et TD8/11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- TD4 et TD6/8/11 (3 tests)
- L2 et TD11 (*Mann-Whitney* et *Kruskal-Wallis*)

Pour les clitiques réf, on observe une opposition entre

- TD4 et TD6/TD11 (*Kruskal-Wallis*)
- L2 et TD4 (*Mann-Whitney* et *Kruskal-Wallis*)/TD6/8/11 (3 tests)

Pour **O prod trad.** (Figure 26), l'opposition pour les clitiques 3acc entre TD4 et TD8 n'était significative pour aucun test, celle pour les clitiques 1acc entre TD4 et TD6 ne l'était pas pour *Kolmogorov-Smirnov* et celle pour les clitiques 1acc entre L2 et TD11 ne l'était pas pour *Mann-Whitney*. En résumé, la mesure de la production permet de distinguer plus robustement les groupes de population si l'on exclut les réponses à enchâssement (qui requièrent la séparation des clitiques sujet et objet). Ceci indique que le regroupement des clitiques ralentit l'acquisition des clitiques dans toutes les catégories (3acc, 1acc et réf).

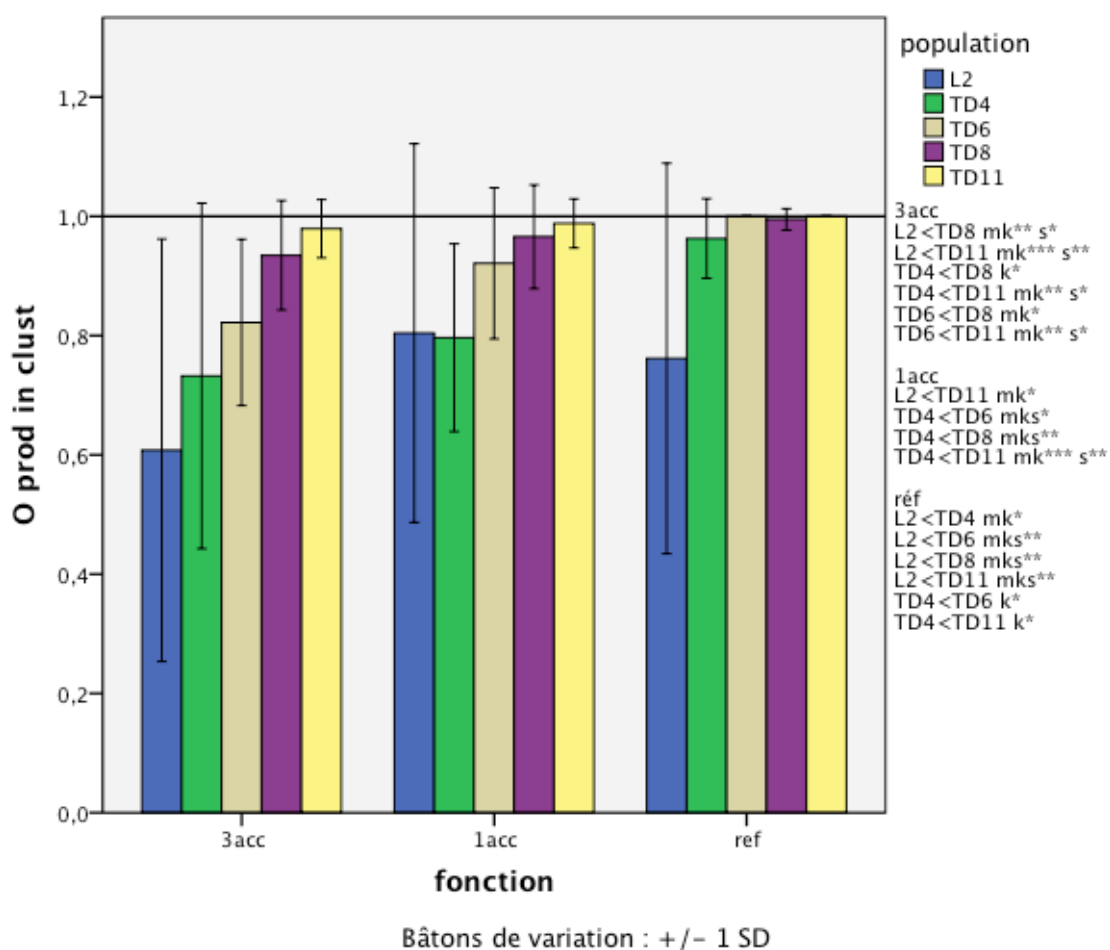


Figure 30. O prod in clust : comparaison des groupes

La mesure de l'omission dans les contextes requérant le regroupement (**O omission in clust**, Figure 31) entretient, elle aussi, une relation d'implication avec **O omission trad.** (Figure 27) : les oppositions significatives dans **O omission trad.** constituent un sous-ensemble propre de celles observées dans **O omission in clust**.

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD11 (*Kruskal-Wallis*)
- TD6 et TD11 (3 tests)
- L2 et TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- TD4 et TD6 (*Kruskal-Wallis*)/TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)
- L2 et TD11 (*Kruskal-Wallis*)

Pour les clitiques réf, on observe une opposition entre

- TD4 et TD6/TD11 (*Kruskal-Wallis*)
- L2 et TD4 (*Kruskal-Wallis*)/TD6/8/11 (3 tests)

On notera qu'une opposition, non significative dans **O omission trad.** (Figure 27), devient ici significative : pour les 1acc, les TD4 se distinguent des TD6 (*Kruskal-Wallis*). Cette observation souligne le statut particulier de la substitution déjà noté précédemment dans cette section.

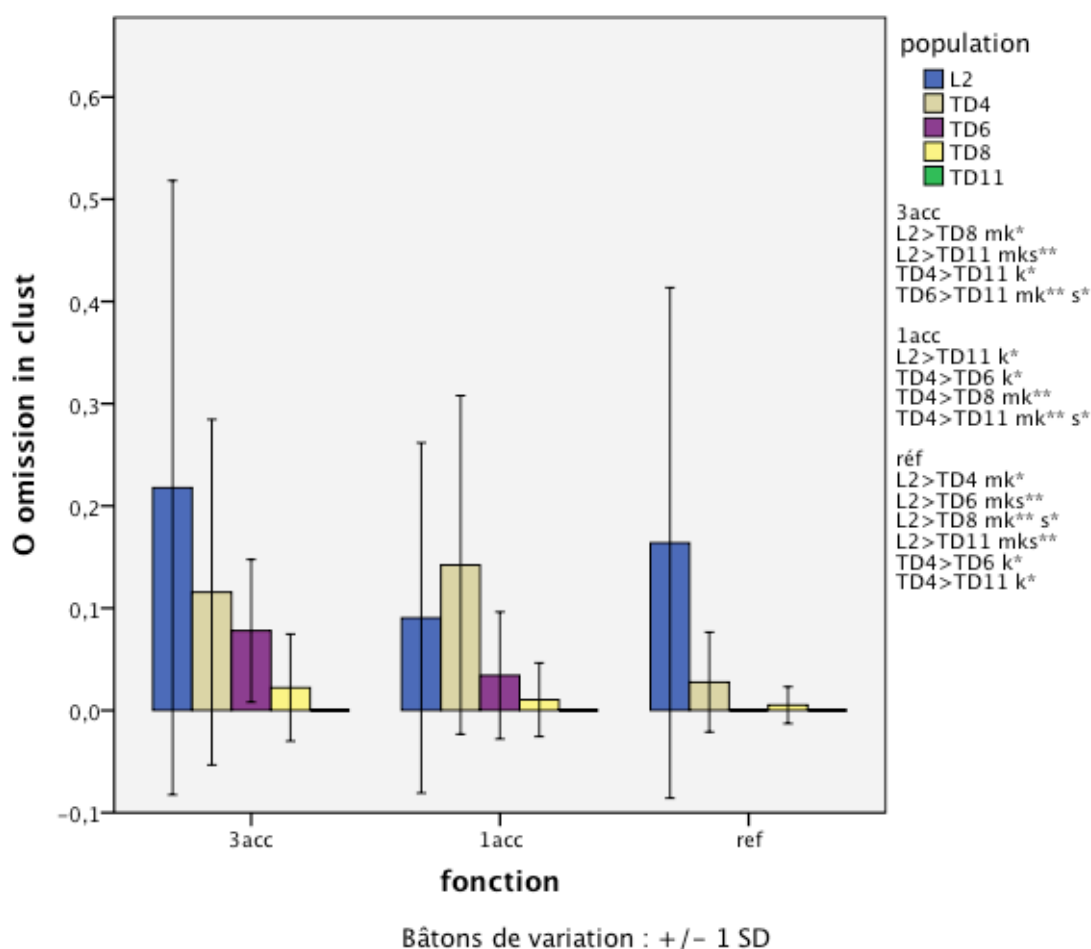


Figure 31. O omission in clust : comparaison des groupes

Pour la mesure de la substitution, il s'avère que l'exclusion des réponses à enchâssement (**DP in clust**, Figure 32) a un effet différent de celui observé en production et omission. Les oppositions significatives dans **DP trad.** ne constituent en effet pas un sous-ensemble propre des oppositions observées dans **DP in clust**.

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD11 (*Kruskal-Wallis*)
- TD6 et TD11 (*Kruskal-Wallis*)
- L2 et TD8 (*Man-Whitney* et *Kruskal-Wallis*)

Pour les clitiques réf, on observe une opposition entre

- L2 et TD4/6/8/11 (*Mann-Whitney* et *Kruskal-Wallis*)

La différence entre **DP trad.** (Figure 28) et **DP in clust** (Figure 32) porte sur les clitiques 3acc :

- l'opposition L2 vs TD8 est significative dans la mesure **DP in clust** (*Mann-Whitney* et *Kruskal-Wallis*), et non significative dans la mesure **DP trad.**
- l'opposition TD6 vs TD11 est significative dans la mesure **DP in clust** (*Kruskal-Wallis*), et non significative dans la mesure **DP trad.**
- en revanche, l'opposition L2 vs TD11, significative dans la mesure **DP trad.** (*Kruskal-Wallis*), ne l'est pas dans **DP in clust**.

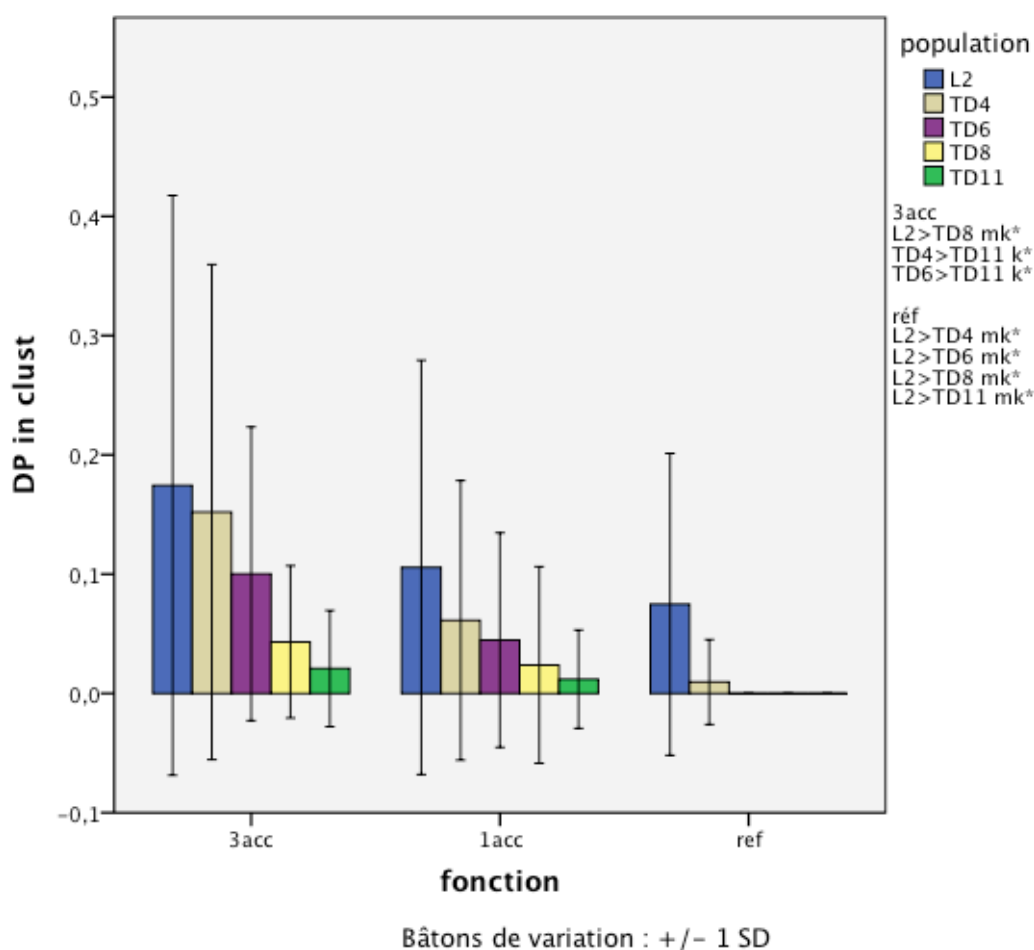


Figure 32. DP in clust : comparaison des groupes

Passons maintenant aux mesures des clitiques objet morphologiquement bien formés (**O ok in clust**, Figure 33 et **O ok+dat in clust**, Figure 34). Cette mesure augmente l'importance du facteur morphologique par rapport à la mesure de production (**O prod in clust**, Figure 30), car elle exclut les réponses réalisant un clitique morphologiquement mal formé. La comparaison entre **O ok in clust** et **O ok+dat in clust** devra nous renseigner sur le statut des réponses à clitique datif grammatical. Je rappelle ici un résultat pertinent établi en section 3.3.2 : l'exclusion des clitiques datifs grammaticaux ayant entraîné des anomalies concernant l'opposition entre fonctions, j'ai conclu qu'il fallait *inclure* les datifs grammaticaux dans les réponses cible. Cette conclusion sera confirmée dans la suite.

Dans **O ok in clust** :

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD6 (*Mann-Whitney et Kruskal-Wallis*)/TD8/11 (3 tests)
- TD6 et TD11 (*Mann-Whitney et Kruskal-Wallis*)
- L2 et TD6/8/11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- TD4 et TD6/8 (3 tests)/TD11 (*Mann-Whitney et Kruskal-Wallis*)

Pour les clitiques réf, on observe une opposition entre

- TD4 et TD8 (*Mann-Whitney et Kruskal-Wallis*)
- L2 et TD8 (*Mann-Whitney et Kruskal-Wallis*)

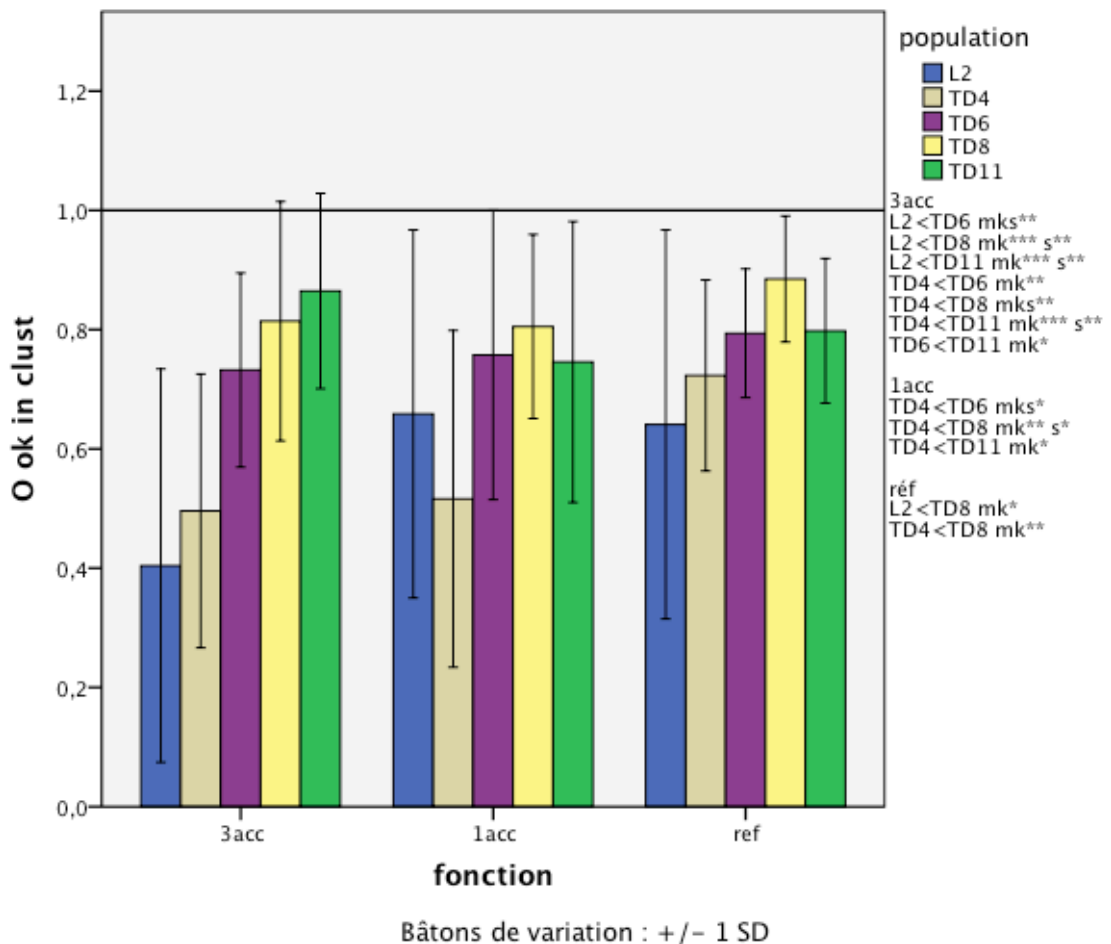


Figure 33. O ok in clust : comparaison des groupes

On notera que les TD11 ont des taux de réussite bas pour les clitiques 1acc et réf, ce qui est plutôt surprenant. En particulier, on constate que, dans la catégorie réf, les TD11 ne se distinguent pas des TD4, alors que les TD8 s'en distinguent bien. Toutes choses

égales par ailleurs, ceci indiquerait une perte de connaissances entre 8 et 11 ans, ce qui est tout à fait inattendu. En fait, tout n'est pas égal par ailleurs : une fois que les datifs grammaticaux sont inclus dans le décompte (i.e. dans **O ok+dat in clust**, Figure 34), le comportement des TD11 redevient normal. En outre, les trajectoires développementales se dessinent plus clairement, et ce, même dans la catégorie la moins affectée par l'absence des datifs, à savoir les 3acc.

Dans **O ok+dat in clust** :

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD6/8/11 (3 tests)
 - sans les datifs, l'opposition TD4 vs TD6 n'est pas significative pour *Kolmogorov-Smirnov*
- TD6 et TD11 (3 tests)
 - sans les datifs, cette opposition n'est pas significative pour *Kolmogorov-Smirnov*
- L2 et TD6/8/11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- TD4 et TD6/8/11 (3 tests) ;
 - sans les datifs, l'opposition TD4 vs TD11 n'est pas significative pour *Kolmogorov-Smirnov*
- TD6 et TD11 (*Kruskal-Wallis*)
 - sans les datifs, cette opposition n'est significative pour aucun test
- L2 et TD4 (3 tests) – notons que ce sont les L2 qui produisent le score plus haut
 - sans les datifs, cette opposition n'est significative pour aucun test
- L2 et TD11 (*Mann-Whitney* et *Kruskal-Wallis*)
 - sans les datifs, cette opposition n'est significative pour aucun test

Pour les clitiques réf, on observe une opposition entre

- TD4 et TD8 (*Mann-Whitney* et *Kruskal-Wallis*)/TD11 (3 tests)
 - sans les datifs, l'opposition TD4 vs TD11 n'est significative pour aucun test

- L2 et TD6 (*Mann-Whitney* et *Kruskal-Wallis*)/TD8/11 (3 tests)
 - sans les datifs,
 - les oppositions L2 vs TD6/TD11 ne sont significatives pour aucun test
 - l'opposition L2 vs TD8 n'est pas significative pour *Kolmogorov-Smirnov*.

La comparaison des résultats obtenus avec et sans datifs montre très clairement que les réponses à datif grammatical doivent être incluses dans le décompte en tant que réponses cible.

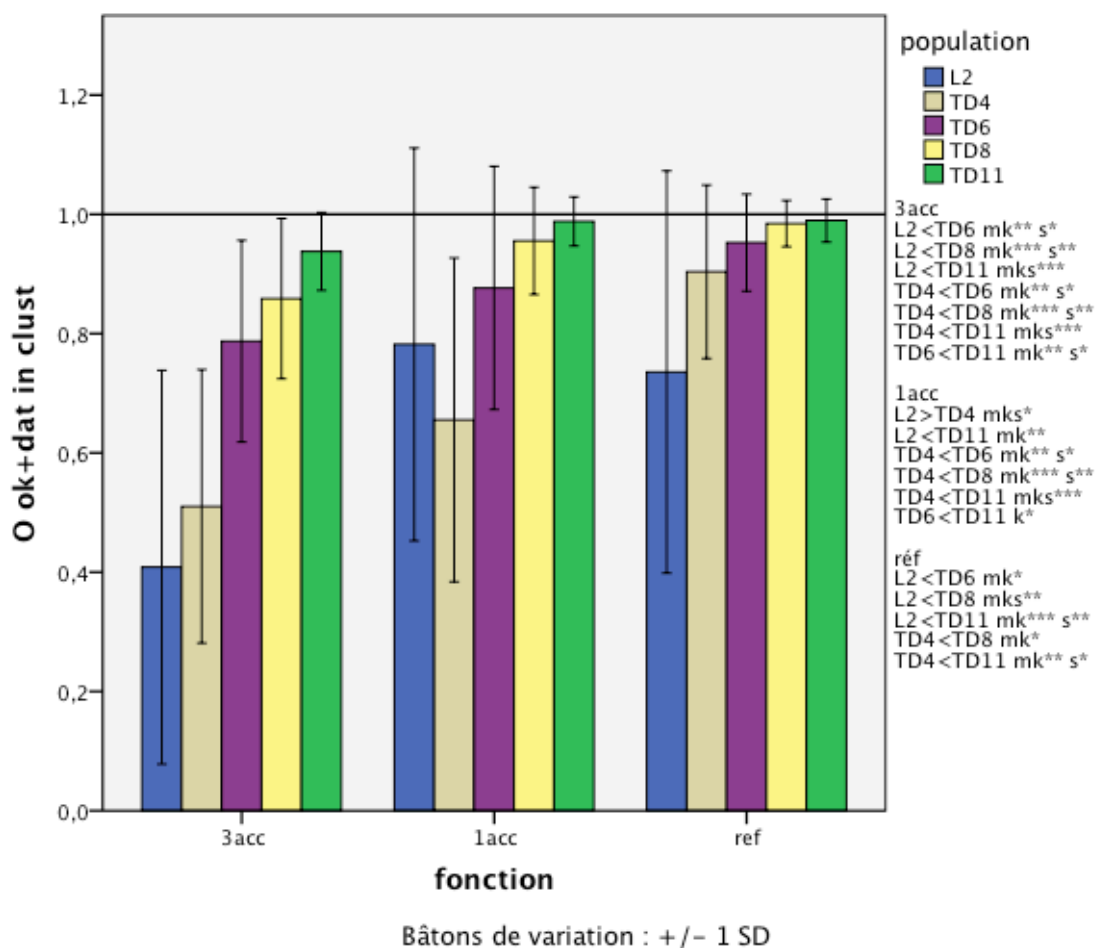


Figure 34. O ok+dat in clust : comparaison des groupes

Comparons maintenant la production (O prod in clust, Figure 30) et la production cible (O ok+dat in clust, Figure 34).

Pour les clitiques 3acc, la production oppose les TD6 aux TD8, ce qui n'est pas le cas de la production cible. En revanche, à la différence de la production, la production cible oppose les TD4 aux TD6. Pour les deux mesures, le développement continue au-delà de 6 ans (opposition TD6 vs TD11), mais il s'accélère à des âges différents : en production (Figure 30), les TD4/6 forment un groupe s'opposant aux TD8/11 ; en production cible (Figure 34), la frontière entre apprenants jeunes et âgés se situe entre les TD4 et les TD6. Les L2 se comportent comme des apprenants jeunes : ils se comportent comme les TD4/6 en production, et s'opposent aux TD6 en production cible.

Ces observations doivent être mises en rapport avec celles faites précédemment sur l'omission. Rappelons que l'omission des 3acc (Figure 31) définit une classe d'enfants jeunes (TD4/6) qui s'oppose aux TD11 tel qu'on l'observe pour la production (Figure 30). J'ai argumenté (section 3.3.2) que l'omission résulte de l'emploi d'un clitique sous-spécifié vide. Le développement d'**O ok+dat in clust** suggère que c'est la progression de l'accès aux formes morphologiquement bien formées qui déclenche l'élimination progressive du clitique vide. L'effet de ce phénomène devient visible en production (**O prod in clust**) et en omission (**O omission in clust**) dans la tranche d'âge suivante.

Pour les clitiques 1acc, le taux de production (**O prod in clust**, Figure 30) est en forte augmentation entre 4 et 6 ans et il n'évolue plus significativement au-delà de 6 ans. En production cible (Figure 34), on observe, là aussi, une forte augmentation entre TD4 et TD6, mais le développement tardif, au-delà de 6 ans, commence à être significatif (TD6 vs TD11). Rappelons ici que les formes morphologiquement bien formées (et donc l'élimination du clitique vide) évoluent plus rapidement en 1acc qu'en 3acc (Figure 17 et Tableau 1) : l'opposition 3acc vs réf se trouve chez les TD4 et les TD6, l'opposition 1acc vs réf n'existe que pour les TD4. Les L2 se comportent comme les TD6/8. Ceci indique que l'âge chronologique est important pour la catégorie 1acc. En fait, les L2 (d'un âge moyen de 9;7) distinguent 3acc vs 1acc en production et en production cible (Tableau 1), avec des taux bien élevés dans la catégorie 1acc.

Pour les clitiques réf, le développement en production est presque achevé à l'âge de 4 ans pour les TD. Seul *Kruskal-Wallis* détecte de faibles oppositions entre TD4 et TD8/11. En production cible, le développement au-delà de 4 ans est significatif pour les trois tests (TD4 vs TD11). Comme pour les clitiques 1acc, les L2 semblent "profiter" du fait que ce

sont les réponses cibles qui sont mesurées : en production, leurs taux sont inférieurs à ceux des TD4 mais en production cible, ils ne se distinguent pas des TD4. En fait, ceci est dû au fait que les TD4 ont un taux bas en production cible.

La comparaison des L2 avec les TD d'âges différents nous permet d'identifier un effet d'âge dans les catégories 1acc et réf. Cet effet sera confirmé lors de l'analyse de la structure interne de la population L2 dans la section 3.3.5.

Passons maintenant aux mesures des groupes de clitiques sujet et objet (**Target SO in clust**, Figure 33 ; **Target SO+dat in clust**, Figure 34). Ces mesures doivent être comparées aux mesures de production cible afin de déterminer l'effet du regroupement des clitiques sur la trajectoire du développement. Par ailleurs, la comparaison entre **Target SO in clust** et **Target SO+dat in clust** confirmera nos conclusions sur le statut des réponses à datif grammatical : les oppositions significatives pour **Target SO in clust** constituent un sous-ensemble propre des oppositions significatives pour **Target SO+dat in clust**. Cependant, l'effet est moins marqué dans la mesure des groupes qu'il ne l'était dans les mesures des clitiques objet. On note en particulier que la catégorie 1acc reste problématique pour les TD11 même si les réponses avec datif sont incluses. Je laisserai l'exploration de cette observation pour une recherche ultérieure. Dans la suite, je détaillerai les résultats des deux mesures dans une seule liste, commençant toujours avec **Target SO+dat in clust**.

Pour les clitiques 3acc, on observe une opposition entre

- TD4 et TD6/8/11 (3 tests)
- TD6 et TD11 (*Mann-Whitney* et *Kruskal-Wallis*)
- L2 et TD6/8/11 (3 tests)

Pour les clitiques 1acc, on observe une opposition entre

- TD4 et TD6/8/11 (3 tests)
- TD6 et TD8 (3 tests)
 - sans les datifs, cette opposition n'est significative pour aucun test
- L2 et TD6/8/11 (3 tests)
 - sans les datifs, l'opposition L2 vs TD6/TD11 n'est pas significative pour *Kolmogorov-Smirnov*

Pour les clitiques réf, on observe une opposition entre

- TD4 et TD6/8/11 (3 tests)
 - sans les datifs, l'opposition TD4 vs TD6/TD11 n'est pas significative pour *Kolmogorov-Smirnov*
- TD6 et TD11 (*Kruskal-Wallis*)
 - sans les datifs, cette opposition n'est significative pour aucun test
- L2 et TD6/8/11 (3 tests)
 - sans les datifs, seule l'opposition L2 vs TD8 est significative (3 tests)

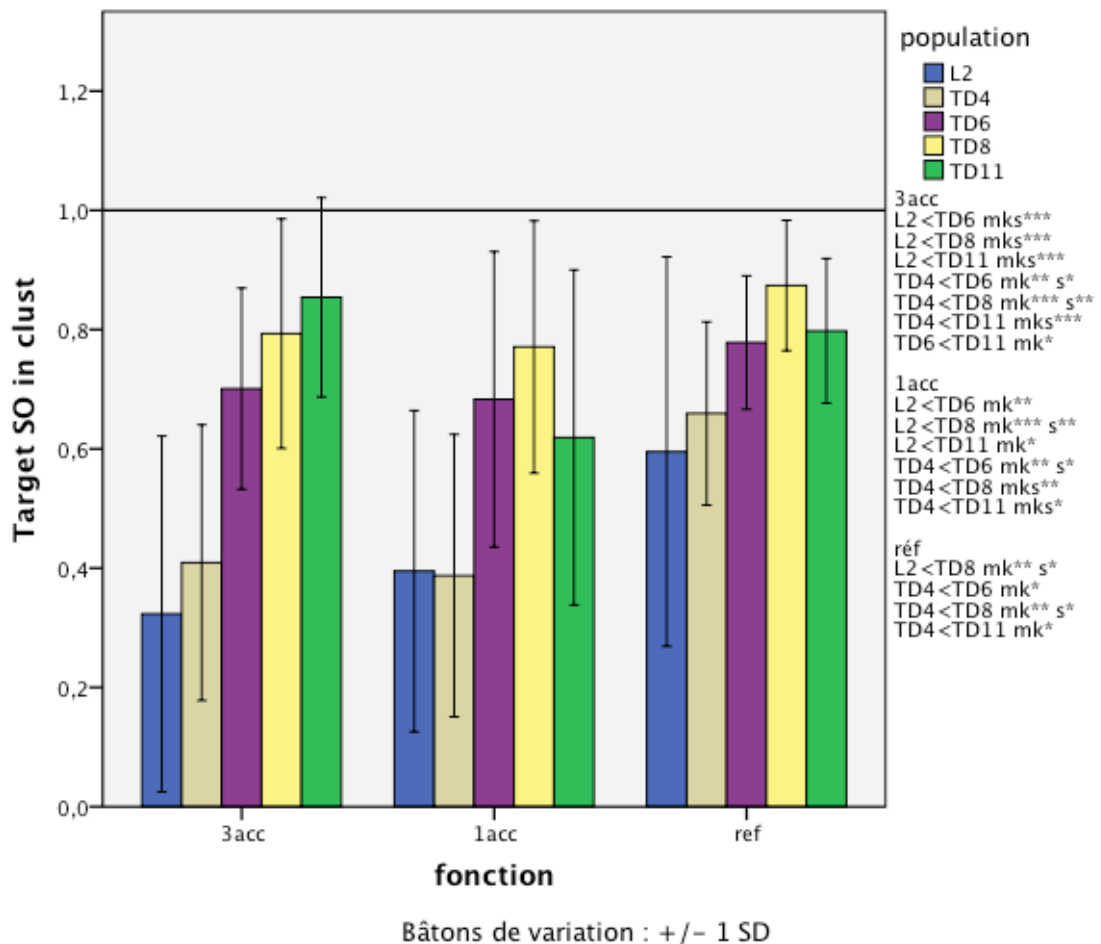


Figure 35. Target SO in clust : comparaison des groupes

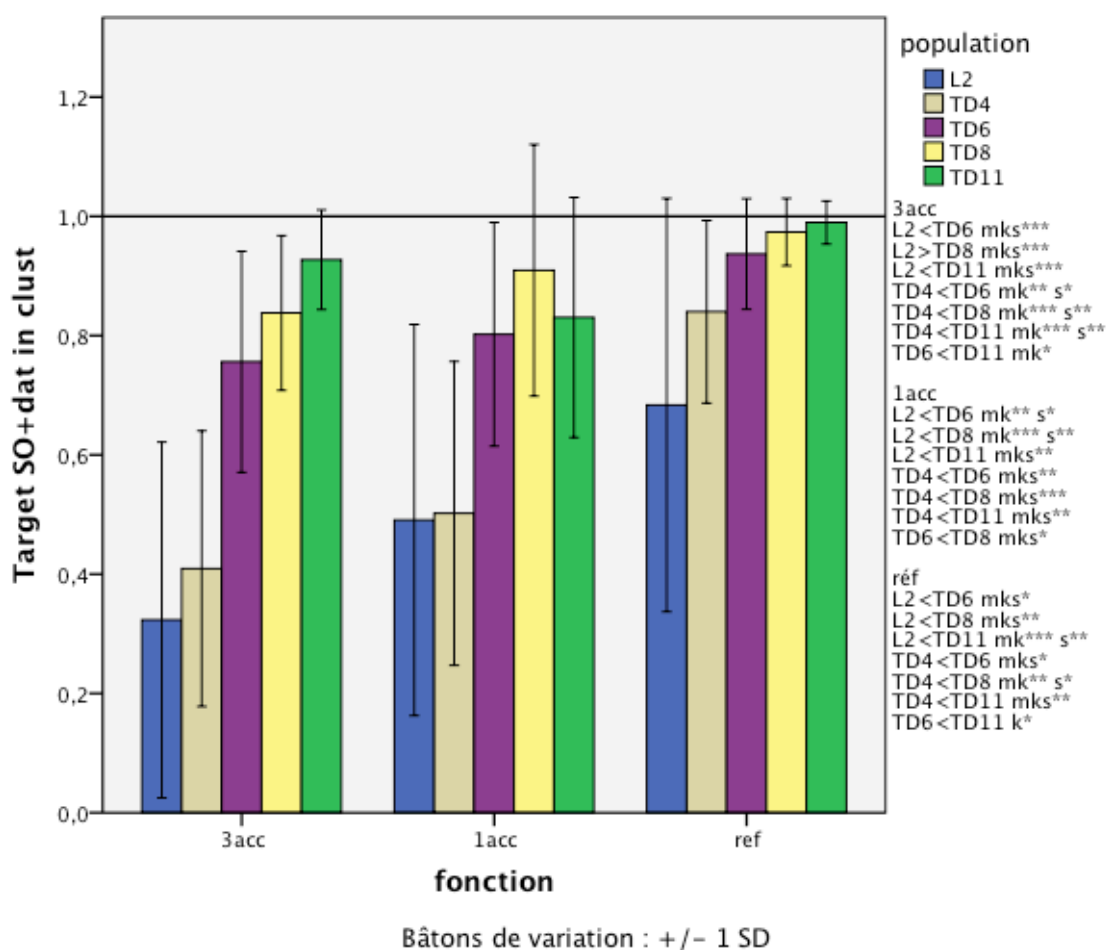


Figure 36. Target SO+dat in clust : comparaison des groupes

La comparaison entre la mesure des groupes (Target SO+dat in clust, Figure 36) et celle de la production cible (O ok+dat in clust, Figure 34) ne met au jour qu'une différence minime dans la catégorie des clitiques 3acc. Pour les deux mesures on observe une augmentation forte entre 4 et 6 ans et un développement tardif, au-delà de 6 ans. Cependant, ce développement est mis en évidence par les 3 tests en O ok+dat in clust, mais il n'est pas significatif pour Kolmogorov-Smirnov en Target SO+dat in clust.

Dans les catégories des clitiques 1acc et réf, on remarque tout d'abord les taux bas des TD11. Rappelons que ce phénomène a déjà été observé pour O ok in clust (Figure 33), mais là, l'inclusion des datifs (O ok+dat in clust, Figure 34) avait ramené le comportement dans la norme. Ici, Target SO+dat in clust (Figure 36) fait monter le taux des TD11 dans la catégorie réf, mais pas dans la catégorie 1acc. Ceci met en évidence que les TD11 ont un problème avec une configuration précise, celle impliquant un clitique sujet de troisième personne et un clitique objet de première personne.

On observe également que les L2 se comportent de façon nettement différente en ce qui concerne les clitiques 1acc. Pour **O ok+dat in clust**, les L2 ont un taux supérieur à celui des TD4 et inférieur à celui des TD11 ; pour **Target SO+dat in clust**, le taux des L2 est inférieur à celui des TD6. La mesure des groupes de clitiques annule donc l'avantage dû à l'âge dont profitent les L2 dans la mesure de la morphologie cible des clitiques objet.

0.1.1.1. Mesures sur l'échantillon enchâssement avec effet de population: O ok in sep et Target SO in sep

Passons maintenant aux contextes requérant la séparation des clitiques. Dans cet échantillon, les données sont nettement moins nombreuses. Les conclusions sont donc moins fortes que celles obtenues dans les catégories précédentes. Cependant, une généralisation robuste émerge : on n'observe presque aucun développement en fonction de l'âge des sujets. (Seules deux mesures conservent quelques oppositions : il s'agit de **O ok in sep** et de **target SO in sep**.) Nous noterons également que les taux de réussite sont très élevés, déjà parmi les TD4 et les L2.

Dans la mesure **O ok in sep** (Figure 37) :

Pour les clitiques 3acc, aucune opposition n'est significative.

Pour les clitiques 1acc, les L2 se distinguent des TD11 (*Kruskal-Wallis*), mais dans la direction opposée à celle attendue : le taux des TD11 est inférieur à celui des L2. Rappelons que nous avons déjà noté à plusieurs reprises que les TD11 ont des difficultés avec la catégorie 1acc.

Pour les clitiques réf, les TD8 se distinguent des TD6 et des L2 (*Kruskal-Wallis*).

Ici encore, ce sont les TD8 qui, étonnamment, ont un taux moins élevé (à savoir 0%) que les TD6 et les L2. En fait, les deux enfants parmi les TD8 qui ont recours à l'enchâssement dans la catégorie des clitiques réf produisent des datifs grammaticaux.

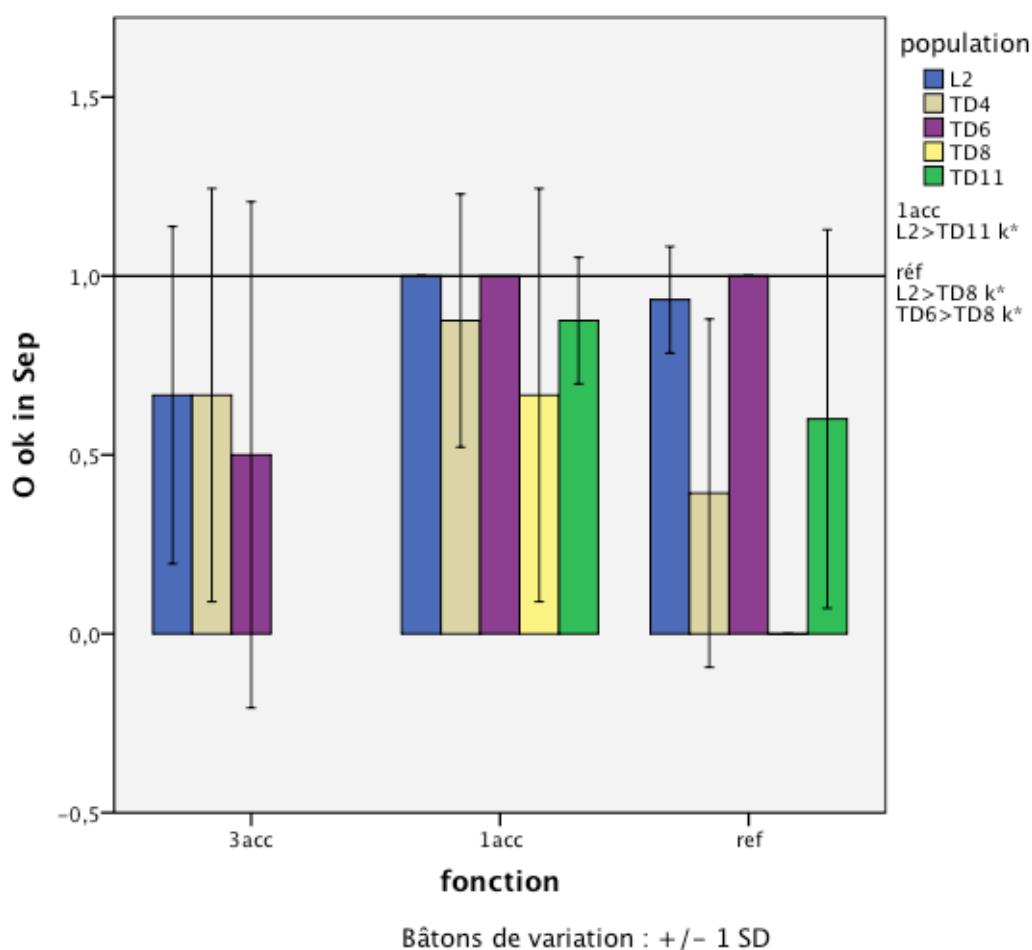


Figure 37. O ok in sep : comparaison des groupes

Pour la mesure **Target SO in sep** (Figure 38) :

Pour les clitiques 3acc et 1acc, aucune opposition n'est significative.

Pour les clitiques réf, les TD8 (avec un taux de réussite zéro) se distinguent des TD6 et des L2 (*Kruskal-Wallis*) pour la raison précédemment citée : les deux enfants parmi les TD8 qui ont recours à l'enchâssement dans la catégorie des clitiques réf produisent des datifs grammaticaux.

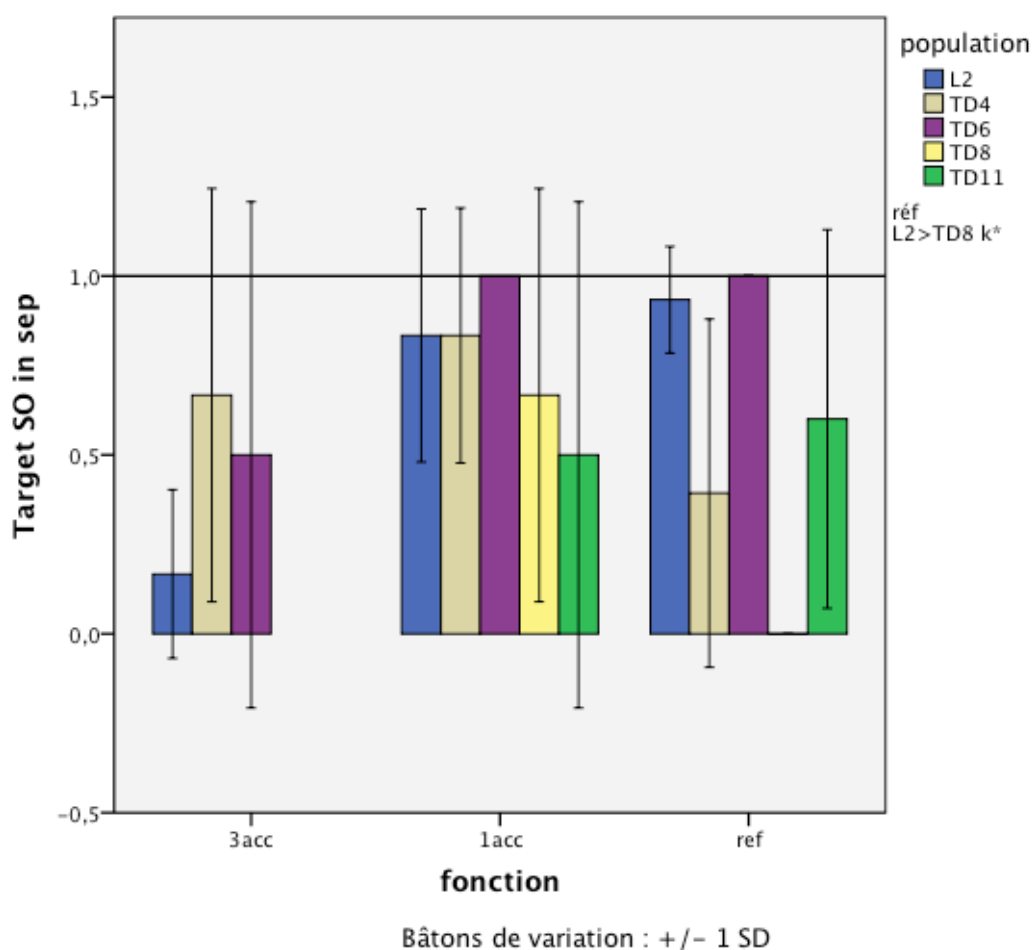


Figure 38. Target SO in sep : comparaison des groupes

3.3.3.3. Les mesures in sep sans effet de population

Les autres mesures (**O prod in sep**, **O ok+dat in sep**, **Target SO+dat in sep**) ne génèrent aucune opposition significative entre les différents groupes de population dans l'échantillon des réponses à enchâssement (requérant la séparation des clitiques). Or, ces mesures sont celles qui évoluent le plus dans l'échantillon de réponses requérant le regroupement des clitiques.

Pour **target SO+dat in sep**, ainsi que pour **O ok+dat in sep**, l'absence d'effet statistique pourrait très bien être imputée au fait que les effectifs de ces catégories sont trop réduits. Cependant, la mesure de production de clitiques objet avec ou sans erreur morphologique et incluant les datifs (**O prod in sep**, Figure 39) fournit des chiffres suffisants à la détection d'oppositions entre les groupes de population. Néanmoins,

aucune opposition n'est détectée. En particulier, les TD4 plafonnent déjà pour cette mesure, et cela même pour les clitiques 3acc. Rappelons ici que les TD4 recourent à l'enchâssement à un taux relativement élevé (Figure 15, Figure 29), ce qui rend leur bonne performance d'autant plus intéressante.

Cette observation souligne une remarque que j'ai déjà faite précédemment : le regroupement des clitiques a une influence sur la trajectoire de l'acquisition. Dans les contextes excluant le regroupement de clitiques, la production des clitiques objet a atteint le niveau de l'état adulte dès l'âge de 4 ans. Par conséquent, les difficultés prolongées avec les clitiques objet du français, telles qu'elles sont signalées dans la littérature, ne peuvent pas être attribuées au statut syntaxique du clitique (qui est identique en contexte racine et en contexte d'enchâssement), mais plutôt au fait que le clitique objet est typiquement réalisé à l'intérieur d'un groupe de clitics.

Cela dit, les données présentées ici n'excluent pas qu'il y ait des effets, dus à la représentation sous-jacente du clitique 3acc, qui se manifestent avant l'âge de 4 ans. Le fait que les 3acc soient moins stables que les 1acc et réf (Figure 39, Figure 40 et Figure 41) en fait va dans ce sens.

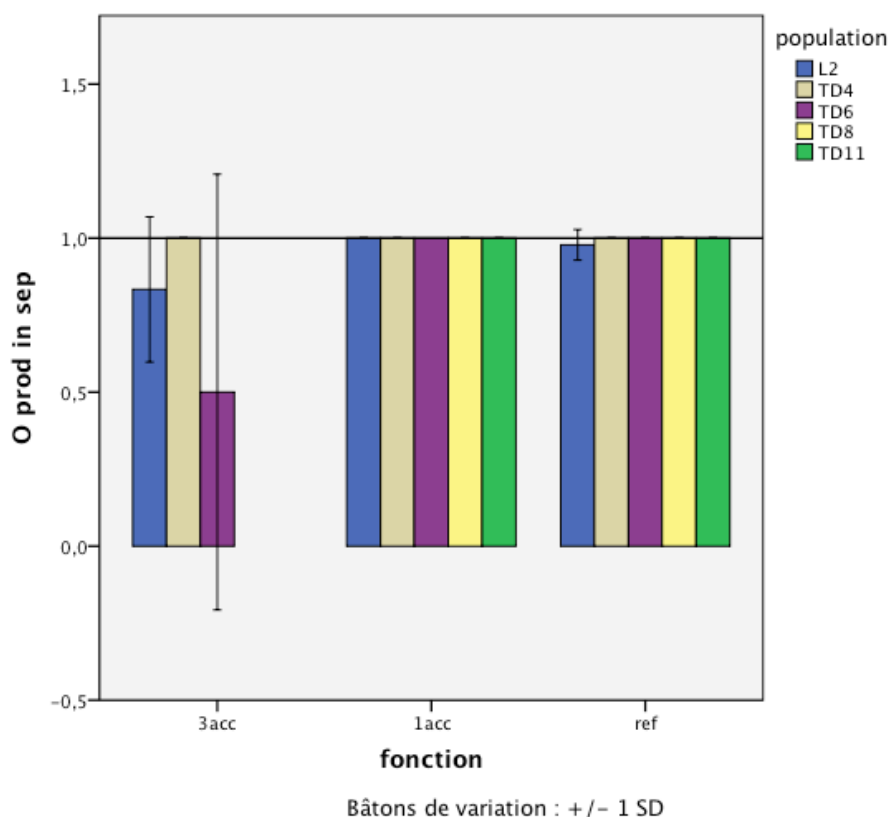


Figure 39. O prod in sep : comparaison des groupes

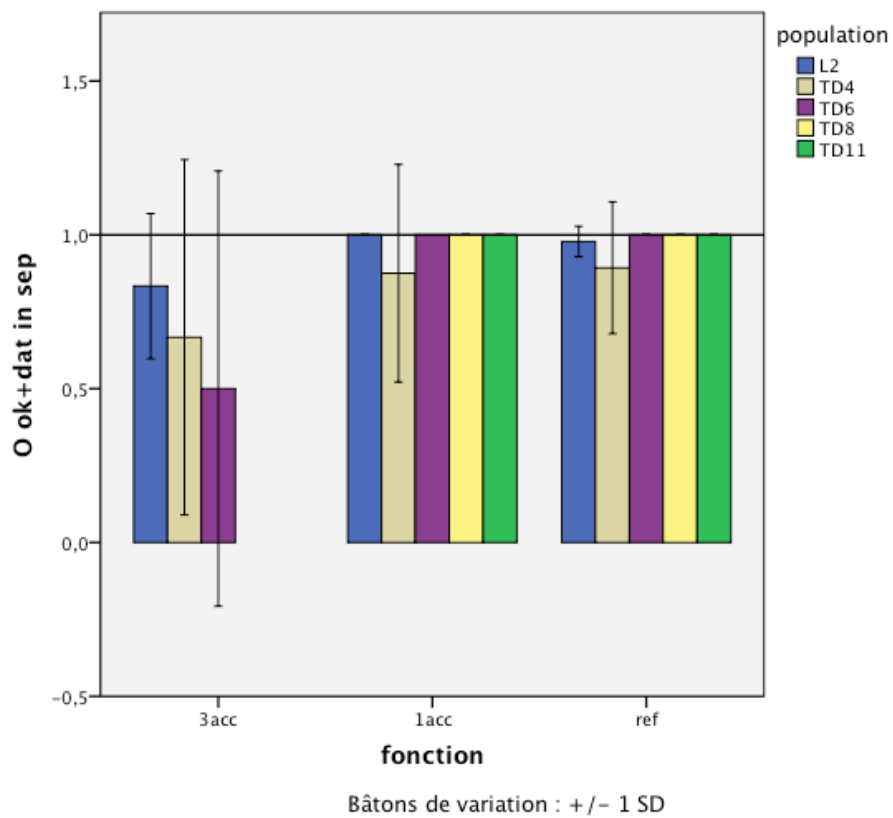


Figure 40. O ok+dat in sep : comparaison des groupes

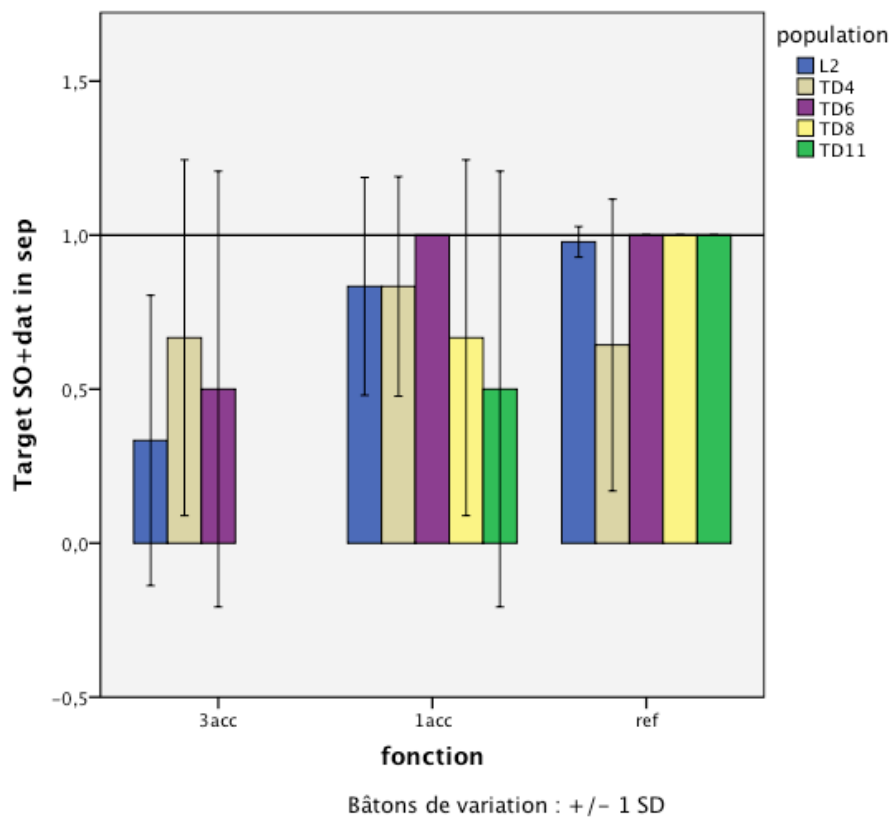


Figure 41. Target SO+dat in sep : comparaison des groupes

3.3.4 Différences entre populations – résumé

Dans le tableau suivant, je résume les différences entre les groupes. J'utiliserai les conventions de notation suivantes :

- "x<y" signifie "x est moins avancé que y" :
 - pour les mesures de production : $x < y$ signifie "x a un score supérieur à celui de y"
 - pour les mesures d'omission et de substitution : $x < y$ signifie "x a un score inférieur à celui de y"
- Les chiffres font référence aux tranches d'âge TD.
- Les parenthèses indiquent le test statistique pour lequel la différence est significative, au cas où il y a désaccord entre les tests :
 - k : significatif pour *Kruskal-Wallis*
 - m : significatif pour *Mann-Whitney*
 - s : significatif pour *Kolmogorov-Smirnov*

Par exemple, "6<8 (mk) " dans la colonne **O prod trad**. signifie que les TD6 ont un score moins élevé que les TD8 (significatif pour *Mann-Whitney* et *Kruskal-Wallis*).

Enfin, les codes couleur sont définis de la façon suivante :

- orange : les TD11 sont plus avancés (<11)
- rouge : les TD8 sont plus avancés (<8)
- bordeaux : les TD6 sont plus avancés (<6)
- violet foncé : les TD4 sont plus avancés (<4)
- vert : les L2 sont plus avancés (<L2)

<u>3acc</u>	O prod trad.	O prod in clust	O omission trad	O omission in clust	DP trad	DP in clust	O ok clust	O ok+dat in clust	Target SO in clust	Target SO+dat clst
TD4		4<8 (k)					4<6 (mk)	4<6	4<6	4<6
							4<8	4<8	4<8	4<8
	4<11	4<11	4<11 (k)	4<11 (k)	4<11 (k)	4<11 (k)	4<11	4<11	4<11	4<11
TD6	6<8 (mk)	6<8 (mk)								
	6<11	6<11	6<11	6<11		6<11 (k)	6<11 (mk)	6<11	6<11 (mk)	6<11
L2							L2<6	L2<6	L2<6	L2<6
	L2<8	L2<8	L2<8 (mk)	L2<8 (mk)			L2<8	L2<8	L2<8	L2<8
	L2<11	L2<11	L2<11	L2<11	L2<11 (mk)	L2<11 (mk)	L2<11	L2<11	L2<11	L2<11

Tableau 2. 3acc : différences entre groupes

<u>1acc</u>	O prod trad.	O prod in clust	O omission trad	O omission in clust	DP trad	DP in clust	O ok clust	O ok+dat in clust	Target SO in clust	Target SO+dat clst
TD4	4<6 (mk)	4<6		4<6 (k)			4<6	4<6	4<6	4<6
	4<8	4<8	4<8 (mk)	4<8 (mk)			4<8	4<8	4<8	4<8
	4<11	4<11	4<11	4<11			4<11 (mk)	4<11	4<11	4<11
TD6								6<11 (k)		6<8
L2									L2<6 (mk)	L2<6
								4<L2	L2<8	L2<8
	L2<11 (k)	L2<11 (mk)	L2<11 (k)	L2<11 (k)				L2<11 (mk)	L2<11 (mk)	L2<11

Tableau 3. 1acc : différences entre groupes

réf	O prod trad.	O prod in clust	O omission trad	O omission in clust	DP trad	DP in clust	O ok clust	O ok+dat in clust	Target SO in clust	Target SO+dat clst
TD4	4<6 (k)	4<6 (k)	4<6 (k)	4<6 (k)					4<6 (mk)	4<6
							4<8 (mk)	4<8 (mk),	4<8	4<8
	4<11 (k)	4<11 (k)	4<11 (k)	4<11 (k)				4<11	4<11 (mk)	4<11
TD6										6<8 (k)
										6<11 (k)
L2	L2<4 (mk)	L2<4 (mk)	L2<4 (mk)	L2<4 (mk)	L2<4 (mk)	L2<4 (mk)				
	L2<6	L2<6	L2<6	L2<6	L2<6 (mk)	L2<6 (mk)		L2<6 (mk)		L2<6
	L2<8	L2<8	L2<8	L2<8	L2<8 (mk)	L2<8 (mk)	L2<8 (mk)	L2<8	L2<8	L2<8
	L2<11	L2<11	L2<11	L2<11	L2<11 (mk)	L2<11 (mk)		L2<11		L2<11

Tableau 4. Réf : différences entre groupes

Je voudrais m'attarder sur quatre ensembles d'observations. Premièrement, l'influence du facteur morphologique sur la trajectoire de l'acquisition des clitiques se trouve confirmée (points 1-4 et 8 ci-dessous). Deuxièmement, nous avons confirmé que les réponses à datif grammatical comptent comme production cible (point 5 ci-dessous). Troisièmement, nous avons observé une particularité des TD11 : ce groupe rencontre des difficultés avec la configuration clitique sujet de troisième personne + clitique objet de première personne (point 6 ci-dessous). Quatrièmement, nous avons observé une spécificité toute particulière de la population L2 dans la catégorie des clitiques réf : la performance dans cette catégorie est fortement retardée chez les L2 alors qu'elle se développe tôt chez les TD (point 7 ci-dessous).

1. L'acquisition de la morphologie des clitiques s'achève tardivement, c'est-à-dire après 6 ans, et ce phénomène est lié (en partie au moins) à la morphologie : plus le poids de la morphologie augmente dans une mesure (voir l'échelle présentée dans la section 3.3.2), plus les TD6 (au milieu des trois tableaux) s'opposent aux enfants plus âgés.

Ainsi, dans la catégorie des clitiques 3acc, les TD6 s'opposent significativement aux enfants plus âgés pour toutes les mesures sauf la substitution (**DP trad.**). Or, c'est justement dans cette dernière mesure que la morphologie des clitiques ne joue aucun rôle. Dans la catégorie des clitiques 1acc, les TD6 ne se distinguent des enfants plus âgés que pour les deux mesures dans lesquelles le poids de la morphologie est le plus élevé (**O ok+dat in clust**, **Target SO+dat in clust**). En outre, cette opposition n'est significative que pour un seul test (*Kruskal-Wallis*) pour **O ok+dat in clust**, mais elle l'est pour les 3 tests pour **Target SO+dat in clust**. Enfin, dans la catégorie des clitiques réf, l'acquisition des clitiques est achevée à 6 ans : on ne relève qu'une seule opposition significative entre TD6 et TD8/11 (dans **Target SO+dat in clust**), et cela pour un seul test (*Kruskal-Wallis*).

2. Je propose qu'il existe un lien causal entre les taux en production cible, en omission et en production. Cette proposition repose sur l'hypothèse que les apprenants jeunes ont accès à un clitique vide sous-spécifié. Tant que cet élément est accessible, l'omission ne constitue pas une violation de grammaticalité et le taux de production reste relativement bas. Puis, au fur et à mesure qu'ils sont acquis, les clitiques entièrement spécifiés (et phonologiquement réalisés) de la langue adulte remplacent le clitique vide. J'ai argumenté que l'augmentation forte de formes cibles (entre 4 et 6 ans pour les 3acc) déclenche l'élimination de la forme vide par le principe *elsewhere* (Kiparsky, 1973). Ainsi j'ai pu expliquer un fait très surprenant : la production cible des clitiques 3acc augmente fortement entre 4 et 6 ans, **avant** la production qui, elle, augmente fortement entre 6 et 8 ans.

3. Le facteur morphologique n'est pas un facteur absolu : il conserve l'ordre d'acquisition des clitiques, réf avant 1acc avant 3acc. En effet, les mesures qui évoluent encore après 6 ans pour les réf constituent un sous-ensemble propre de celles qui évoluent après 6 ans pour les 1acc, et ces dernières constituent un sous-ensemble propre des mesures évoluant encore après 6 ans pour les 3acc.

4. Le regroupement des clitiques diminue la performance et ralentit l'acquisition des clitiques.

Ainsi, dans l'échantillon des réponses à enchâssement, requérant la séparation des clitiques, on observe des taux de réussite très élevés chez les enfants les plus jeunes

(Figure 37, Figure 38, Figure 39, Figure 40, Figure 41) et l'absence d'opposition significative intéressante entre les groupes.

Cette observation sur un échantillon réduit est confirmée par l'examen d'un échantillon suffisamment important, celui constitué par la comparaison de **O prod trad.** et **O prod in clust** : l'inclusion des réponses à enchâssement dans **O prod trad.** fait augmenter le taux de réussite des TD4. Par conséquent, si l'on exclut les réponses à enchâssement (**O prod in clust**) :

- pour les clitiques 3acc (Tableau 2) : l'opposition entre TD4 et TD8 devient significative (*Kruskal-Wallis*).

- pour les clitiques 1acc (Tableau 3) : l'opposition entre TD4 et TD6 devient significative (3 tests), alors qu'elle ne l'est pas pour *Kolmogorov-Smirnov* si l'on compte les réponses à enchâssement (**O prod trad.**).

Par ailleurs, le regroupement des clitiques a un effet clair dans la catégorie 1acc. Dans la presque totalité des mesures en 1acc, les L2 bénéficient d'un avantage par rapport aux TD4 en raison de leur âge. Cet avantage disparaît dans **Target SO+dat in clust.** : la mesure de la grammaticalité des groupes de clitiques fait baisser le taux des L2.

5. Les réponses à datif grammatical se comportent comme des réponses cible, et leur exclusion entraîne des anomalies.

Mon raisonnement a la forme suivante. Admettons que si l'on compte un ensemble plus large de réponses comme réponses cibles d'une mesure, alors l'état final (à savoir un taux de 100%) est atteint plus tôt. Admettons également pour le moment que les datifs grammaticaux sont des réponses erronées mettant en évidence une stratégie d'évitement des clitiques accusatif. En tant que stratégie d'évitement, l'emploi des datifs devrait diminuer au fur et à mesure que les clitiques accusatif sont acquis. Si l'on inclut les datifs dans une mesure, on devrait donc constater que le développement s'achève plus tôt ; si on les exclut, on devrait observer un développement prolongé. Or, on observe le contraire. Pour les clitiques 3acc (Tableau 2), les 3 tests statistiques établissent solidement un développement tardif (une opposition entre TD6 et TD11) pour **O ok+dat in clust** et pour **Target SO+dat in clust**. Si les datifs sont exclus (*i.e.* dans **O ok in clust** et **Target SO in clust**), cette opposition n'est plus significative pour le test de *Kolmogorov-Smirnov*. Pour les clitiques 1acc (Tableau 3 – **O ok/+dat in clust**, **Target SO/+dat in clust**), on observe un développement tardif (TD6 vs TD8/11)

significatif avec les datifs, qui disparaît si les datifs sont exclus. Il en est de même des clitiques réfs (Tableau 4 – **Target SO/+dat in clust**). J'en conclus que les réponses à datif ne doivent pas être écartées en tant qu'erreurs résultant d'une stratégie d'évitement. Au contraire, elles doivent être considérées comme réponses cibles.

6. Notons cependant que les TD11 présentent une anomalie qui ne disparaît pas avec l'inclusion des datifs. Dans **Target SO+dat in clust**, les TD11 présentent un taux de clitiques 1acc bas. Ceci met en évidence que cette tranche d'âge a un problème avec la configuration « clitique sujet de troisième personne + clitique objet de première personne ». Je laisse l'exploitation et l'approfondissement de cette observation pour une recherche ultérieure.

7. La population L2 (dans la dernière partie des Tableau 2, Tableau 3, Tableau 4) se distingue nettement de la population TD en ce qui concerne la catégorie des clitiques réf.

Les L2 ont une performance similaire à celle des TD4/6 pour la catégorie 3acc ; pour la catégorie 1acc, la performance des L2 est similaire à celle des TD6/8. Par extension de l'opposition entre les fonctions observée chez les TD (Tableau 1), on devrait s'attendre à ce que la performance des L2 dans la catégorie réf soit encore plus élevée. Or, on observe que les L2 ont des taux inférieurs même à ceux des TD4 (en violet foncé dans le Tableau 4) pour les mesures de production, d'omission et de substitution. Cette observation invite à spéculer sur le statut grammatical des réfléchis dans les deux langues concernées, le français et l'anglais. En fait, (Marelj, 2004, Reinhart et Siloni, 2005) proposent une classification typologique opposant les langues qui forment les prédicats réfléchis en syntaxe à celles qui les forment dans le lexique (voir aussi (Haiden, 2005a)). L'anglais dériverait les réfléchis dans le lexique, le français en syntaxe, d'où des paires telles que "Max shaved." et "Max s'est rasé." Si les anglophones apprenant le français en tant que L2 transfèrent la dérivation lexicale des prédicats réfléchis en français, alors on s'attend à ce qu'ils produisent peu de pronoms réfléchis, et à ce qu'ils présentent des taux d'omission (Figure 27, Figure 31) et de substitution (Figure 28, Figure 32) élevés, ce qui est vrai.

Notons ici que ce type d'omission doit être distingué de celui observé pour les catégories 3acc et 1acc. En fait, la dérivation lexicale d'un prédicat réfléchi rend ce prédicat syntaxiquement intransitif. La stratégie du clitique vide n'est donc pas accessible dans cette configuration. La situation est différente pour la substitution des

clitiques réfléchis par un DP plein. Dans ce cas, il s'agit d'un évitement de la réflexivisation : la structure avec complément DP n'est pas réfléchie, mais transitive.

8. Le facteur morphologique observé précédemment chez les TD (cf. plus haut dans cette section et en section 3.3.2) joue uniformément dans toutes les mesures et dans toutes les catégories chez les L2 : plus le poids de la morphologie augmente dans une mesure, plus les L2 se comportent comme les TD jeunes.

Ainsi, on observe, pour les clitiques 3acc (Tableau 2), que les L2 se comportent comme les TD6 dans les mesures de production (**O prod trad.**, **O prod in clust**) et d'omission (**O omission trad.**, **O omission in clust**), mais qu'ils ont des taux inférieurs à ceux des TD6 dans les mesures de production cible (**O ok/+dat in clust**, **Target S0/+dat in clust**). Pour les clitiques 1acc (Tableau 3), les L2 se comportent comme les TD8 dans toutes les mesures d'omission, de production et de production cible à l'exception de **Target S0/+dat in clust**, dans laquelle ils présentent un taux inférieur à celui des TD6 ! Seuls les réfléchis (Tableau 4) inversent la tendance, mais ceci ne signifie pas que les taux des L2 augmentent. En fait, c'est le retard des enfants TD dans les mesures les plus morphologiques (**O ok/+dat in clust**, **Target S0/+dat in clust**) qui masque la difficulté qu'ont les L2 avec les réfléchis.

3.3.5 La structure de la population L2

La population L2 est caractérisée par quatre variables numériques : l'âge chronologique de l'apprenant au moment du test (age), l'âge de l'apprenant au moment du début de l'exposition au français (onset), la durée de l'exposition au français au moment du test (length) et le niveau en langue maternelle. Le niveau en langue maternelle a été déterminé à l'aide du CELF-IV. Afin de m'assurer d'obtenir un résultat fiable, j'ai calculé les corrélations pour le score standardisé (celf) et l'écart-type (celfET). Ces calculs ont donné le même résultat. Je ne donnerai donc désormais que les corrélations basées sur le score standardisé (celf).

Mon objectif dans cette section est de dégager la structure de la population L2. A cette fin, je présenterai les corrélations entre les 4 variables de classement que je viens de mentionner et les mesures précédemment discutées. J'ai appliqué les tests non paramétriques suivants : le TAU-b de *Kendall* et le Rho de *Spearman* (le test de *Pearson*

est exclu car les données ne sont pas distribuées normalement). Notons d’emblée que ces tests ne mettent en évidence aucune corrélation avec l’âge au début de l’exposition au français (onset). Les résultats pour cette variable ne seront donc pas présentés dans la suite. En outre, aucune corrélation n’est significative dans l’échantillon des réponses à enchâssement, celles-ci ne seront donc plus mentionnées.

Dans les graphiques, les coefficients de corrélation et le niveau de significativité sont indiqués selon les conventions déjà utilisées précédemment :

- les tests sont indiqués par une lettre (k pour *Kendall*, s pour *Spearman*),
- cette lettre est suivie du coefficient de corrélation, suivi de 1 à 3 astérisques (* pour $p < 0,05$; ** pour $p < 0,01$; *** pour $p < 0,001$).
- R^2 indique, pour chaque ligne d’ajustement, sa fiabilité : plus ce chiffre s’approche de 1, plus la courbe s’ajuste aux données.

3.3.5.1. Mesures sur la totalité des réponses pertinentes

La production de clitiques objets dans toutes les réponses pertinentes (**O prod trad.**) est corrélée à la durée d’exposition (Figure 42) dans les trois catégories (3acc, 1acc, réf). Pour les 1acc et réf, mais pas pour les 3acc, elle est également corrélée à l’âge (Figure 43). En revanche, pour les 3acc, et non pour les 1acc et réf, elle est corrélée au niveau en L1 (celf, Figure 44).

Les clitiques 3acc sont donc corrélés à la durée d'exposition et au niveau en L1 ; les clitiques 1acc et réf sont corrélés à la durée et à l’âge.

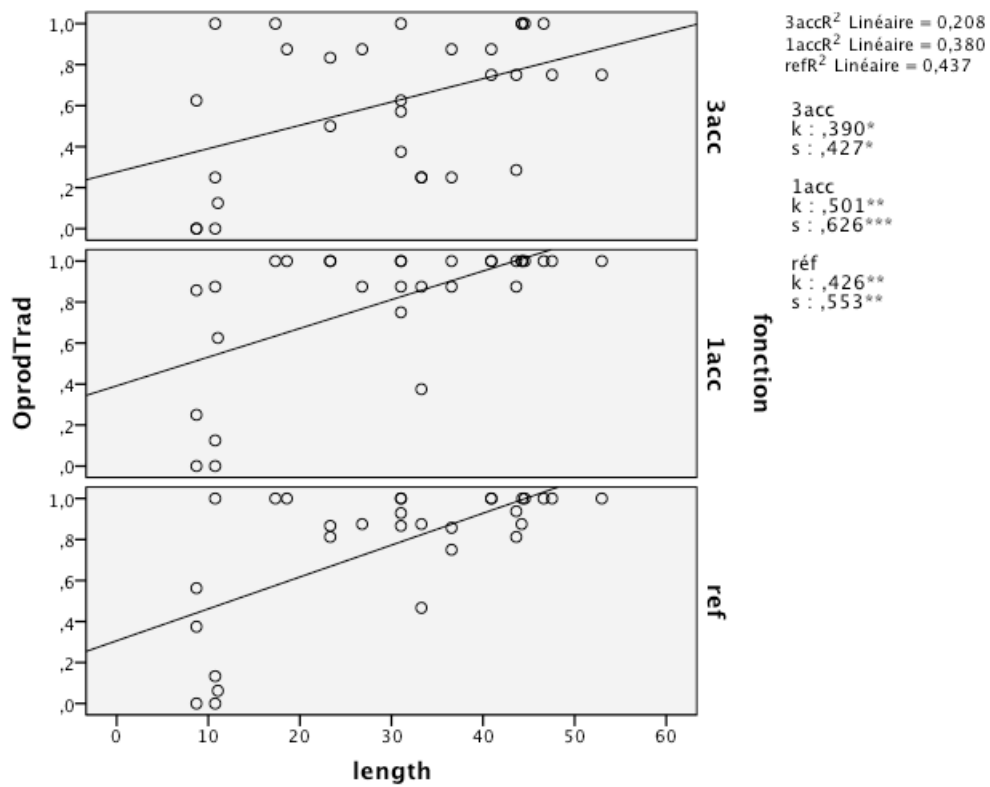


Figure 42. O prod trad. et durée

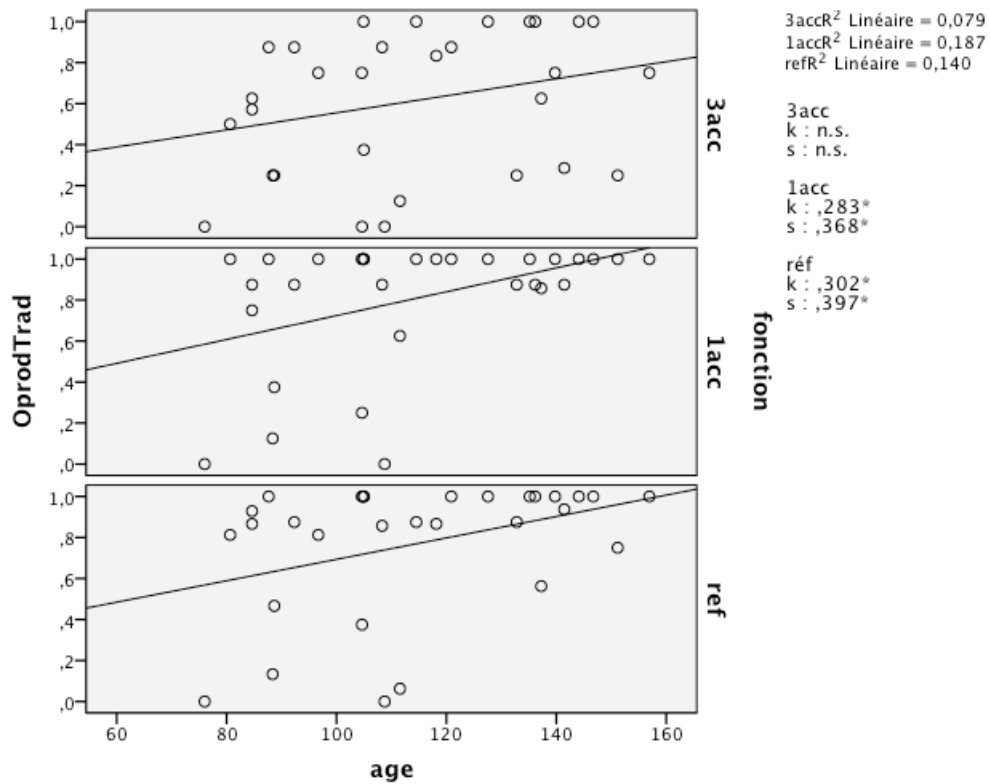


Figure 43. O prod trad et âge

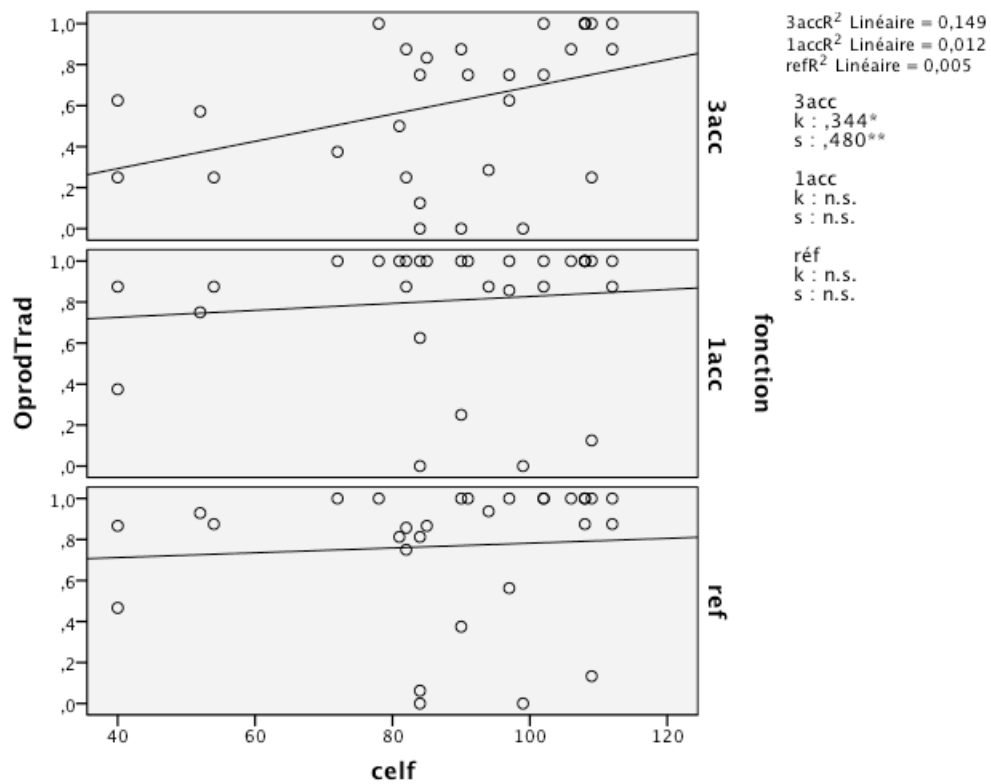


Figure 44. O prod trad et celf

L'omission des clitiques (**O omission trad**) est corrélée à la durée d'exposition pour les clitiques 1acc et réf (Figure 45). Aucune autre corrélation n'est significative.

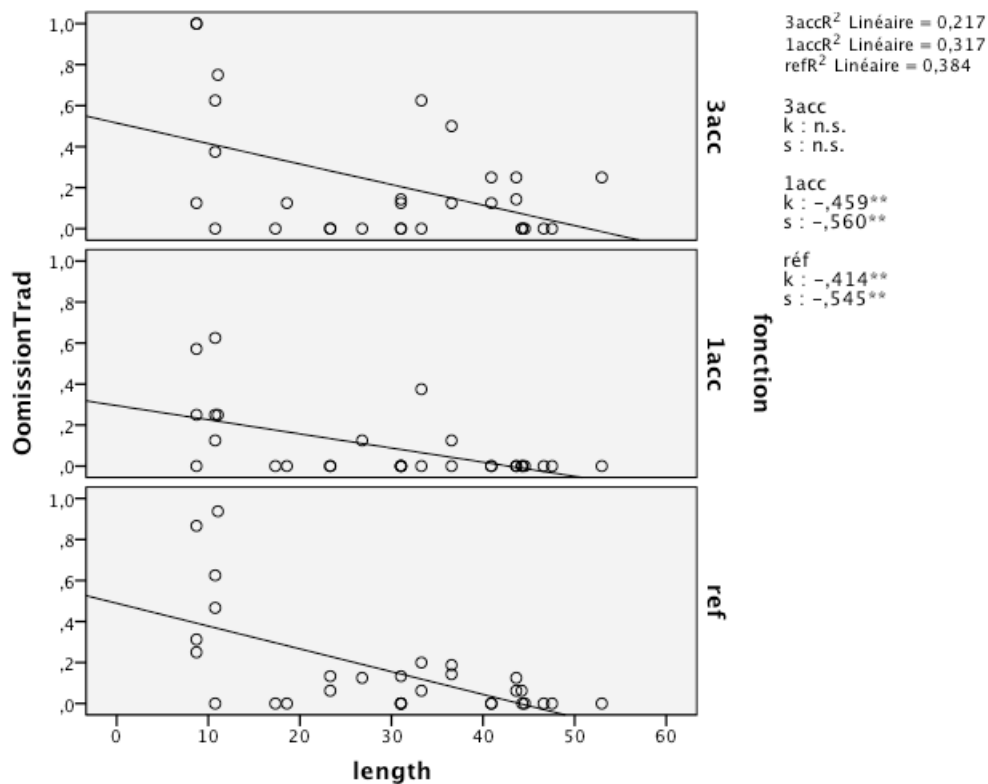


Figure 45. O omission trad et durée

La substitution du clitique objet par un élément tonique (**DP trad.**) est corrélée à la durée d'exposition (Figure 46). Le test de *Spearman* détecte une corrélation entre les clitiques réfléchis et l'âge (Figure 47). Les clitiques 3acc ne sont corrélés ni à la durée d'exposition, ni à l'âge, mais au niveau en L1 (corrélation détectée par les deux tests, Figure 47). Cette observation sera cruciale dans la suite (section 3.3.5.3) pour l'identification des enfants présentant un risque de trouble d'acquisition.

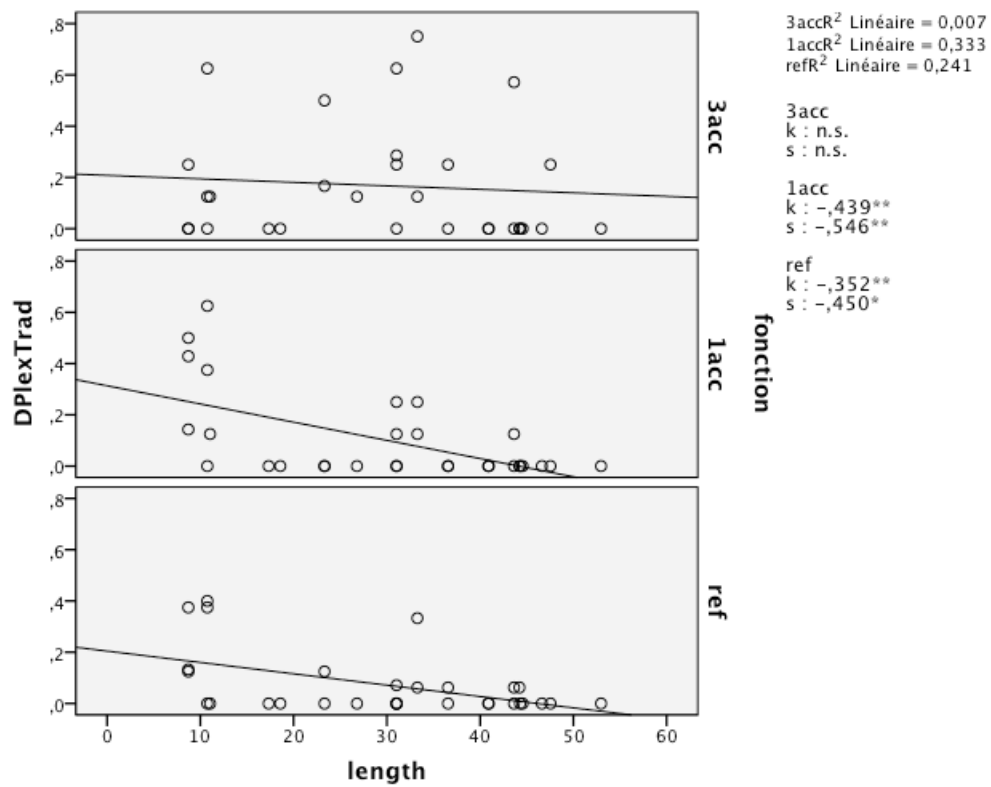


Figure 46. DP trad et durée

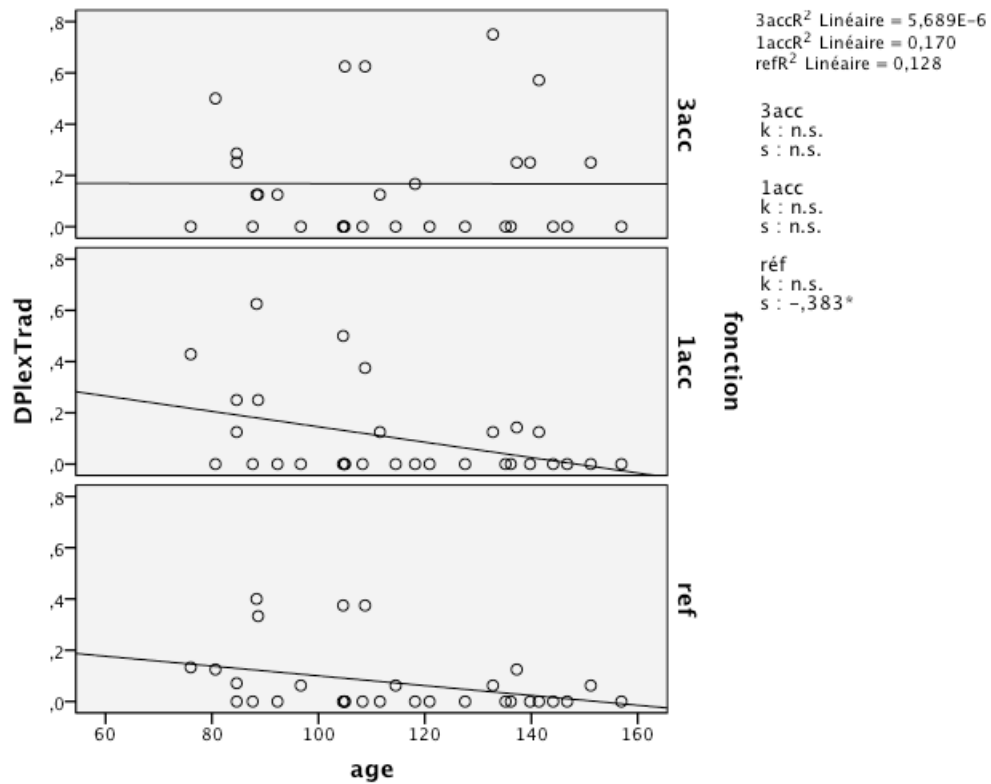


Figure 47. DP trad et âge

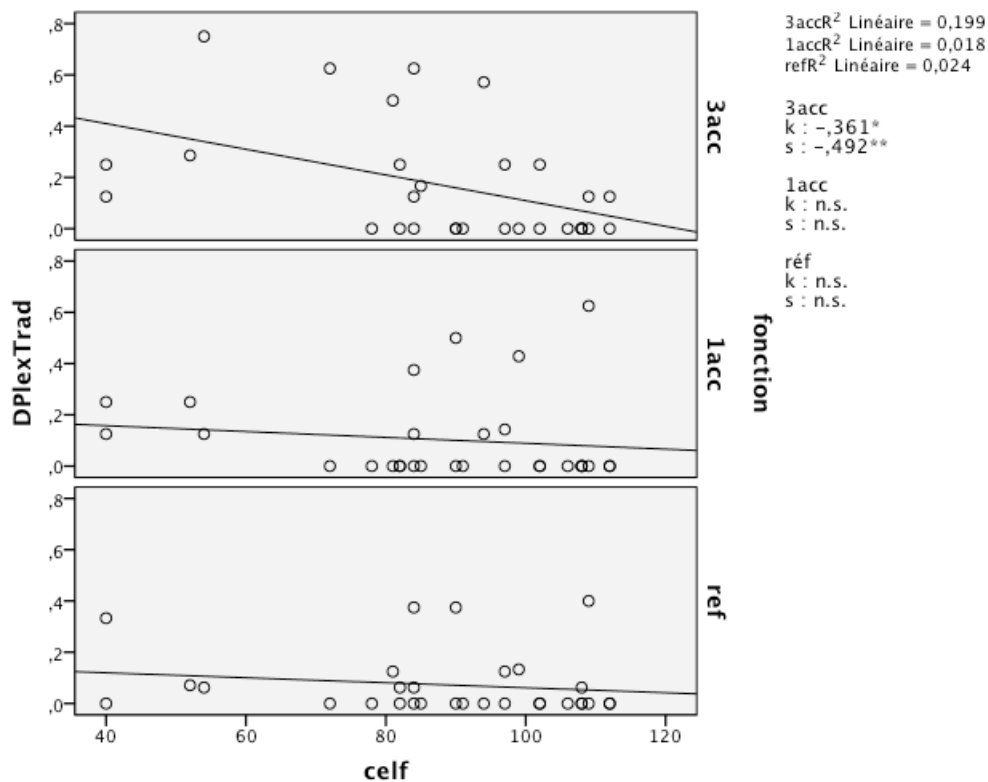


Figure 48. DP trad et celf

3.3.5.2. Mesures sur les réponses requérant un regroupement des clitiques

Pour la mesure de la production dans les contextes requérant le regroupement des clitiques (**O prod in clust**), on observe une corrélation entre les trois catégories et la durée d'exposition. Les corrélations sont identiques à celles observées pour la mesure de la production dans l'échantillon entier (**O prod trad.** ; comparer Figure 42 et Figure 49).

Les clitiques réf sont, en outre, corrélés à l'âge. Les clitiques 1acc sont corrélés à l'âge pour **O prod trad.**, mais ils ne le sont pas pour **O prod in clust** (comparer Figure 43 et Figure 50).

Enfin, l'effet du niveau en L1 est identique à celui observé pour **O prod trad.** : les 3acc sont corrélés à ce niveau tandis que les 1acc/réf ne le sont pas (comparer Figure 44 et Figure 51).

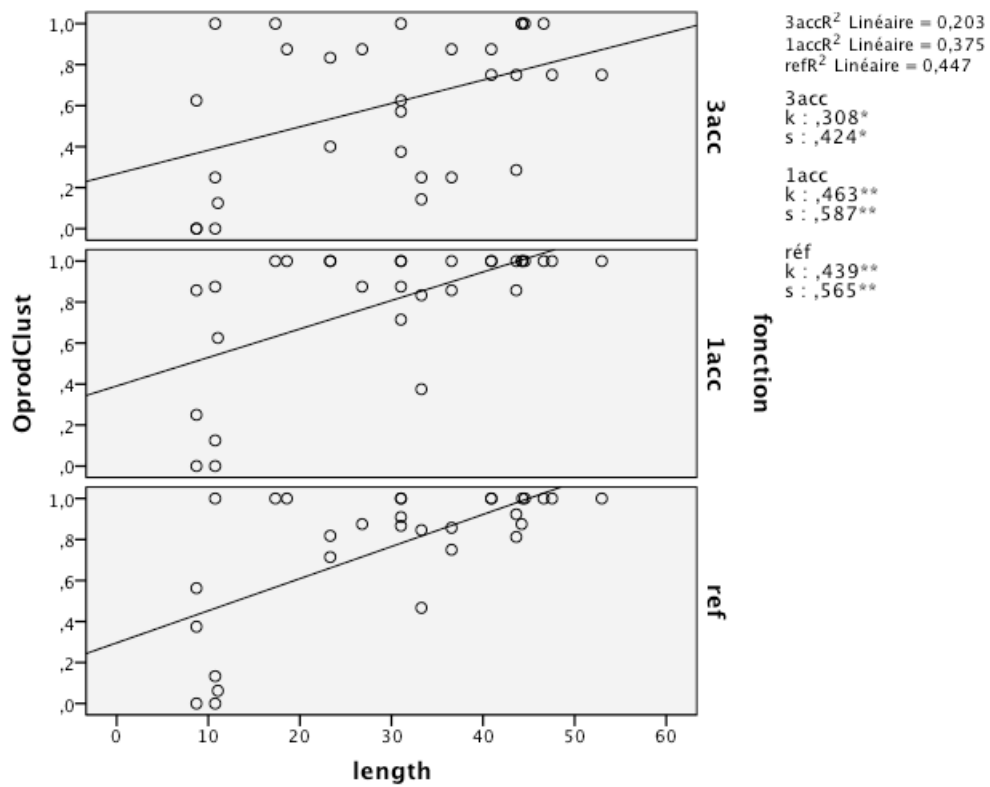


Figure 49. O prod in clust et durée

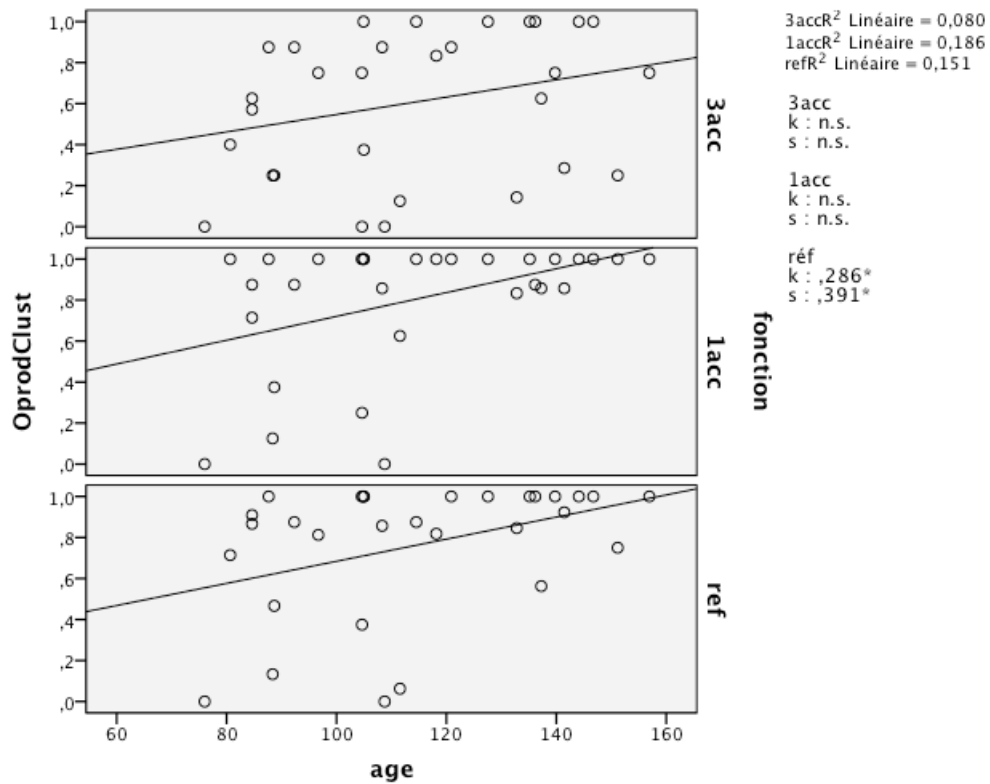


Figure 50. O prod in clust et âge

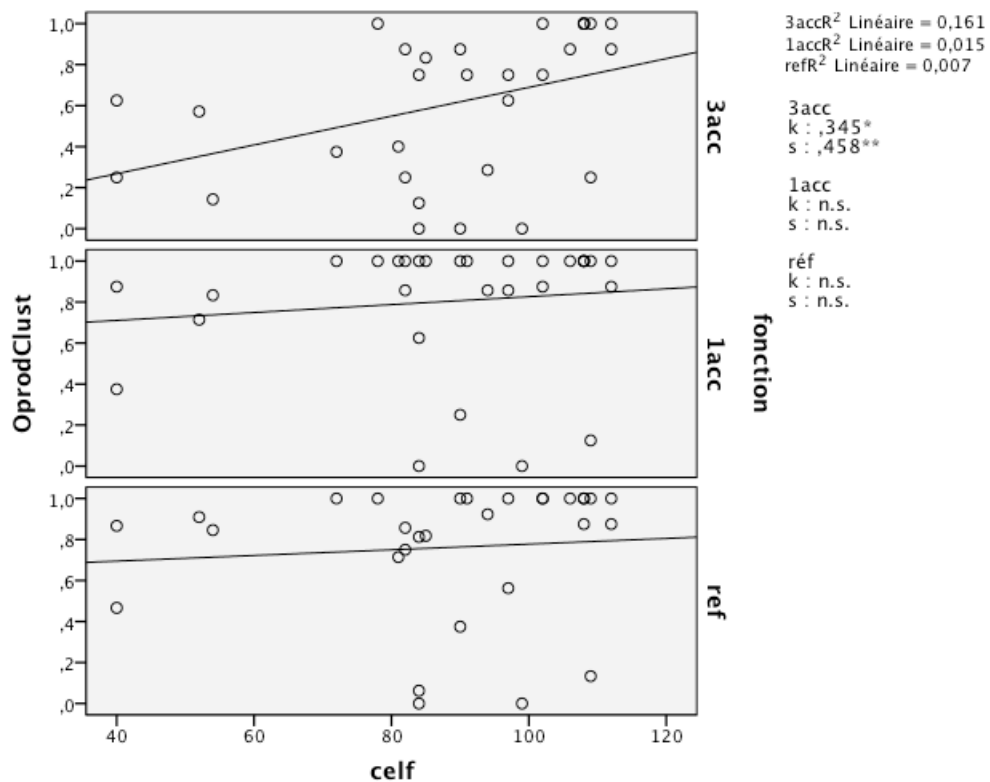


Figure 51. O prod in clust et celf

Pour la mesure de l'omission (**O omission in clust**), on observe une corrélation entre les clitiques 1acc/réf et la durée d'exposition. Ce schéma est identique à celui observé pour **O omission trad.** (comparer Figure 45 et Figure 52). On retiendra en particulier que l'omission des 3acc ne diminue pas significativement avec la durée d'exposition, même si la ligne d'ajustement (Figure 52) le suggère.

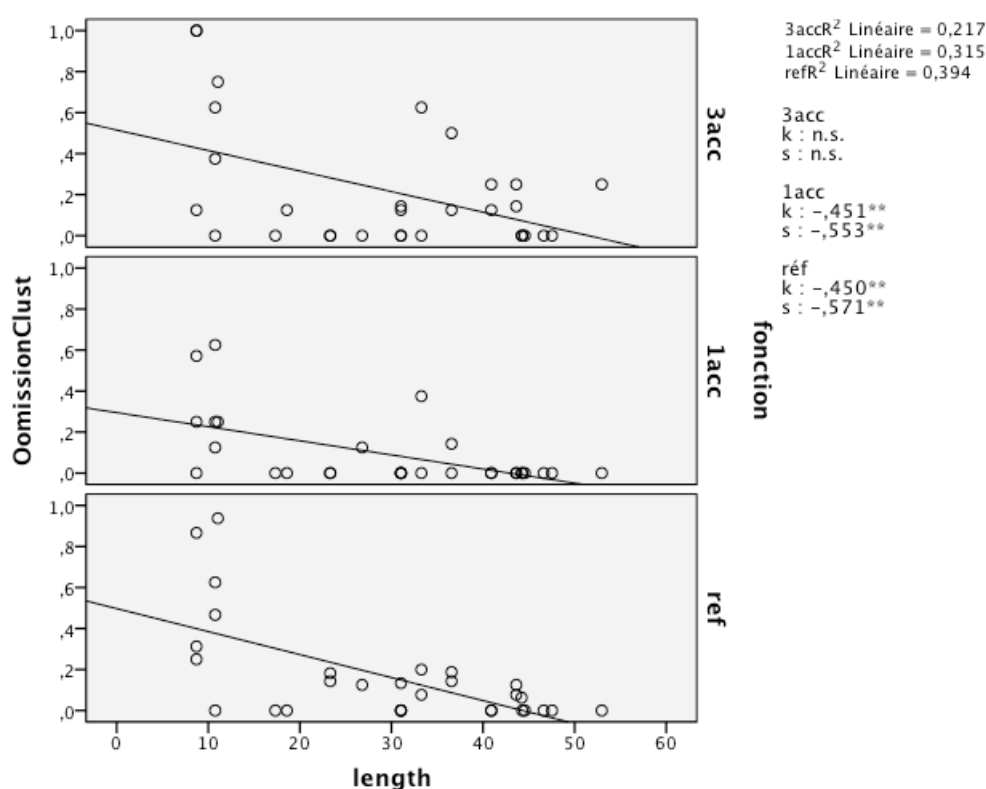


Figure 52. O omission in clust et durée

Pour la substitution du clitique par un élément tonique (**DP in clust**) on remarque, là aussi, l'absence de corrélation entre 3acc et la durée d'exposition ; en revanche, les 1acc/réf sont corrélés à la durée d'exposition. Ce schéma est identique à celui observé pour **DP trad.** (comparer Figure 46 et Figure 53).

Le test de *Spearman* repère en outre une corrélation entre les clitiques réf et l'âge (corrélation qui n'est pas significative pour *Kendall*). A nouveau, ce schéma est identique à celui observé pour **DP trad.** (comparer Figure 47 et Figure 54).

Enfin, les 3acc sont corrélés au niveau en L1, tout comme ils l'étaient pour **DP trad.** (comparer Figure 48 et Figure 55).

On retiendra enfin que la substitution des 3acc ne dépend ni de la durée d'exposition, ni de l'âge ; en revanche, elle est (inversement) corrélée au niveau en L1. Ce fait est intéressant : il indique en effet que la substitution des 3acc est un candidat privilégié au statut de marqueur de troubles d'acquisition lors de l'acquisition du français en tant que L2.

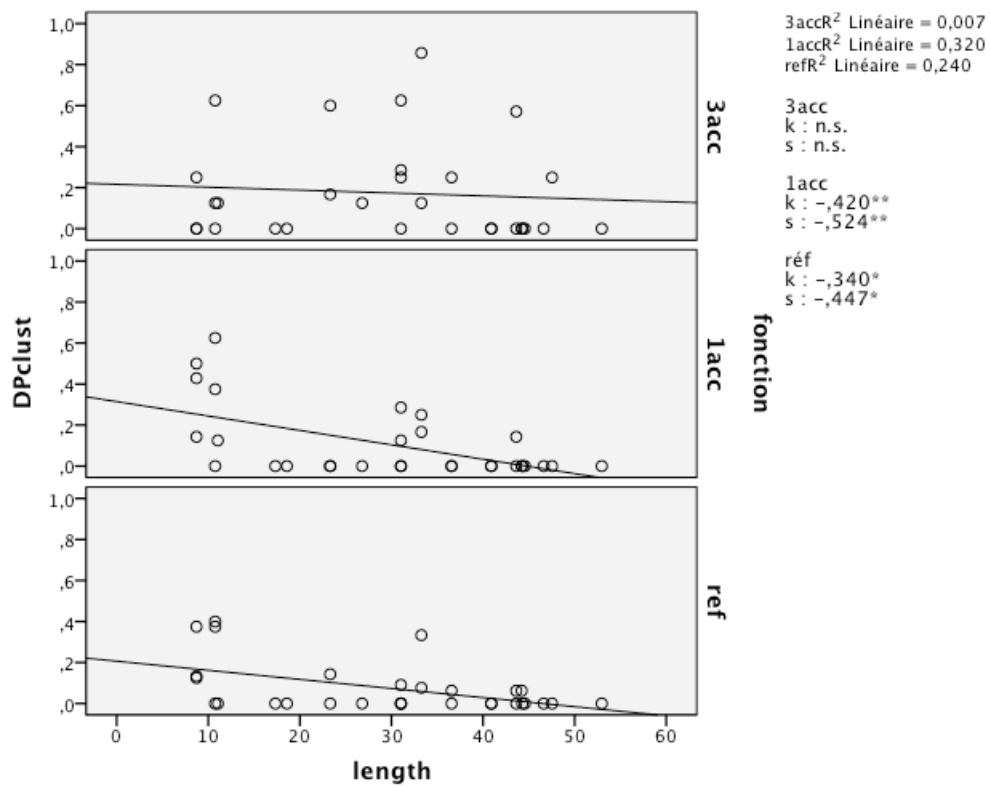


Figure 53. DP in clust et durée

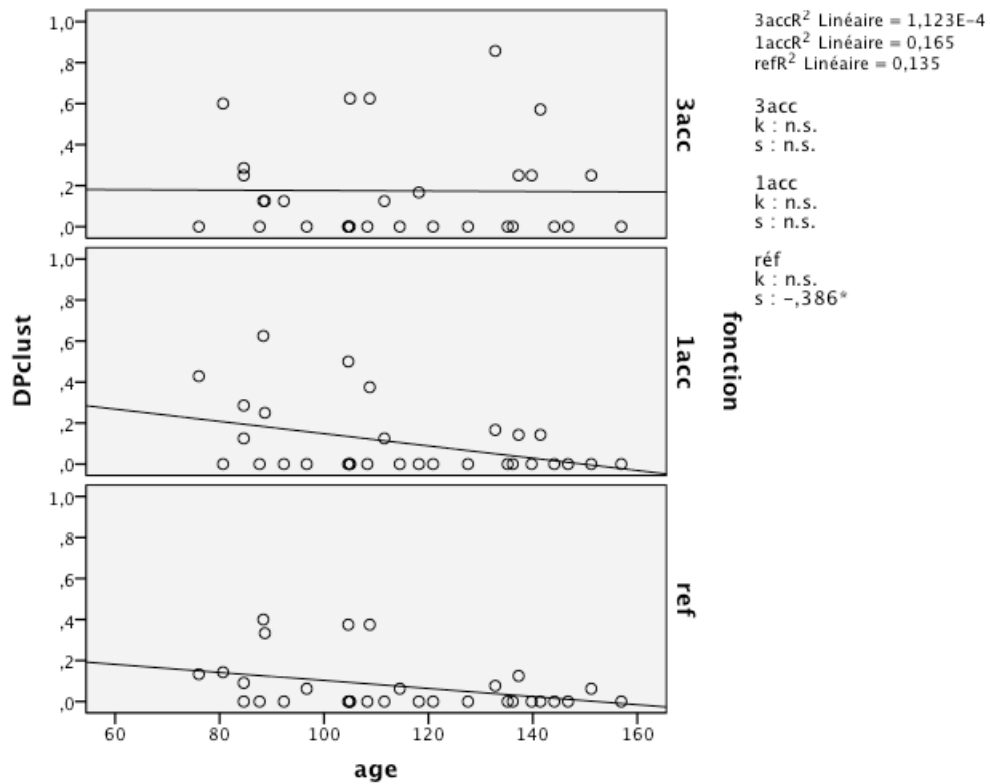


Figure 54. DP in clust et âge

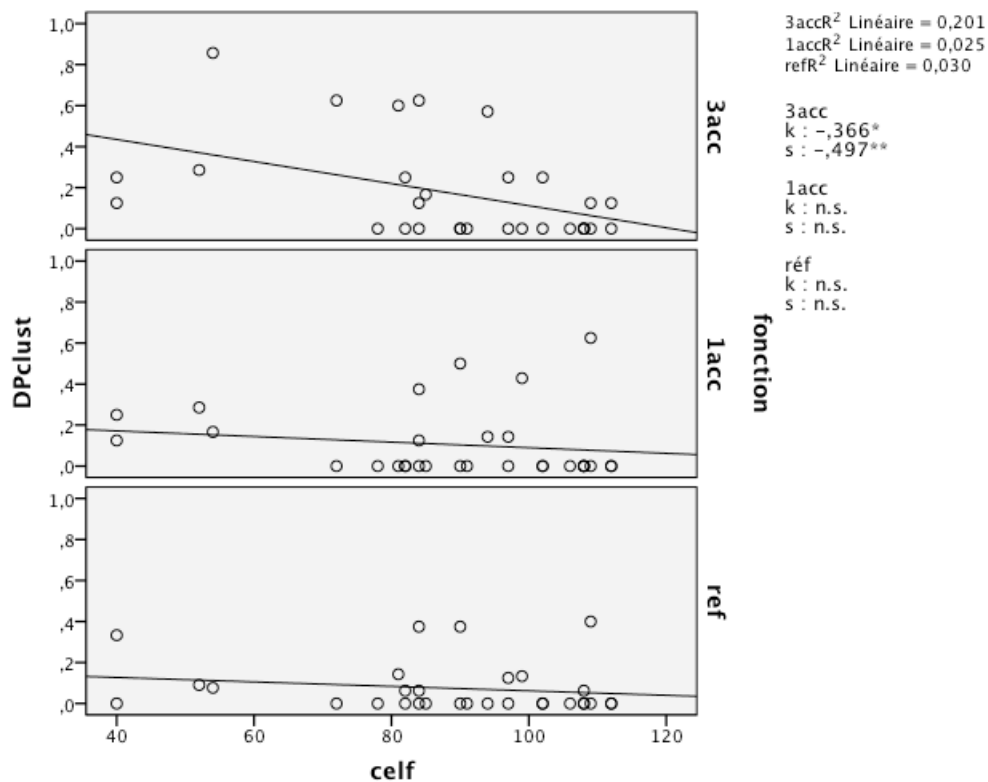


Figure 55. DP in clust et celf

O ok+dat in clust met au jour les mêmes corrélations que **O prod trad.** et **O prod in clust** :

- le clitique 3acc est corrélé à la durée d'exposition (comparer Figure 42, Figure 49 et Figure 56) et au celf/niveau en L1 (comparer Figure 44, Figure 51 et Figure 59)
- les clitiques 1acc/réf sont corrélés à la durée d'exposition (comparer Figure 42, Figure 49 et Figure 56) et à l'âge (comparer Figure 43, Figure 50 et Figure 58)

L'exclusion des datifs grammaticaux (**O ok in clust**, Figure 57) a pour effet

- de rendre la corrélation entre 1acc et âge non significative pour le test de *Kendall*.
- de rendre les autres corrélations moins nettes

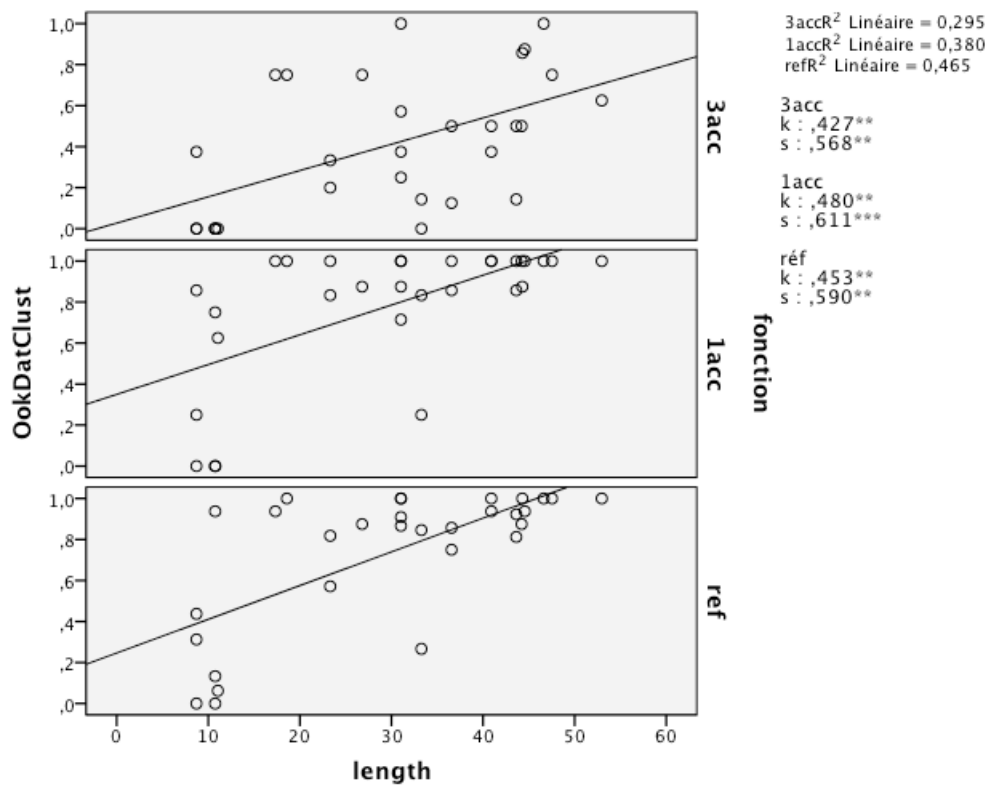


Figure 56. O ok+dat in clust et durée

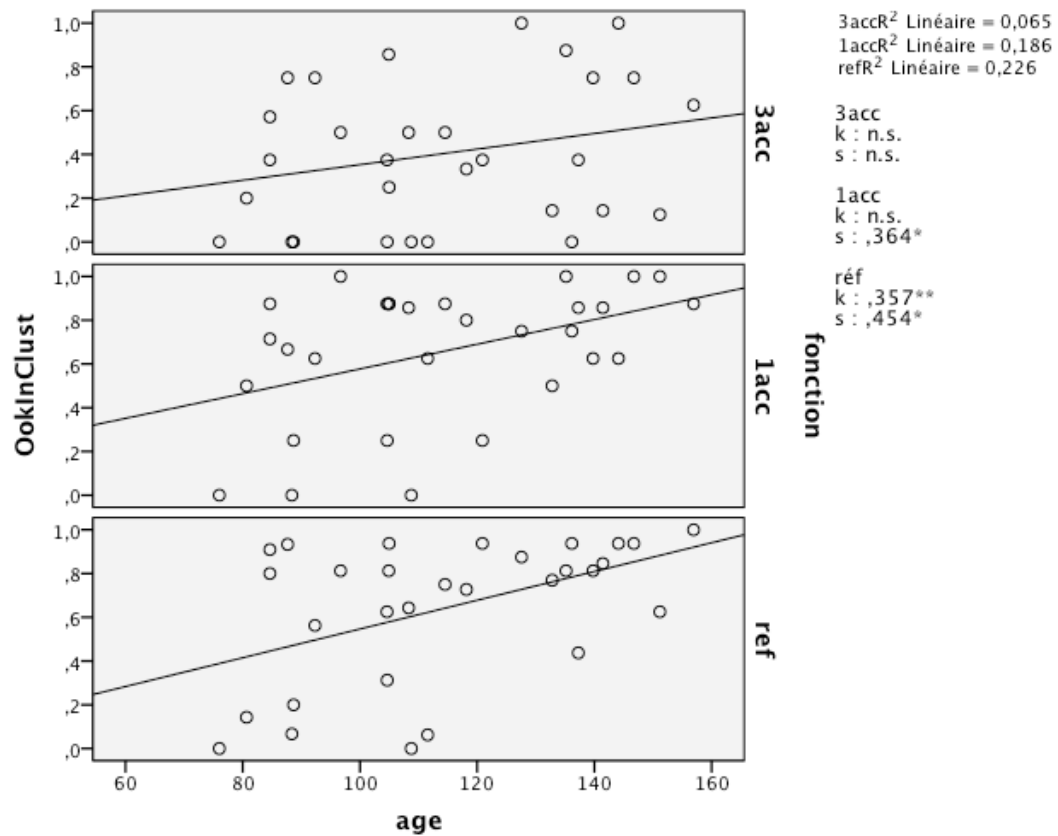


Figure 57. O ok in clust et âge

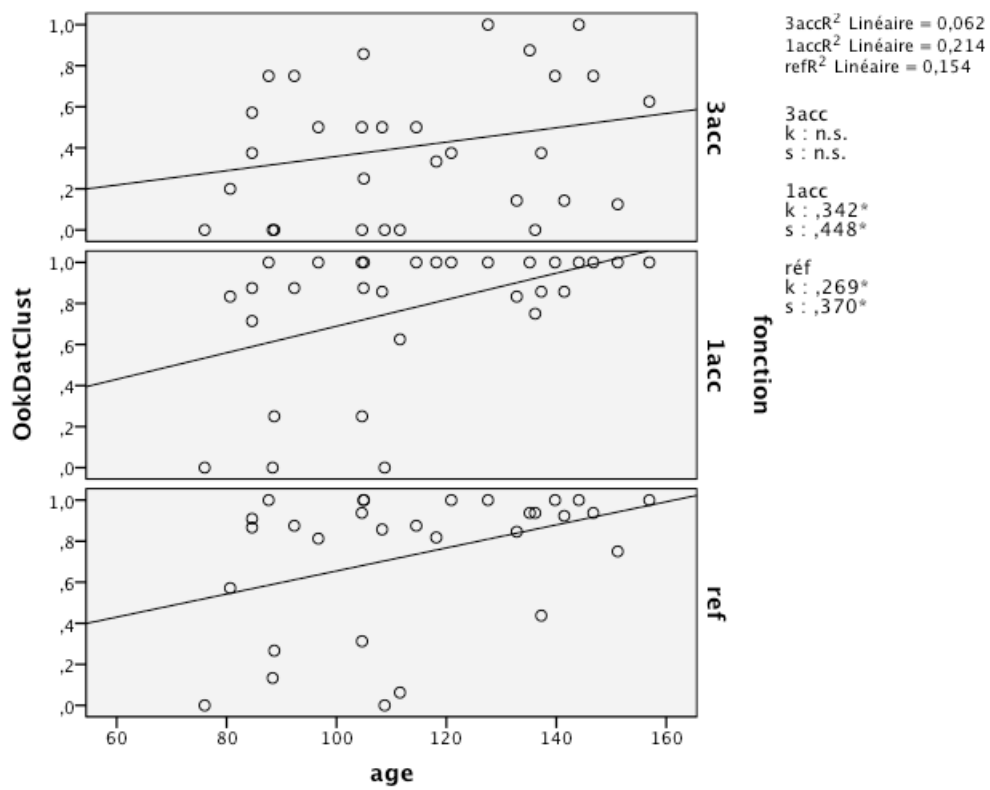


Figure 58. O ok+dat in clust et âge

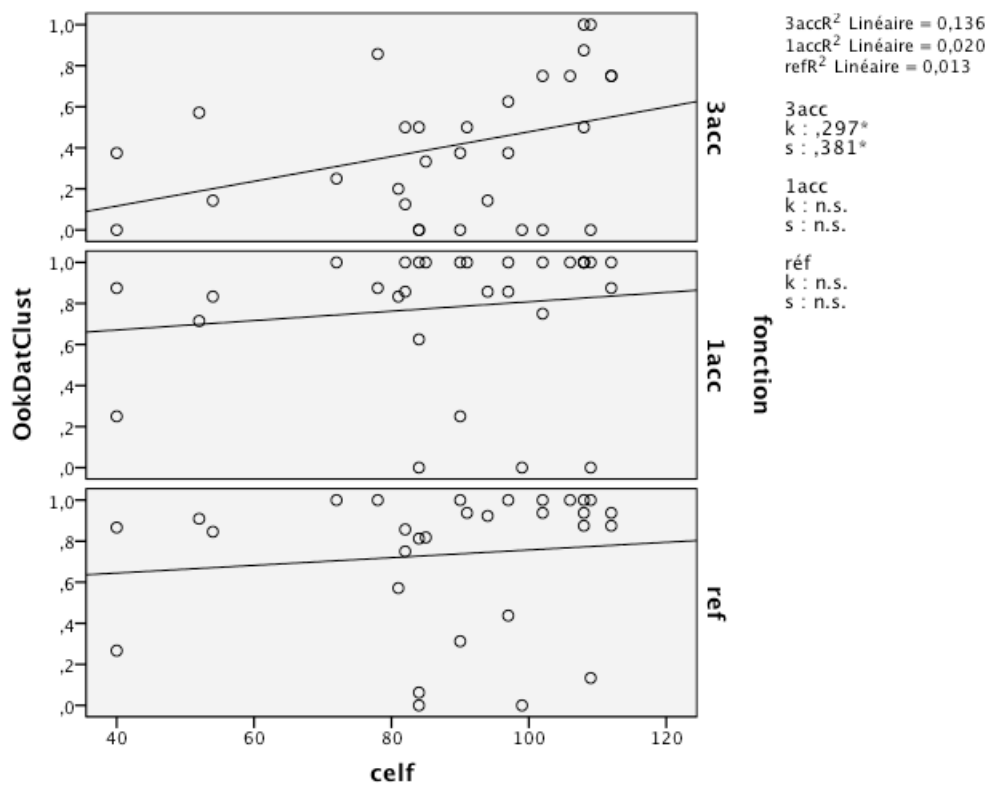


Figure 59. O ok+dat et celf

Les mesures des groupes de clitiques cible (**Target SO in clust**, Figure 60, Figure 62 ; **Target SO+dat in clust**, Figure 61, Figure 63) produisent les effets suivants :

- Les 3 catégories (3acc, 1acc, réf) sont corrélées à la durée d'exposition, aussi bien dans la mesure sans datifs que dans celle avec datifs (Figure 60, Figure 61).
- Les clitiques réf sont corrélés à l'âge, les 3acc ne le sont pas. Les 1acc le sont parfois (Figure 62, Figure 63).

Le comportement des clitiques 1acc présente par ailleurs d'autres particularités : pour **Target SO in clust** (Figure 62) les 1acc sont corrélés à l'âge, et cela pour les deux tests appliqués. Si l'on inclut les datifs grammaticaux dans le décompte (**Target SO+dat in clust**, Figure 63), la corrélation disparaît, et cela pour les deux tests. Dans les mesures de production (**O prod trad.**, Figure 43, **O prod in clust**, Figure 50), les 1acc sont corrélés à l'âge. Pour la mesure de la production cible, l'exclusion des datifs avait eu l'effet inverse : la corrélation entre 1acc et âge est significative pour les deux tests dans **O ok+dat in clust** (Figure 58) ; elle ne l'est que pour *Spearman* dans **O ok in clust** (Figure 57). Cette observation doit être mise en rapport avec le comportement des TD11 dans la catégorie 1acc. Rappelons (section 3.3.4) que l'exclusion des datifs de la mesure de la production cible (**O ok in clust**) fait baisser le taux de réussite des TD11. L'inclusion des datifs dans cette mesure, rend les TD11 « normaux ». Pour la mesure des clusters grammaticaux (**Target SO in clust**), l'inclusion des datifs n'a pas cet effet. Maintenant, on observe que, chez les L2, l'inclusion des datifs dans la mesure des groupes fait disparaître la corrélation entre 1acc et l'âge. Ces observations indiquent, à mon sens, que les L2 les plus âgés ont le même problème avec la catégorie 1acc que celui que nous venons de citer. Puisque l'inclusion des datifs augmente le taux des L2 jeunes et non celui des L2 âgés, la corrélation avec l'âge disparaît.

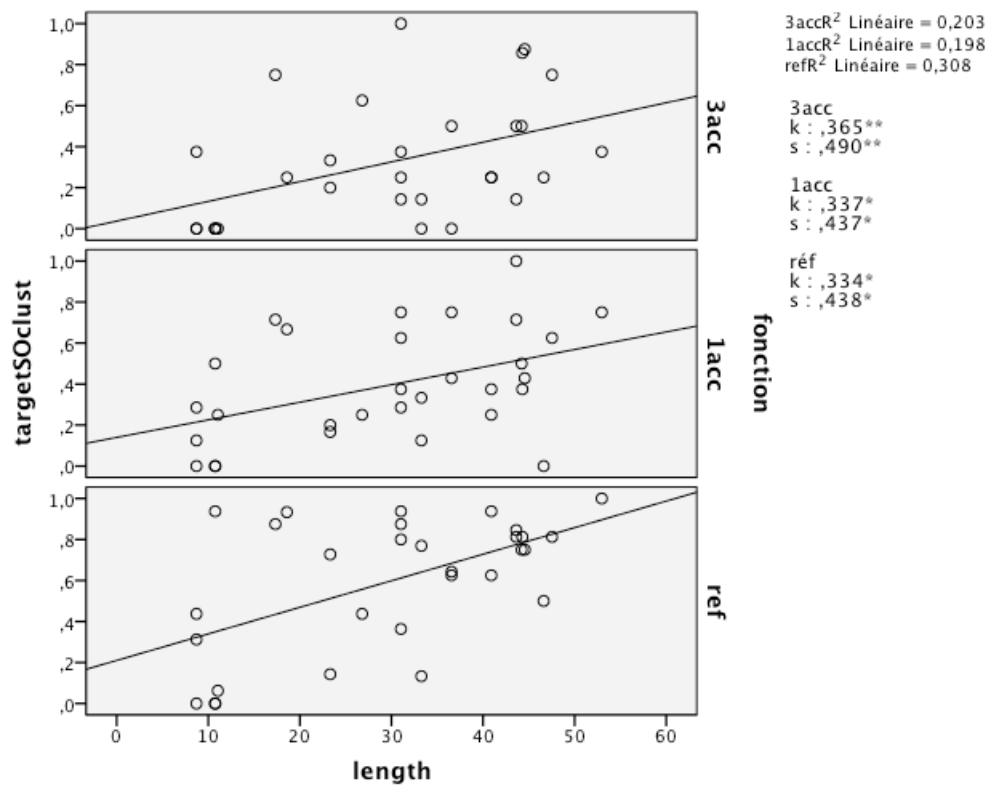


Figure 60. Target SO in clust et durée

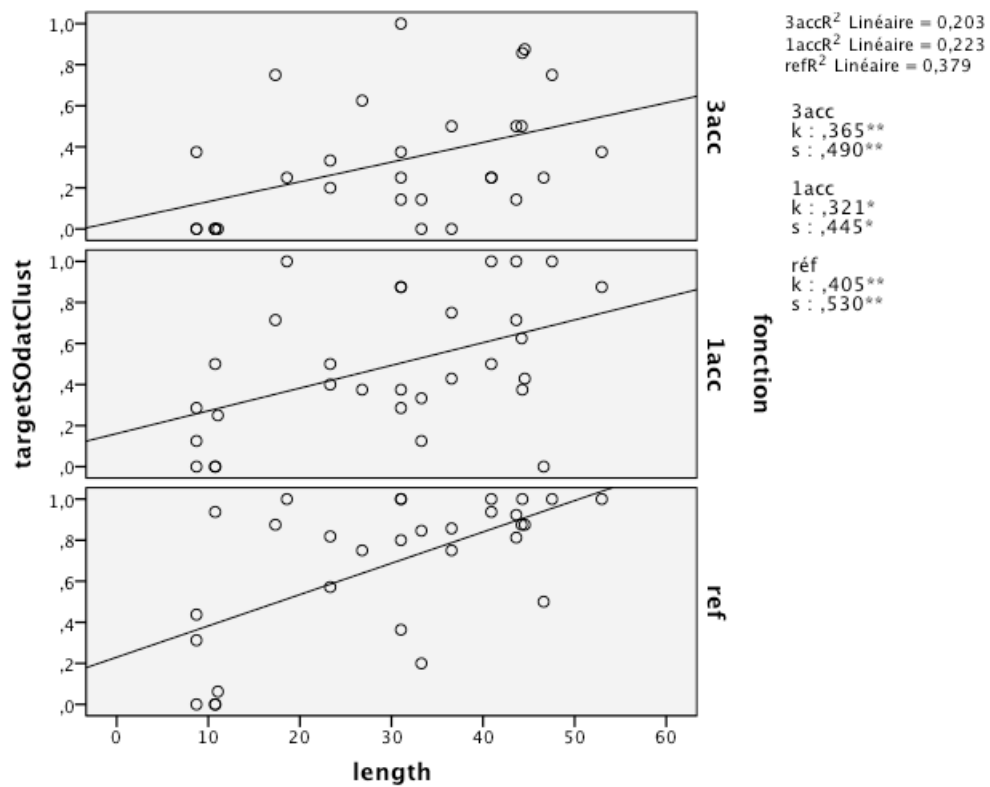


Figure 61. Target SO+dat in clust et durée

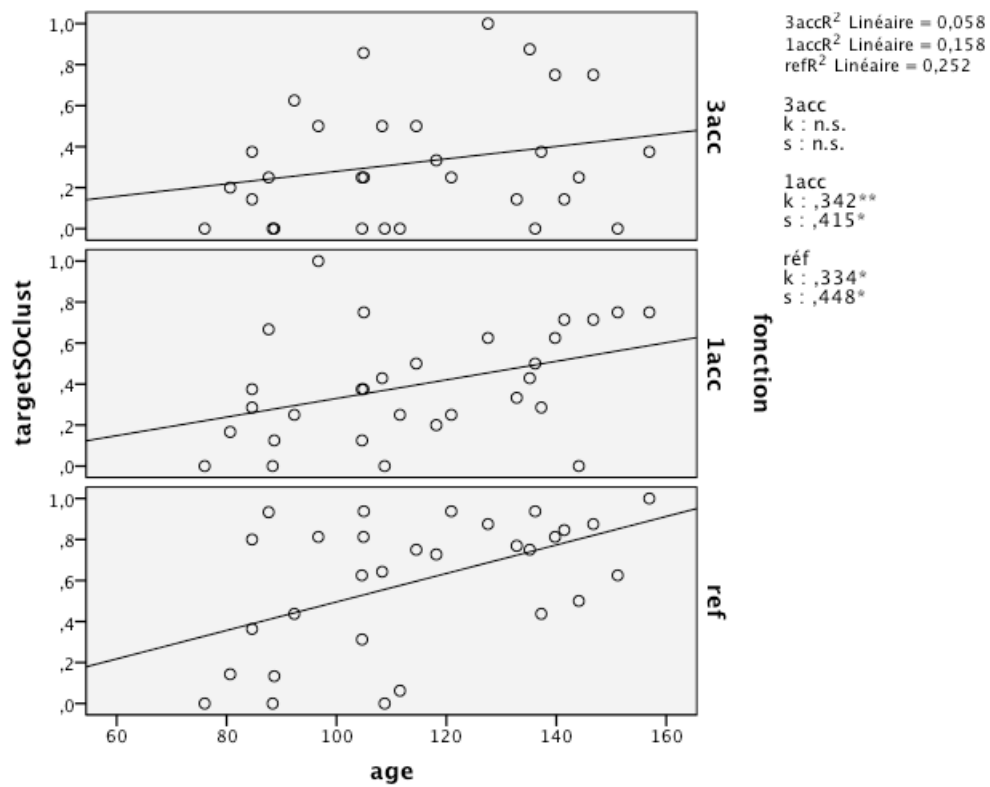


Figure 62. Target SO in clust et âge

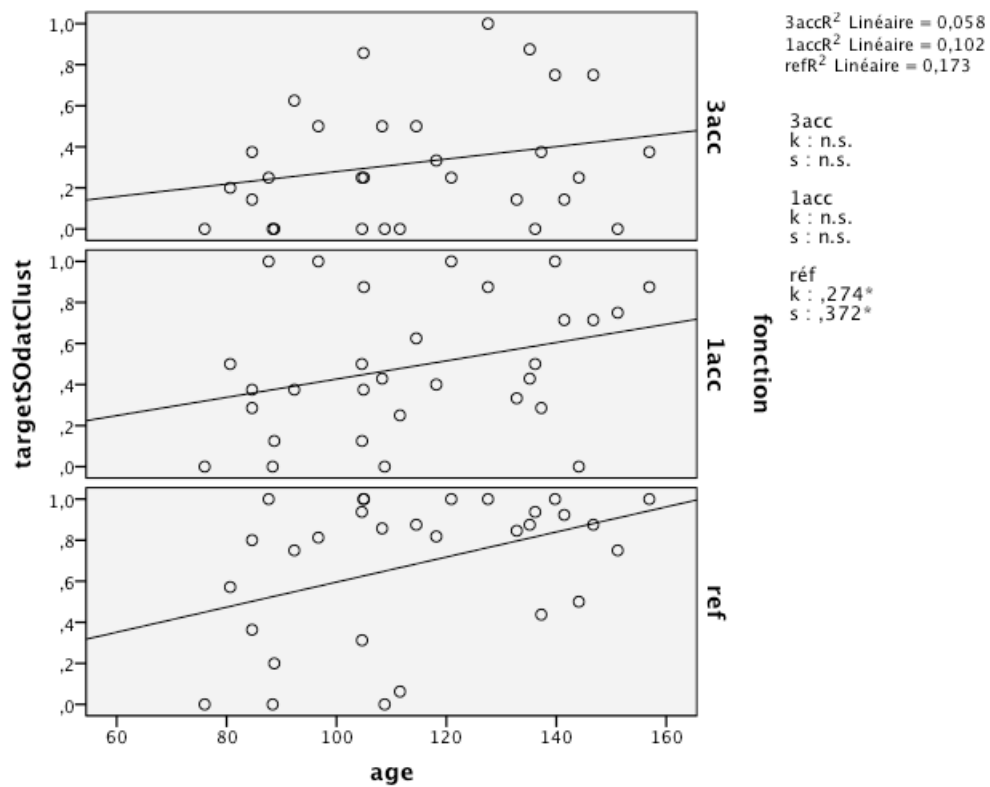


Figure 63. Target SO+dat in clust et âge

Notons enfin un fait remarquable : la disparition de la corrélation entre 3acc et le niveau en L1. Cette corrélation est caractéristique de toutes les mesures de production (**O prod trad.**, Figure 44, **O prod in clust**, Figure 51) et de production cible (**O ok+dat in clust**, Figure 59). Pour **Target SO/+dat in clust**, il n'existe aucune corrélation avec le niveau en L1.

Ce dernier point suggère que la grammaticalité des formes morphologiques ne nous renseigne pas sur le niveau de langue des apprenants L2 du français. Cette conclusion doit être mise en relation avec les conclusions tirées précédemment sur le développement de la morphologie des clitiques chez les TD monolingues. J'ai argumenté en particulier (sections 3.3.2, 3.3.4) que le choix de la forme morphologique correcte constitue un facteur indépendant du développement des clitiques. Ce facteur s'applique de manière uniforme à toutes les catégories, il ralentit le développement et il peut masquer les spécificités d'une population. En particulier, le retard des L2 par rapport aux TD dans la catégorie des réfléchis est moins net ($L2 < TD6$) si l'on mesure la grammaticalité des groupes que si l'on mesure la production du clitique objet ($L2 < TD4$). Ces observations sont corroborées par la structure interne de la population L2.

3.3.5.3. Interprétation des corrélations et identification des individus présentant un risque de trouble de développement

Le Tableau 5 résume les corrélations relevées dans cette section.

	O prod trad	O prod in clust	O omission trad	O omission in clust	DP trad	DP in clust	O ok in clust	O ok+dat in clust	Target so in clust	Target so+dat clust
3acc	durée celf	durée celf			celf	celf	durée celf	durée celf	durée	durée
1acc	durée âge	durée	durée	durée	durée	durée	durée âge (s)	durée âge	durée âge	durée
réf	durée âge	durée âge	durée	durée	durée âge (s)	durée âge (s)	durée age	durée âge	durée âge	durée âge

Tableau 5. L2 - corrélations

Dans cette section, nous avons mis à jour deux généralisations importantes.

1. Il existe une asymétrie entre les clitiques 3acc d'une part, et les clitiques 1acc/réf d'autre part, concernant les corrélations avec le niveau en langue maternelle (variable *celf*). Seule la catégorie 3acc présente de telles corrélations. Les catégories 1acc et réf ne les présentent jamais. Cette observation confirme que les clitiques 3acc sont des marqueurs de troubles (cf. (Jakubowicz et Tuller, 2008)).

2. Dans les mesures d'omission et de substitution, il n'existe pas de corrélation avec les variables de temps (*durée d'exposition* et *âge*). Ce fait est particulièrement intéressant dans la mesure de substitution, car cette mesure est bien corrélée au niveau en langue maternelle. Rappelons également que la mesure de substitution est très stable dans le développement L1 (voir l'introduction à la section 3.3.3). J'ai imputé cette stabilité au fait que cette mesure est indépendante du facteur morphologique (section 3.3.2) : la production d'un DP plein est une stratégie d'évitement du clitique. Si ce raisonnement est correct, alors la mesure de substitution nous permet de percevoir les difficultés qu'un individu rencontre avec la catégorie 3acc, et cela indépendamment du facteur morphologique.

Pour conclure, je voudrais suggérer deux types de marqueur de troubles. Le premier type serait un marqueur absolu, c'est-à-dire un marqueur indépendant des variables de temps (âge et durée d'exposition), mais dépendant de la compétence sous-jacente. Ce type de marqueur serait évidemment bienvenu pour l'identification de troubles dans le contexte d'acquisition L2. Je propose que la substitution des clitiques 3acc est un tel marqueur. Ensuite, on pourrait identifier un second type de marqueur de trouble, un marqueur relatif, qui dépend de la compétence sous-jacente, mais également d'une ou plusieurs variables de temps. Un marqueur relatif devrait être ajusté en fonction de la/des variables de temps afin de donner une indication sur la compétence sous-jacente. Je propose que les mesures de production et de production cible des clitiques 3acc sont des marqueurs relatifs de trouble.

Pour évaluer cette proposition, il faudrait comparer deux populations appariées en âge, âge du début d'exposition et durée d'exposition : l'une L2 saine et l'autre à pathologie avérée. Cette tâche est évidemment hors de portée dans le cadre de ce chapitre, mais des projets comparables sont poursuivis dans le cadre de l'action COST IS0804 "Language Impairment in a multilingual society. Linguistic Patterns and the Road to assessment." (porteur : S. Armon-Lotem, Bar-Ilan Univ.) Une tâche moins ambitieuse, mais néanmoins impossible dans le cadre de la présente étude, consisterait à évaluer les tests effectués ici sur la population L2 un an après les passations discutées dans ce chapitre. Cette tâche sera effectuée dans le cadre du projet ANR CCAAL BLAN 08-3-311192 "Complexité et compensation en acquisition atypique du langage" (porteur : L. Tuller, Univ. Tours).

A ce jour, je peux cependant indiquer la façon dont les marqueurs de trouble que je propose identifient les individus hors norme. Considérons tout d'abord le marqueur absolu, la substitution. La Figure 64 représente le niveau en langue maternelle sur l'axe des abscisses, et les taux pour **DP trad.** et **DP in clust** sur l'axe des ordonnées. La ligne de référence à 50% isole un groupe de 4 individus produisant des taux de substitution bien au-dessus de la norme. Il s'agit des points de données 60, 61, 59 et 37.

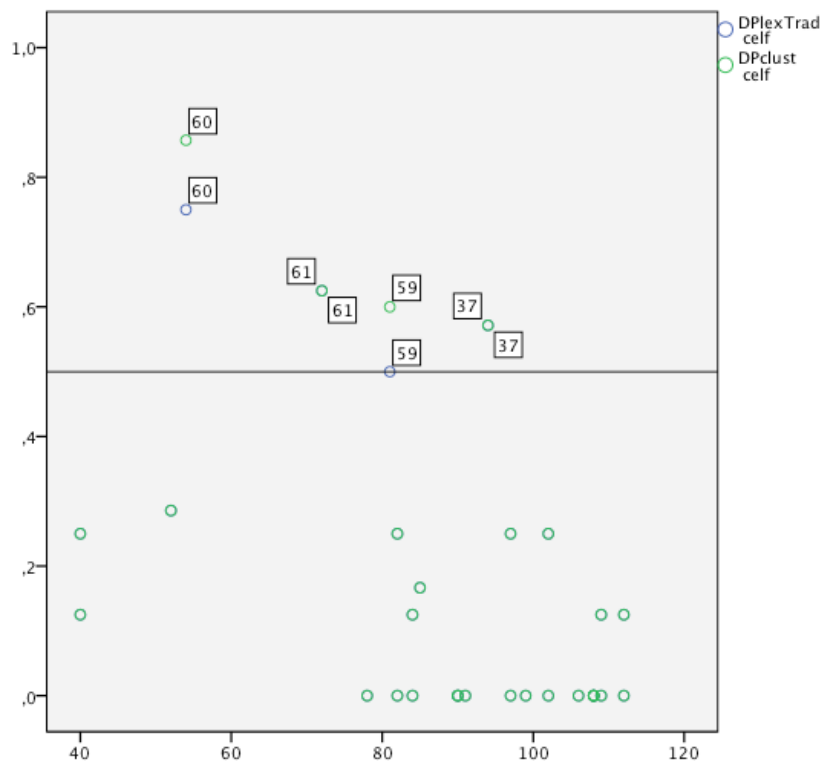


Figure 64. La substitution en tant que marqueur de trouble

Considérons maintenant les marqueurs relatifs, i.e. les mesures de production et de production cible. Afin d'identifier les individus hors norme, j'ai établi des graphiques représentant la durée d'exposition sur l'axe des abscisses, et la mesure en question sur l'axe des ordonnées. J'ai ensuite tracé une ligne d'ajustement avec un intervalle de confiance de 85%. Les points de données se situant au dessous de cet intervalle seront considérés hors norme. Pour **O prod trad.**, le point de données 37 est (de justesse) en dehors de l'intervalle de confiance, 34, 42 et 60 sont (de justesse) dedans. Pour **O prod in clust**, 37 rentre dans l'intervalle de confiance, mais 60 en sort. Rappelons que 37 et 60 s'étaient déjà signalés par leur taux élevé en substitution (Figure 64).

Je conclus que les marqueurs proposés dans cette section isolent deux individus comme présentant un fort risque de trouble : il s'agit des points de données 37 et 60.

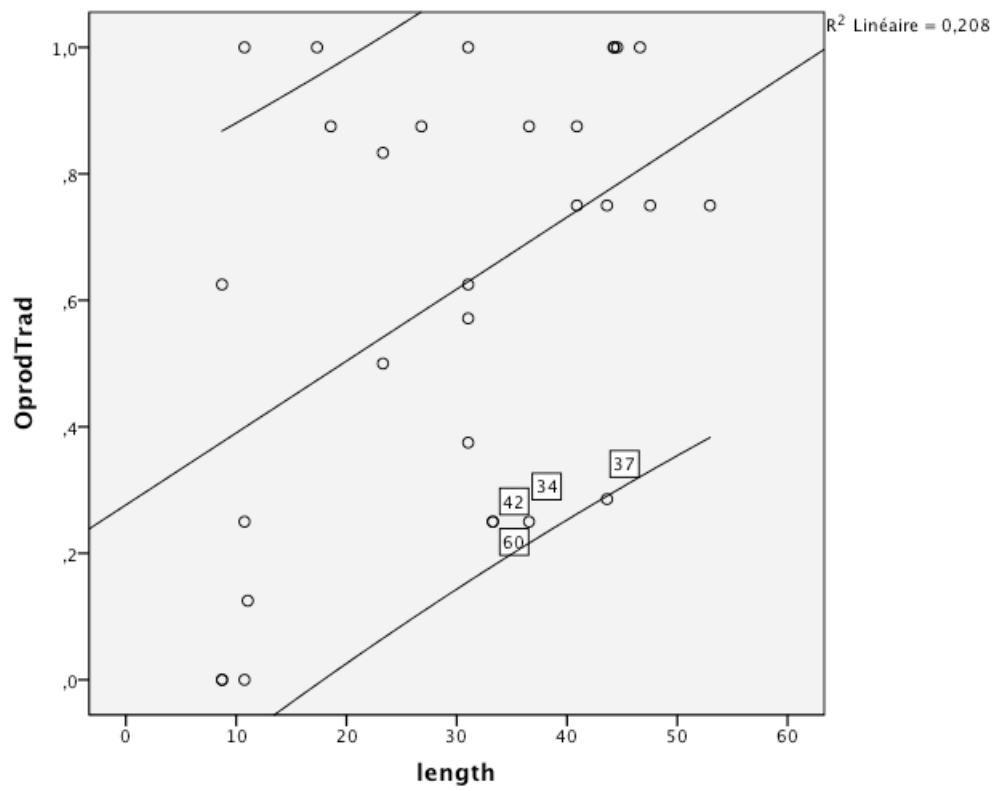


Figure 65. O prod trad. en tant que marqueur de trouble

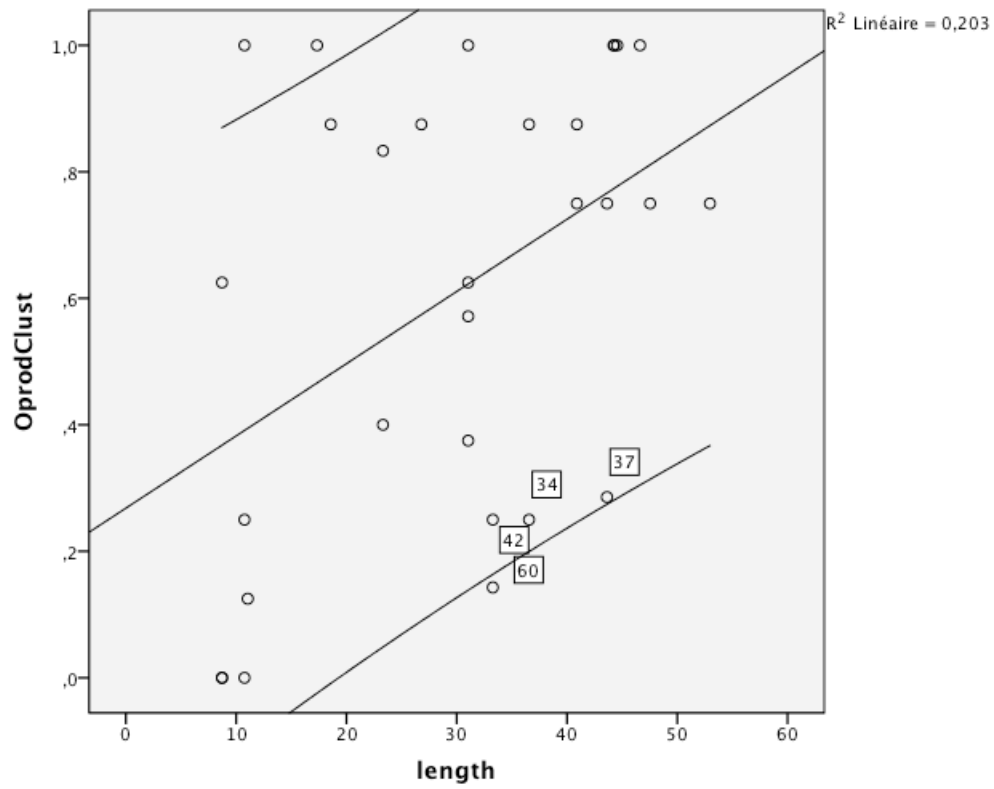


Figure 66. O prod in clust. en tant que marqueur de trouble

Pour la mesure de la production cible (Figure 67, Figure 68), 37 reste en dehors de l'intervalle de confiance et 42 en sort, mais 60 y entre. Notons que 60 et 37 avaient été isolés comme présentant un risque sur la base de leurs taux en substitution, mais 42 ne l'avait pas été. Si l'on admet, comme nous l'avons fait précédemment dans cette étude, que la mesure de la grammaticalité morphologique peut masquer la maîtrise d'une configuration sous-jacente, nous devrions considérer que 60 présente plus de risques que 42, même si 60 est bien à l'intérieur de l'intervalle de confiance pour la mesure de la production cible des clitiqes 3acc.

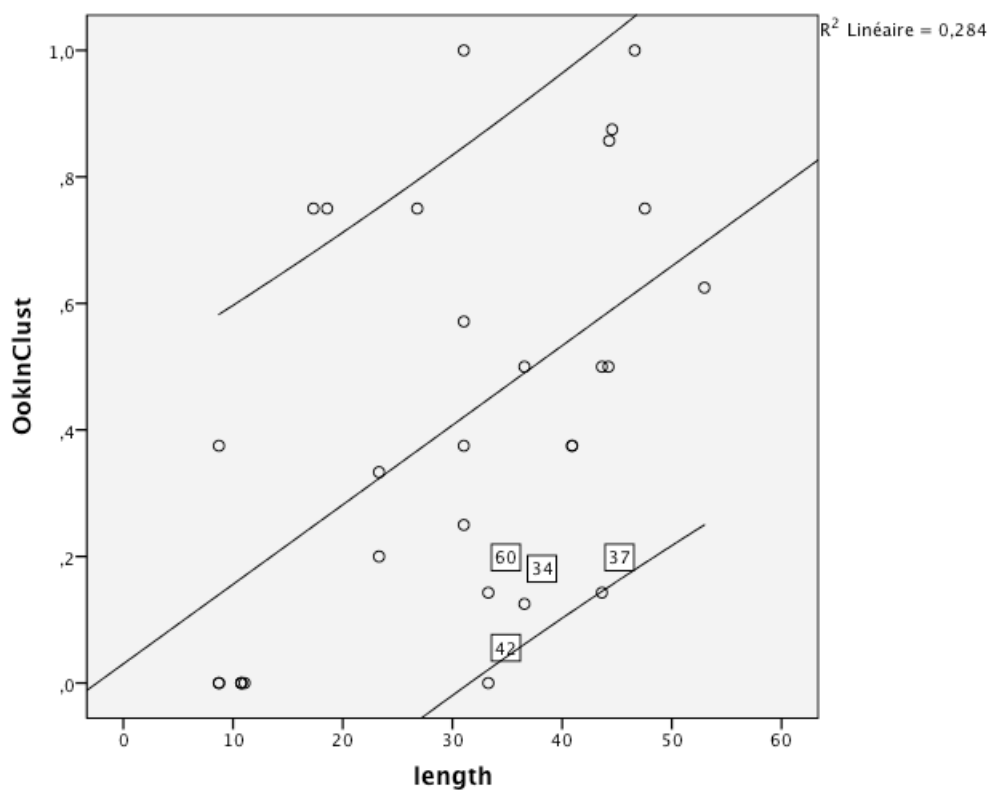


Figure 67. O ok on clust. en tant que marqueur de trouble

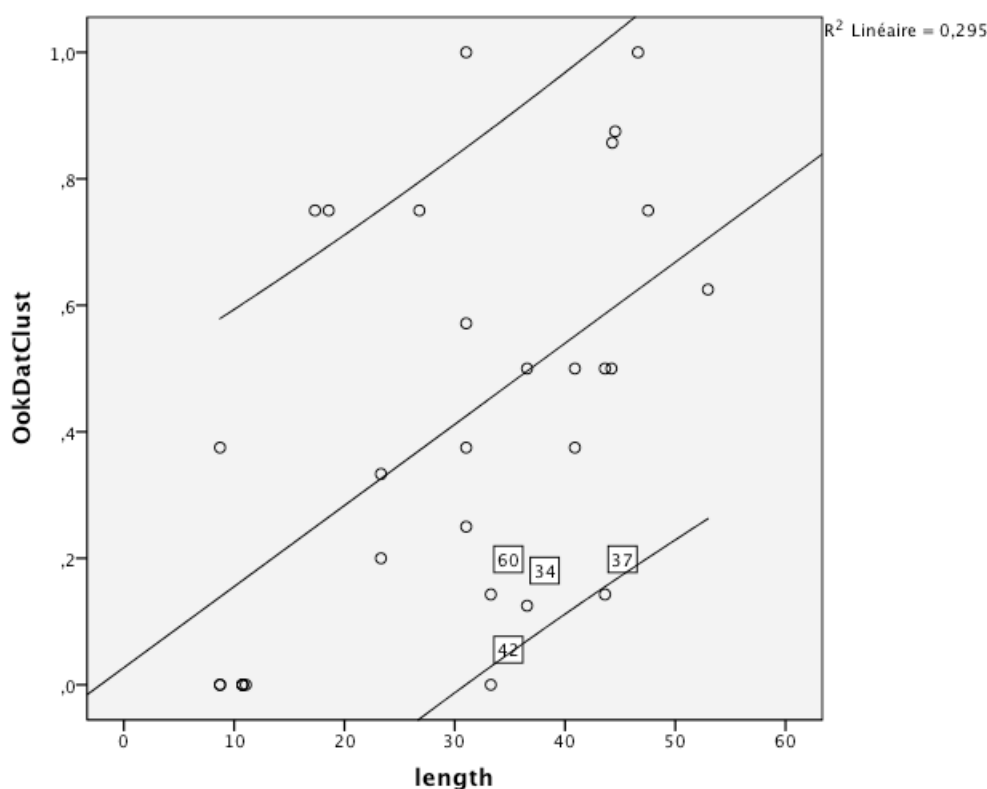


Figure 68. O ok+dat. en tant que marqueur de trouble

3.4 Conclusion

Ce dernier chapitre a été consacré à l'acquisition des pronoms clitiques du français. Mon constat de base était que les études sur l'acquisition des clitiques objets donnent des taux de production très variés pour des populations cependant comparables. Afin d'isoler les facteurs pouvant expliquer ces différences, j'ai proposé un nouvel échantillonnage des données et de nouvelles mesures, et j'ai étudié le développement des oppositions, plutôt que des moyennes. Ainsi, mes résultats peuvent être comparés à ceux obtenus dans d'autres études, et ce, même si les moyennes ne sont pas nécessairement comparables.

Je résume maintenant brièvement les résultats principaux de mon travail, sous la forme de huit points.

1. J'ai isolé trois étapes dans le développement des clitiques objets, définies par l'opposition entre les fonctions 3acc, 1acc et réf (pour le constat d'une opposition entre ces trois fonctions, voir (Tuller et al., 2011)) :

- a. la première étape est caractérisée par l'opposition entre acc et réf (3acc et 1acc présentent des taux de réussite inférieurs à réf).
 - b. la deuxième étape est caractérisée par l'opposition entre 3acc et réf (au cours de cette étape, 1acc ne se distingue ni de 3acc, ni de réf).
 - c. la dernière étape est caractérisée par l'absence d'opposition entre les fonctions.
- 2. Les trois étapes précédentes (a, b, c) sont observées à des âges TD différents selon la mesure (production, omission, production cible etc). Il en suit qu'il n'existe pas de développement unique "du clitique". Le développement des clitiques objets est affecté par plusieurs facteurs se manifestant de manière différente selon la mesure et la catégorie étudiée (i.e. 3acc, 1acc et réf).
- 3. L'épel morphophonologique (entre autres, le regroupement des clitiques) constitue un facteur important dans l'évolution de la production des clitiques. Ce facteur est stable à travers les catégories et les mesures : il conserve l'ordre d'acquisition des clitiques (réf avant 1acc avant 3acc). Il agit de la manière suivante sur les TD (a, b) et sur les L2 (c, d) :
 - a. plus on mesure la conformité morphologique des clitiques exprimés aux clitiques cibles, plus les oppositions entre fonctions ressemblent à celles typiques d'un état précoce du développement.
 - b. plus on fait abstraction de la réalisation morphologique des clitiques, plus les oppositions entre fonctions ressemblent à celles typiques d'un état de développement avancé.
 - c. plus on mesure la conformité morphologique des clitiques exprimés aux clitiques cibles, plus les L2 ressemblent aux TD jeunes.
 - d. plus on fait abstraction de la réalisation morphologique des clitiques, plus les L2 se distinguent des TD jeunes.

De façon générale, le développement de la morphologie est lent, et il s'achève tardivement.

4. J'ai argumenté qu'il existe un lien causal entre les taux en production cible, en omission et en production. Cette proposition repose sur l'hypothèse que les apprenants jeunes ont accès à un clitique vide sous-spécifié. Tant que cet élément est accessible, l'omission ne constitue pas une violation de grammaticalité et le taux de production reste relativement bas. Puis, au fur et à mesure qu'ils sont acquis, les clitiques entièrement spécifiés (et phonologiquement réalisés) de la langue adulte remplacent le clitique vide. J'ai argumenté que la forte augmentation de formes cibles (entre 4 et 6 ans pour les 3acc) déclenche l'élimination de la forme vide par le principe *Elsewhere* (Kiparsky, 1973). Ainsi j'ai pu expliquer un fait très surprenant : la production cible des clitiques 3acc augmente fortement **avant** la production (entre 4 et 6 ans, et entre 6 et 8 ans respectivement).
5. Le principe *Elsewhere* intervient dans l'épel morphophonologique, et régit le choix d'une forme cible parmi un ensemble de compétiteurs. Cette hypothèse nous permet de comprendre pourquoi le développement morphologique est lent et s'achève tardivement (cf. point 3. ci-dessus). En effet, selon le modèle développé par (Grodzinsky et Reinhart, 1993, Reinhart, 2006), le traitement d'ensembles de référence (requis par le principe *Elsewhere*) excède les capacités de traitement des enfants, en particulier leur mémoire de travail ("[...] reference-set computation exceeds children's processing ability, specifically the limitations of their working memory" (Reinhart, 2006) : 200). Nous nous attendons alors à ce que l'épel morphophonologique soit affecté par les contraintes de traitement, et donc à ce que son acquisition soit lente et tardive.
6. Les datifs grammaticaux doivent être comptés comme des réponses cibles. S'ils sont exclus d'une variable en effet, on obtient des anomalies allant jusqu'à donner l'impression d'une perte de connaissances vers l'âge de 11 ans.
7. La population L2 se distingue nettement de la population TD en ce qui concerne la catégorie des clitiques réfléchis : les L2 ont des taux de réussite très bas dans cette catégorie. J'ai argumenté que ce phénomène est lié au paramètre régissant la formation des prédicats réfléchis (en syntaxe ou bien dans le lexique, (Reinhart et Siloni, 2005)). L'anglais dériverait les réfléchis dans le lexique, le français les

dériverait en syntaxe. Si les anglophones apprenant le français en tant que L2 transfèrent la dérivation lexicale des prédicats réfléchis au français, alors on s'attend à ce qu'ils produisent peu de pronoms réfléchis, et à ce qu'ils présentent des taux élevés d'omission (i.e. de formation lexicale) et de substitution (i.e. d'évitement de la réflexivisation).

8. Dans la population L2, les corrélations observées entre les mesures des clitiques et les variables de classification (âge, durée d'exposition, niveau en langue maternelle/celf) varient en fonction de la catégorie du clitique. En particulier, le clitique 3acc se distingue des clitiques 1acc et réf par ses corrélations avec le niveau en L1. Ceci confirme que les clitiques 3acc sont des marqueurs de trouble (Jakubowicz et Tuller, 2008). Or, dans le contexte de l'acquisition L2, l'intérêt de ce marqueur est réduit : en effet, la performance d'un sujet à un moment donné dépend également de variables parfois difficiles à contrôler, telles que la durée et la qualité de l'exposition. Il est donc souhaitable de mettre au jour des marqueurs indépendants de la durée d'exposition. J'ai été en mesure d'identifier une telle mesure : la substitution du clitique 3acc par un DP plein (cette mesure n'évolue pas significativement avec le temps d'exposition à la L2 et elle est corrélée au niveau en langue maternelle). J'ai proposé que cette mesure soit un marqueur absolu de troubles dans le contexte de l'acquisition L2. Dans mon étude, cette mesure isole en effet un sous-groupe bien défini du reste de la population, ce qui suggère que ce sous-groupe partage une propriété statistiquement pertinente.

4. Références bibliographiques

- Abney, Steven Paul. 1987. The English Noun Phrase in its Sentential Aspect, MIT.
- Appleyard, David. 1991. The role of tone in some Cushitic languages. *York Papers in Linguistics* 15:5-32.
- Arnould, Antoine, et Lancelot, Claude. 1660. *Grammaire générale et raisonnée*. Paris: Pierre le Petit.
- Aske, Jon. 1989. Path predicates in English and Spanish: a closer look. In *Proceedings of the fifteenth annual meeting of the Berkeley Linguistics Society*, eds. Kira Hall, Michael Meacham et Richard Shapiro, 1-14. Berkeley: Berkeley Linguistics Society.
- Ball, W. A. 1973. The perception of causality in the infant: Developmental Program, Department of Psychology, University of Michigan.
- Baron-Cohen, Simon. 1994. How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/ Current Psychology of Cognition* 13:513-552.
- Baron-Cohen, Simon. 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Bendjaballah, Sabrina. 1999. Trois figures de la structure interne des gabarits, Paris 7: PhD.
- Bendjaballah, Sabrina. 2001. The negative preterite in Kabyle Berber. *Folia Linguistica* 34:185-223.
- Bendjaballah, Sabrina, et Haiden, Martin. 2002. Templatic morphology in the German verbal system. Ms. CNRS-Univ. Lille 3; Aston University, Birmingham.
- Bendjaballah, Sabrina, et Haiden, Martin. 2003. Templatic Architecture. *Recherches Linguistiques de Vincennes* 32:157-168.
- Bendjaballah, Sabrina. 2005. Longueur phonologique des voyelles en kabyle. *Etudes et documents berbères* 22:47-69.
- Bendjaballah, Sabrina, et Haiden, Martin. 2005a. Berber Prepositions, Templates, and the PF Interface. Paper presented at GLOW, Genève.
- Bendjaballah, Sabrina, et Haiden, Martin. 2005b. Präpositionen und Komplementierer in Chemini Berber (Taqbaylit). *Folia Linguistica* 39:319-361.

- Bendjaballah, Sabrina, et Haiden, Martin. 2005c. The Grammar of Prepositions in Berber (Taqbaylit). Ms. STL.
- Bendjaballah, Sabrina, et Haiden, Martin. 2008. A Typology of Emptiness in Templates. In *Sounds of Silence*, eds. Jutta Hartmann, Vera Hegedus et Henk C. van Riemsdijk, 23-59. Amsterdam, Oxford: Elsevier.
- Bendjaballah, Sabrina, et Haiden, Martin. submitted. The derivation of concatenative and non-concatenative exponents.
- Bendjaballah, Sabrina, et Haiden, Martin. to appear. The representational anomalies of floating markers: light prepositions in Taqbaylit of Chemini. In *Principles of Linearization*, eds. Theresa Biberauer et Ian Roberts. Berlin: Mouton de Gruyter.
- Biro, Szilvia, et Leslie, Alan. 2007. Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science* 10:379-398.
- Bloom, Paul. 2000. *How children learn the meanings of word*. Cambridge MA: MIT Press / Bradford books.
- Borer, Hagit. 1984. *Parametric Syntax*. Dordrecht: Foris.
- Borer, Hagit. 1994. The Projection of Arguments. 19-47.
- Borer, Hagit. 1998. Deriving the Passive without Theta Grids. In *Morphology and Its Relation to Phonology and Syntax*, eds. Stephen G. Lapointe, Diane K. Brentari et Patrick M. Rarrell, 60-99: CSLI Publications.
- Borer, Hagit. 2005. *Structuring Sense*. Oxford: Oxford University Press.
- Castañeda, Hector-Neri. 1967. Comments on Donald Davidson's 'The Logical Form of Action Sentences'. In *The Logic of Decision and Action*, ed. Nicholas Rescher, 104-112. Pittsburgh: University of Pittsburgh Press.
- Chaker, Salem. 1983. *Un parler berbère d'Algérie (syntaxe)*. Aix en Provence: Publications de l'Université de Provence.
- Chillier Zesiger, Laurence, Arabatzi, Marina, Baranzini, Lara, Cronel-Ohayon, Stéphanie, et Thierry, D. 2006. The acquisition of French pronouns in normal children and in children with SLI. Genève: Université de Genève.
- Chomsky, Noam, et Halle, Morris. 1968. *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Berlin: de Gruyter.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

- Clark, Eve. 1985. *The acquisition of Romance, with special reference to French*. Hillsdale: Lawrence Erlbaum Associates.
- Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.
- Croft, William. 1998. Event Structure in Argument Linking. In *The Projection of Arguments. Lexical and Compositional Factors*, eds. Miriam Butt et Wilhelm Geuder, 21-64. Stanford: CSLI Publications.
- Csibra, Gergely. 2008. Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition* 107:705-717.
- Davidson, Donald. 1967. The logical form of action sentences. In *Essays on Actions and Events*. Oxford: Oxford University Press (1980).
- Davidson, Donald. 1970. Mental events. In *Essays on actions and events*, ed. Donald Davidson, 207-224. Oxford: Oxford University Press (1980).
- Dell, François, et Elmedlaoui, Mohamed. 1989. Clitic Ordering, Morphology and Phonology in the Verbal Complex of Imdlawn Tashlhiyt Berber. *Langues Orientales, Anciennes, Philologie et Linguistique* 2:165-194.
- Dell, François, et Elmedlaoui, Mohamed. 1991. Clitic Ordering, Morphology and Phonology in the Verbal Complex of Imdlawn Tashlhiyt Berber. Part II. *Langues Orientales, Anciennes, Philologie et Linguistique* 3:77-104.
- Dennett, Daniel. 1987. *The Intentional Stance*. Cambridge, MA.: Bradford Books/MIT Press.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67:547-619.
- Dupuy, Lucile. 2009. Acquisition des clitiques objet en développement atypique : l'hypothèse de la complexité morphologique Ecole d'orthophonie, Univ. Tours.
- Emonds, Joseph. 1985. *A unified theory of syntactic categories*. Dordrecht: Foris.
- Fisher, Cynthia, Gertner, Yael, Scott, Rose M., et Yuan, Sylvia. 2010. Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science* 1:143-149.
- Fodor, Jerry. 1983. *The Modularity of Mind*. Cambridge, Massachusetts: MIT Press.
- Fodor, Jerry, et Lepore, Ernie. 1998. The Emptiness of the Lexicon: Reflections on James Pustejovsky's *The Generative Lexicon*. *Linguistic Inquiry* 29:269-288.
- Galand, Lionel. 1957. Un cas particulier de phrase non-verbale: 'l'anticipation renforcée' et l'interrogation en berbère. In *Mémorial André Basset*, 27-37. Paris: Maisonneuve.

- Gennari, Silvia P., Sloman, Steven A., Malt, Barbara C., et Fitch, W. Tecumseh. 2002. Motion events in language and cognition. *Cognition* 83:49-79.
- Gergely, Gyorgy, Nadasdy, Zoltan, Csibra, Gergely, et Biro, Szilvia. 1995. Taking the intentional stance at 12 months of age [1995/8]. *Cognition* 56:165-193.
- Gergely, Gyorgy, et Csibra, Gergely. 1997. Teleological reasoning in infancy: The infant's naive theory of rational action: A reply to Premack et Premack [1997/5/2]. *Cognition* 63:227-233.
- Gibson, Eleanor J., Owsley, Cynthia, J., et Johnston, Joan. 1978. Perception of invariants by five-month-old infants: Differentiation of two types of motion. *Developmental Psychology* 14:407-415.
- Gleitman, Lila R. 1990. The structural sources of verb meaning. *Language Acquisition* 1:3-55.
- Goldsmith, John. 1979. Autosegmental Phonology, MIT: PhD.
- Grodzinsky, Yosef, et Reinhart, Tanya. 1993. The Innateness of Binding and Coreference. *Linguistic inquiry* 24:69.
- Gropen, Jess, Pinker, Steven, Hollander, Michelle, et Goldberg, Richard. 1991. Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*:153-195.
- Grüter, Theres. 2006. Object clitics and null objects in the acquisition of French, McGill University.
- Guerssel, Mohamed. 1987. The Status of the Lexical Category Preposition in Berber: Implications for the Nature of the Construct State. In *Studies in Berber Syntax*, ed. Ken Hale, 159-190. Cambridge, Mass.: Center for Cognitive Science, MIT.
- Guerssel, Mohamed, et Lowenstamm, Jean. 1990. The derivational morphology of the Classical Arabic verbal system. Ms. UQAM, Paris 7.
- Guerssel, Mohamed. 1992. On the Case System of Berber. *Canadian Journal of Linguistics* 37:175-195.
- Haas, M. R. 1940. Ablaut and its function in Muskogee. *Language* 16:141-150.
- Haiden, Martin. 2005a. *Theta Theory*. vol. 78: Studies in Generative Grammar. Berlin: Mouton de Gruyter.
- Haiden, Martin. 2005b. Concatenation and Interpretation. In *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, eds. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz et Jan Koster, 157-165. Berlin: Mouton de Gruyter.

- Haiden, Martin. 2005c. Case 117: Verb Particle Constructions. In *The Blackwell Companion of Syntax*, eds. Martin Everaert et Henk C. van Riemsdijk. Oxford: Blackwell.
- Haiden, Martin. 2006. Review of Stefan Müller (2002): *Complex Predicates*. *Journal of Comparative Germanic Linguistics* 9:143-160.
- Haiden, Martin. 2007. Sous-specification et projection argumentale dans le *Theta System*. *Recherches Linguistiques de Vincennes* 36:153-188.
- Haiden, Martin. 2008a. Linearization in Bare Prosodic Structure. *Poznań Studies in Contemporary Linguistics* 44:401-421.
- Haiden, Martin. 2008b. Linearization in Bare Prosodic Structure. *Poznan Studies in Contemporary Linguistics* 44:401-421.
- Haiden, Martin. 2009. On Bare Prosodic Structure and the Spell-Out of Features. In *Explorations of Phase Theory: Interpretation at the interfaces*, ed. Kleanthes Grohmann, 67-94. Berlin: de Gruyter.
- Haiden, Martin. 2010. Thematic roles and modality. Ms. Univ. Tours et INSERM U930 "Imagerie et cerveau".
- Haiden, Martin. à paraître. The content of semantic roles. In *The Theta System*, eds. Martin Everaert, Marjana Marelj et Tal Siloni. Oxford: Oxford University Press.
- Halle, Morris, et Marantz, Alec. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, eds. Ken Hale et Samuel Jay Keyser, 111-176. Cambridge MA: MIT Press.
- Hamann, Cornelia, Rizzi, Luigi, et Frauenfelder, Uli. 1996. On the acquisition of Subject and object clitics in French. In *Generative Perspective on Language Acquisition*, ed. Harald Clahsen, 309-334. Amsterdam: John Benjamins.
- Hauser, Marc D., Chomsky, Noam, et Fitch, W. Tecumseh. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298:1569-1579.
- Heider, Fritz, et Simmel, Marianne. 1944. An experimental study of apparent behavior. *American Journal of Psychology* 57:243-259.
- Hirschfeld, Lawrence A., et Gelman, Susan A. 1994. *Mapping the Mind. Domain Specificity in Cognition and Culture*. Cambridge: Cambridge University Press.
- Idrissi, Ali. 2000. Towards a Root-and-Template Approach to Shape-Invariant Morphology, UQAM: PhD.

- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, Massachusetts: MIT Press.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 1996. The Architecture of the Linguistic-Spatial Interface. In *Language and Space*, eds. Paul Bloom, Mary A. T. Peterson, Lynn Nadel et Merrill F. Garrett, 1-30. Cambridge, MA.: MIT Press.
- Jaeggli, Osvaldo. 1982. *Topics in Romance Syntax*. Dordrecht: Foris.
- Jaeggli, Osvaldo. 1986. Three Issues in the Theory of Clitics: Case, Doubled NPs, and Extraction. In *The Syntax of Pronominal Clitics.*, ed. Hagit Borer, 15-42. New York: Academic Press.
- Jakubowicz, Celia, Nash, Lea, Rigaut, Catherine, et Loic-Gérard, Christophe. 1998. Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition* 7:113-160.
- Jakubowicz, Celia, et Tuller, Laurice. 2008. Specific language impairment in French. In *Studies in French applied linguistics*, ed. Dalila Ayoun, 97-133. Amsterdam: John Benjamins.
- Jakubowicz, Celia, et Nash, Lea. to appear. Why Accusative Clitics are Avoided in Normal and Impaired Language Development. In *Essays in Syntax, Morphology and Phonology in SLI*, eds. Celia Jakubowicz, Lea Nash et Ken Wexler. Cambridge, Mass.: MIT Press.
- Kaye, Jonathan, Lowenstamm, Jean, et Vergnaud, Jean-Roger. 1985. The Internal Structure of Phonological Elements: A Theory of Charm and Government. *Phonology Yearbook* 2:305-328.
- Kaye, Jonathan, Lowenstamm, Jean, et Vergnaud, Jean-Roger. 1990. Constituent Structure and Government in Phonology. *Phonology* 7:305-328.
- Kayne, Richard S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, Mass.: MIT Press.
- Kiparsky, Paul. 1973. Elsewhere in phonology. In *A Festschrift for Morris Halle*, eds. Stephen Anderson et Paul Kiparsky, 93-106. New York: Holt, Rinehart and Winston.
- Kotovskiy, Laura, et Baillargeon, Renée. 2000. Reasoning about collisions involving inert objects in 7.5-month-old infants. *Developmental Science* 3:344-359.

- Landau, Barbara, et Gleitman, Lila R. 1985. *Language and experience: evidence from the blind child*. Cambridge, MA.: Harvard University Press.
- Leben, Will. 1978. The representation of tone. In *Tone: a linguistic survey*, ed. V. Fromkin: Academic Press.
- Lee, Joanne N., et Naigles, Letitia R. 2008. Mandarin learners use syntactic bootstrapping in verb acquisition. *Cognition* 106:1028-1037.
- Leslie, Alan. 1984. Infant perception of a manual pick-up event. *British Journal of Developmental Psychology* 2:19-32.
- Leslie, Alan, et Keeble, Stephanie. 1987. Do six-month-old infants perceive causality? *Cognition* 25:265-288.
- Leslie, Alan. 1994. ToMM, ToBy, and Agency: Core architecture and domain specificity. In *Mapping the Mind. Domain Specificity and Culture*, eds. Lawrence A. Hirschfeld et Susan A. Gelman, 119-148. New York: Cambridge University Press.
- Leslie, Alan, Friedman, Ori, et German, Tim P. 2004. Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences* 8:528-533.
- Levin, Beth, et Rappaport Hovav, Malka. 1995. *Unaccusativity. At the Syntax-Lexical Semantics Interface*. Cambridge, Mass.: MIT Press.
- Levin, Beth, et Rappaport Hovav, Malka. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Lidz, Jeffery, Gleitman, Henry, et Gleitman, Lila. 2003. Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition* 87:151-178.
- Lipschutz, Seymour, et Lipson, Marc. 2001. *Theory and Problems of Linear Algebra: Schaum's Outline*. New York: McGraw-Hill.
- Lowenstamm, Jean, et Kaye, Jonathan. 1986. Compensatory Lengthening in Tiberian Hebrew. In *Studies in Compensatory Lengthening*, eds. Leo Wetzels et Engin Sezer, 97-132. Dordrecht: Foris.
- Lowenstamm, Jean. 1996. CV as the only syllable type. In *Current trends in phonology*, eds. Jacques Durand et Bernard Laks, 419-441. Manchester: ESRI.
- Lowenstamm, Jean. 2000. The No-Straddling Effect and its Interpretation. In *Research in Afroasiatic Grammar*, eds. Jacqueline Lecarme, Jean Lowenstamm et Uri Shlonsky, 183-198. Amsterdam/Philadelphia: Benjamins.
- Lowenstamm, Jean. 2003. À propos des gabarits. *Recherches Linguistiques de Vincennes* 32:7-30.

- MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ.: Lawrence Erlbaum Associates.
- Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *University of Pennsylvania Working Papers in Linguistics* 4:201-225.
- Marelj, Marjana. 2004. *Middles and Argument Structure across Languages*. Utrecht: LOT.
- McCarthy, John. 1983. Consonantal Morphology in the Chaha Verb. Paper presented at WCCFL.
- McCarthy, John, et Prince, Alan. 1993. Prosodic Morphology: Constraint Interaction and Satisfaction: Rutgers University Center for Cognitive Science.
- Merchant, Jason. 2002. Swiping in Germanic. In *Studies in Comparative Germanic Syntax*, eds. C. Jan-Wouter Zwart et Werner Abraham, 295-321. Amsterdam: John Benjamins.
- Mettouchi, Amina. 2005. Contribution à l'étude de la prédication. De la sémantique énonciative à la typologie., INALCO: HDR.
- Michotte, Albert. 1945. *La perception de la causalité*. Louvain: Publications Universitaires de Louvain.
- Miller, Philip, et Sag, Ivan. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15:573-639.
- Montgomery, Derek E., et Montgomery, Demaris A. 1999. The influence of movement and outcome on young children's attributions of intention. *British Journal of Developmental Psychology* 17:245-261.
- Muentener, Paul. 2009. The origin and development of causal reasoning, Harvard University.
- Muentener, Paul, et Carey, Susan. 2010. Infants' causal representations of state change events. *Cognitive Psychology* online preview: doi:10.1016/j.cogpsych.2010.02.001.
- Naigles, Letitia R. 1988. Syntactic bootstrapping as a procedure for verb learning, University of Pennsylvania.
- Naigles, Letitia R. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17:357-374.
- Naigles, Letitia R., Fowler, Anne, et Helm, Atessa. 1992. Developmental shifts in the construction of verb meanings. *Cognitive Development* 7:403-427.

- Newman, Paul. 1990. *Nominal and verbal plurality in Chadic*. vol. 12: Publications in African languages and linguistics. Dordrecht: Foris.
- Ostler, Nicholas. 1980. A theory of case linking and agreement, Indiana University Linguistics Club.
- Ouhalla, Jamal. 1993. Subject-Extraction, Negation and the Anti-Agreement Effect. *Natural language & linguistic theory* 11:477.
- Ouhalla, Jamal. 2005. Clitic-placement, grammaticalization and reanalysis in Berber. In *The Oxford Handbook of Comparative Syntax*, eds. Guglielmo Cinque et Richard Kayne, 607-638. Oxford: Oxford University Press.
- Papafragou, Anna, Hulbert, Justin, et Trueswell, John. 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 108:155-184.
- Parsons, Terence. 1990. *Events in the Semantics of English: A Study in Sub-Atomic Semantics*. Cambridge, Mass.: MIT Press.
- Perez-Leroux, Ana, Pirvalescu, Mihaela, et Roberge, Yves. 2008. Null objects in child language; Syntax and the lexicon. *Lingua* 118:370-398.
- Pinker, Stephen. 1984. *Language learnability and language development*. Cambridge: MIT Press.
- Pinker, Stephen. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA.: MIT Press.
- Postal, Paul. 1972. On some rules that are not successive-cyclic. *Linguistic Inquiry* 3:211-222.
- Premack, David. 1990. The infant's theory of self-propelled objects [1990/7]. *Cognition* 36:1-16.
- Prince, Alan, et Smolensky, Paul. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell.
- Prunet, Jean-Francois, Beland, Renee, et Idrissi, Ali. 2000. The Mental Representation of Semitic Words. *Linguistic inquiry* 31:609-648.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, Mass.: MIT Press.
- Rapp, Irene. 1997. *Partizipien und semantische Struktur: zu passivischen Konstruktionen mit dem 3. Status*. vol. 54: Studien zur deutschen Grammatik. Tübingen: Stauffenberg.
- Reinhart, Tanya. 2000. The Theta System: Syntactic Realization of Verbal Concepts. *OTS Working Papers in Linguistics*.
- Reinhart, Tanya. 2003. The Theta System - An Overview. *Theoretical Linguistics* 28.

- Reinhart, Tanya, et Siloni, Tal. 2005. The Lexicon-Syntax Parameter: reflexivization and other arity operations. *Linguistic Inquiry* 36:389-436.
- Reinhart, Tanya. 2006. *Interface Strategies*. Cambridge, MA: MIT Press.
- Roberge, Yves. 1990. *The syntactic recoverability of Null Arguments*. Montreal: McGill - Queen's University Press.
- Rochat, Philippe, Morgan, Rachel, et Carpenter, Melinda. 1997. Young infants' sensitivity to movement information specifying social causality. *Cognitive Development* 12:537-561.
- Rose, Sharon. 1997. Theoretical issues in comparative Ethio-Semitic phonology and morphology, McGill.
- Ruffman, Ted, Garnham, Wendy, Import, Arlina, et Connolly, Dan. 2001. Does Eye Gaze Indicate Implicit Knowledge of False Belief? Charting Transitions in Knowledge. *Journal of Experimental Child Psychology* 80:201-224.
- Russell, Bertrand. 1905. On Denoting. *Mind & language* 5:479-493.
- Saxe, Rebecca, et Carey, Susan. 2006. The perception of causality in infancy. *Acta Psychologica* 123:144-165.
- Saxe, Rebecca, Tzelnik, Tania, et Carey, Susan. 2007. Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology* 43:149-158.
- Scheer, Tobias, et Ségéral, Philippe. 2005. What lenition and fortition tells us about Gallo-Romance Muta cum Liquida. In *Romance Languages and Linguistic Theory 2003*, eds. Twan Geerts, Haike Jacobs et Ivo van Ginneken, 235-267. Amsterdam: Benjamins.
- Scheer, Tobias. 2010. How morpho-syntax talks to phonology. A survey of extra-phonological information in phonology since Trubetzkoy's Grenzsignale. Ms., To appear at Mouton de Gruyter, Berlin.
- Schlottman, Anne, et Surian, Luca. 1999. Do 9-month-olds perceive causation-at-a-distance? *Perception* 28:1105-1113.
- Ségéral, Philippe. 2005. Racine indo-européenne et théorie phonologique. Quelques remarques. In *7e Rencontres Internationales RFP*. Aix-en-Provence.
- Shlonsky, Ur. 2004. Enclisis and proclisis. In *The Structure of CP and IP. The cartography of syntactic structures*, ed. Luigi Rizzi, 329-353. New York: Oxford University Press.
- Southgate, Victoria, Johnson, Mark H., et Csibra, Gergely. 2008. Infants attribute goals even to biomechanically impossible actions. *Cognition* 107:1059-1069.

- Spelke, Elizabeth, Phillips, Ann T., et Woodward, Amanda. 1995. Infants' knowledge of object motion and human action. In *Causal cognition: a multidisciplinary debate*, eds. Dan Sperber, David Premack et Ann J. Premack, 44-78. Oxford: Clarendon Press.
- Sportiche, Dominique. 1996. Clitic constructions. In *Phrase structure and the lexicon*, eds. Jan Rooryck et Laurie Zaring, 213-267. Dordrecht: Kluwer.
- Stechow, Arnim von. 2007. Schritte zu Satzsemantik II. Ms. Univ. Tübingen.
- Talmy, Leonard. 1975. Semantics and Syntax of Motion. In *Syntax and Semantics*, ed. J. P. Kimball, 181-238. New York: Academic Press.
- Talmy, Leonard. 1976. Semantic Causative Types. In *Syntax and Semantics*, ed. M. Shibatani, 43-116. New York: Academic Press.
- Talmy, Leonard. 1983. How Language Structures Space. In *Spatial orientation: Theory, Research and Application*, eds. H. Pick et L. Acredolo. New York: Plenum.
- Talmy, Leonard. 1988. Force Dynamics in Language and Thought. *Cognitive Science* 12:49-100.
- Talmy, Leonard. 2000. *Toward a cognitive semantics*. Cambridge, Mass.: MIT Press.
- Tomasello, Michael. 1992. *First verbs : a case study of early grammatical development*. Cambridge England ; New York, NY: Cambridge University Press.
- Tuller, Laurice, Audollent, C., Delage, Hélène, et Monjauze, Cécile. 2004. Protocole expérimental de Production des Pronoms Clitiques (PPPC). Tours: Laboratoire Langage et handicap.
- Tuller, Laurice, Henry, Célia, et Barthez, Marie-Anne. 2007. SLI in adolescence: avoiding complexity. Tours: Univ. François Rabelais.
- Tuller, Laurice, Delage, Hélène, Monjauze, Cécile, Pillar, Ane-Gaëlle, et Barthez, Marie-Anne. 2011. Clitic pronoun production as a measure of atypical language development in French. *Lingua* 121:423-441.
- Uriagereka, Juan. 1995. Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic inquiry* 26:79-124.
- Varlokosta, et al, et. in prep. A cross linguistic study of the acquisition of clitic pronoun production: COST action A33.
- Verkuyl, Henk. 1993. *A Theory of Aspectuality. The Interaction between Temporal and Atemporal Structure*: Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.

- Wexler, Ken. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106:23-79.
- Wexler, Ken. 2002. Lenneberg's Dream: Learning, Normal Language Development and Specific Language Impairment. In *Language competence across populations: Towards a definition of Specific Language Impairment*, eds. J. Schaffer et Y. Levy: Erlbaum.
- White, Lydia. 1996. Clitics in L2 French. In *Generative perspectives on language acquisition*, ed. Harald Clahsen, 335-368. Amsterdam: John Benjamins.
- Wimmer, Heinz, et Perner, Josef. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13:103-128.
- Yuan, Sylvia, et Fisher, Cynthia. 2009. "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science* 20:619-626.
- Zesiger, Pascal, Chillier Zesiger, Laurence, Arabatzi, Marina, Baranzini, Lara, Cronel-Ohayon, Stéphanie, Franck, Julie, Frauenfelder, Ulrich, Hamann, Cornelia, et Rizzi, Luigi. in press. The acquisition of pronouns by French children: A parallel study of production and comprehension. *Applied psycholinguistics*.