

Université Côte d'Azur
ED 86 - SHAL

Variabilité et acquisition du français hexagonal :
Comment la syntaxe vient aux enfants

Katerina Palasis
Maître de conférences en Sciences du langage

Habilitation à Diriger des Recherches – CNU 07
Volume 1 : Note de synthèse

28 novembre 2024

Jury :

Lena BAUNAZ, Maître de conférences HDR, Université Côte d'Azur

Caterina DONATI, Professeure, Université Paris Cité

Amanda EDMONDS, Professeure, Université Côte d'Azur

Richard FAURE, Professeur, Université de Tours

Karen LAHOUSSE, Professeure, KU Leuven

Marie-Thérèse LE NORMAND, Directrice de Recherche, INSERM,
Université Paris Cité

Fanny MEUNIER, Directrice de recherche, CNRS, Université Côte
d'Azur (garante)

Remerciements

Merci à Fanny Meunier, qui a accepté de se porter garante pour ce travail et qui, en tant que Directrice du Laboratoire BCL jusqu'au moment de la rédaction de ce texte, a fourni un environnement propice et motivant pour la recherche.

Merci aux collègues de l'équipe de recherche 'Acquisition, Apprentissage, Adaptation' (AAA) et d'autres collègues chercheurs et enseignants-chercheurs du Laboratoire BCL pour tous les échanges scientifiques stimulants et humains bienveillants.

Merci à Richard Faure pour toutes ces années de collaborations sur les questions *wh-* en français et pour ces belles publications conjointes.

Merci également aux rapporteurs et membres du jury qui ont accepté ces tâches. Votre accompagnement est un honneur et un plaisir pour moi.

Enfin, merci à Stéphane, Juliette et Rémi pour tout ce que vous m'apportez au quotidien.

Table des matières

Résumé	6
Abstract	7
Abréviations	8
Introduction	9
Chapitre I : Domaines d’investigation et cadres théoriques	13
I. Acquisition du langage	14
1. Trajectoire développementale typique.....	14
1.1 Capacités de traitement des informations linguistiques	15
1.2 Production du langage.....	17
2. L1/L2P/L2.....	18
2.1 Contextes de transmission : implicite vs explicite	18
2.2 Modalités : orale vs écrite	19
2.3 Types de mémoire : non-déclarative vs déclarative	20
2.4 Âges initiaux d’exposition : la question de la période critique	21
3. Modèles de l’acquisition du langage.....	22
3.1 Approche innéiste.....	24
3.2 Approche basée sur l’usage	26
3.3 Questions oui/non au cœur du débat	28
3.4 Approche « GU + statistiques »	30
4. Positionnement scientifique	33
II. Apport linguistique	34
1. LAE	35
1.1 Caractéristiques et fonction dans le processus acquisitionnel	35
1.2 Facteurs sociolinguistiques de variation	36
2. Français hexagonal	38
2.1 Doublement vs dislocation	40
2.2 Statut morphosyntaxique du clitique nominatif.....	42
2.3 Négation simple vs discontinue	45
2.4 Langue à sujet nul ?	47
2.5 Comment analyser les explétifs ?	49
2.6 Structures interrogatives	51
3. Théorisations sur la variation morphosyntaxique.....	55

3.1	Variationnisme.....	55
3.2	Diglossie.....	56
4.	Positionnement scientifique	59
Chapitre II : Travaux réalisés.....		60
I.	Spécificités de l'approche proposée.....	60
1.	Création d'une base de données du français enfantin	60
2.	Cohérence interne du système enfantin.....	65
II.	Formes du sujet : morphologie et syntaxe	66
1.	Phénomène du sujet nul enfantin.....	68
1.1	Apport des données.....	69
1.2	Contribution à la théorie	72
1.3	Questions ouvertes.....	73
2.	Statut morphosyntaxique du clitique nominatif.....	74
2.1	Application des tests aux données enfantines	75
2.2	Retour sur la négation préverbale.....	83
2.3	Perspectives diatopique et diachronique	85
3.	Synthèse	87
3.1	Clitique nominatif : marqueur préverbal d'accord	87
3.2	Français enfantin : langue à sujet nul	89
3.3	Différents types d'explétifs.....	89
III.	Questions <i>wh</i>- : morphosyntaxe et sémantique	90
1.	Chez l'enfant.....	91
1.1	Méthode	91
1.2	Facteur 1 : émergence de « qu'est-ce que »	94
1.3	Facteur 2 : prévalence du présentatif « c'est »	95
1.4	Facteur 3 : deixis et non-présupposition	97
1.5	Synthèse	98
2.	Français informel : langue à <i>wh</i> -in-situ	98
3.	Chez l'adulte : trait [Exclusion].....	100
IV.	Hypothèse de la diglossie	102
1.	Français L1	103
2.	Français L2P.....	107
3.	Synthèse	110

Bilan et perspectives	112
Références bibliographiques	116

Résumé

Le premier objectif de ce texte d'Habilitation à Diriger des Recherches est de proposer une synthèse actualisée de recherches en linguistique théorique, cognitive, sociolinguistique, psycholinguistique, neurolinguistique et linguistique computationnelle sur les processus qui sous-tendent **l'acquisition du langage** et sur le **français hexagonal non-standard/informel/oral** produit par les adultes (Chapitre I). Puis, à partir d'un panorama des principales approches théoriques qui formalisent les processus acquisitionnels (innéisme, empirisme, GU + statistiques) et la variation morphosyntaxique du français (variationnisme, diglossie), mes travaux sont situés dans une **approche intégrative de l'acquisition (GU + données)** et **l'hypothèse de la diglossie**.

Le second objectif est de rapporter une sélection de mes travaux de recherche publiés depuis ma thèse en 2009 grâce au traitement d'une **base de données enfantines collectée pendant trois ans auprès d'une classe de maternelle** en France¹, et de proposer quelques ajouts inédits (Chapitre II). Les recherches portent sur l'acquisition d'aspects cruciaux de la **morphosyntaxe du français oral** : (i) les différentes formes que prend le **sujet** dans les phrases enfantines, avec des problématiques majeures, telles que le **sujet nul enfantin** et le **statut morphosyntaxique du clitique nominatif**, et (ii) les **questions wh-**, avec les questionnements sur les facteurs qui sous-tendent la position des éléments interrogatifs et sur les aspects développementaux de ces structures. Dans le cadre de la **diglossie**, le français informel est décrit comme une langue à sujet nul avec des marqueurs préverbaux d'accord, et comme une langue à *wh*-in-situ avec un effet de la sémantique du verbe sur le déplacement ex-situ. Le français informel représente une grammaire initiale (G1), acquise spontanément par tous les enfants francophones et complétée subséquentement par une seconde grammaire (G2, français standard/formel) dont la maîtrise varie beaucoup d'un locuteur à un autre.

¹ <http://childes.talkbank.org/access/French/Palasis.html>

Abstract

The first objective of this *Habilitation à Diriger des Recherches* is to provide an up-to-date synthesis of research in theoretical linguistics, cognitive linguistics, sociolinguistics, psycholinguistics, neurolinguistics and computational linguistics on the processes underlying **language acquisition** and on the **non-standard/informal/oral hexagonal French** produced by adults (Chapter I). Then, based on an overview of the main theoretical approaches that formalise acquisitional processes (nativism, empiricism, UG + statistics) and the morphosyntactic variation of French (variationism, diglossia), my work is situated within an **integrative approach to acquisition (UG + data)** and the **diglossia hypothesis**.

The second objective is to report on a selection of my research published since my thesis in 2009, thanks to the processing of a **database of children's utterances collected over three years from a kindergarten class** in France², and to propose some unpublished additions (Chapter II). The research focuses on the acquisition of crucial aspects of the **morpho-syntax of spoken French**: (i) the different forms of the **subjects**, with major issues such as the **null-subject phenomenon** and the **morpho-syntactic status of the nominative clitic**, and (ii) **wh-questions**, with the factors underlying the position of interrogative items and the developmental aspects of these structures. In the context of **diglossia**, informal French is described as a null-subject language with preverbal agreement markers, and as a *wh*-in-situ language with an effect of verb semantics on ex-situ fronting. Informal French represents an initial grammar (G1), acquired spontaneously by all French-speaking children and subsequently supplemented by a second grammar (G2, standard/formal French) whose mastery varies greatly from one speaker to another.

² <http://childes.talkbank.org/access/French/Palasis.html>

Abréviations

X;X	Âges donnés en années;mois
CP	Syntagme du complémenteur (<i>Complementizer phrase</i>) : point de départ de toute proposition dans la théorie générative, héberge notamment les items interrogatifs « ex-situ »
DP	Syntagme du déterminant (<i>Determiner phrase</i>) : héberge le déterminant et le syntagme nominal
G1	Grammaire 1 : première grammaire acquise par l'enfant (hypothèse de la diglossie)
G2	Grammaire 2 : deuxième grammaire acquise ou apprise en fonction de l'environnement linguistique (hypothèse de la diglossie)
GS	Grande section de maternelle (âges du corpus : 4;5-5;11)
GU	Grammaire Universelle (approche innéiste du langage)
IP	Syntagme de la flexion (<i>Inflectional phrase</i>) : héberge les marques de temps, mode, aspect et personne
LAE	Langage adressé aux enfants (différent du langage adressé aux adultes)
MS	Moyenne section de maternelle (âges du corpus : 3;6-4;11)
PS	Petite section de maternelle (âges du corpus : 2;5-4;0)
TPS	Toute petite section de maternelle (âges du corpus : 2;3-3;1)
VP	Syntagme verbal (<i>Verbal phrase</i>) : héberge le verbe et ses compléments

Introduction

Même les modèles les plus innéistes de l'acquisition du langage concèdent une place incompressible à l'environnement linguistique de l'enfant dans le processus du développement du langage. La démonstration la plus simple consiste à rappeler que la langue acquise par un jeune enfant est toujours celle de son environnement linguistique précoce et quotidien. Ainsi, deux aspects complémentaires sont à prendre en compte dans les travaux de recherche sur l'acquisition du langage : l'étude du système linguistique dans lequel baigne le jeune enfant et l'étude des processus acquisitionnels déclenchés par cet environnement linguistique. Il est alors très stimulant pour un chercheur de découvrir que ces deux domaines représentent des sources de nombreux débats et qu'ils suscitent une recherche riche, foisonnante, avec de nombreux questionnements non seulement en linguistique, mais également en psycholinguistique, neurolinguistique et psychologie cognitive, pour ne citer que quelques-uns des domaines qui s'intéressent à l'acquisition du langage. Le sous-titre de cette habilitation est d'ailleurs un hommage appuyé à l'ouvrage de Bénédicte De Boysson-Bardies (1996), *Comment la parole vient aux enfants. De la naissance jusqu'à deux ans*, qui a été pour moi un ouvrage fondateur dans ma découverte enthousiaste de la psycholinguistique à partir de ma formation initiale en linguistique théorique (plutôt générativiste).

Ce texte d'Habilitation à Diriger des Recherches a deux objectifs principaux, répartis dans les deux chapitres du Volume I. Le premier objectif est de proposer une synthèse actualisée et commentée de recherches sur le développement du langage et sur l'étude du français hexagonal contemporain utilisé par des adultes à l'adresse d'adultes et par des adultes à l'adresse d'enfants. Le numéro spécial de la revue *Journal of Child Language* édité par Ambridge & Wagner (2021) sur les formalisations actuellement en compétition pour rendre compte du développement linguistique infantin et la compilation des travaux de Ashby (2023) sur la variation morphosyntaxique en français oral montrent à quel point ces deux domaines de recherches sont foisonnants. Cette revue de littérature

permet de regrouper des travaux que l'on trouve généralement dans des supports séparés, puisqu'elle met en regard des travaux en linguistique théorique, cognitive, sociolinguistique, psycholinguistique, neurolinguistique et linguistique computationnelle notamment.

Le second objectif est de rapporter mes travaux de recherche réalisés depuis ma soutenance de thèse en 2009 (Palasis, 2009; publication en 2010b) et de proposer quelques ajouts inédits, rendus possibles par le recul et le travail de synthèse que représente cet ouvrage. Les recherches présentées portent principalement sur l'acquisition de plusieurs aspects cruciaux de la morphosyntaxe du français oral par des enfants francophones natifs entre trois et six ans. Toutes les thématiques ont été traitées grâce à une base de données créée pendant mes années de doctorat et qui, depuis et encore à l'heure actuelle, permet une extraction ciblée de productions de seize enfants pendant leurs trois années de maternelle. Les recherches principales portent sur : (i) les différentes formes que prend le sujet dans les phrases enfantines, avec toutes les problématiques afférentes qui représentent des débats importants en recherche (par exemple, le sujet nul enfantin ou le statut du clitique nominatif), (ii) les questions *wh-*, avec les questionnements sur les facteurs qui sous-tendent la position des éléments interrogatifs et sur les aspects développementaux de ces structures, et (iii) l'hypothèse de la diglossie en France, qui peut être examinée grâce à la connaissance fine de ce que ce produisent les locuteurs francophones natifs, pendant leurs premières années de vie et à l'âge adulte.

Le Volume II de l'ouvrage comporte un CV synthétique, un descriptif rédigé de mes activités, une liste exhaustive de mes travaux, et une sélection parmi mes publications effectuées depuis ma thèse en 2009. Il présente ainsi six articles publiés dans des revues internationales à comité de lecture (1, 4, 6, 7, 8, 9) et trois actes de colloques à comité de lecture (2, 3, 5) :

1. Palasis, K. (2010). Introducing New French Child Data: Thoughts on their Gathering and Coding. *Corpus*, 9 'La syntaxe de corpus', 33-51, <https://journals.openedition.org/corpus/1801>

2. Palasis, K. (2011). Français spontané et français normé : bien plus que deux variétés d'un même français. Implications acquisitionnelles et didactiques. In O. Bertrand & I. Schaffner (Eds.), *Variétés, Variations & Formes du français* (pp. 55-72). Palaiseau: Les Éditions de l'École Polytechnique, <https://www.researchgate.net/profile/Katerina-Palasis/publications>

3. Palasis, K. (2012). Economy Principles within Child Speech: When the Nominative Clitic does not Surface. In S. Ferré, P. Prévost, L. Tuller, & R. Zebib (Eds.), *Selected Proceedings of the Romance Turn IV Workshop on the Acquisition of Romance Languages* (pp. 261-283). Newcastle: Cambridge Scholars Publishing, <https://www.researchgate.net/profile/Katerina-Palasis/publications>

4. Palasis, K. (2013). The case for diglossia: Describing the emergence of two grammars in the early acquisition of Metropolitan French. *Journal of French Language Studies*, 23(1), Numéro spécial 'L'hypothèse d'une diglossie en France', Massot, Benjamin & Paul Rowlett (Eds.), 17-35, <https://doi.org/10.1017/S0959269512000348>

5. Palasis, K. (2014). Étude morphosyntaxique des pronoms sujets à partir d'un corpus de maternelle. In R. Mir-Samii (Ed.), *Du sujet et de son absence dans les langues* (pp. 259-270). Le Mans: Université du Maine, <http://doi.org/10.13140/2.1.1586.9440>

6. Palasis, K. (2015). Subject clitics and preverbal negation in European French: Variation, acquisition, diatopy and diachrony. *Lingua*, 161, 125-143, <http://dx.doi.org/10.1016/j.lingua.2014.11.012>

7. Palasis, K., Faure, R. & Lavigne, F. (2019). Explaining variation in *wh*-position in child French: A statistical analysis of new seminaturalistic data. *Language Acquisition*, 26(2), 210-234, <https://doi.org/10.1080/10489223.2018.1513004>

8. Faure, R. & Palasis, K. (2021). Exclusivity! *Wh*-fronting is not optional *wh*-movement in Colloquial French. *Natural Language & Linguistic Theory*, 39(1), 57-95, <https://doi.org/10.1007/s11049-020-09476-w>
9. Palasis, K., Faure, R. & Meunier, F. (2023). *Wh*-in-situ in child French: Deictic triggers at the syntax-semantics interface. *Journal of French Language Studies*, 33(3), 273-298, <https://doi.org/10.1017/S0959269523000030>

Chapitre I : Domaines d'investigation et cadres théoriques

Ce premier chapitre expose des notions fondamentales et un état de l'art commenté de recherches dans les deux domaines qui sous-tendent le développement du langage : les processus acquisitionnels (cf. Section I) et l'environnement linguistique qui déclenche ces processus (cf. Section II).

Plus précisément, les Sections I.1 et I.2 synthétisent un certain nombre de caractéristiques qui rendent compte de l'éventail des trajectoires développementales chez l'enfant, des toute premières perceptions du langage à la production de la syntaxe complexe, et de l'acquisition de la langue native à l'apprentissage des langues en contexte scolaire. L'objectif de ces sections est de situer et décrire les particularités de ce qui représente le cœur de mes recherches détaillées ensuite dans le Chapitre II, c'est-à-dire l'acquisition de la morphosyntaxe en langue native (L1). Puis, la Section I.3 propose un panorama des principales théories linguistiques actuelles qui ont pour objectif de formaliser ces trajectoires développementales. Cette partie expose ainsi pourquoi tant de débats perdurent dans le domaine et attire l'attention sur les dernières tendances de ces recherches.

Dans un second temps, la Section II décrit l'apport linguistique qui nourrit le développement du langage, en s'intéressant notamment au langage adressé aux enfants (LAE, cf. Section II.1) puis aux caractéristiques de l'environnement linguistique des enfants francophones en France (cf. Section II.2). Il est en effet également fondamental pour la qualité de la recherche en acquisition d'être au fait des travaux sur les différents facteurs qui peuvent impacter les processus acquisitionnels (la catégorie socio-professionnelle des parents par exemple). La Section II se termine par une partie plus théorique, car le français a la particularité de présenter de nombreux phénomènes de variation dans sa morphosyntaxe, ce qui a conduit les chercheurs à en proposer différentes formalisations (cf. Section II.3).

I. Acquisition du langage

L'intérêt pour le développement du langage chez l'enfant est relativement récent dans l'histoire des sciences et la publication généralement rapportée comme marquant le point de départ de cet intérêt est celle de Darwin (1877) qui décrit la production des premiers mots de son propre fils. Le XX^e siècle a ensuite vu un nombre croissant de scientifiques s'intéresser à l'acquisition du langage, à commencer par des psychologues américains auxquels on doit des collectes massives de données enfantines qui ont permis de mettre en place les premières normes développementales du langage typique (par exemple, Smith, 1926). Ces collectes initiales de productions spontanées ont ensuite été complétées par des collectes de données expérimentales (à partir du fameux *Wug* de Berko, 1958) et neurophysiologiques, grâce au développement des techniques d'imagerie cérébrale et d'oculométrie notamment (par exemple, Dehaene-Lambertz, 2017 ; Franck, Millotte, Posada & Rizzi, 2013).

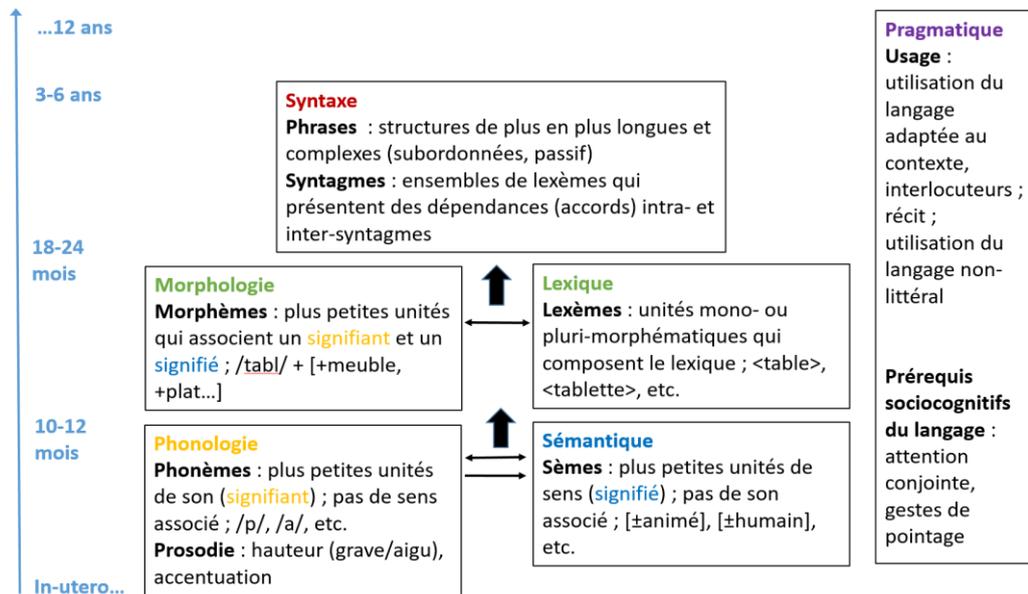
La trajectoire développementale typique est ainsi relativement bien documentée à l'heure actuelle, tant en perception du langage qu'en production (cf. Section 1). À partir de ces données, l'attention des chercheurs s'est également portée sur la diversité des situations développementales en fonction notamment du contexte et de la modalité de transmission des informations linguistiques, du type de mémoire impliqué dans le traitement de ces informations, et de l'âge d'exposition (cf. Section 2). Enfin, la recherche a également eu pour objectif de formaliser les processus qui sous-tendent le développement linguistique en accordant des places respectives très diverses à la biologie de l'être humain et à l'apport environnemental (cf. Section 3).

1. Trajectoire développementale typique

Le langage est composé d'un ensemble de composantes de forme (phonologie, morphologie, syntaxe), de contenu (sémantique, lexique) et d'usage (pragmatique). La Figure 1 ci-après situe ces composantes les unes par rapport aux

autres dans la chronologie du développement typique du langage et décrit brièvement les unités manipulées au niveau de chaque composante.

Figure 1 : développement typique du langage oral



Cette habilitation décrira principalement les composantes morphologique et syntaxique, souvent nommées conjointement « morphosyntaxe » notamment parce que les processus morphologiques de flexion (accords en genre, nombre, personne, etc.) sont déclenchés par la mise en relation de constituants lors de la création des syntagmes et des phrases au niveau de la composante syntaxique. Cette composition génère en effet des relations de dépendances intra- et inter-syntagmes, adjacentes et/ou non-adjacentes qui conditionnent les accords. Les phrases « [S_gLe S_gbébé] [S_gdort] » et « [S_gLe S_gbébé] [P_ldes P_lvoisins] [S_gdort] » illustrent ces différents types de dépendances avec les traits de Nombre [Singulier] et [Pluriel].

1.1 Capacités de traitement des informations linguistiques

Les travaux expérimentaux et d'imagerie cérébrale de ces dernières décennies ont mis en avant les capacités remarquables des jeunes enfants à traiter certaines informations linguistiques de leur environnement (par exemple, Kuhl, 2004 et Pearl & Goldwater, 2016 pour des vues d'ensemble). L'acquisition du langage démarre ainsi dès la période intra-utérine (à partir de la 28ème semaine de

grossesse), période pendant laquelle le fœtus entend et discrimine des aspects phonologiques de sa ou ses langues environnantes (Querleu, Renard, Versyp, Paris-Delrue & Crèpin, 1988 ; Das, Jana, Arora & Sengupta, 2020 ; Nallet & Gervain, 2021). Puis, dès la naissance, le nourrisson est sensible à l'ordre des items dans une séquence linguistique (Benavides-Varela & Gervain, 2017). Au cours de sa première année, l'enfant développe ensuite une sensibilité aux dépendances linguistiques adjacentes dans le continuum phonique, ce qui lui permet de traiter notamment les informations phonologiques sur la base des probabilités transitionnelles entre deux ou plusieurs phonèmes ou syllabes (Saffran, Aslin & Newport, 1996 ; Kuhl, 2004 ; Pelucchi, Hay & Saffran, 2009 ; de la Cruz-Pavía & Gervain, 2023). Puis, au cours de sa seconde année, l'enfant développe des capacités qui lui permettent de traiter les dépendances linguistiques non-adjacentes au niveau de la phrase (par exemple, « [Fém]la maison] [est] [Fém]Construite] » ; van der Kant et al., 2020), ce qui permet d'émettre et de tester l'hypothèse que les informations linguistiques sont non seulement temporellement ordonnées mais également syntaxiquement hiérarchisées (Perfors, Tenenbaum & Regier, 2011 ; Dehaene, Meyniel, Wacongne, Wang & Pallier, 2015 ; Ding, Melloni, Zhang, Tian & Poeppel, 2016 ; Goucha, Zaccarella & Friederici, 2017 ; Shi, Legrand & Brandenberger, 2020 ; Coopmans, De Hoop, Kaushik, Hagoort & Martin, 2022).³

De nombreux travaux mettent également en avant la capacité précoce de l'enfant à distinguer les morphèmes lexicaux (noms, verbes, adjectifs, notamment) des morphèmes grammaticaux (pronoms et déterminants, notamment), ce qui le guide dans la segmentation du continuum phonique en phrases, propositions et syntagmes (Shi, Cutler, Werker & Cruickshank, 2006 ; Christophe, Millotte, Bernal & Lidz, 2008 ; Kedar, Casasola, Lust & Parnet, 2017 ; revue dans Dye, Kedar & Lust, 2019). Ainsi, au fil des mois et à partir de son environnement linguistique, l'enfant détecte, mémorise et généralise les informations pertinentes au développement

³ La Section 3.3 ci-après montrera que les approches théoriques de l'acquisition du langage divergent sur la question de la hiérarchisation, notamment discutée à partir de la formation des questions oui/non en anglais.

de sa morphosyntaxe (Marcus, Vijayan, Bandi Rao & Vishton, 1999 ; Friederici, Mueller & Oberecker, 2011 ; Xu, Chodorow & Valian, 2023), bien avant le moment où il entame la période de la production de ses premières phrases vers l'âge de deux ans.

1.2 Production du langage

Tout en poursuivant l'analyse de plus en plus fine de ses données environnementales, l'enfant développe ses capacités de production. La première année du jeune enfant est essentiellement consacrée à la production d'un premier inventaire phonologique (vocalisations puis babillage) pour arriver à la production des premiers mots en fin de première année. La deuxième année voit les inventaires phonologique et lexical s'étoffer. La fin de la deuxième année et la troisième année sont ensuite consacrées au développement de la morphosyntaxe, avec les premières combinaisons de constituants, essentiellement lexicaux dans un premier temps (noms, verbes : « mangé purée »), puis progressivement enrichies par la production des morphèmes grammaticaux (déterminants, pronoms, auxiliaires : « j'ai mangé la purée » ; Prévost, 2009 ; Le Normand & Thai-Van, 2022). Vers trois ans, l'enfant dispose d'un inventaire phonologique étendu (mais encore à compléter par les phonèmes les plus complexes à articuler, notamment les fricatives telles que [ʃ] <ch> pour le français), il dispose d'un stock lexical d'environ 500 mots, et il produit des phrases assez complètes morphologiquement mais qui restent encore syntaxiquement « simples », c'est-à-dire qui ne comportent qu'une seule proposition (Kern, 2019).

La syntaxe « complexe » émerge aux alentours de 4 ans avec l'utilisation progressive de la coordination, de la subordination, des structures clivées et de la voix passive. L'enfant passe alors également de la production de phrases majoritairement isolées au récit (Canut, 2014). Enfin, comme indiqué dans la Figure 1 ci-dessus, la composante pragmatique se développe en parallèle à toutes ces étapes, puisqu'elle intervient dès les premières semaines de vie avec l'attention conjointe et se développe jusqu'à l'adolescence avec la maîtrise du

langage non-littéral, en lien avec la maturation plus tardive des lobes frontaux (Bohn & Frank, 2019 ; Clark, 2004 ; Dardier, 2004).

Les travaux rapportés dans le Chapitre II porteront essentiellement sur le développement de la morphosyntaxe chez des enfants entre 2;3 et 5;11. Par ailleurs, la trajectoire développementale esquissée dans la Section 1 est celle d'un enfant typique en contact précoce avec un environnement linguistique. Elle est donc représentative d'une situation acquisitionnelle bien spécifique parmi un ensemble de possibilités d'acquisition ou d'apprentissage d'une ou de plusieurs langues. Ces différentes situations sont distinguées dans la Section 2.

2. L1/L2P/L2

Un locuteur ne peut développer une compétence « native », c'est-à-dire optimale, pour une ou plusieurs langues que dans certaines conditions. Nous sommes nombreux à avoir rencontré des difficultés plus ou moins importantes lorsque nous avons appris une langue dans le contexte scolaire, alors que par ailleurs nous utilisons notre langue « maternelle » sans efforts particuliers. Un certain nombre de facteurs expliquent cette différence entre ce que nous appellerons la langue première (L1) et les autres langues acquises ou apprises tout au long de la vie, de façon précoce (L2P) ou plus tardive (L2). Quatre de ces facteurs sont explicités dans les sections suivantes et récapitulés dans le Tableau 1 ci-après : le contexte de transmission des informations linguistiques, la modalité de leur transmission, le type de mémoire impliqué dans la rétention et le traitement de ces informations, et l'âge d'exposition initial à un environnement linguistique.

2.1 Contextes de transmission : implicite vs explicite

La synthèse de l'Organisation internationale de la Francophonie (2022) indique que la répartition des locuteurs francophones à travers le monde et la façon dont le français a été acquis/appris par ces différents locuteurs ont beaucoup évolué au fil des années. Ainsi, le français ne représente plus la langue première (L1) pour de nombreux locuteurs francophones à travers le monde. Cela signifie que la transmission des informations linguistiques pour de nombreux locuteurs s'est

faite de façon largement explicite (scolaire), en général après l'âge de 6 ans, âge auquel l'enfant connaissait et pratiquait déjà au moins une autre langue sur une base quotidienne. Dans ce cadre, le français est alors qualifié de langue seconde (L2) et est considéré comme « appris ».

Les travaux de recherche rapportés dans le Chapitre II porteront essentiellement sur le développement du français lorsque cette langue est considérée comme « acquise », ce qui correspond à une situation d'immersion précoce et quotidienne dans un cadre familial, préscolaire (avant l'âge de 6 ans). Le mode de transmission des informations linguistiques est alors implicite et le français acquis représente une langue première (L1) pour l'enfant. Lorsque le français est la seule langue acquise, il s'agit de monolinguisme, mais les études actuelles montrent que cette configuration diminue face à des situations de plus en plus fréquentes de bilinguisme (2L1), c'est-à-dire deux langues acquises précocement dans le cadre familial, préscolaire (en général, lorsque les deux parents ont chacun une L1 différente). Enfin, il existe également un cas de figure intermédiaire entre les cas de L1 et L2 décrits ci-dessus. Il s'agit de la L2 précoce (L2P), lorsque l'enfant acquiert une ou plusieurs L1 dans le cadre familial et qu'il est exposé, également de façon précoce (avant l'âge de six ans), à une langue en dehors du foyer, en général la langue du pays dans lequel vit la famille, par l'intermédiaire de sorties quotidiennes, de garde en crèche et de scolarisation en maternelle (de Houwer, 2019). Ce cas de figure sera étudié dans la Section IV.2 du Chapitre II.

2.2 Modalités : orale vs écrite

La distinction entre modalité orale et modalité écrite est également fondamentale dans le domaine de l'acquisition du langage. Il existe actuellement entre 5 000 et 7 000 langues à travers le monde. Ces langues sont majoritairement des systèmes linguistiques avec une modalité orale et parfois avec une modalité signée (Jaber, Donati & Geraci, 2022). Dans les deux cas, il s'agit de la première modalité à laquelle l'enfant est exposé et il s'agit donc de la modalité qui déclenche les processus d'acquisition L1 et L2P mentionnés dans la Section 2.1 ci-dessus.

Dans la chronologie du développement linguistique de l'enfant, la modalité écrite apparaît toujours après la modalité orale ou signée (comme dans l'histoire de l'évolution du langage et des langues, l'oralité a émergé bien avant l'écriture).⁴ De plus, à la différence de la compréhension et de la production orales d'une langue, les équivalents écrits (lecture et écriture) se développent par apprentissage, c'est-à-dire grâce à une transmission explicite (en général scolaire) des informations nécessaires à la maîtrise du système graphique. Par ailleurs, la maîtrise de l'écrit ne vient pas systématiquement compléter la maîtrise du système oral. Hilton (2022, p. 79) par exemple rapporte que 14 % des adultes de plus de 18 ans à travers le monde n'ont pas appris une modalité écrite. La modalité orale (ou signée) est donc la modalité initiale, commune à tous les locuteurs dans le processus de l'acquisition du langage (hors pathologies et circonstances exceptionnelles), alors que la modalité écrite est subséquente et non-systématique. Nous retrouverons cette distinction lorsque nous évoquerons l'hypothèse de la diglossie dans la Section IV du Chapitre II.

2.3 Types de mémoire : non-déclarative vs déclarative

En termes de types de mémoires impliqués dans le traitement et la rétention des informations linguistiques, la distinction entre composantes de forme (phonologie, morphologie et syntaxe) et composantes de contenu (sémantique, lexique ; cf. Section 1 ci-dessus) se superpose avec la distinction entre L1 et L2P d'une part vs L2 d'autre part (Paradis, 2009 ; cf. récapitulatif dans le Tableau 1 ci-dessous). En effet, le modèle d'Ullman (2016) notamment propose que ces deux dichotomies proviennent de la distinction entre la mémoire non-déclarative (celle des procédures et des généralisations) et la mémoire déclarative (celle des connaissances et des souvenirs ; selon le modèle de Squire, 1992). Très schématiquement, l'acquisition L1-L2P des composantes de forme est plutôt attribuée à la mémoire non-déclarative (détection et mémorisation des régularités

⁴ La publication récente de Boë et al. (2019), grâce à des recherches pluridisciplinaires entre linguistes, bio-acousticiens et primatologues, vient de reculer la datation de l'émergence de la parole chez l'être humain (sous forme de protolangage vocalique) de 200 000 à environ 20 millions d'années. L'émergence de l'écriture est plus consensuellement estimée à il y a environ 5 000 ans.

et des dépendances présentes dans l'environnement linguistique mentionnées en Section I.1.1), alors que l'apprentissage de la L2 dans son ensemble, ainsi que celui du lexique et du système graphique quelle que soit la langue sont plutôt attribués à la mémoire déclarative (traitement et mémorisation de connaissances explicites).

Par ailleurs, les recherches les plus récentes montrent un impact différent de l'âge puis du vieillissement sur ces deux types de mémoires. Par exemple, Tóth-Fáber, Nemeth & Janacsek (2023) rapportent que l'âge n'affecte pas des aspects de la mémoire non-déclarative, notamment la capacité de traitement statistique des informations linguistiques, déjà bien développée dès les premières années de vie (Finn, Kharitonova, Holtby & Sheridan, 2019 ; Jung, Walther & Finn, 2021 ; Forest, Schlichting, Duncan & Finn, 2023). *A contrario*, les capacités d'apprentissage sous-tendues par la mémoire déclarative se développent progressivement chez l'enfant puis l'adolescent, avant d'atteindre un plateau et de décliner à partir de l'âge adulte (Finn et al., 2016 ; Ullman, 2020). Les informations linguistiques ne sont donc pas traitées de façon identique tout au long de la vie et l'âge initial d'exposition apparaît alors comme un facteur crucial dans la typologie des situations d'acquisition ou d'apprentissage.

2.4 Âges initiaux d'exposition : la question de la période critique

La partition entre mémoires non-déclarative et déclarative rejoint ainsi la question de l'âge initial d'exposition à un système linguistique : les composantes de forme sont *acquises* (mémoire non-déclarative) en L1 grâce à une exposition précoce, mais *appries* (mémoire déclarative) en L2 par une exposition plus tardive (cf. Tableau 1 ci-dessous). En amont des résultats les plus récents en neurolinguistique, de nombreuses recherches avaient montré que le traitement de certaines informations linguistiques se fait de façon différenciée en fonction de l'âge (Johnson & Newport, 1989 ; Eubank & Gregg, 1999 ; Paradis, 2004 ; Gerken, Quam & Goffman, 2019, entre autres). Ainsi, il existe des périodes optimales de traitement, appelées « périodes critiques », notamment pour la phonologie (avant l'âge d'un an ; cf. Werker, 1994 ; Werker & Hensch, 2015) et la morphosyntaxe

(avant l'âge de quatre-six ans ; cf. Meisel, 2009 ; Veríssimo, Heyer, Jacob & Clahsen, 2018 ; mais cf. Hartshorne, Tenenbaum & Pinker, 2018 qui proposent un seuil critique bien plus tardif pour la syntaxe à 17,4 ans, corroboré par les tout récents résultats de Menks et al., 2024 qui rapportent un changement significatif dans le traitement neuronal d'informations grammaticales lors d'une exposition à un système nouveau à 15,4 ans). En revanche, il est généralement proposé qu'il n'y a pas de période critique pour la composante sémantique/lexicale (Slabakova, 2006). Ces différentes trajectoires développementales sont récapitulées dans le Tableau 1 ci-après.

Tableau 1 : trajectoires développementales

	Acquisition	Apprentissage
Phonologie	L1/L2P	L2
Morphologie	L1/L2P	L2
Syntaxe	L1/L2P	L2
Lexique		L1/L2P/L2
Système écrit		L1/L2P/L2
Mémoire	Non-déclarative	Déclarative
Période critique	Oui	Non

Les travaux rapportés dans le Chapitre II seront principalement consacrés à l'acquisition de la morphosyntaxe entre deux et six ans en L1 et en L2P. La Section 2 ci-dessus a explicité les raisons pour lesquelles le terme d'« acquisition » est utilisé et distingué de celui d'« apprentissage » dans ces travaux. La Section 3 va à présent détailler comment les linguistes, psychologues, psycholinguistes et psychologues cognitifs ont formalisé les processus qui sous-tendent l'acquisition de la L1 chez le jeune enfant et montrer pourquoi des modèles différents continuent à coexister dans le paysage de la recherche en acquisition du langage.

3. Modèles de l'acquisition du langage

C'est surtout à partir des années 1930 que des scientifiques ont commencé à s'intéresser à la description des processus de l'acquisition de la L1. Ainsi, Bloomfield (1933), linguiste distributionnaliste, et Skinner (1957), psychologue comportementaliste, ont décrit le langage et son développement dans le cadre de la psychologie comportementale, modèle prégnant aux États-Unis depuis les

années 1920 et basé sur le conditionnement environnemental (Watson, 1924). Dans ce cadre, tout apprentissage est considéré comme relevant uniquement d'apports de l'environnement, de l'expérience (ce qui rejoint l'idée bien connue et plus ancienne de l'absence de préexistence de connaissances chez le jeune enfant).⁵ L'acquisition de la L1 est alors perçue comme l'apprentissage d'« habitudes de production et de réponses au cours des premiers années de vie [...] renforcées par la répétition » (Bloomfield, 1933, p. 31).

L'entrée en scène du linguiste Noam Chomsky (1957, 1959), avec la Grammaire générative et sa critique très dure de l'ouvrage de Skinner (1957), va constituer le début d'une opposition scientifique forte entre « empiristes » et « innéistes ». Cette opposition est connue sous le nom de « première révolution cognitive ». Chomsky va en effet introduire l'idée que l'acquisition du langage n'est pas entièrement déterminée par les apports extérieurs, environnementaux, mais qu'elle repose en partie sur des apports internes, biologiques. La « seconde révolution cognitive » verra le jour dans les années 1970 lorsque des scientifiques s'opposeront, de nouveau fortement, aux aspects innéistes introduits par la Grammaire générative. On assiste alors à la naissance de la Linguistique cognitive, avec des modèles nouveaux de grammaires (dont la Grammaire cognitive de Langacker, 1987, 1988 et la Grammaire de construction de Goldberg, 1995) et une approche de l'acquisition basée sur l'usage et en lien avec la cognition sociale (à partir notamment des travaux de Michael Tomasello, 1992b, psychologue cognitif).

Après deux « révolutions cognitives » et malgré soixante-dix ans de recherches tant théoriques qu'expérimentales (cf. l'étude princeps de Crain & Nakayama, 1987), l'opposition scientifique entre empiristes et innéistes perdure et semble même plus active que jamais. La littérature rapporte encore à l'heure actuelle un « paradoxe interdisciplinaire » dans les théorisations génératives qui, malgré des déclarations d'intention de tendre vers l'interdisciplinarité, restent

⁵ Cf. par exemple les propos d'Aristote ([1994]) et de Locke (1690) opposés à ceux de Platon ([1999]), et les termes de « tabula rasa », « ardoise vierge », « coquille vide ».

quelque peu opaques pour les non-spécialistes (Leivada, 2020). Dans le domaine spécifique de l'acquisition du langage, le numéro spécial de la revue *Journal of Child Language* (Ambridge & Wagner, 2021) illustre la vivacité actuelle du débat entre ces différentes approches. Les sections qui suivent entrent dans le détail de quelques concepts qui fondent l'approche innéiste (Section 3.1), l'approche basée sur l'usage (Section 3.2) et une approche intermédiaire, plus récente, qui combine les deux (Section 3.4).

3.1 Approche innéiste

La recension très sévère de l'ouvrage de Skinner (1957) par Chomsky (1959) marque une rupture nette avec l'approche comportementaliste du langage et représente le point de départ des recherches « innéistes » sur l'acquisition du langage.⁶ Dans les années qui suivent, Chomsky (1965, 1968, 1975, 1981), pour ne citer que quelques-unes de ses très nombreuses publications, va proposer une formalisation du traitement du langage et de son acquisition basée sur le fait qu'il considère que l'être humain est génétiquement programmé pour parler, comme il l'est pour voir, et que cet apport biologique, commun à tous, permet un développement du langage considéré comme rapide en dépit d'un environnement linguistique décrit comme qualitativement insuffisant (l'argument de la « Pauvreté du stimulus » ; cf. Belletti, 2017 entre autres, et Section 3.4 ci-après sur un étayage statistique de cette notion). Cette « Faculté de langage », spécifiquement dédiée au traitement du langage dans les premiers modèles, se présente sous un état initial, commun à tous, en amont de toute exposition linguistique, appelé « Grammaire Universelle » (GU). L'ambition de la Grammaire générative est d'explicitier la compétence linguistique implicite du locuteur. Dans ce cadre, les générativistes utilisent un formalisme très abstrait, qui postule notamment que les phrases disposent d'une structure sous-jacente hiérarchisée, que le cerveau génère cette structure initiale à laquelle il applique ensuite des modifications pour générer la structure de surface effectivement produite

⁶ Chomsky (1959, p. 27) : "The magnitude of the failure of this attempt to account for verbal behaviour serves as a kind of measure of the importance of the factors omitted from consideration, and an indication of how little is really known about this remarkably complex phenomenon."

(hypothèse étayée depuis par van der Kant et al., 2020 et d'autres travaux mentionnés en Section 1.1.1 ci-dessus). La Grammaire générative défend également l'hypothèse de « la continuité de la grammaire », c'est-à-dire que les jeunes enfants démarrent leur acquisition en manipulant les mêmes entités abstraites que les adultes (catégories grammaticales, syntagmes, structures syntaxiques, déplacements de constituants, etc.).

L'acquisition du langage a été notamment formalisée par un modèle qui a suscité beaucoup de travaux depuis les années 1980 : « Principes et Paramètres ». Les Principes sont les contraintes communes à toutes les langues du monde ; par exemple, la « dépendance à la structure » stipule que les opérations grammaticales s'appliquent aux syntagmes hiérarchisés dans la structure sous-jacente, et non pas aux constituants ordonnés linéairement dans la réalisation de surface (Chomsky, 1975 ; Crain & Nakayama, 1987 ; Yang, Crain, Berwick, Chomsky & Bolhuis, 2017). Les Paramètres sont les points de variation binaires entre les langues du monde ; par exemple, les langues du monde sont soit à « sujet nul » comme l'italien et le grec, soit à « sujet non-nul » comme l'anglais et le français (cf. par exemple Rizzi, 1982, 1997 et des synthèses dans Camacho, 2013 ; D'Alessandro, 2015). Formulés initialement comme résultant d'un regroupement de traits, ces « macro-paramètres » ont ensuite été progressivement détaillés pour devenir des « micro-paramètres » hiérarchisés les uns par rapport aux autres (cf. par exemple Baker, 1996, 2005 ; Lahousse, 2022 sur le français).

Les modèles proposés depuis les années 2000 ont ensuite évolué dans leur formalisme et sur certaines propositions de fond avec une volonté d'être plus bio-réalistes, notamment sur la part spécifiquement dédiée au langage, qui à partir de Chomsky (2005) est considérablement réduite au profit de « principes non spécifiques à la faculté de langage », tels que le traitement statistique des données environnementales et l'utilisation de propriétés cognitives plus générales (Yang et al., 2017). L'approche bilinguistique de la Grammaire générative trouve alors des partisans en neurosciences cognitives (cf. Di Sciullo & Boeckx, 2011 ; Berwick, Friederici, Chomsky & Bolhuis, 2013, entres autres). L'implémentation neuronale

de la hiérarchisation syntaxique fait ainsi l'objet de nombreux travaux (cf. Goucha et al., 2017 pour une revue, et les autres références mentionnées en I.1.1 ci-dessus) et les Principes et les Paramètres sont décrits dans le cadre des sciences cognitives. Dehaene (2018, p. 64 et 86) par exemple décrit les Principes comme « un jeu d'hypothèses abstraites héritées de [l'] évolution [du cerveau], qu'il projette sur le monde extérieur. [...]. Apprendre, c'est partir d'un jeu d'hypothèses *a priori* » et les Paramètres comme « apprendre consiste à sélectionner, parmi toutes les grammaires abstraites, celle qui s'ajuste le mieux aux données. [...] c'est un excellent modèle de ce que fait l'enfant : scientifique en herbe, il formule des théories et les confronte au monde extérieur. »

Néanmoins, certains générativistes ont proposé de réduire le nombre de ces paramètres. Longobardi (2018) par exemple propose un total de neuf types de paramètres (« *schemata* ») au lieu des plusieurs dizaines de paramètres initialement discutés (cf. Ayoun, 2003 pour une vue d'ensemble). Le Paramètre du Sujet nul peut alors être intégré dans le type « grammaticalisation » ou « catégories nulles » et le Paramètre *wh-* dans « force ». ⁷ D'autres linguistes proposent d'abandonner le système des paramètres en raison de leurs redondances et de la surgénération de possibilités par rapport à la variation effectivement observée dans les langues du monde (Boeckx & Leivada, 2013 ; Lightfoot, 2017 ; Leivada & Murphy, 2022). L'absence de paramètres devient alors un point commun de ces nouveaux modèles basés sur des notions de la Grammaire générative et une toute autre approche « basée sur l'usage ».

3.2 Approche basée sur l'usage

La quintessence de l'opposition entre innéistes et empiristes est parfaitement résumée dans les propos de Gopnik & Bonawitz (2015) :

« Comment pouvons-nous acquérir des représentations abstraites à partir de données concrètes ? La réponse traditionnelle des innéistes consiste à dire que l'apprentissage est une illusion. Nous naissons

⁷ Ces deux paramètres seront détaillés dans les Sections II.2.4 et 2.6, puis dans le Chapitre II.

équipés avec des représentations abstraites. La réponse opposée des empiristes dit que l'abstraction est une illusion. La connaissance est tout simplement issue d'associations effectuées sur la base d'un traitement statistique de notre environnement.⁸ »

Le modèle basé sur l'usage de Tomasello (1992a, 2000, 2003, 2006), pour ne citer que quelques-unes de ses publications, se situe aux antipodes du modèle innéiste sur un certain nombre de postulats fondateurs de chacune des deux approches. Le formalisme est radicalement différent et ne recourt ainsi à aucune abstraction. Il n'y a pas de structures hiérarchisées, pas de sous-jacence, pas de catégories grammaticales initiales, pas de déplacements, pas de Paramètres, etc. et le modèle postule une discontinuité entre la grammaire de l'enfant et celle de l'adulte (« les toutes premières représentations linguistiques des enfants ne sont pas des catégories et des règles semblables à celles des adultes », Ambridge & Lieven, 2015, p. 478). Par ailleurs, le langage est considéré comme similaire à d'autres capacités cognitives comme l'activité motrice, et il est géré par des principes généraux de traitement (catégorisation, mémorisation, etc.). Comprendre une phrase, c'est traiter linéairement une séquence d'informations.

Dans ce cadre, Tomasello (2003, p. 9) décrit l'acquisition du langage par les enfants comme l'acquisition des « dizaines de milliers de conventions linguistiques utilisées par leur entourage ». Entre un et deux ans, l'enfant utilise essentiellement des « structures figées lexicalisées » sans syntaxe (appelées également « constructions » ; par exemple, « je+veux+ça » ou « le+chat » sont considérés comme des blocs inanalysés qui associent un sens à une forme sonore). Il n'y a pas de contrainte de taille sur ces constructions ; elles peuvent ainsi varier de l'équivalent d'un morphème à celui d'une phrase complète. Puis, entre deux et trois ans, l'enfant exploite des gabarits dont une partie devient progressivement abstraite en termes de catégories sémantiques, pas grammaticales (par exemple,

⁸ Texte original : « How can we learn abstract representations from concrete data? The classic nativist response to this question is to say that learning is an illusion. We come equipped with abstract representations. The alternative empiricist response is to say that abstraction is an illusion. Knowledge is just a collection of associations derived from the statistics of our environment. »

« je+ACTION+ça », cf. Pine & Lieven, 1997 sur le modèle « slot and frame »). Les facteurs déterminants qui permettent d’aller vers l’abstraction sont la fréquence dans l’environnement linguistique et l’analogie (Theakston & Lieven, 2017 ; Ambridge, 2020). L’acquisition est décrite comme « lente et fragmentaire » et les catégories et fonctions grammaticales adultes émergent progressivement à partir de l’âge de trois ans environ.

3.3 Questions oui/non au cœur du débat

L’approche basée sur l’usage diverge ainsi fortement de l’approche innéiste sur un nombre important de points fondamentaux concernant les mécanismes qui soutiennent l’acquisition du langage et sur la place que l’environnement linguistique occupe dans le processus. Alors que les innéistes postulent que l’apport linguistique est qualitativement insuffisant au regard de ce que l’enfant est capable de produire (l’argument de la Pauvreté du stimulus mentionné dans la Section 3.1 ci-dessus), les constructivistes le considèrent comme fondateur.

L’acquisition d’une structure syntaxique est alors particulièrement débattue entre les deux approches : les questions oui/non à double auxiliaire en anglais (*Is the man [who is tall] is_i in the room ?* est l’un des exemples initiaux de Chomsky, 1975).⁹ Cette structure est emblématique pour au moins trois raisons, décrites en (1), qui semblent représenter un paradoxe acquisitionnel.

(1) Paradoxe des questions oui/non à double auxiliaire :

- i. Cette structure est rare dans l’environnement linguistique de l’enfant : Legate & Yang (2002) par exemple ne rapportent aucune question oui/non à double auxiliaire dans un corpus de 29 540 questions adressées à deux jeunes enfants et uniquement quelques rares questions *wh-* complexes (0,06 % des questions ; par exemple *What_j is_i the animal [that says cockadoodledoo] is_i what_j ?* « quel est l’animal qui dit cocorico ? »).

⁹ « L’homme [qui est grand] est-il dans la pièce ? » ; les crochets à l’intérieur de la phrase indiquent la présence d’une proposition enchâssée (ici, une subordonnée relative) ; les items barrés et indicés indiquent des déplacements selon les conventions de la Théorie des copies en Grammaire générative.

- ii. En revanche, une structure fréquente dans l'environnement linguistique des jeunes enfants est celle des questions oui/non simples (par exemple : *Can_i you ~~can~~_i eat ?* « peux-tu manger ? »). Cette structure simple semble indiquer que pour transformer une phrase déclarative en phrase interrogative, il faut déplacer le premier verbe de la déclarative vers la première position dans l'interrogative. Cette règle transformationnelle fait appel à des connaissances grammaticales (la catégorie « verbe ») et à la linéarité des items dans la phrase (« premier »).
- iii. Pourtant, des résultats expérimentaux ont montré que des enfants natifs entre trois et cinq ans ne produisent *jamais* une question avec déplacement du premier verbe dans des cas de questions complexes (**Is_i the man [who is_i tall] is in the room ?* ; Crain & Nakayama, 1987 ; Crain, Koring & Thornton, 2017).

Ces résultats expérimentaux alimentent ainsi les débats entre les deux approches. Les générativistes postulent que les enfants appliquent une règle de déplacement dépendante de la structure syntaxique sous-jacente (faisant partie des Principes mentionnés en Section 3.1 ci-dessus). Le constituant déplacé est alors le verbe de la proposition principale et non celui de la proposition subordonnée (« structure trumps linear order », Chomsky, Gallego & Ott, 2019). Aux antipodes de cette approche innéiste, les constructivistes postulent que les enfants produisent des questions oui/non par l'utilisation d'un nombre réduit de gabarits, par exemple *Can I ACTION ?*, sans l'application de règles abstraites de déplacement (Dabrowska, 2000) et discutent donc les conclusions des générativistes sur la prééminence de la structure syntaxique. Ambridge, Rowland & Pine (2008) notamment relient l'absence d'erreurs de type linéaire (**Is_i the man [who is_i tall] is in the room ?*) avec l'absence de séquences dans l'environnement linguistique de l'enfant du type *who tall*. En effet, des questions qui, bien qu'agrammaticales, créent des séquences valides dans la proposition relative (**Can_i babies [who ~~can~~_i crawl] can walk ?*) sont produites dans leur expérimentation (N = 16, 4,55 % de l'ensemble des réponses). Les auteurs contestent ainsi les conclusions des générativistes et évoquent les capacités des

enfants à traiter les probabilités transitionnelles (propriétés distributionnelles de surface) dans leur environnement linguistique (cf. Section 1.1 ci-dessus).

Ainsi, les deux approches ne semblent pas réconciliables quant à la formalisation des processus qui sous-tendent l'acquisition de la L1. En revanche, des rapprochements se sont opérés ces dernières années autour de la question de l'importance des données environnementales (Waterfall, Sandbank, Onnis & Edelman, 2010) et de l'utilisation de modèles de traitement du langage naturel (Yedetore, Linzen, Frank & McCoy, 2023). Ces approches sont subsumées sous l'appellation « GU + statistiques » dans la Section 3.4.

3.4 Approche « GU + statistiques »

Très schématiquement, les critiques adressées à l'approche innéiste portent souvent sur sa déconnection des données linguistiques effectivement perçues et produites, les générativistes revendiquant clairement l'étude de la compétence linguistique sous-jacente du locuteur-auditeur idéal (cf. Section 3.1 ci-dessus). À l'inverse, les critiques adressées à l'approche basée sur l'usage portent souvent sur l'absence de créativité attribuée aux jeunes enfants, les constructivistes postulant clairement qu'ils réutilisent essentiellement le langage de leur environnement (cf. Section 3.2). Ces deux approches « classiques » représentent ainsi deux points de vue antagonistes entre lesquels d'autres hypothèses ont progressivement émergé depuis les années 2000. Les travaux de Yang (2004, p. 453) par exemple illustrent cette voie intermédiaire qui combine contraintes *a priori* et traitement des données linguistiques :

« [...] le cadre P&P [Principes & Paramètres] restera une composante indispensable dans toute théorie de l'acquisition des langues, mais à condition que le « triggering model » [...] soit abandonné au profit de théories modernes de l'apprentissage où l'apprentissage est probabiliste. »

Dans un premier temps, certains concepts fondateurs de la Grammaire générative sont mis à l'épreuve de données statistiques. Ainsi, Legate & Yang

(2002) examinent l'argument de la Pauvreté du stimulus en termes de fréquences de certaines structures dans des productions adultes et en termes d'âge d'acquisition chez les enfants (cf. également Section 3.3 ci-dessus sur l'acquisition des questions oui/non). Les auteurs mettent alors en avant des disparités statistiques majeures entre certaines formes peu fréquentes chez l'adulte qui sont pourtant acquises précocement (par exemple, le placement correct de la négation par rapport au verbe non-fléchi en français dès la deuxième année ; cf. Section II.2.3 puis Chapitre II) et d'autres formes plus fréquentes mais acquises plus tardivement (par exemple, la disparition tardive, vers trois-quatre ans, du phénomène du sujet nul enfantin dans les langues à sujet non-nul ; cf. Section II.1 du Chapitre II). Legate & Yang (2002, p. 161) arrivent alors à la conclusion que l'hypothèse d'un apprentissage basé uniquement sur un traitement statistique de l'environnement linguistique ne rend pas compte de la réalité des processus acquisitionnels et que c'est « une connaissance innée de GU » qui permet d'expliquer ces disparités distributionnelles importantes entre productions adultes et enfantines.¹⁰

Yang (2004, 2012), entre autres publications, propose également de revoir la théorie des Principes & Paramètres à la lumière de données statistiques. L'auteur présente un modèle de l'acquisition de la grammaire, certes basé sur des mécanismes d'apprentissage généraux (de type probabiliste) mais contraint par GU (« UG with statistics », p. 453). Ce modèle est « variationniste » car il postule une variation initiale, sous forme de compétition entre plusieurs grammaires disponibles concomitamment et initialement chez le jeune enfant (en raison notamment de données parfois contradictoires ; par exemple : « il ne faut pas » vs « faut pas » donnent des indications différentes en termes de présence vs absence du sujet et de réalisation de la négation). Le modèle propose ensuite une sélection qui entraîne la disparition progressive de la grammaire la moins représentée dans l'environnement linguistique. L'approche établit également une corrélation entre la quantité de données non-ambigües dans l'input de l'enfant et l'âge d'acquisition

¹⁰ Lidz & Gagliardi (2015) proposent également une discussion très intéressante et statistiquement étayée sur la pauvreté du stimulus et sur la disparité entre données adultes et enfantines.

de chaque paramètre. Ce modèle présente les avantages de rendre compte de l'aspect progressif du développement linguistique et de la présence concomitante de formes divergentes chez le jeune enfant (par exemple, sujets nuls et sujets non-nuls jusqu'à trois-quatre ans). Le modèle reste néanmoins innéiste dans la mesure où il postule que les grammaires initialement en concurrence sont toutes des grammaires de langues humaines, donc contraintes par GU.

Ce modèle variationniste n'est pas le seul à proposer une approche « GU + statistiques ». Pearl (2021) utilise ce terme pour désigner un ensemble de propositions qui reposent sur cette même approche. Ce sont des modèles cognitifs computationnels qui implémentent une stratégie d'apprentissage à partir des contraintes de GU et des probabilités basées sur des statistiques environnementales, en acquisition de la morphologie et de la syntaxe notamment. Pearl (2021) décrit chacun des modèles et évalue leurs prédictions à la lumière de données acquisitionnelles (entre autres sur les paramètres V2, sujet nul et position de l'item *wh-*) et les résultats sont généralement probants et prometteurs pour de futures investigations dans ce cadre.

Concernant le débat entre structure et linéarité soulevé par les questions oui/non et détaillé dans la Section 3.3 ci-dessus, Yedetore et al. (2023) ont évalué la capacité de deux modèles computationnels à fournir des questions oui/non complexes à partir d'un large corpus de données réelles produites à destination d'enfants (cf. Section II.1 ci-dessous sur le Langage adressé aux enfants). Après entraînement avec 10 870 paires de déclaratives/interrogatives, la structure la plus fréquemment fournie par les deux modèles s'avère agrammaticale et consiste à copier le premier verbe en tête de phrase et à effacer le verbe principal (par exemple, **Has_i the person [who has_i seen this boy] ~~did~~ try ?*), alors que la forme grammaticale nécessite de copier et effacer le verbe principal (*Did_i the person [who has seen this boy] ~~did~~_i try ?*). Les auteurs en concluent que les données seules ne suffisent pas pour reconstituer la règle dépendante de la structure syntaxique sous-jacente nécessaire à la formation des questions oui/non complexes et que leurs résultats « soutiennent l'hypothèse selon laquelle l'apprenant a besoin d'un

biais hiérarchique » pour formuler la règle correcte, rejoignant ainsi l'argument de la Pauvreté du stimulus posé par les générativistes il y a une cinquantaine d'années et d'autres travaux tout récents en neurosciences (Coopmans et al., 2022).

4. Positionnement scientifique

La Section I avait pour objectif de proposer une synthèse sur le vaste domaine de recherches que représente l'acquisition du langage, tant au niveau de la description de la trajectoire développementale typique de l'enfant et des différentes situations d'acquisition ou d'apprentissage jusqu'à l'âge d'environ six ans (Sections 1 et 2), que des différentes approches théoriques qui formalisent les processus qui sous-tendent le développement linguistique précoce (Section 3). La Section 4 présente ma position de chercheure dans ce paysage théorique.

Après un peu plus d'un siècle de recherches, l'acquisition du langage reste un domaine d'investigations foisonnant et débattu, notamment parce qu'il a intéressé au fil des années un nombre de spécialités de plus en plus grand et qu'il a également impulsé de nombreuses recherches pluridisciplinaires. Ainsi, des travaux en psychologie, linguistique, psycholinguistique, linguistique cognitive, neurolinguistique et linguistique computationnelle notamment jalonnent cette Section I. Pourtant, et malgré les apports les plus récents des neurosciences, des dissensions fortes continuent à exister concernant les aspects innés du traitement linguistique (cf. par exemple la Section I.3.3 ci-dessus sur les questions oui/non et le numéro spécial de la revue *Journal of Child Language* de Ambridge & Wagner, 2021).

Ma position dans ce débat est celle développée dans la Section I.3.4 « GU + statistiques ». Ces recherches m'ont convaincue que le langage humain se développe grâce aux capacités computationnelles fines de l'enfant pour traiter les informations linguistiques environnantes (notamment, le traitement statistique) et grâce à une forme ou une autre de contraintes/biais cognitifs qui guident l'enfant dans ce traitement. J'adhère ainsi notamment à la formulation de Yang

(2004, p. 455) qui prône une approche intégrative de l'acquisition du langage¹¹, d'autant plus que des travaux récents en neurosciences étayent cette approche (Gervain, 2015 ; Shi et al., 2020 ; Pearl, 2021 ; Sanchez-Alonso & Aslin, 2022 ; Yedetore et al., 2023, entre autres) et que la dichotomie entre inné vs acquis est même dorénavant parfois décrite comme « obsolète » (Marino & Gervain, 2019, p. 389).

Les travaux présentés dans le Chapitre II représentent ainsi la rencontre entre certains fondements théoriques de la Grammaire générative et un apport empirique grâce à une base de données enfantines collectée pendant trois ans dans une classe de maternelle en France.

II. Apport linguistique

Quel que soit le cadre théorique adopté, l'étude des caractéristiques de l'apport linguistique à l'enfant pendant la période d'acquisition (essentiellement de la période prénatale jusqu'à l'âge de six ans, cf. Sections I.1 et I.2 ci-dessus) représente la seconde question cruciale en recherche sur l'acquisition du langage. Sans surprise, les recherches sur cette problématique ont surtout émergé dans des cadres non-génératifs et les travaux en psycholinguistique de Susan Ervin (1964) sur la question du rôle de l'imitation dans la production infantile sont notamment considérés comme fondateurs. L'auteure a comparé deux conditions de production (spontanée vs imitée) chez des enfants de deux-trois ans et a conclu que « L'imitation dans des conditions optimales, celles d'un rappel immédiat, n'améliore pas la grammaire », les phrases présentant des omissions de morphèmes grammaticaux dans les deux conditions (p. 172).

Les recherches se sont ensuite attachées à décrire les caractéristiques linguistiques de l'environnement infantile sous l'appellation de « langage adressé

¹¹ « Both endowment and learning are important to language acquisition – and both linguists and psychologists can take comfort in this synthesis. » (Yang, 2004, p. 455)

à l'enfant » (LAE), son impact sur le processus acquisitionnel, ainsi que les facteurs non-linguistiques qui font que le LAE varie potentiellement d'un enfant à un autre. Ces points ont beaucoup été étudiés dans le cadre de l'acquisition de l'anglais et sont détaillés dans la Section 1 ci-après. La Section 2 entre ensuite dans le détail de l'apport linguistique dans le cadre de l'acquisition du français hexagonal.

1. LAE

De nombreuses études se sont penchées sur la question du langage adressé à l'enfant (LAE) depuis notamment Ferguson (1964) et Snow (1972). Il s'agit du langage utilisé par un adulte à l'adresse d'un enfant, que cet adulte soit la mère de l'enfant ou pas. Dans la littérature anglophone, différentes appellations sont utilisées en fonction des caractéristiques de l'enfant et/ou de l'adulte : *infant-directed speech* (IDS), *child-directed speech* (CDS), *caregiver speech*, *nursery talk*, *motherese*, *parentese*, etc. Dans un premier temps, les études se sont intéressées aux caractéristiques générales de ce moyen de communication et se sont interrogées sur sa fonction dans le processus acquisitionnel.

1.1 *Caractéristiques et fonction dans le processus acquisitionnel*

Toutes les études s'accordent sur le fait que le LAE présente des caractéristiques phonologiques, morphologiques, lexicales et syntaxiques différentes du langage adressé par un adulte à un autre adulte. Newport, Gleitman & Gleitman (1977) rapportent une articulation plus claire, une intonation exagérée, un vocabulaire restreint, des thématiques qui portent sur le « ici et maintenant », des énoncés plus courts, beaucoup plus de phrases interrogatives et impératives et beaucoup moins de phrases déclaratives. Veneziano (2000) note moins de formes verbales composées et moins de pronoms de troisième personne.

Intuitivement, on pourrait penser que la fonction du LAE est de faciliter l'acquisition du langage en proposant une version simplifiée de la langue environnante au jeune enfant. Newport et al. (1977) ont réfuté cette hypothèse (« Motherese is not a syntax-teaching language », p. 124) et ont proposé que la forme spécifique du LAE aurait plutôt pour fonction de communiquer de façon

efficace et rapide pour transmettre des instructions parfois vitales en cas de danger immédiat pour le tout jeune enfant. Cette analyse est connue sous le nom de « l'hypothèse communicative » (Pine, 1994, Veneziano, 2000). Les caractéristiques spécifiques du LAE ne sont donc pas analysées comme ayant une vocation pédagogique. Les énoncés ne deviennent pas plus complexes avec l'âge notamment. Ce qui évolue le plus et qui permet de prévoir l'âge de l'enfant en fonction du LAE sont les variables liées au contexte : par exemple, à partir de dix-huit/vingt-quatre mois, on constate une augmentation de la fréquence des noms d'objets absents du contexte perceptif (Khomsî, 1982). Dans ce cadre, sa forme est alors plutôt considérée comme relevant d'une adaptation (inconsciente, peut-être instinctive) de l'adulte aux capacités de traitement et d'attention plus limitées chez les jeunes enfants que chez les adultes.¹²

1.2 Facteurs sociolinguistiques de variation

Dans un second temps, de nombreux travaux se sont intéressés à la possible variabilité qualitative et quantitative du LAE en fonction de facteurs sociolinguistiques. Le LAE a ainsi souvent été examiné en lien avec la catégorie socio-professionnelle (CSP) de la famille (professions, revenus et diplômes des parents). Il est alors très intéressant de noter que les résultats sont consensuels en ce qui concerne le développement du lexique, mais qu'ils semblent diverger sur le développement de la syntaxe. Rappelons que le lexique est appris (connaissances déclaratives), alors que la syntaxe est acquise (processus non-déclaratifs ; cf. Tableau 1 en fin de Section 2 ci-dessus).

L'apprentissage du lexique est ainsi généralement décrit comme étant fortement contraint par l'apport environnemental et ce dès la première année de l'enfant (Bergelson & Aslin, 2017 ; Gurgand et al., 2023). Les études bien connues de Hart & Risley (1995, 2003) rapportent des correspondances avec la CSP de la

¹² L'empan verbal, c'est-à-dire la capacité à rappeler une liste donnée de mots, évolue avec l'âge. Vers trois ans et demi, l'empan est d'environ trois mots (Gillet, Hommet & Billard, 2000 p. 164). À l'époque des premiers travaux sur le LAE, on pensait que l'empan adulte était de sept items, plus ou moins deux (G. A. Miller, 1956), mais cette capacité a été revue à la baisse. Actuellement, on trouve le chiffre de quatre, plus ou moins un, avec une discussion sur le développement d'autres capacités cognitives, telles que l'attention (Cowan, 2001).

famille et une réanalyse récente de leurs données pointe vers l'importance plus spécifique du niveau d'instruction des parents (Rindermann & Baumeister, 2015). Le niveau de diplomation de la mère a également été relevé comme une variable significative dans un autre domaine qui relève des connaissances déclaratives, l'apprentissage de la lecture (Negro & Genelot, 2009). La correspondance avec la CSP est également mentionnée concernant le développement de la syntaxe entre trois et cinq ans, avec des enfants de CSP élevées produisant des énoncés plus complexes que des enfants de CSP moins élevées (Levine et al., 2020). Huttenlocher, Vasilyeva, Cymerman & Levine (2002) établissent un lien plus général entre environnement linguistique et développement syntaxique vers l'âge de cinq ans dans des tâches de compréhension et de production de phrases complexes, leurs résultats montrant une correspondance non seulement avec la production familiale mais également avec l'apport scolaire. Certaines études rapportent également des interactions entre la CSP, la taille du vocabulaire et les effets d'une exposition à certaines structures syntaxiques. Luo, Masek, Alper & Hirsh-Pasek (2022) rapportent par exemple que le bénéfice de l'exposition à des questions *wh-* est moins élevé chez des enfants avec un inventaire lexical plus réduit et que cet effet est plus fort pour les CSP les moins élevées. Bates & Pearl (2019) ont également examiné les structures interrogatives à destination d'enfants entre trois et cinq ans provenant de différentes CSP. Contrairement aux résultats précédents, cette étude rapporte que ces données sont quantitativement et qualitativement similaires. Les résultats semblent donc moins tranchés concernant l'acquisition de la syntaxe.

Ces quelques exemples n'épuisent pas la liste des variables sociolinguistiques dans l'environnement des enfants, mais ils attirent l'attention sur leur importance et sur toute la difficulté à cerner la dimension multifactorielle de l'environnement. De plus, la recherche s'est intéressée à de nouveaux facteurs ces dernières années, notamment la variation interindividuelle (Dąbrowska, 2018 ; Turker, Seither-Preisler & Reiterer, 2021) et l'impact de l'environnement scolaire précoce (Dai et

al., 2022).¹³ Par ailleurs et de toute évidence, une autre variable cruciale est la langue (ou les langues) présente(s) dans l'environnement du jeune enfant. Des premiers cris du nouveau-né à la connectivité neuronale chez l'adulte, de nombreuses études montrent des données différenciées en fonction de la langue environnante (par exemple, Mampe, Friederici, Christophe & Wermke, 2009 et Wei et al., 2023). La Section 2 va donc décrire l'environnement linguistique particulier des jeunes enfants francophones en France.

2. Français hexagonal

D'après le rapport de l'Organisation internationale de la Francophonie (2022), les « locuteurs quotidiens de français » se répartissent sur les cinq continents, essentiellement dans des pays de l'Afrique subsaharienne et de l'Océan Indien (47,4 %) et de l'Europe (31,2 %). Cette répartition géographique entraîne inévitablement une diversité des systèmes linguistiques appelés de façon générique « français ». Ainsi, même dans des pays limitrophes de la France tels la Belgique et la Suisse, la langue pratiquée et acquise dans chacun de ces pays présente des différences phonologiques, lexicales et morphosyntaxiques bien documentées (Ashby, 1988 ; Walter, 1988 ; De Cat, 2005). Le terme de « français européen » parfois utilisé (Culbertson, 2010 ; Fonseca-Greber & Waugh, 2003 ; Martineau & Mougeon, 2003) ne semble donc pas adopter une granularité suffisamment fine par rapport aux questions d'acquisition et il convient de considérer chaque variété de français séparément afin de neutraliser cette variation diatopique (cf. Gadet, 2003 sur les facteurs de variation en sociolinguistique). La Section 2 portera donc principalement sur le français utilisé et acquis en France continentale que nous appellerons « français hexagonal ».

Le français hexagonal peut présenter des variations d'usage importantes en fonction du registre plus ou moins formel adopté par le locuteur (diaphasie), qui souvent mais pas systématiquement se superpose avec la modalité orale ou écrite (diamésie). Ces deux facteurs de variation dans l'usage sont eux-mêmes

¹³ Avec notamment en France, le projet DyLNet qui vise à étudier les liens entre développement linguistique et socialisation à la maternelle (<https://dylnet.suniv-grenobsle-alpes.fr/en>).

dépendants d'un certain nombre de variables liées à l'usager, telles que son âge, son groupe social et son niveau d'instruction mentionnés dans la section sur le LAE (diastratie). Ces nombreux facteurs se projettent dans le français oral hexagonal contemporain et génèrent une morphosyntaxe très diverse. Les exemples (2) rapportent trois points de variation qui font l'objet de très nombreuses études en (socio-)linguistique.

(2) Points de variation cruciaux en français oral hexagonal :

- i. L'utilisation d'un pronom nominatif (*je, tu, il, etc.*) en cooccurrence avec un syntagme nominal : « Ne dites pas *Mon père, il est malade*, au lieu de *Mon père est malade* », préconisait déjà Chifflet (1659) dans son *Essay d'une parfaite grammaire de la langue française*. Cette cooccurrence soulève la question très débattue du statut morphosyntaxique du pronom et interroge sur la fonction sujet des items cooccurrents.
- ii. La production de la négation simple postverbale (par exemple : *pas*), au lieu de la négation complète discontinuée (*ne... pas*).
- iii. La variation dans la position des items interrogatifs, avec une surutilisation de la position in-situ au détriment de la position ex-situ qui est potentiellement accompagnée d'une inversion du verbe et du sujet (par exemple : *Tu fais quoi ?* vs *Que fais-tu ?*).

C'est à partir des années 1970-1980 que les (socio-)linguistes ont commencé à examiner ces phénomènes de variation en français oral, à décrire un registre « non-standard » et à s'interroger sur le statut de ce registre par rapport au registre « standard » généralement associé à l'écrit et au « bon usage » (Blanche-Benveniste, 1983, 1991, 1994, 2000 ; Cappeau, 2021 ; Lambrecht, 1981 ; Zribi-Hertz, 1994). Les Sections 2.1 à 2.6 ci-après détaillent les phénomènes linguistiques présentés en (2) et documentent leur utilisation en LAE. La question de la formalisation de cette variation sera traitée dans la Section 3 ci-dessous.

2.1 Doublement vs dislocation

Parmi tous les phénomènes qui présentent des divergences entre le français standard et le français oral, la production cooccurrence d'un syntagme nominal et d'un pronom nominatif semble représenter la structure la plus étudiée. Plusieurs raisons expliquent cet intérêt fort et persistant de la part des grammairiens et des (socio-)linguistes (Ashby, 1980, 1984 ; Blanche-Benveniste, 2000 ; Cabredo Hofherr, 2004 ; Culbertson, 2010 ; Hulk, 1991 ; Kayne, 1975 ; Lambrecht, 1981 ; Liang, 2023 ; Morin, 1979 ; Zribi-Hertz, 1994, pour ne citer que quelques-uns des très nombreux travaux). Tout d'abord, cette cooccurrence, souvent appelée « dislocation », est présente en français depuis le XVII^e siècle, tant à l'oral que dans la littérature. Le syntagme nominal est « disloqué », c'est-à-dire qu'il ne fait pas partie de la valence verbale ; il n'est donc pas sujet du verbe et c'est le pronom qui endosse cette fonction. La dislocation est donc généralement admise, *modulo* quelques commentaires comme celui de Chifflet (1659) mentionné en (i), avec pour fonction une mise en relief du syntagme nominal.

La configuration qui suscite toutes les polémiques est celle du « (re)doublement » qui associe également les deux éléments mais sans dislocation du syntagme nominal. Cabredo Hofherr (2004) par exemple caractérise le doublement (opposé à la dislocation) sur la base d'une absence de mise en relief de l'élément doublé, la possibilité que l'élément doublé soit indéfini (*Personne il a rien dit*) et la position de l'élément doublé dans les interrogatives après l'item *wh-* (*Pourquoi Jean il parle comme ça ?*). Un argument majeur qui étaye l'analyse en doublement (donc sans emphase) est l'argument empirique. Les données orales adultes collectées à partir des années 1980 indiquent des taux élevés entre 60 % et 74 % de doublement des syntagmes nominaux par un pronom (Ashby, 1982 ; Culbertson, 2010 ; Lambrecht, 1984 ; Liang, 2023) et un taux assez bas entre 5,9 % et 20 % de syntagmes nominaux seuls (avec une valeur plutôt générique ; par exemple *Le matériel se composait de...*, Blanche-Benveniste, 1994, 2003a). La fonction distinctive de la mise en relief est donc perdue dans la configuration du doublement en raison de son taux d'utilisation élevé. L'emphase est également

exclue dans le cas d'un syntagme indéfini (on ne peut mettre l'emphase que sur une entité définie, c'est-à-dire clairement identifiée par le locuteur et son auditoire) et la position interne du syntagme nominal après l'élément *wh-* n'évoque pas non plus une mise en relief (contrairement à la position externe : *Jean, pourquoi il parle comme ça ?*).¹⁴ Il semble alors que l'on puisse distinguer français standard et oral sur la base des structures disponibles dans chaque variété : (i) la dislocation (structure avec emphase) est présente dans les deux variétés, (ii) le doublement (structure sans emphase) est spécifique au français oral car en français standard, l'absence d'emphase est véhiculée par l'utilisation du syntagme nominal seul ou du pronom seul (*Jean parle* ou *Il parle*).

Les corpus collectés au fil des années ont montré un taux élevé de doublement du syntagme nominal en français oral (entre 60 % et 74 % des syntagmes, Ashby, 1982 ; Culbertson, 2010 ; Lambrecht, 1984 ; Liang, 2023). Il s'agit de données adultes produites par des adultes à destination d'autres adultes. Culbertson (2010) s'est également intéressée au taux de doublement dans le LAE sur la base de l'étude des données de Demuth & Tremblay (2008) (« Corpus de Lyon » : 106 000 énoncés par quatre enfants entre 0;11 et 4;06 et leurs mères). Culbertson (2010, p. 115) rapporte ainsi un taux de doublement très élevé en LAE : 81 % (N = 889/1 103), avec une variation d'une mère à une autre (73 %-94 %). Legendre, Culbertson, Barrière, Nazzi & Goyet (2010) ont également étudié le LAE en français hexagonal (54 000 énoncés adressés à cinq enfants entre 1;2 et 2;6).¹⁵ Le taux de doublement des syntagmes nominaux est également très élevé, allant de 81 % à 100 %. Dans les deux corpus, le taux de doublement en LAE est donc très élevé (73 %-100 %) et globalement supérieur au taux de doublement en français oral adulte-adulte (60 %-74 %).

¹⁴ La prosodie des dislocations et des doublements est également une source de débats laissée de côté ici, mais cf. par exemple Rossi (1999), De Cat (2007), Culbertson (2010) et Legendre, Barrière, Goyet & Nazzi (2010).

¹⁵ Ensemble de données de Bassano & Mendes-Maillochon (1994), Champaud, Hunkeler (2005), Plunkett (2002), disponibles sur la plateforme CHILDES (MacWhinney, 2000a), <http://childes.talkbank.org/>.

Si l'on admet l'existence du doublement¹⁶, il reste alors une question majeure à traiter. En effet, le doublement, contrairement à la dislocation, implique que le syntagme nominal fait partie de la valence verbale, c'est-à-dire qu'il exerce la fonction de sujet du verbe. La question majeure est donc : quel est le statut du pronom nominatif dans cette configuration ?

2.2 Statut morphosyntaxique du clitique nominatif

Les pronoms présents dans les structures avec doublement sont des pronoms « clitiques ». Ce terme renvoie à une caractéristique phonologique et signifie que ces éléments n'ont pas d'accentuation propre. Un élément clitique ne peut donc émerger dans le continuum phonique qu'en s'appuyant sur un hôte qui dispose d'un accent, en français le verbe. Ce phénomène de « proclise » ou d'« enclise », en fonction de la place du clitique par rapport à son hôte, exclut la possibilité d'utiliser un clitique seul et explique ainsi l'utilisation d'un équivalent « tonique » dans certains contextes (par exemple : la réponse **Je* à la question *Qui a fait ça ?* étant exclue, le locuteur doit utiliser l'équivalent tonique *Moi* ; Kayne, 1975). « Nominatif » indique que sa forme morphologique renvoie à la fonction syntaxique de sujet. Les taux élevés de doublement indiqués dans la Section 2.1 impliquent, de fait, un taux élevé de pronoms clitiques nominatifs en français oral, car ces pronoms sont non seulement cooccurrents des syntagmes nominaux, mais peuvent également être utilisés seuls (*(Le chien) il aboie*). Blanche-Benveniste (1994) rapporte par exemple que, dans le corpus mentionné précédemment avec 5,9 % de syntagmes nominaux non-doublés, on trouve 12,8 % de syntagmes nominaux doublés et 81,3 % de clitiques utilisés seuls, soit un taux global de présence du clitique nominatif de 94,1 %.

Ce taux très élevé dans ce corpus et de nombreux autres, qu'il s'agisse de français oral hexagonal ou de français oral de Suisse, du Québec ou de l'Ontario a mené les chercheurs à proposer une analyse non pas syntaxique mais morphologique du clitique (Auger, 1994 ; Creissels, 1995 ; Culbertson, 2010 ;

¹⁶ Cette structure n'est pas reconnue par tous (cf. par exemple Brandi & Cordin, 1989, Kayne, 2000, Rizzi, 1986b).

Fonseca-Greber, 2000 ; Jakubowicz & Rigaut, 1997 ; Kaiser, 1994 ; Liang, 2023 ; P. H. Miller, 1992 ; P. H. Miller & Monachesi, 2003 ; Nadasdi, 1995 ; Pierce, 1992 ; Roberge, 1990 ; Zribi-Hertz, 1994, entre autres ; cf. Heap & Roberge, 2001 pour un historique du débat). Cela signifie que le clitique nominatif n'est pas un « pronom » au sens étymologique du terme (il ne remplace pas un nom), mais qu'il est un préfixe verbal qui porte les marques de l'accord avec le verbe. Le clitique *il* par exemple porte ainsi les traits [Personne 3], [Masculin], [Singulier]. Cette analyse rapproche les clitics nominatifs du français des clitics nominatifs de certains dialectes du nord de l'Italie qui sont largement décrits comme des marqueurs d'accord préverbaux (Brandi & Cordin, 1989 ; Pescarini, 2022 ; Poletto, 2000 ; Rizzi, 1986b, entre autres). De plus, cette hypothèse introduit un parallèle entre la structure du syntagme verbal (VP) et celle du syntagme déterminatif (DP). Dans le DP, l'ordre fixe déterminant-nom entraîne une présentation strictement ordonnée des informations grammaticales d'abord, avec les traits de [Définitude], [Genre] et [Nombre] du déterminant, suivies par les informations lexicales sous la forme d'un nom (par exemple, *le chat*). L'analyse morphologique du clitique nominatif entraîne que, dans le VP également, les informations grammaticales sont introduites d'abord, avec les traits de [Personne], [Genre] et [Nombre] du clitique, suivies par les informations lexicales sous la forme d'un verbe (par exemple, *i-dort*). Rappelons qu'à l'oral, les flexions nominales et verbales graphiques suffixées (<-s>, <-t>, etc.) ne sont généralement pas audibles. Ces questions de l'ordre fixe et de l'apport uniquement préverbal des informations grammaticales rejoignent le point développé en (iii) ci-dessous.

Dans ce cadre, un certain nombre de tests morphosyntaxiques ont été proposés afin de mettre à l'épreuve des données cette hypothèse de l'analyse morphologique du clitique en français oral par rapport aux dialectes du nord de l'Italie. Si le clitique est un marqueur d'accord préverbal, un certain nombre de configurations morphosyntaxiques, qui sont exclues en français standard, sont attendues en français oral. Celles-ci sont rapportées en (3).

(3) Configurations morphosyntaxiques si le clitique est morphologique :

- i. Le clitique est systématiquement présent avec un verbe fléchi, même dans une suite de verbes juxtaposés ou coordonnés, par exemple *Il-va ouvrir la porte et il-va rentrer* (Culbertson, 2010).
- ii. Le clitique peut apparaître indifféremment avec un syntagme nominal défini ou indéfini¹⁷, par exemple *Personne il-a rien dit* (Zribi-Hertz, 1994).
- iii. Le clitique est toujours préverbal, ce qui exclut la structure avec inversion verbe-clitique dans les interrogatives, par exemple **Le bébé dort-il ?* vs *Le bébé i(l)-dort ?* ou *l(l)-dort le bébé ?*.
- iv. La proximité phonologique entre clitique et verbe entraîne des ajustements du clitique à la phonologie du verbe, notamment l'élision du clitique *il* en *i(l)* devant un verbe à initiale consonantique, par exemple *Les p(e)tits chiens i(ls)-savent pas compter* (Culbertson, 2010).

De très nombreuses études ont proposé des illustrations des caractéristiques (i)-(iv) dans des corpus adultes, étayant ainsi l'hypothèse du clitique marqueur d'accord préverbal en français oral et distinguant alors l'éventail des possibles en français oral d'une part et en français standard d'autre part (Coveney, 2002 ; Culbertson, 2010 ; Lambrecht, 1981 ; Sportiche, 1999 ; Zribi-Hertz, 1994, entre autres). Concernant les points (i) et (iii) ci-dessus et le LAE, Culbertson (2010, p. 101) rapporte un taux de répétition du clitique nominatif en coordination de 98.4 % (N = 181/184) et aucune inversion verbe-clitique dans les interrogatives.

Jusqu'à présent, nous avons illustré la question du marqueur d'accord préverbal avec les clitics nominatifs, qui sont les éléments les plus mentionnés dans le débat car les plus présents dans les énoncés du français oral. Ce débat ne couvre néanmoins pas la totalité des éléments sujets et a temporairement laissé de côté la question des pronoms interrogatifs et relatifs qui peuvent également occuper la fonction sujet (par exemple : *Qui a fait ça ? L'enfant qui joue*). À ma connaissance, il n'y a pas de travaux sur la question du statut de ces éléments en

¹⁷ Sont considérés comme indéfinis les syntagmes dont le déterminant porte uniquement une valeur quantifiante (par exemple : *un, une, des, tous, chaque, aucun*) et les pronoms indéfinis (par exemple : *personne, rien*) (Grevisse & Goosse, 2008, p. 738).

français adulte hexagonal. Les recherches existantes portent sur le français du Québec (Auger, 1994, 1995) et de Suisse (Fonseca-Greber, 2013, 2018) et ces travaux apportent également des arguments en faveur de l'analyse morphologique de ces éléments interrogatifs et relatifs. En effet, Auger (1994) propose que les différentes formes du pronom interrogatif sujet en français québécois (*qui qui, qui est-ce qui, qui c'est qui*) étayent l'hypothèse du doublement dans la mesure où toutes les formes présentent deux items *qui*. Concernant les propositions relatives, la non-élision de *qui* devant voyelle et la cooccurrence possible du clitique nominatif après *que* semblent également représenter des arguments en faveur d'une présence récurrente d'un marqueur préverbal d'accord en français québécois sous la forme de ces items interrogatifs et relatifs. Les tests appliqués par Fonseca-Greber (2018) au pronom relatif *qui* en français de Suisse vont également dans le sens d'une analyse en tant que marqueur préverbal d'accord dans les propositions relatives (*vs il* dans les propositions principales).

Néanmoins, l'analyse morphologique du clitique nominatif en français hexagonal ne fait pas l'unanimité. Certains auteurs apportent en effet des arguments en faveur de l'hypothèse inverse, c'est-à-dire celle de l'analyse syntaxique du clitique (par exemple, De Cat, 2005 ; Stark, 2014). Dans ce cadre, un argument central est la présence de constituants placés entre le clitique nominatif et le verbe fléchi, qui empêchent alors de postuler que le clitique est un préfixe verbal. L'élément le plus discuté dans la littérature est l'adverbe négatif préverbal *ne*.¹⁸

2.3 Négation simple vs discontinue

Le français standard dispose d'une négation « discontinue », c'est-à-dire que les marqueurs de la négation sont répartis de part et d'autre du verbe fléchi (*ne... pas*), contrairement à d'autres langues romanes qui disposent soit d'un marqueur

¹⁸ Il existe également un vaste débat sur le statut morphosyntaxique des clitiques accusatifs (cf. par exemple Kaiser, 1994 ; Bahtchevanova & van Gelderen, 2016 pour une argumentation en faveur de l'analyse morphologique de ces éléments préverbaux).

préverbal (italien), soit d'un marqueur postverbal (valdotain, Zanuttini, 1997). Néanmoins, les données recueillies en (socio-)linguistique montrent qu'en français oral ou dans des productions qui reproduisent l'oralité, telles que les pièces de théâtre, l'omission de *ne* ne cesse d'augmenter et de se propager au sein des différentes classes sociales surtout depuis le XIX^e siècle (Martineau & Mougeon, 2003). Les données du français oral contemporain montrent ainsi des taux d'omission de la particule préverbale *ne* particulièrement élevés (Armstrong & Smith, 2002 ; Ashby, 1976, 1981, 1982, 2001 ; Barra-Jover, 2004 ; Coveney, 1996, 2002 ; Hansen & Malderez, 2004 ; Iglesias & Larrivée, 2014), avec une variabilité notamment en fonction du degré de familiarité avec l'interlocuteur (Armstrong & Smith, 2002 : 98,9 % ; Ashby, 1982 : 84 % ; Coveney, 2002 : 81,2 %).

Par ailleurs, les corpus oraux adultes ont révélé une correspondance entre l'absence du marqueur *ne* et la présence du clitique nominatif dans la même proposition, et ce depuis le XVIII^e siècle (Martineau & Mougeon, 2003). Pour le XX^e siècle, Ashby (1977b, p. 62) par exemple rapporte un taux d'omission de *ne* de 55 % en présence d'un clitique nominatif alors que ce taux est de 12,1 % en présence d'un syntagme nominal. Ashby (1981) et Martineau & Mougeon (2003) proposent alors que la transformation du statut du clitique nominatif de sujet syntaxique en marqueur préverbal d'accord représente l'un des facteurs qui a possiblement favorisé l'omission de *ne* (un autre facteur étant le renforcement de *pas*, cf. Hirschbühler & Labelle, 1994). Iglesias & Larrivée (2014) mentionnent également le contexte de la proposition relative comme favorisant l'omission de *ne*.

En ce qui concerne le LAE, les deux mêmes caractéristiques émergent. D'une part, la variabilité en fonction de la dyade locuteur-interlocuteur est également notable : par exemple, dans les travaux de Buson, Nardy, Rousset & Zhang (2022) en milieu scolaire, les enseignantes et les assistantes de maternelle ont davantage tendance à « surveiller leur discours » à l'adresse des enfants. Leurs taux d'omission du *ne* sont ainsi plus élevés en discours adulte-adulte (93,9 % et 98,5 %) qu'en LAE (70,6 % et 93,8 %) et les taux d'omission sont significativement

différents en LAE entre enseignantes et assistantes ($p < 0,001$). D'autre part, Culbertson (2010, p. 95) rapporte une variabilité de l'omission de *ne* en LAE en fonction de la forme du sujet. Ainsi, le taux d'omission est de 93,7 % (N = 4 897/5 229) après un clitique nominatif (avec ou sans syntagme nominal cooccurrent) et de 16,7 % (N = 4/24) après un syntagme nominal sans clitique nominatif, et la différence est statistiquement significative ($p < 0.0001$).

Les données du français adulte oral contemporain en France montrent ainsi que les clitiques nominatifs sont extrêmement fréquents et que la particule préverbale négative *ne*, qui représente un argument fort en défaveur de l'analyse morphologique des clitiques nominatifs, est globalement peu présente. L'analyse morphologique des clitiques nominatifs du français oral est donc une possibilité. Elle implique alors un autre débat car, lorsque ces marqueurs ne sont pas doublés par un syntagme nominal sujet, il n'y a pas de sujet réalisé dans la phrase (par exemple, *i-mange*) comme dans certains dialectes du nord de l'Italie (exemple du trentin : *el-magna*, Rizzi, 1986b). L'analyse morphologique du clitique implique alors de distinguer le statut du français oral de celui du français standard par rapport à la présence obligatoire ou non d'un sujet grammatical, et de considérer le français oral comme les dialectes du nord de l'Italie, c'est-à-dire comme une langue à « sujet nul ». Cependant, avant de proposer une telle analyse, il convient de considérer quelques points de variation au sein des dialectes du nord de l'Italie (Pescarini, 2020 ; Poletto, 2000), qui rendent l'analyse encore plus complexe pour le français oral.

2.4 Langue à sujet nul ?

Les dialectes du nord de l'Italie sont bien connus dans la littérature pour plusieurs raisons : (i) de façon globale (mais nuancée ci-après), ces dialectes doublent leurs pronoms toniques, leurs syntagmes nominaux définis et indéfinis et leurs pronoms relatifs et interrogatifs avec un clitique, (ii) ce clitique est analysé de façon assez consensuelle comme un marqueur préverbal d'accord, et (iii) malgré l'omniprésence du clitique, ces dialectes sont généralement considérés comme des langues à sujet nul, au même titre que l'italien standard (Brandi & Cordin,

1989 ; Poletto, 2000 ; Rizzi, 1986b). Le parallèle établi entre clitiques nominatifs dans ces dialectes et ceux du français oral entraîne alors l'hypothèse que le français oral pourrait également être une langue à sujet nul.

Néanmoins, le phénomène du doublement du sujet dans les dialectes nord-italiens n'est pas un phénomène homogène dans l'aire dialectale. Poletto (1993, 2000) propose alors de formaliser la variation observée entre dialectes en les classant en quatre types, en fonction du nombre de contextes qui autorisent le doublement et en établissant des implications entre les quatre types identifiés (cf. (4) et Manzini & Savoia, 2005 et Pescarini, 2020 pour des représentations cartographiques de la variation).

(4) Contextes de doublement et implications entre contextes (Poletto 1993, 2000) :

- i. Les dialectes qui ne doublent qu'un seul type de sujet doublent toujours les pronoms toniques (Type 1, par exemple : vénitien central).
- ii. Si les syntagmes nominaux définis sont doublés par un clitique dans un dialecte, alors tous les pronoms toniques le sont également dans cette même langue (Type 2, par exemple : trentin et lombard du nord).
- iii. Si les syntagmes nominaux indéfinis sont doublés par un clitique, alors tous les syntagmes nominaux définis et tous les pronoms toniques le sont également (Type 3, par exemple : milanais et lombard de l'est).
- iv. Si les variables dans les contextes *wh*- (c'est-à-dire les structures relatives, interrogatives et clivées) sont doublées par un clitique, alors le doublement est obligatoire avec tous les autres types de sujets (Type 4, par exemple : friulien, piémontais et ligurien).

Les travaux sur la réalisation du clitique nominatif préverbal dans les dialectes nord-italiens proposent ainsi une étude fine de la variation dans cette aire dialectale, contigüe avec celle du français hexagonal. Les dialectes du nord de l'Italie disposent d'un clitique préverbal qui est assez consensuellement analysé en tant que marqueur d'accord préverbal bien qu'il n'émerge pas dans tous les contextes possibles dans tous les dialectes. Sur la base des données du français

informel rapportées dans la Section 2, un rapprochement a été proposé entre le fonctionnement de ces dialectes et celui du français informel, et Culbertson (2010) par exemple a placé ce dernier dans le Type 2 décrit en (4) ci-dessus. La proposition de Culbertson (2010) sera rediscutée à la lumière des données enfantines dans la Section II.2 du Chapitre II.

2.5 Comment analyser les explétifs ?

Les constituants « explétifs » sont des éléments non-référentiels, c'est-à-dire qu'ils ne renvoient à aucune entité (étymologiquement, explétif signifie « qui remplit inutilement »). Dans les langues à sujet non-nul comme le français standard, les explétifs sont réalisés avec la même forme phonique que le pronom sujet de troisième personne singulier masculin (par exemple : *il pleut, il faut, il semble*, etc.). Dans la théorisation générative, la vocation d'un explétif est de satisfaire le Principe de projection étendu (EPP, Chomsky, 1981 ; Svenonius, 2002) qui stipule que toute phrase doit avoir un sujet syntaxique, minimalement réalisé sous la forme d'un pronom explétif dans les langues à sujet non-nul, ou non réalisé sous la forme de *pro_{expl}* dans les langues à sujet nul (par exemple, en grec *pro_{expl} vrèchi* « pleut », mais cf. Cognola & Casalicchio, 2018 pour une discussion sur certains explétifs dans des langues à sujet nul). De plus, il existe différents sous-types d'explétifs. Rizzi (1986a) notamment a distingué les éléments purement explétifs (par exemple avec *falloir*) de ceux qui accompagnent les verbes météorologiques qu'il appelle des « quasi-argumentaux » (cf. également Ruwet, 1986).

Les dialectes nord-italiens, étant considérés comme des langues à sujet nul, ne devraient donc pas présenter d'explétifs réalisés. Or, la présence de clitiques explétifs dans certains de ces dialectes vient falsifier cette prédiction et représente donc un problème vis-à-vis de la partition entre langues à sujet nul et langues à sujet non-nul (Pescarini, 2014, 2020). Néanmoins, et parallèlement aux différents types de dialectes présentés en (4) à partir des travaux de Poletto (1993, 2000), Pescarini (2014) a proposé une échelle d'implications entre les différents contextes d'apparition de l'explétif dans ces dialectes. Les différents contextes

sont au nombre de cinq : les verbaux modaux du type *falloir*, les constructions impersonnelles avec *si/se* « on », les verbes du type *sembler*, les verbes existentiels du type *il y a*, et les verbes météorologiques. Ainsi, le clitique explétif n'émerge avec un verbe modal (du type *falloir*) que s'il émerge dans tous les autres contextes. Néanmoins, si l'on admet que le clitique explétif est un marqueur d'accord préverbal comme les autres dans les langues à sujet nul, sa fonction reste encore à clarifier dans la mesure où les verbes qui appellent un explétif sont des verbes défectifs, essentiellement uniquement fléchis à la troisième personne du singulier. Ainsi, lorsqu'un locuteur entend *faut que...*, même en l'absence du clitique, il sait qu'il s'agit d'une flexion à la troisième personne du singulier. La présence de l'explétif semble donc inutile, que ce constituant soit analysé comme un pronom syntaxique ou comme un marqueur d'accord. Il pourrait alors simplement s'agir dans les deux cas d'un élément utilisé pour régulariser un paradigme sans aucun apport référentiel, comme son nom l'indique.

Il est alors très intéressant de noter que Culbertson & Legendre (2014) ont étudié l'impact de différents contextes sur le jugement de grammaticalité de locuteurs adultes du français hexagonal face à des phrases avec rétention ou omission de l'explétif. Les variables testées étaient la fréquence de la structure, le type d'explétif (non-argumental vs quasi-argumental) et la catégorie sémantico-syntaxique du verbe (météorologique : *il pleut*, modal : *il faut*, non-modal : *il semble* et existentiel : *il y a*, cf. Travis, 1984). Culbertson & Legendre (2014) ont ainsi montré que la fréquence de la structure impacte significativement l'acceptabilité du verbe utilisé sans explétif (par exemple, *y a un embouteillage*), et que le type de verbe participe également au résultat, les modaux (par exemple, *vaut mieux...*) étant significativement plus acceptés sans explétif que les autres catégories. Les auteures proposent donc une échelle d'acceptabilité de l'omission : modal > existentiel > non-modal > météorologique. Il y a alors une proximité avec l'échelle d'implications de Pescarini (2014), reprise dans Pescarini (2020, p. 114), qui indique pour les dialectes du nord de l'Italie que les explétifs sont omis dans plus de dialectes différents avec les verbes modaux, puis le nombre de dialectes qui omettent l'explétif baisse avec les verbes non-modaux, les verbes

existentiels, et enfin, aucun dialecte n'omet l'explétif avec un verbe météorologique.

Des langues différentes semblent ainsi posséder des points communs concernant la présence/omission du clitique explétif en fonction du type de verbe. Il sera intéressant d'étudier les aspects développementaux des implications décrites par Poletto (2000) et Pescarini (2014) et mises en regard l'une de l'autre dans Pescarini (2020, p. 127) chez des enfants francophones dans le cadre du débat sur le statut morphosyntaxique du clitique explétif et le statut du français oral en tant que langue à sujet nul ou langue à sujet non-nul (cf. Section II.3.3 du Chapitre II).

2.6 Structures interrogatives

Les structures interrogatives en français représentent également un domaine d'investigations extrêmement riche et foisonnant, touchant à des domaines très divers tels que la prosodie, la morphologie, la syntaxe, la sémantique et la pragmatique. Des ouvrages et numéros de revues récents témoignent de ce foisonnement des recherches (Béguelin, Coveney & Guryev, 2018 ; Guryev, Dekhissi & Bonan, 2024 ; Larrivée & Guryev, 2021).

Parmi tous les types d'interrogatives possibles, nous nous concentrerons sur les interrogatives « verbales » (*c'est qui ?*) opposées aux « averbales » (*qui ça ?*), « directes » (*qu'est-ce que tu lis ?*) opposées aux « indirectes » (*je me demande ce que tu lis*) et « *wh-* » (*où tu vas ?*) opposées aux « oui/non » (*tu viens ?*) et « alternatives » (*tu pars le matin ou l'après-midi ?*). En effet, les questions verbales directes *wh-* possèdent à elles seules plusieurs facteurs de variation qui génèrent une diversité de structures, dont les principales sont illustrées dans le Tableau 2 avec le clitique nominatif *tu*. Cette diversité de possibilités rend alors le français remarquable parmi les langues du monde, notamment par la possibilité de produire des questions *wh-* soit « in-situ » (*wh-* à droite du verbe fléchi, dans la position où est générée la réponse) soit « ex-situ » (*wh-* à gauche du verbe

fléchi)¹⁹, car c'est une alternative rare (2,5 % des 902 langues répertoriées par Dryer, 2013), les autres langues disposant de l'une ou l'autre possibilité uniquement.²⁰

Tableau 2 : variations et agrammaticalités des structures *wh-* avec clitique

	Ex-situ	In-situ
Clitique-Verbe	<i>où tu vas ?</i>	<i>tu vas où ?</i>
<i>Est-ce que</i>	<i>où est-ce que tu vas ?</i>	<i>*est-ce que tu vas où ?</i>
Inversion Verbe-Clitique	<i>où vas-tu ?</i>	<i>*vas-tu où ?</i>
Clivée	<i>où c'est que tu vas ?</i>	<i>c'est où que tu vas ?</i>

De par sa rareté parmi les langues du monde, la variabilité in-situ/ex-situ du français a suscité de très nombreuses recherches avec des études qui mettent en avant des facteurs possibles en prosodie, morphologie, syntaxe, sémantique ou pragmatique. La Grammaire générative s'est notamment emparée de la problématique dans le cadre de sa formalisation paramétrique de la variation inter-linguistique. Le « Critère *Wh-* », Principe universel qui régit donc toutes les langues du monde, spécifie que l'opérateur *wh-* et la tête C, qui porte le trait positif [+*wh*] lorsqu'une phrase est interrogative, doivent être dans une relation spécifieur-tête (Rizzi, 1991/1996, 2006). La variation entre les différentes langues du monde est ensuite formalisée via le « Paramètre *Wh-* », qui stipule que la relation spécifieur-tête est obtenue :

- soit par déplacement « visible » de l'opérateur, c'est-à-dire qu'il y a déplacement du syntagme interrogatif en syntaxe et en Forme logique depuis sa position de base (où est générée la réponse à la question) vers le Spécifieur de CP ; cette alternative génère les questions à *wh-ex-situ* (comme en anglais) ;

¹⁹ La théorie générative considère que la position ex-situ est obtenue par déplacement depuis la position in-situ, alors que les approches basées sur l'usage ne postulent aucun déplacement (divergences similaires à celles rapportées pour les questions oui/non, cf. Section 1.3.3).

²⁰ Notons que de nombreux dialectes nord-italiens appartiennent également à cette catégorie « mixte » (Baunaz & Bonan, 2023, fn 4 ; Bonan, 2021).

- soit par déplacement « furtif » de l'opérateur, c'est-à-dire qu'il y a déplacement du syntagme interrogatif en Forme Logique uniquement ; cette alternative génère les questions à *wh*-in-situ (comme en mandarin).

Le défi théorique soulevé par le français est que cette langue permet les deux configurations et les chercheurs ont alors proposé de formaliser les contraintes sous-jacentes qui régissent la variation, en termes syntaxiques, sémantiques, prosodiques ou pragmatiques notamment (Aoun, Hornstein & Sportiche, 1981 ; Baunaz, 2011, 2016 ; Baunaz & Patin, 2011 ; Baunaz & Bonan, 2023 ; Bošković, 2000 ; Chang, 1997 ; Cheng & Rooryck, 2000 ; Coveney, 1989 ; Déprez, Syrett & Kawahara, 2012, 2013 ; Hamlaoui, 2011 ; Larrivée, 2019 ; Mathieu, 2004 ; Munaro, Poletto & Pollock, 2001 ; Munaro & Pollock, 2005 ; Myers, 2007 ; Obenauer, 1994 ; Poletto & Pollock, 2005, 2020 ; Quillard, 2000 ; Reinhart, 1998 ; Shlonsky, 2012 ; Vergnaud & Zubizarreta, 2005, entre autres publications).

La diversité des formes des interrogatives en français est également décrite dans le cadre des registres du français standard et oral (Ashby, 1977a ; Coveney, 1997 ; Lambrecht, 1981 ; Larrivée & Guryev, 2021 ; Pohl, 1965, entre autres). Parmi les facteurs de variation mentionnés dans le Tableau 2, l'inversion dans les questions *wh*- est assez consensuellement décrite comme relevant du français standard car très peu présente dans les corpus du français oral hexagonal (Coveney, 2002 : 6,6 % ; De Cat, 2007 : 2 %) et les données du LAE renvoient un taux encore plus bas (Culbertson, 2010 : 0,8 %, N = 66/7 862 questions *wh*- avec un clitique, par exemple : *Que va-t-il manger ?*). Apparue au XII^e siècle afin de focaliser le constituant interrogatif, la forme *est-ce que* a progressivement perdu cette fonction focalisatrice et est décrite comme remplaçant l'inversion verbe-clitique, en lien avec le changement du statut des clitics vers le XV^e siècle, qui sont passés de formes toniques à formes clitiques (Zumwald Küster, 2018). Contrairement à l'inversion, l'insertion de *est-ce que* dans une question *wh*- est souvent perçue par les grammaires contemporaines comme relevant d'une langue « non surveillée » (Zumwald Küster, 2018, p. 96). Par ailleurs, cette insertion est parfois associée à des fonctions pragmatiques spécifiques (questions rhétoriques

ou émotionnellement chargées notamment ; par exemple : *qu'est-ce que je veux dans la vie ?* ; Donaldson, 2016, p. 493), et il a été remarqué que *est-ce que* émerge essentiellement avec les items interrogatifs *que* et *quand* (Donaldson, 2016 ; Hamlaoui, 2011).

Enfin, concernant l'alternance ex-situ/in-situ, la variante in-situ est généralement associée au français non-standard/informel/oral. Cette structure ne cesse de gagner du terrain au détriment de la variante ex-situ, essentiellement depuis le XIX^e siècle (Guryev & Larrivée, 2021 ; Baunaz & Bonan, 2023) et son augmentation en français oral est souvent associée à la baisse de l'utilisation de l'inversion verbe-clitique qui entraînait l'utilisation de *wh-ex-situ*. Néanmoins, la baisse de l'inversion verbe-clitique ne suffit pas à expliquer l'augmentation du ratio de in-situ, car l'insertion de *est-ce que*, décrite comme relevant de l'oralité, impose la position *wh-ex-situ*. Le nombre de structures différentes qui permet un *wh-ex-situ* est donc supérieur à celui qui autorise un *wh-in-situ* à l'oral (en excluant la structure avec inversion : 3 vs 2, cf. Tableau 2).

Les données de terrain montrent des taux de in-situ très différents d'un corpus à un autre (Ashby, 1977a : 45,9 % ; Coveney, 1995 : 10,5 % ; Coveney, 2002 : 16,1 %). Les études de LAE rapportent également des taux de in-situ divers (Strik & Pérez-Leroux, 2011 : 38,1 % ; Geveler & Müller, 2015 : 51 % ; Becker & Gotowski, 2015 : 16,6 %). Cet ensemble de données permet néanmoins de constater que, malgré la progression diachronique du ratio de *wh-in-situ*, les productions adultes orales contemporaines montrent encore une majorité de *wh-ex-situ*.²¹ Un calcul des ratios dans les données LAE de Strik & Pérez-Leroux (2011) illustre également la prévalence des structures avec *wh-ex-situ* parmi les questions *wh-* adressées aux enfants : 47,7 % sont des *wh-est-ce que* (N = 74/155), 14,2 % sont des *wh-ex-situ* SV (N = 22) et seules 38,1 % sont des SV *wh-in-situ* (N = 59).

²¹ Avec une réserve néanmoins sur les toute dernières années au cours desquelles Baunaz & Bonan (2023) montrent que les *wh-in-situ* sont devenus majoritaires (62 % dans les années 2010).

Le défi qui s'ouvre alors aux chercheurs en acquisition est d'expliquer pourquoi les enfants qui sont en contact avec ce français oral majoritairement ex-situ produisent davantage de *wh*-in-situ. Cette question sera traitée dans la Section III.1 du Chapitre II. La variabilité de la position de l'élément *wh*- en français adulte sera ensuite discutée à la lumière des données enfantines et une nouvelle proposition sera rapportée pour les données adultes à l'interface de la syntaxe et de la sémantique (cf. Sections III.2 et 3).

3. Théorisations sur la variation morphosyntaxique

Comme dans le domaine de l'acquisition du langage (cf. Section I.3 ci-dessus), les recherches sur la variation morphosyntaxique du français ont donné naissance à différentes théorisations. Deux approches principales et antagonistes ont pris en charge cette problématique : le variationnisme et la diglossie.

3.1 Variationnisme

À partir des années 1960, les travaux de William Labov ont tout particulièrement mis en avant les liens qui pouvaient exister entre la distribution et l'évolution de différentes variantes au sein d'un système linguistique et les facteurs sociaux à l'origine de cette variation. Les travaux de Labov (1963) signent ainsi la naissance de la sociolinguistique et du « variationnisme » qui consiste à analyser la variation linguistique en termes de variables sociolinguistiques, définies comme étant des « alternatives significativement différentes au niveau socio-stylistique mais qui ont la même fonction dans le discours, telles que l'omission variable de *ne* [...] ». (Coveney, 2011, p. 51).²²

La variabilité de différentes formes occupant une même fonction est alors notamment décrite en termes de fréquences et de cooccurrences des différentes variantes. Coveney (2011, p. 75) par exemple synthétise les variations observées en français pour l'utilisation du sujet seul (*vs* son doublement), la présence de *ne*

²² Texte original : “[...] sociolinguistic variables – sociostylistically significant, alternating ways of saying the same thing, such as the variable omission of *ne* [...].”

(vs son absence), l'utilisation du pronom sujet *nous* (vs on) et la production de *l'on* (vs on) dans son corpus de Picardie (dix-huit heures d'enregistrements auprès de trente informateurs) et rapporte que les fréquences de chaque variante et de leurs cooccurrences varient d'un locuteur à un autre et sont également différentes de celles du groupe considéré dans sa totalité. L'auteur ajoute que des variantes formelles et informelles peuvent être combinées à l'intérieur d'une même phrase et que leur émergence dépend d'une combinaison de facteurs linguistiques, stylistiques et pragmatiques. Blanche-Benveniste & Martin (2010), Gadet (1996, 2007) et Stark & Binder (2021), entre autres, décrivent également la variation du français hexagonal dans ce cadre.

Les approches variationnistes sont ainsi décrites comme étant plutôt « probabilistes » dans le sens où elles mettent en avant un ensemble de possibilités à partir desquelles chaque locuteur effectue un choix en fonction de facteurs sociolinguistiques (Cornips, 2015). Ces approches adoptent donc un modèle de description plutôt basé sur l'usage et sont ainsi à rapprocher des modèles « basés sur l'usage » (cf. Section 1.3.2 ci-dessus pour l'acquisition de la L1). Cette approche sociolinguistique de la variation s'oppose à l'approche « diglossique » détaillée dans la section suivante.

3.2 Diglossie

Le terme « diglossie » a été mis en avant par Ferguson (1959) dans le cadre de ses recherches sur la variation en arabe, grec moderne, suisse allemand et créole haïtien. Ferguson (1959, p. 325) a défini la diglossie comme étant une situation linguistique « où deux variétés d'une langue existent côte à côte dans toute la communauté, chacune jouant un rôle précis ». ²³ Ainsi, la variété « haute » de la langue est activée dans les situations formelles orales et écrites, alors que la variété « basse » est utilisée dans un cadre plus informel et quotidien, à l'oral et à l'écrit. La variété haute peut être considérée comme plus prestigieuse par les locuteurs lorsqu'ils ont conscience de l'existence des deux variétés. Cependant,

²³ Texte original : “[...] where two varieties of a language exist side by side throughout the community, with each having a definite role to play.”

cela n'est pas toujours le cas et l'existence/l'utilisation de la variété basse peut parfois être réfutée par certains locuteurs (Ferguson, 1959, p. 330).

Dans la mesure où le français est une langue qui connaît une très grande variabilité en phonologie, morphologie, syntaxe et sémantique, comme nous avons pu l'illustrer pour la morphosyntaxe ci-dessus, Blanche-Benveniste (2003b, p. 317) en est venue à poser la question suivante : « Les Français seraient-ils atteints de diglossie ? ». Certains linguistes ont répondu positivement à l'hypothèse d'une diglossie en France, notamment Barra-Jover (2004, 2010), Hamlaoui (2011), Massot (2008, 2010), Massot & Rowlett (2013) et Zribi-Hertz (2011) pour le français contemporain, et Larrivée (2022) pour le français médiéval, étayant alors les propos de Ferguson (1959, p. 332) sur le fait que « la diglossie persiste habituellement au moins plusieurs siècles ».

Ferguson (1959) précise que les différences les plus marquantes entre les variétés basse et haute d'une langue sont d'ordre grammatical. Pour le français contemporain hexagonal, Massot (2010) rapporte notamment les différentes formes du sujet (cooccurrences ou non), de pronoms (*on* vs *nous*) et d'omissions (explétifs), les deux formes de la négation et les formes de l'interrogation. N'observant pas de mélanges des deux variétés au sein d'un même énoncé dans ses données, Massot (2010, p. 90), contrairement à Coveney (2011) mentionné ci-dessus, propose alors que « Pour chaque énoncé produit par un locuteur éduqué, celui-ci s'engage pour l'énoncé entier dans l'une de ses deux compétences [...] ». Cette généralisation, bien que très forte, possède néanmoins l'avantage de pouvoir être mise à l'épreuve d'autres données.

La cohérence grammaticale interne à chaque énoncé est également présente, mais de façon implicite, lorsque Culbertson (2010) propose de rendre compte de l'utilisation des deux formes de la négation en LAE. En effet, nous avons vu un certain nombre de corpus dans les sections précédentes et aucun ensemble de données ne renvoie des taux de 0 vs 100 %, pour aucun phénomène linguistique. Culbertson (2010, p. 97) propose alors d'expliquer le taux de rétention du *ne* en présence d'un clitique nominatif en LAE (6 %) par le fait que « les locuteurs

maintiennent des grammaires qui encodent aussi bien le registre du français standard que celui du français informel dans leur langue [...] »²⁴ et l’auteure démontre statistiquement qu’un modèle linguistique qui « mélange deux grammaires » rend mieux compte des données qu’un modèle qui ne postule pas la présence de deux grammaires ($p < 0.001$, test du taux de vraisemblance de Lehmann, 1986).

L’hypothèse de la diglossie est également très intéressante dans le cadre de l’acquisition du langage, car Ferguson (1959) en a dessiné des contours très précis en termes développementaux. Ainsi, Ferguson (1959, p. 331) a distingué les processus de développement des deux variétés chez les locuteurs. La variété basse est « acquise » par tous les enfants de la communauté linguistique par une transmission implicite des informations linguistiques dans le cadre du bain linguistique ordinaire et quotidien, alors que la variété haute est moins présente dans l’environnement quotidien du jeune enfant et est largement transmise dans le cadre scolaire. L’hypothèse de la diglossie permet alors de formuler deux prédictions fortes et testables en termes de processus acquisitionnels. Ces éléments sont rapportés en (5).

(5) Hypothèse de la diglossie et prédictions développementales :

- i. La variété basse est maîtrisée par tous les locuteurs de la communauté linguistique, alors que la variété haute est utilisée avec des degrés de compétence différents en fonction de l’âge du locuteur (préscolaire ou pas) puis de son niveau de scolarisation. Tous les locuteurs de la communauté linguistique ne sont donc pas forcément diglosses.
- ii. Les deux variétés se développent de façon séquentielle : la variété basse est systématiquement acquise avant la variété haute. En fonction de l’âge de l’enfant, on devrait donc pouvoir constater des étapes de l’acquisition/apprentissage de la variété haute suite à l’acquisition initiale de la variété basse.

²⁴ Texte original : “I argue that rare cases of *ne* found with subject clitics indicate that speakers maintain grammars encoding both the Standard French and Colloquial French registers of their language [...].”

Ces hypothèses seront testées et discutées à la lumière des données enfantines, non seulement auprès d'une population d'enfants francophones L1, mais également à partir de données de deux enfants dont le français est une L2P (cf. Section IV du Chapitre II).

4. Positionnement scientifique

La Section II a montré l'importance de connaître les caractéristiques de l'environnement des enfants pour étudier l'acquisition du langage. Les recherches dans ce domaine ont en effet objectivé que les facteurs de variation sont nombreux et divers, de l'environnement socioéconomique de l'entourage quotidien à l'environnement linguistique très variable en fonction des locuteurs et des situations de communication (cf. Section 1). De plus, le français est une langue qui présente un nombre non-négligeable de variantes dans sa composante morphosyntaxique (formes diverses du sujet, de la négation et des interrogatives, cf. Section 2).

Le jeune enfant est donc confronté à des informations linguistiques qui présentent en permanence une grande variabilité en fonction du locuteur et de la situation de communication. Les travaux assez récents sur l'hypothèse de la diglossie appliquée au français (notamment la thèse de Massot, 2008, cf. Section 3.2 ci-dessus) ont donné un cadre qui pouvait être mis à l'épreuve de données, ce que j'ai fait avec de nouvelles données enfantines, afin de tester non seulement cette hypothèse de la diglossie mais également les aspects développementaux décrits par Ferguson (1959). Les questions de l'acquisition de la variation morphosyntaxique et de sa formalisation dans le cadre de l'hypothèse de la diglossie sont ainsi au centre des travaux rapportés dans le Chapitre II.

Chapitre II : Travaux réalisés

Ce second chapitre vise à détailler les travaux de recherche publiés depuis ma thèse en 2009 (Palasis, 2009; publication en 2010b) et à les commenter au regard de nouvelles données et publications plus récentes. Des ajouts inédits jalonnent donc également ce chapitre. La plupart des études portent sur le français hexagonal contemporain et les thématiques principales sont le développement de la morphosyntaxe de ce français dans sa modalité orale chez des enfants natifs entre 2;3 et 5;11 (cf. Sections II et III) et la formalisation de la variation dans ce français hexagonal (cf. Section IV). Neuf publications entre 2010 et 2023 ont été choisies pour illustrer cette synthèse et sont présentées dans le Volume 2 par ordre chronologique (p. 18/226 *sqq.*).²⁵

Mais commençons par le commencement : le constat que la grammaire générative, modèle théorique dans lequel j'ai reçu ma formation initiale, ne s'appuie pas sur un nombre important de données de terrain (cf. Section I).

I. Spécificités de l'approche proposée

Bien qu'initialement ancrés dans le modèle générativiste, mes travaux sur le développement du français s'en détachent nettement à deux niveaux : la place fondatrice attribuée aux données de terrain (cf. Section 1) et l'hypothèse que les différences entre productions enfantines et adultes s'expliquent par des différences entre leurs systèmes linguistiques respectifs, et non pas par une déficience dans la compétence ou la performance des enfants par rapport aux adultes (cf. Section 2).

1. Création d'une base de données du français enfantin

Cf. pages 19-40/226 du Volume 2 : Palasis, K. (2010b). *Introducing New French Child Data: Thoughts on their Gathering and Coding. Corpus, 9, 'La syntaxe de corpus', 33-51, <https://journals.openedition.org/corpus/1801>*

²⁵ Les paginations dans le format N/226 dans ce volume renvoient à la pagination globale du Volume 2 (et non à la pagination originale de chaque article).

Dès mes toute premières lectures sur l'acquisition du langage dans le cadre de la Grammaire générative, je suis restée perplexe quant à ce qui m'a semblé être un décalage important mais assumé entre un arsenal théorique très sophistiqué et un support empirique très réduit. Par exemple, Demuth (1996) préconise de collecter des données auprès de « plus qu'un seul enfant » et Friedemann & Rizzi (2000) qualifient de « grand corpus » un ensemble de données recueillies auprès de six enfants. L'état des lieux des bases de données enfantines francophones proposé dans Palasis (2010b, pp. 149-153) rapporte ainsi que de nombreuses recherches fondatrices en Grammaire générative portent sur un ensemble restreint et récurrent de données provenant de un à quatre enfants enregistrés dans les années 1970-1980 en France et au Québec (Nathalie et Daniel par Lightbown, 1977, Philippe par Suppes, Smith & Léveillé, 1973, et Grégoire par Champaud, 1988). Les travaux de Crisma (1992) sur les questions *wh*- et ceux de Côté (2001) sur les clitiques nominatifs reposent ainsi exclusivement sur les données de Philippe. Friedemann (1993) utilise les données de Philippe et de Grégoire pour son étude sur la position du sujet. Les productions des quatre enfants sont utilisées par Pierce (1992) dans son étude comparative entre syntaxe du français et de l'anglais, par Déprez & Pierce (1993) sur la négation, et par Ferdinand (1996) sur le développement des catégories fonctionnelles. De nouvelles données ont été collectées dans les années 1990-2000 (le « corpus de Genève » par Frauenfelder, Rasetti & Rizzi, 1993-2002, le « corpus de York » par Plunkett & De Cat, 2002, et le « corpus de Lyon » par Demuth & Tremblay, 2008, mentionné par Culbertson, 2010), mais de nouveau il s'agit de petits groupes d'enfants (plus ou moins quatre) qui représentent des systèmes francophones différents (suisse, belge, québécois ou hexagonal).²⁶

Ainsi, malgré les apports indéniables de ces différentes collectes de données, il m'a semblé qu'elles présentaient des variables que je souhaitais exclure d'une base de données, et qui ne répondaient donc pas au cahier des charges qui a fait l'objet d'une réflexion approfondie afin de conduire mes propres recherches

²⁶ Les données des corpus Léveillé, Champaud et De Cat sont disponibles sur la plateforme CHILDES, <http://chilides.talkbank.org/> (MacWhinney, 2000a).

(Palasis, 2010b, p. 153-171 ; Palasis, 2010a, p. 19-26/226 ; Palasis, 2011a), notamment concernant les points repris en (6) ci-après (et cf. Section II.1 du Chapitre I sur le LAE).

(6) Cahier des charges pour la collecte de données enfantines :

- i. Représentativité : ce point nécessite une collecte de données auprès d'un groupe plus important d'enfants afin de contrôler les variables individuelles et diastratiques (notamment liées à la catégorie socioprofessionnelle des parents et le niveau de diplomation de la mère. Les travaux rapportés en Section II.1 du Chapitre I semblent néanmoins indiquer que ces dernières ont davantage d'impact sur les processus déclaratifs (apprentissage du lexique et de la lecture) que sur les non-déclaratifs (acquisition de la morphosyntaxe).
- ii. Neutralisation de la variation diatopique : ce point impose de contrôler l'origine géographique des collectes et de n'utiliser que des données hexagonales (cf. page 38 et Ashby, 1988 et Walter, 1988, entre autres).
- iii. Neutralisation de la variation diachronique : une génération sépare les données de Philippe recueillies en 1971 et celles des corpus de York et de Lyon recueillies dans les années 2000 ; on ne peut donc pas exclure que ces productions présentent peut-être des différences liées à l'évolution de la langue orale en général et du LAE en particulier (cf. page 35).
- iv. Neutralisation des variables méthodologiques : Jakubowicz (1996) et Baude (2006) par exemple rapportent des effets potentiels des différences dans les méthodes de recueil, de transcription, du matériel, du cadre théorique, etc.
- v. Proximité avec les données : recueillir ses propres données entraîne une connaissance beaucoup plus fine des données primaires à partir desquelles le chercheur va travailler (préconisée notamment par De Cat, 2002 et Baude, 2006).

Ces réflexions méthodologiques m'ont conduite à créer ma propre base de données de productions « semi-spontanées » (interactions enfant-enfant et

enfant-chercheure ; Eisenbeiss, 2010) recueillies pendant trois ans auprès d'une classe de maternelle en France. Le Tableau 3 rapporte les caractéristiques principales de cette collecte.

Tableau 3 : base de données du français en maternelle

	Petite section (PS)	Moyenne section (MS)	Grande section (GS - 5/10 sessions)	Totaux
Dates	2006/2007	2007/2008	2008/2009	
Heures	25	20	20 (10)	65
Sessions	13	10	10 (5)	33
Enregistrements	Audio & vidéo	Audio & vidéo	Audio & vidéo	
Adultes	3	1	1	
Énoncés adultes	12 891	9 291	4 947	27 129
Enfants (tous/L1 français)	20/17	19/17	18/16	
Tranches d'âge	2;5-4;0	3;6-4;11	4;5-5;11	2;5-5;11
Énoncés enfants (tous)	15 992	14 348	11 505	41 845
Énoncés 16 enfants (français L1)	13 320	12 242	10 501	36 063
Énoncés 2 enfants (français L2P)	858	1 106	1 004	2 968

Dans un second temps, les étapes de la transcription et du codage ont également été importantes pour construire un socle empirique de qualité. Grâce aux outils de la plateforme CHILDES (MacWhinney, 2000a, 2000b), les données recueillies ont été transcrites orthographiquement puis codées à deux niveaux : morphosyntaxique (%mor) et productions non-cibles (%err), comme le montre l'exemple extrait de la troisième année en (7) et l'exemple (4) p. 30/226. Concernant les codes %mor, ces exemples illustrent que cette base de données utilise d'une part des codes morphosyntaxiques fournis par défaut par CHILDES et d'autre part des codes spécifiquement créés lors de la constitution de la base de données, notamment concernant les pronoms (en gras dans l'exemple (7) et cf. p. 31-34/226 sur l'élaboration du codage). Tous les codes %err quant à eux sont spécifiques à cette base de données (cf. également Palasis, 2010b, p. 208-215 et 372-377). Ils permettent d'extraire les agrammaticalités enfantines de ce corpus entier, par année, par enfant, par composante du langage, par type d'erreur, etc. et d'étudier un phénomène linguistique de façon longitudinale et/ou transversale

(par exemple, l'évolution du phénomène du sujet nul enfantin, cf. Section II.1 ci-dessous).

(7) Exemple de transcription et de codage (08-09X60c, LOU, 5;11) :

LOU: euh Matteo tu pourrais me donner la famille <de le> [*] docteur ?

%mor: co|euh n:prop|Matteo, **pro:cli:d:nom|tu&2S,**

v:mdl|pouvoir&COND&12s, **pro:cli:dat|me&1S,** v|donner-INF,

det:art|la&f&sg, n|famille&f prep|de, det:art|le&m&sg,

n|docteur&m ?

%err: de le = du \$MOR \$LOS \$SBL

Le Tableau 4 illustre le niveau de détail du codage avec les pronoms clitique et tonique de première personne du singulier *je* et *moi* (simplification du Tableau 3, p. 31/226). Le codage permet d'extraire les pronoms en fonction de leur forme (clitique vs tonique), leur cas (nominatif, accusatif, etc.), leur doublement (seul vs doublé) et la position du pronom tonique lorsqu'il double un clitique (gauche vs droite). Ce codage permet notamment d'étudier le détail des formes du sujet grammatical (cf. Section II ci-dessous).

Tableau 4 : codage spécifique des pronoms

Formes	Codes & exemples	
<i>j(e)</i>	a	pro:cli:nom je&1S <i>j'</i> ai perdu l(e) chien .
	b	pro:cli:d:nom je&1S attends <i>je</i> l' ai <i>moi</i> .
<i>moi</i>	c	pro:ton:nom moi&1S <i>moi</i> 0 [*] veux ça les cartes .
	d	pro:ton:dg:nom moi&1S <i>moi j(e)</i> suis jaune !
	e	pro:ton:dd:nom moi&1S <i>j'</i> ai perdu <i>moi</i> .

Il me semble alors que cette base de données propose un fondement empirique solide à partir duquel les chercheurs peuvent examiner le développement du français hexagonal entre 2;5 et 5;11 et présenter une description complète et détaillée du système linguistique de l'enfant francophone natif juste avant son entrée à l'école primaire. Les deux premières années de la

base de données sont partagées sur la plateforme CHILDES (Palasis, 2012a).²⁷ La troisième année, ayant été transcrite et codée en 2022, n'est pas encore partagée et n'a pas encore été pleinement exploitée. Mes propres recherches reposent donc souvent sur l'extraction et l'analyse des deux premières années (et d'un autre petit corpus de 1 072 énoncés d'enfants en toute petite section de maternelle entre 2;3 et 3;1, recueilli avant ma thèse ; Palasis, 2005), avec quelques travaux récents qui exploitent les trois années (Palasis, 2022 ; Palasis, Faure & Meunier, 2023). Par ailleurs, la base de données sur CHILDES a commencé à être exploitée par d'autres chercheurs (par exemple, Becker & Gotowski, 2015 sur les questions *wh-* et Martin, Nie, Alexiadou & Guasti, 2022 sur les constructions causatives) et la totalité des données est partagée sur demande. Ce partage m'a également permis d'intégrer un projet de recherche international porté par Karen Lahousse (KU Leuven) et intitulé *The emergence of syntactic complexity. The first language acquisition of cleft sentences as a window onto the syntax-prosody-information structure interface* (Lahousse, Jourdain, Thonissen & Palasis, 2023 ; Wensink, Lahousse, de Cat, Palasis & Busson, 2024).

2. Cohérence interne du système enfantin

Dans le cadre générativiste, les différences entre productions enfantines et productions adultes sont souvent décrites comme émanant d'un « déficit » de la compétence ou de la performance des enfants par rapport à celle des adultes. Ainsi, les recherches rendent généralement compte des productions enfantines non-cibles par un paramétrage initial erroné (par exemple, Hyams, 1986), des structures syntaxiques enfantines incomplètes notamment en termes de projections fonctionnelles (Radford, 1990 ; Crisma, 1992 ; Rizzi, 1992 ; Ferdinand, 1996 ; Friedmann, Belletti & Rizzi, 2021), ou par un déficit de capacités de production, de computation, et/ou de mémoire de travail (Jakubowicz & Rigaut, 1997; Jakubowicz, 2011 ; Goodluck & Kazanina, 2020). L'apport de données

²⁷ <http://chilides.talkbank.org/access/French/Palasis.html> (dans la version en ligne, les codes %mor par défaut ont été automatiquement réattribués).

nouvelles permet alors d'envisager la description du développement linguistique sous un angle différent.

À partir de l'observation des productions enfantines décrites dans la Section 1 ci-dessus, mes hypothèses possèdent la particularité d'attribuer les spécificités des productions enfantines à des caractéristiques de leur système linguistique, considéré en tant que système à part entière fonctionnant avec une cohérence interne. Le système enfantin certes évolue puisqu'il est en phase de développement (ce que Cohen, 1924 décrivait comme une succession de « petits systèmes »), mais cela ne lui enlève pas pour autant la possibilité de disposer d'une cohérence interne et d'une dynamique endogène à chaque étape (Palasis, 2010b, p. 75-83). Pour reprendre une comparaison utilisée par Vercelli & Piattelli-Palmarini (2009), « nous avons tous deux jambes mais des démarches différentes » et ma contribution consiste à montrer que « différent » ne signifie pas « déficient/incomplet/immature/etc. ». Mes hypothèses sur le sujet nul enfantin (en Section II.1 ci-dessous) et les structures des questions *wh*- (en Section III), qui présentent des divergences avec les formes et structures adultes, illustrent ce propos.

II. Formes du sujet : morphologie et syntaxe

La base de données du français en maternelle présentée en Section I.1 ci-dessus permet d'extraire des données longitudinales sur les formes que prend le sujet chez des enfants francophones natifs pendant trois années (2;5-5;11). Le Tableau 5 rapporte les cinq formes utilisées par les seize enfants du groupe dans leurs phrases déclaratives (N = 24 296) et les exemples en (8) illustrent ces différentes formes (Palasis, 2022).²⁸

²⁸ La question des pronoms relatifs et interrogatifs sera abordée dans la Section 2.

Tableau 5 : formes du sujet en maternelle (français L1)

		PS-N	PS-%	MS-N	MS-%	GS-N	GS-%	TOT-N	TOT-%
a	Clitique seul	5027	64,0	6427	71,6	5691	76,2	17145	70,6
b	Clitique + DP	2492	31,7	2440	27,2	1692	22,7	6624	27,3
c	Sujet nul	300	3,8	60	0,7	14	0,2	374	1,5
d	DP seul	37	0,5	44	0,5	70	0,9	151	0,6
e	Elision (coord)	0	0,0	1	0,0	1	0,0	2	0,0
	TOTAL	7856	100,0	8972	100,0	7468	100,0	24296	100,0

(8) Exemples de sujets (français L1, grande section de maternelle) :

- a. là on dirait qu'elle marche. (CAR, 08-09II10b, 5;2)
- b. Papi et Mamie i(ls) sont venus à ma maison. (ELE, 08-09I5a, 5;4)
- b'. ça c' est un micro. (LIZ, 08-09I4a, 5;0)
- c. Q [*] faut bien respirer. (WIL, 08-09II7a, 4;9)
- d. là son copain à Juliette tape des pieds et des mains. (NIN, 08-09X56c, 5;4)
- e. Maman prépare un gâteau au chocolat fourré de fraises et de bonbons <et goûte> [/] et _ le goûte. (KEL, 08-09I5c, 5;4, lecture)

Le Tableau 5 montre que le constituant de loin le plus fréquemment utilisé est le clitique nominatif : cet élément peut émerger seul (a : 64 %-71,6 %-76,2 %) ou en cooccurrence avec un DP (syntagme nominal ou pronom tonique ; b-b' : 31,7 %-27,2 %-22,7 %). Les totaux des deux configurations (95,7 %-98,8 %-98,9 %) montrent que le clitique nominatif est présent dans quasiment toutes les phrases déclaratives des enfants pendant toute la période. L'utilisation de ce constituant est donc robuste et stable. En revanche, les structures d. et e. sont très rares dans le système enfantin et le restent pendant toute la période (0,5 %-0,5 %-0,9 %). Ce qui est intéressant est que les structures d. et e. sont emblématiques du français standard. Le détail de ces différentes formes du sujet chez le jeune enfant (a-b-b' vs d-e) a fait l'objet de plusieurs publications depuis ma thèse, qui sont rapportées dans la Section 2. Enfin, la structure c. du Tableau 5 (3,8 %-0,7 %-0,2 %) représente une absence totale de sujet, sous quelque forme que ce soit, et est considérée comme agrammaticale en français standard. Il s'agit d'un phénomène développemental très documenté dans la littérature acquisitionnelle connu sous l'appellation du « phénomène du sujet nul enfantin » (cf. Section 1 ci-après).

1. Phénomène du sujet nul enfantin

Cf. pages 59-82/226 du Volume 2 : Palasis, K. (2012b). *Economy Principles within Child Speech: When the Nominative Clitic does not Surface*. In S. Ferré, P. Prévost, L. Tuller, & R. Zebib (Eds.), *Selected Proceedings of the Romance Turn IV Workshop on the Acquisition of Romance Languages* (pp. 261-283). Newcastle: Cambridge Scholars Publishing, <https://www.researchgate.net/profile/Katerina-Palasis/publications>

Le phénomène du sujet nul enfantin, dont quelques exemples sont rapportés en (9), est probablement le phénomène acquisitionnel qui a fait couler le plus d'encre dans la littérature générative, notamment dans le cadre de la Théorie des Principes & Paramètres depuis les travaux fondateurs de Hyams (1986) sur l'anglais enfantin (cf. Section I.3.1 du Chapitre I).

(9) Phénomène du sujet nul enfantin en français (grande section de maternelle) :

- a. 0 [*] faut courir le [/] le premier. (WIL, 08-09I2b, 4;9, 0 = il explétif)
- b. 0 [*] me rappelle plus. (MAT, 08-09I4a, 4;11, 0 = je)
- c. 0 [*] veux commencer que l(e) tennis pa(r)c(e)+que j(e) fais que du tennis. (MAS, 08-09IV21b, 4;7, 0 = je)

On doit cet intérêt des chercheurs au fait que ce phénomène présente une série de défis, rapportés en (10) ci-après.

(10) Défis du phénomène du sujet nul enfantin :

- i. Expliquer pourquoi les enfants dont l'environnement linguistique regorge de verbes conjugués précédés d'un sujet omettent toute forme de sujet dans un nombre non-négligeable de leurs phrases (environ 25% à 30% vers 2;5, Rasetti, 1996).
- ii. Proposer une hypothèse qui rende compte non seulement des omissions mais qui soit également compatible avec les 70% de sujets présents.
- iii. Présenter une explication qui soit également compatible avec l'aspect développemental du phénomène, c'est-à-dire que les omissions diminuent progressivement pour disparaître vers l'âge de quatre ans.
- iv. Idéalement, rendre également compte de l'aspect inter-linguistique du phénomène, dans la mesure où ces omissions sont bien documentées non seulement en anglais et en français mais également en allemand, danois

(Hamann & Plunkett, 1998), basque (Ezeizabarrena, 2003), russe (Gordishevsky & Schaeffer, 2002), etc.

Les recherches publiées sur cette problématique depuis ma thèse représentent une première illustration de la rencontre entre données considérées en tant que système avec une dynamique endogène et théorie générative.

1.1 Apport des données

Dans Palasis (2012b), les données de la petite section de maternelle sont exploitées (2;5-4;0) et confrontées aux données de la toute petite section précédemment analysées (2;3-3;1, Palasis, 2005 ; Palasis & Oliviéri, 2007). Ces travaux mettent en avant des particularités de l'architecture du système verbal enfantin comparé au système adulte et des caractéristiques développementales entre les deux années de maternelle.

Concernant les différences entre enfants et adultes, il est notable que, pour les deux années, les paradigmes verbaux enfantins sont essentiellement composés de la première et de la troisième personne du singulier, avec seulement 1,2 % puis 1,4 % des formes verbales au pluriel. La seconde différence concerne l'inventaire verbal très réduit chez les jeunes enfants, le verbe *être* représentant à lui seul 48,7 % du système en toute petite section. Le croisement de ces deux caractéristiques génère alors un système avec un ratio de formes verbales morphologiquement distinctives beaucoup plus élevé chez les jeunes enfants que chez les adultes qui utilisent cinq à six personnes différentes (le *nous* a tendance à disparaître à l'oral) et de nombreux verbes du premier groupe qui présentent un nombre dominant de formes non-distinctives dans chaque paradigme (par exemple, *je-tu-il-ils* [mãʒ] « mange/s/nt »).²⁹ Ainsi, 72,9 % des formes verbales sont distinctives en toute petite section (notamment, [sui] « suis » vs [e] « est », cf. Tableau 1, p. 62/226). Il est alors proposé que c'est l'architecture particulière de ce système verbal enfantin qui crée une configuration propice au phénomène

²⁹ Les verbes du premier groupe représentent 87 % du système verbal français et 47 % des occurrences verbales en termes de fréquence (New, Pallier, Ferrand & Matos, 2001, cité par Legendre, Barrière, et al., 2010).

du sujet nul (22,4 % en toute petite section). Cette hypothèse fait la prédiction que lorsque les paradigmes verbaux et l'inventaire verbal s'étoffent dans le cadre du développement linguistique de l'enfant, le phénomène du sujet nul régresse.

L'analyse des données d'enfants un peu plus âgés (petite section : 2;5-4;0) montre des évolutions entre les deux années, notamment le développement du pronom *tu* (de 2,9 % à 6,3 % entre les deux années) et une diversification du système verbal : le verbe *être* qui était omniprésent (48,7 %) perd du terrain (36,7 %) au profit d'autres verbes tels que *aller*, *pouvoir*, *falloir* et *mettre* (cf. Tableau 2, p. 63/226). Ces verbes, bien que ne relevant toujours pas du premier groupe de conjugaison, présentent néanmoins davantage de formes identiques (par exemple, [pø] « peut/peut » pour les trois formes du singulier). Ces deux développements entraînent alors une modification de l'architecture du système au niveau du ratio des formes verbales distinctives qui chute de 72,9 % à 14,01 % entre les deux années. Si le taux de sujets nuls est effectivement lié à celui des formes verbales distinctives dans le système, alors on s'attend à une chute des sujets nuls dans une proportion similaire à celle des verbes (cinq fois moins de formes distinctives). Les données semblent confirmer l'hypothèse car, en petite section, le taux de sujets nuls non seulement chute (de 22,4 % à 3,9 %), mais il chute dans une proportion similaire à celle des formes distinctives (cinq fois moins). L'hypothèse d'un lien de cause à effet entre formes verbales morphologiquement distinctives et sujets nuls peut alors être maintenue. Il me semble que l'hypothèse proposée dans Palasis (2012b) et récapitulée dans le Tableau 6 ci-après résout alors les défis (i) et (iii) mentionnés en (10) ci-dessus sur l'omission de sujets malgré leur présence dans l'environnement linguistique et sur les aspects développementaux de l'explication fournie.

Tableau 6 : Hypothèse d'un lien entre formes verbales et sujets nuls

	TPS	PS
Âges	2;3-3;1	2;5-4;0
Formes verbales distinctives (%)	72,9	14,0
Sujets nuls (%)	22,4	3,9

Concernant le défi (ii) en (10), l'extraction fine des formes verbales enfantines, par verbe, par personne et par enfant, met en avant une caractéristique supplémentaire du système verbal enfantin (cf. Tableaux 9 à 12, p. 70-73/226). Les résultats en petite section montrent en effet que le phénomène du sujet nul n'affecte qu'une seule forme à l'intérieur de chaque paradigme (formes en gras, Tableau 9, p. 70/226). Le paradigme du verbe *vouloir* par exemple comporte les trois formes du singulier (1sg, 2sg, 3sg), mais le phénomène du sujet nul n'affecte que la forme 1sg. Pour le verbe *être*, qui est également produit aux trois formes du singulier, seule la 3sg est affectée. Les personnes affectées varient donc d'un verbe à un autre, mais l'analyse met en lumière un point commun à ces omissions : elles affectent systématiquement la personne la plus fréquente dans chaque paradigme en fonction de chaque enfant. Les enfants maintiennent donc le pronom avec les formes les moins fréquentes dans chaque paradigme et ce processus est dynamique, c'est-à-dire que lorsqu'un nouveau verbe ou un nouveau temps entre dans le système d'un enfant, le processus s'applique à l'identique (cf. Tableau 14, p. 77/226 : lorsque Antoine commence à produire le verbe *vouloir* à l'imparfait aux trois personnes du singulier, il omet uniquement le pronom avec la forme verbale la plus fréquente).

Enfin, dans Palasis (2012b), il est également remarqué que certains verbes enfantins ne sont pas du tout affectés par le phénomène du sujet nul (par exemple, *aller*). Les données montrent alors que le maintien des pronoms avec la forme la plus fréquente (3sg) peut être lié avec la diversité des pronoms de 3sg utilisés avec *aller* (*il, elle, on, ça, wh-* interrogatifs et relatifs), qui est en contraste avec l'uniformité des pronoms avec *être* (*ce* accompagnant à lui seul 76 % des formes, cf. Graphiques 1 et 2 p. 75/226). La non-omission du pronom avec *aller* semble alors pouvoir s'expliquer par rapport à l'apport supplémentaire d'information de chaque pronom (masculin, féminin, voire pluriel avec *on*, etc.). Ces deux derniers points (un seul pronom omis par paradigme et certains paradigmes intacts) semblent alors résoudre le défi (ii) en (10) ci-dessus, en

expliquant pourquoi le système génère tout de même une majorité de formes verbales accompagnées d'un pronom (environ 70 % des formes entre 2;5 et 3;0).³⁰

Dans Palasis (2022), je mentionne brièvement les taux de sujets nuls dans les années de maternelle qui suivent (moyenne et grande sections). La base de données renvoie des taux de sujets nuls de plus en plus bas au fil des années (0,7 % puis 0,2 %). Ces taux illustrent ainsi la disparition quasi-totale du phénomène dès la moyenne section (3;6-4;11).

1.2 Contribution à la théorie

L'hypothèse, issue de l'étude détaillée de données enfantines, d'un lien de cause à effet entre l'architecture particulière du système verbal enfantin et le phénomène du sujet nul a également été traitée au niveau théorique dans Palasis (2012b, p. 75-78/226), dans le cadre de la théorie générative (cf. I.3.1 dans le Chapitre I).

L'omission du clitique, qui n'apporte aucune information grammaticale supplémentaire en termes de [Personne] lorsque la forme verbale est morphologiquement distinctive, suggère un principe sous-jacent d'économie. Ce principe d'économie s'applique dans les premiers paradigmes à deux formes distinctives où aucun clitique n'est réalisé (par exemple, [sɥi] « suis » vs [e] « est »), puis dans les paradigmes plus étoffés à trois formes où deux clitics suffisent à distinguer les trois formes qui ne sont plus morphologiquement distinctives (par exemple, [pø] « peux », [tɥpø] « tu peux », [ipø] « i(l) peut »). Les notions d'« économie » et de « moindre effort » sont récurrentes dans la théorisation générative qui stipule que la computation syntaxique et les déplacements sont cognitivement coûteux et qu'ils sont réalisés uniquement s'ils sont nécessaires (cf. Chomsky, 1981, 1995, 1998, 1999, entre autres). Ces

³⁰ Le point (iv) sur l'aspect inter-linguistique avait été abordé dans Palasis & Oliiviéri (2007) : Benveniste (1946) avait noté que les formes verbales 3sg sont souvent différentes des autres formes dans les langues du monde (un exemple emblématique étant celui de l'anglais +s). Si les enfants passent par une étape commune pendant laquelle leurs paradigmes verbaux sont initialement réduits aux personnes 1sg et 3sg, alors l'hypothèse peut être étendue à d'autres langues.

principes d'économie computationnelle sont également souvent évoqués dans la littérature acquisitionnelle (cf. Gelderen, 2008, 2009 ; Jakubowicz, 2011; Radford, 2000, entre autres).

L'hypothèse défendue dans Palasis (2012b) rejoint ces propos en décrivant le phénomène du sujet nul enfantin comme une illustration supplémentaire des principes d'économie. L'économie est appliquée de façon ciblée à son propre système linguistique par l'enfant au niveau informationnel : économie quantitative (omission du clitique avec la forme verbale la plus fréquente dans chaque paradigme) et qualitative (le clitique n'est pas nécessaire pour interpréter le trait de [Personne] sur le verbe). Avant de pouvoir franchir le pas et de postuler une économie dérivationnelle, il faudrait d'abord définir précisément ce que ces clitiques représentent dans le système enfantin (et peut-être plus largement dans le système du français oral hexagonal contemporain) et quel type de constituant ils représentent dans la dérivation. Cette problématique est traitée dans le cadre de l'analyse des différentes formes du sujet en Section 2 ci-dessous et le phénomène du « sujet » nul enfantin est reformulé dans la synthèse (cf. Section II.3).

1.3 Questions ouvertes

Deux questions principales sont laissées en suspens à ce stade suite à l'hypothèse d'un lien de cause à effet entre l'architecture particulière du système verbal enfantin et le phénomène du sujet nul. Il s'agit des questions de l'effet de fréquence et du développement de l'inventaire verbal des enfants (cf. Section 1.1 ci-dessus).

Tout d'abord, la question de l'effet de la fréquence d'un item sur un système linguistique pourra être développée. En effet, la fréquence est considérée comme un facteur important dans le processus acquisitionnel : elle est notamment fondamentale dans les approches basées sur l'usage (cf. Section I.3.2 du Chapitre I) et a également été intégrée dans les approches « GU + statistiques » (cf. Section I.3.4). Néanmoins, il est important de préciser que la fréquence largement

discutée dans la littérature acquisitionnelle est celle mesurée dans les données qui composent l'environnement linguistique de l'enfant, généralement appelée *input frequency* (cf. notamment Ambridge, Kidd, Rowland & Theakston, 2015, p. 242), que nous traduirons par « fréquence perçue ». La fréquence mesurée dans la production de l'enfant (*output frequency*, que nous appellerons « fréquence générée ») est très peu discutée dans la littérature et, lorsqu'elle l'est, l'effet de fréquence considéré est une « meilleure performance avec des séquences qu'ils [les enfants] produisent plus souvent » (Ambridge et al., 2015, p. 242). Le lien de cause à effet entre l'architecture particulière du système verbal enfantin et le phénomène du sujet nul suggère un autre effet de fréquence : celui de l'application d'un principe d'économie, qui au lieu de rapprocher le système enfantin du système adulte (qui présente de nombreuses redondances) l'en distingue en éliminant une redondance de marque de [Personne] devant le verbe lorsque la forme verbale est morphologiquement distinctive. La première série de questions ouvertes est : est-ce cognitivement réaliste ? Existe-t-il d'autres phénomènes dans le système enfantin qui illustrent ce processus ? Pouvons-nous créer des contextes expérimentaux qui pourraient tester cet effet particulier de la fréquence générée ?

La seconde question ouverte à ce stade concerne le développement de l'inventaire verbal des enfants. Depuis la publication de Palasis (2012b) qui propose que le phénomène du sujet nul est lié à un système verbal quantitativement réduit chez le jeune enfant, la question de l'influence de l'architecture particulière de ce système a de nouveau surgi dans mes travaux à l'occasion de l'étude de la position des items *wh-* dans les interrogatives enfantines (Palasis, Faure & Lavigne, 2019; Palasis et al., 2023). La Section III ci-après reviendra sur ce point.

2. Statut morphosyntaxique du clitique nominatif

Cf. pages 103-114/226 du Volume 2 : Palasis, K. (2014). Étude morphosyntaxique des pronoms sujets à partir d'un corpus de maternelle. In R. Mir-Samii (Ed.), *Du sujet et de son absence dans les langues* (pp. 259-270). Le Mans: Université du Maine, <http://doi.org/10.13140/2.1.1586.9440>

Cf. pages 115-134/226 du Volume 2 : Palasis, K. (2015). Subject clitics and preverbal negation in European French: Variation, acquisition, diatopy and diachrony. *Lingua*, 161, 125-143. <http://dx.doi.org/10.1016/j.lingua.2014.11.012>

Comme mentionné dans la Section II.2.2 du Chapitre I, le statut des clitiques nominatifs (*je, tu, il*, etc.) en français oral adulte est débattu depuis plusieurs dizaines d'années. Heap, Oliviéri & Palasis (2017) ont proposé un chapitre de synthèse sur le débat (arguments syntaxiques vs marqueurs d'accord préverbaux), sur la comparaison entre les clitiques en français oral adulte et ceux des dialectes du nord de l'Italie, sur les tests mis en place par les linguistes en vue de trancher la question, sur les aspects diachroniques du statut et sur l'importance de l'analyse de la négation. Mes recherches sont entrées dans le détail de ces différents points grâce à l'étude des données enfantines de maternelle (Palasis, 2014, 2015, 2022 ; Oliviéri, Kaiser, Palasis, Zimmermann & Faure, 2020).

2.1 Application des tests aux données enfantines

Dans un premier temps, cette section complète les configurations mentionnées dans la Section II.2.2 du Chapitre I concernant les données adultes. Il s'agit des tests que les linguistes ont proposés afin de déterminer si les clitiques nominatifs sont des arguments syntaxiques ou des marqueurs d'accord préverbaux. Ces tests ont été mis en place initialement afin de trancher la question du statut des clitiques nominatifs dans les dialectes du nord de l'Italie (cf. notamment Poletto, 1993, 2000). Dans mes travaux, ils sont appliqués aux données enfantines, au fur et à mesure du traitement des données de maternelle, de la toute petite section (2;3-3;1) à la grande section de maternelle (4;5-5;11 ; cf. Section I.I.1 ci-dessus). Le récapitulatif dans le Tableau 7 est une synthèse inédite de l'ensemble des données de maternelle, de la toute petite à la grande section (N = 25 198 ; Palasis, 2005, 2010b, 2014, 2015, 2022 ; Oliviéri et al., 2020).

Tableau 7 : étude longitudinale complète des sujets enfants en maternelle

		TPS (2;3-3;1)		PS (2;5-4;0)		MS (3;6-4;11)		GS (4;5-5;11)	
		N	%	N	%	N	%	N	%
a	Clitique seul	459	50,9	5027	64,0	6427	71,6	5691	76,2
b	Clitique + DP	203	22,5	2492	31,7	2440	27,2	1692	22,7
c	Sujet nul	240	26,6	300	3,8	60	0,7	14	0,2
d	DP seul	0	0,0	37	0,5	44	0,5	70	0,9
e	Elision (coord)	0	0,0	0	0,0	1	0,0	1	0,0
	TOTAL	902	100,0	7856	100,0	8972	100,0	7468	100,0
a+b	TOTAL	662	73,4	7519	95,7	8867	98,8	7383	98,9

Test 1 : taux de présence du clitique

Le premier test porte sur la présence systématique d'un clitique nominatif avec un verbe fléchi. En effet, l'hypothèse d'un statut morphologique n'est soutenable que si les clitics sont présents massivement et indépendamment de la configuration syntaxique. S'ils sont marqueurs d'accord, on s'attend à les trouver de façon égale dans les propositions principales, subordonnées, juxtaposées et coordonnées (critère (3i) page 44). Les totaux a+b dans le Tableau 7 renvoient un taux élevé en TPS (73,4 %) et des taux très élevés en PS, MS et GS (95,7 %-98,8 %-98,9 %). La différence entre le premier et les autres taux est explicable par le taux de sujets nuls en TPS. Les taux indiquent une présence massive dans les énoncés déclaratifs enfantins, ce qui implique une présence égale dans tous les types de propositions illustrés en (11) ci-après, même lorsqu'un syntagme nominal est présent (cf. également les exemples (4)-(7), p. 107/226).

- (11) Présence massive des clitics nominatifs avec les verbes fléchis :
- elle joue elle danse elle met la musique et elle dort. (LUS, 3;10)
 - i nage i nage i nage et après il est jusque sur l(e) bateau là. (DYL, 4;2)
 - pa(r)c(e)+que pour pas que le gâteau i(l) colle. (LIZ, 5;0, 08-09I4c)
 - et ma maman elle dit que ça c' est les p(e)tits caniches. (LOU, 5;6, 08-09IV23b)

La présence quasi-systématique d'un clitique nominatif avec un verbe fléchi permet d'envisager l'analyse de ces éléments en français enfantin de manière similaire à celle des clitics dans les dialectes nord-italiens, c'est-à-dire des

marqueurs d'accord et non des pronoms syntaxiques (cf. Section II.2.2 du Chapitre I).

Test 2 : nature des éléments cooccurrents

Le second test porte sur le type d'élément co-occurent avec le clitique nominatif (cf. (3ii) p. 44 et les implications de Poletto, 1993, 2000 p. 48). Les éléments cruciaux dans ce test sont les éléments indéfinis³¹ car ces derniers ne peuvent apparaître qu'en position sujet et non disloquée (cf. p. 40 sur la dislocation vs le doublement). Lorsqu'un élément indéfini est doublé par un clitique, le clitique ne peut donc pas apparaître en position sujet, ce qui est un argument en faveur de son analyse en tant que marqueur d'accord. Le problème majeur soulevé par ce test est la rareté des éléments indéfinis en position sujet car, en général, ils sont plutôt introduits au niveau du prédicat (par exemple, *le garçon entend quelqu'un* vs *?quelqu'un entend le garçon*). La structure informationnelle des jeunes enfants ne déroge pas à la règle et 98,6 % des sujets en petite et moyenne sections sont définis (Palasis, 2015, p. 122/226). Néanmoins, lorsqu'un syntagme nominal indéfini émerge en position sujet (MS : N = 43), les données montrent qu'il est majoritairement doublé par un clitique (76,7 %), comme illustré en (12).

(12) Syntagmes nominaux indéfinis doublés par un clitique :

- a. et même qu'un Pokemon il a des crocodiles . (MAS, 3;9, 07-08V)
- b. parc(e) que personne i(l) m(e) l'a dit . (MAT, 4;5, 07-08IX)
- c. ben un squelette il est mort hein . (LSN, 5;1, 08-09IV19b)

Un parallèle intéressant est à noter ici avec les échelles d'implications de Poletto (1993, 2000). Dans les dialectes du nord de l'Italie, lorsqu'un dialecte double les syntagmes nominaux définis, alors il double également tous les pronoms toniques (cf. (4ii) p. 48). Dans le système des enfants étudiés (MS, Palasis,

³¹ Sont considérés comme indéfinis les syntagmes dont le déterminant porte uniquement une valeur quantifiante (par exemple : *un, une, des, tous, chaque, aucun*) et les pronoms indéfinis (par exemple : *personne, rien*) (Grevisse & Goosse, 2008, p. 738).

2015, p. 122-123/226), les mêmes implications existent : tous les enfants de MS qui doublent les syntagmes nominaux définis en position sujet (N = 17/17) doublent également les pronoms toniques. Le taux de doublement de ces éléments va de 88 % à 100 % en fonction de l'enfant. L'implication (4iii) de Poletto (1993, 2000) concerne les syntagmes nominaux indéfinis : si ces derniers sont doublés dans un dialecte, alors les syntagmes nominaux définis et pronoms toniques le sont également. Le système enfantin semble fonctionner avec la même implication car les enfants qui produisent des syntagmes nominaux indéfinis en position sujet (N = 12/17) et les doublent (N = 11/17) doublent également les syntagmes nominaux définis et les pronoms toniques. Le taux de doublement des indéfinis est plus hétérogène en fonction de l'enfant, de 50 % à 100 %.

Enfin, l'implication (4iv) concerne les variables *wh-* : si un élément relatif, interrogatif ou clivé est doublé, alors tous les autres types de sujets le sont aussi. Cette dernière implication est très intéressante à étudier en français hexagonal car les adultes francophones ne produisent pas du tout ce type de doublement dans les relatives (par exemple : **le chien qui il aboie* ; cf. Labelle, 1988). Cette absence totale dans l'environnement linguistique permet alors de faire la prédiction que les enfants francophones ne devraient pas doubler les variables *wh-* sauf si dans leur grammaire le clitique est un marqueur préverbal d'accord (qui n'est donc pas sensible à l'environnement syntaxique). Par ailleurs, en français québécois, la non-élision du pronom relatif *qui* devant voyelle est interprétée par Auger (1994) comme un argument en faveur de l'analyse morphologique du [i] non-élide, et deux des trois formes interrogatives avec deux *qui* mises en avant par Auger (1994) pour défendre l'hypothèse morphologique en français québécois (cf. Section II.2.2 du Chapitre I) existent en français hexagonal (*qui est-ce qui, qui c'est qui*).

Contrairement à la prédiction énoncée pour les pronoms relatifs, les données enfantines de moyenne section renvoient un taux de doublement des pronoms relatifs de 2,6 %, ce qui est beaucoup plus bas que les autres taux de doublement

mais qui représente tout de même 13,7 % de la totalité des items doublés (cf. Graphique 3, p. 108/226). Les exemples p. 108/226 sont repris en (13) a. et b. ci-dessous. Ce type de doublement concerne quatre enfants sur les dix-sept du groupe étudié, avec des taux entre 9,1 % et 20 % en fonction de l'enfant. La forme du pronom relatif *qui* est majoritairement non-élidée devant voyelle (cf. (13) c. et d.) et 11/17 enfants présentent un taux de non-élision de 100 % en moyenne section.

(13) Pronoms relatifs doublés par un clitique :

- a. ma mère qui il a dit qu'y aura des médailles. (MAS, 3;9)
- b. le petit renard qui il a faim. (MAI, 3;2)
- c. j'ai vu un cochon qui est parti. (ENZ, 4;1)
- d. non c' est toi qui es toc toc . (LUS, 08-0911a)

Concernant les pronoms interrogatifs, tous les enfants de moyenne section utilisent majoritairement des formes qui présentent deux *qui* (79,5 % des interrogatives sujets, comme illustré en (14)) et seuls cinq enfants sur les dix-sept produisent la forme simple de *qui* (20,5 %).

(14) Formes complexes du pronom interrogatif *qui* :

- a. c'est qui qui a la vache ? (ELE, 4;9)
- b. qui c'est qui était là ? (DYL, 3;8)
- c. qui est-ce qui m'a ouvert ça ? (ENZ, 4;1)

Les cooccurrences mises en avant par ces données montrent que les enfants doublent : (i) leurs pronoms toniques (99,9 % en MS), ce qui correspond au Type 1 des dialectes nord-italiens ; (ii) leurs syntagmes nominaux définis (97,5 % en MS), ce qui correspond au Type 2 ; (iii) une grande partie des syntagmes nominaux indéfinis (76,7 % en MS), ce qui correspond au Type 3 ; (iv) une grande partie de leurs items interrogatifs (79,5 % de formes complexes de *qui*) et possiblement relatifs (100 % de non-élision de *qui* pour 11/17 enfants), ce qui correspond au Type 4. Si nous suivons les implications de Poletto (1993, 2000), le système de doublement des jeunes enfants présente alors de grandes similitudes avec celui

des dialectes de Type 3 (milanais, lombard de l'est), voire Type 4 (friulien, piémontais, ligurien). Culbertson (2010, p. 121) avait proposé que le français informel adulte appartienne au Type 2 ; je propose que le français enfantin appartienne au Type 3 ou 4 (Palasis, 2015, p. 123/226).

Cette hypothèse est renforcée par un argument géographique : la répartition diatopique des dialectes nord-italiens du Type 1 au Type 4, suit une progression qui va d'est en ouest, pour arriver aux abords de la frontière franco-italienne avec les dialectes du Piémont et de Ligurie. Il se trouve que les données de maternelle que j'étudie proviennent du sud-est de la France, dans la continuité géographique de ce modèle de dispersion. De nombreuses données recueillies dans le nord de la France et en Belgique montrent que les systèmes plus au nord sont différents (De Cat, 2002, 2007 ; Villeneuve & Auger, 2013). La Section II.2.3 ci-dessous rapportera les résultats d'une étude d'une zone géographique intermédiaire, celle des dialectes nord-occitans (Olivieri et al., 2020).

Test 3 : position du clitique par rapport au verbe fléchi

Les analyses syntaxique et morphologique des clitics nominatifs font des prédictions différentes quant à la position relative du clitique par rapport au verbe fléchi. Lorsque le clitique est analysé comme un véritable pronom, argument syntaxique du verbe, celui-ci peut émerger à gauche ou à droite du verbe, par proclise ou enclise suite à une inversion du verbe et du clitique (par exemple, *il dort* ou *dort-il ?*). En revanche, l'analyse morphologique du clitique en tant que marqueur préverbal limite la position du clitique à la proclise (cf. (3iii) page 44). Le test de position permet donc de nourrir le débat sur le statut de cet élément chez les locuteurs du français informel. Culbertson (2010, p. 101) par exemple avait rapporté une absence totale d'inversion verbe-clitique dans les interrogatives en LAE, argumentant ainsi en faveur de l'analyse morphologique du clitique.

Ce test de position appliqué aux données enfantines donne des résultats également très tranchés puisque, sur les trois années de maternelle, la base de données ne renvoie qu'un très maigre total de huit clitics postverbaux produits

par trois enfants sur les seize du groupe. Quelques exemples sont présentés en (15). La position de ces huit clitiques oblige à les analyser en tant que pronoms syntaxiques, mais elle n'empêche pas de considérer les 23 761 clitiques préverbaux restants (cf. Tableau 7 ci-dessus, PS-GS) comme possibles marqueurs d'accord.

(15) Clitiques enfantins postverbaux :

- a. euh: 0 [*] s(e) passe t il ? (VIC, 06-07IV19d)
- b. où es tu Maman Ours ? (CAR, 06-07VII30b)
- c. le vois tu ? (KEL, 08-09IV24c, lecture)

Par ailleurs, les contextes de production de ces huit clitiques postverbaux ont été examinés et il semble que leur émergence puisse être expliquée par des contextes bien particuliers qui entraînent la production d'un français plus formel : la production de comptines, de chansons ou la narration, voire la lecture, à partir de livres (Palasis, 2011b, 2013). L'émergence de ce français plus formel chez certains enfants du groupe sera détaillée dans la section sur la diglossie (cf. Section IV).

Test 4 : élision de [il] devant consonne

Les données sociolinguistiques collectées auprès d'adultes francophones ont montré des ajustements phonologiques des clitiques nominatifs au contexte en français informel mais pas en français standard. Le clitique *il* a notamment été largement étudié dans ce cadre (Morin, 1979 ; Ashby, 1984 ; Culbertson, 2010 pour une synthèse). L'ajustement phonologique par rapport au contexte subséquent entraîne une variation des formes de *il* qui émerge sous la forme pleine [il] devant voyelle et sous la forme réduite [i] devant consonne (79,6 %, N = 2 093/2 629 ; Ashby, 1984). Cet ajustement, présent en français informel uniquement, est attendu dans le cadre de l'analyse morphologique du clitique (cf. point iv page 44). Le test d'alternance de formes a donc également été appliqué à la base de données enfantines de maternelle au fur et à mesure du traitement des données (Palasis, 2010b, 2013, 2014, 2015, 2022) et un récapitulatif inédit des

formes du clitique *il* pour les trois années de maternelle est proposé dans le Tableau 8 ci-après avec des exemples en (16).

Tableau 8 : ajustement du clitique *il* au contexte phonologique

	PS (2;5-4;0)		MS (3;6-4;11)		GS (4;5-5;11)	
	N	%	N	%	N	%
[il] / _V	483	35,7	861	40,0	514	43,2
[i] / _V	0	0,0	0	0,0	0	0,0
[i] / _C	857	63,3	1281	59,5	653	54,8
[il] / _C	14	1,0	11	0,5	24	2,0
TOTAL	1354	100,0	2153	100,0	1191	100,0

(16) Formes du clitique *il* :

- a. [il] / _V : il est à moi. (WIL, 2;9)
- b. [i] / _C : l'escargot i dort. (LUS, 2;9)
- c. [il] / _C : il ne chante plus. (MAI, 3;2, chant)

Les données montrent que la règle d'ajustement de la forme de *il* en fonction du contexte subséquent est également active et robuste chez les jeunes enfants : leurs clitiques *il* sont toujours pleins devant voyelle et la quasi-totalité de leurs *il* est élidée devant consonne (98,4 %-99,1 %-96,5 % ; cf. Tableau 9). Tous les enfants étudiés pendant les trois années (16/16) appliquent cette alternance [il] / _V et [i] / _C, ce qui va dans le sens de l'analyse morphologique du clitique.

Tableau 9 : ajustement de *il* devant consonne

	PS (2;5-4;0)		MS (3;6-4;11)		GS (4;5-5;11)	
	N	%	N	%	N	%
[i] / _C	857	98,4	1281	99,1	653	96,5
[il] / _C	14	1,6	11	0,9	24	3,5
TOTAL	871	100,0	1292	100,0	677	100,0

Le Tableau 9 indique néanmoins qu'il existe des exceptions à cette règle et que leur ratio double entre la petite et la grande section de maternelle (1,6 %-3,5 %). Les données indiquent également que seuls douze enfants sur seize produisent la forme non-élidée de *il* devant consonne et que cette forme rare mais en augmentation est peut-être elle-même conditionnée par un autre contexte,

syntactique et non plus phonologique, car il semblerait qu'il s'agisse notamment de l'émergence du clitique préverbal *ne*, détaillée dans la Section 2.2 ci-dessous.

2.2 Retour sur la négation préverbale

Les données du français adulte ont montré des taux d'omission de la particule préverbale *ne* particulièrement élevés et une correspondance entre l'omission de *ne* et la présence du clitique nominatif en français oral en général et en LAE en particulier (93,7 % d'omission après un clitique nominatif ; cf. Section II.2.3 du Chapitre I). Rappelons que la présence de *ne* est un argument fort présenté contre l'analyse morphologique du clitique car lorsque *ne* est présent en syntaxe, il sépare le clitique du verbe (De Cat, 2005, 2007 ; Culbertson, 2010 ; récapitulatif de la problématique dans Palasis, 2015, p. 118-119/226). La présence de *ne* représente alors également un test dans le cadre du débat sur le statut morphosyntaxique des clitiques nominatifs.

Les sections précédentes ont montré que le clitique nominatif est omniprésent dans les phrases déclaratives enfantines (98,9 % en grande section ; cf. Tableau 7). On peut alors s'attendre à un taux très bas de présence de *ne*, si les enfants suivent le modèle des adultes. Les données de maternelle nous éclairent sur ce point (Palasis, 2013, 2015, 2022) et un récapitulatif des ratios des formes simples et discontinues de la négation verbale infantine est proposé dans le Tableau 10, avec des exemples en (17). L'adverbe négatif postverbal *pas* est utilisé dans la quasi-totalité des énoncés négatifs pendant les trois années (88,9 %). Les enfants produisent également quelques *plus* (7,7 %) et quelques très rares occurrences de *jamais* (1,7 %), *que* (1,6 %) et *rien* (0,1 %).

Tableau 10 : formes de la négation

	PS (2;5-4;0)		MS (3;6-4;11)		GS (4;5-5;11)	
	N	%	N	%	N	%
Simple	870	98,3	931	97,9	923	98,8
Discontinue	15	1,7	20	2,1	11	1,2
TOTAL	885	100,0	951	100,0	934	100,0

Les données montrent des taux très bas de production de *ne* qui accompagnent ces adverbes postverbaux pour les trois années (Total N =

46/2770 ; 1,7 %-2,1 %-1,2 %). Les taux d'omission rejoignent alors les taux d'omission les plus élevés relevés chez les adultes (Armstrong & Smith, 2002 : 98,9 %) et ils dépassent ceux observés en LAE hors contexte scolaire (Culbertson, 2010 : 93,7 % ; cf. Section II.2.3 du Chapitre I).

- (17) Formes simples et discontinues de la négation verbale en grande section :
- a. j'ai pas fait l(e) papillon. (NIN, 5;2, 08-09VIII45b)
 - b. j(e) me rappelle plus. (LAN, 4;8, 08-09I2c)
 - c. ah non je n' ai pas. (CAR, 5;8, 08-09X58b)
 - d. alors y+a marqué alors "<ah ah tu> [/] ah ah tu ne mont(e)ras pas". (KEL, 5;6, 08-09IV24c, lecture)

Par ailleurs, les taux de production de *ne* sont stables sur les trois années, ce qui montre qu'il n'y a pas d'augmentation de l'utilisation de la négation discontinue chez ce groupe d'enfants au fil de ces années et qu'elle reste donc globalement inférieure à celle des adultes. Une autre caractéristique de la négation discontinue est la variabilité individuelle, entre enfants d'une part et dans le développement de chaque enfant d'autre part. En effet, seuls quatre enfants sur seize produisent le *ne* préverbal en petite section. Ce total passe à dix en moyenne section mais redescend à six en grande section. Cela signifie que certains enfants qui avait produit le *ne* préverbal en MS ne l'ont plus produit en GS et, des quatre enfants qui avaient produit la forme en PS, aucun ne l'a produite en GS. Enfin, quatre enfants ne la produisent jamais pendant les trois années d'enregistrements. Les données montrent donc que la production de la négation discontinue, forme emblématique du français standard, est loin d'être acquise en maternelle.

D'autres résultats tout récents de maternelle (Buson et al., 2022 ; 61 élèves) renvoient des taux de production de la particule préverbale *ne* également très bas (PS : 1,9 % et 2,6 % ; MS : 2,1 %, 1 % et 0,9 % ; GS : 3,2 %). Les auteurs rapportent que l'hypothèse d'une augmentation de la production de *ne* avec l'âge n'est pas non plus validée dans ce corpus dans la mesure où l'écart significatif entre MS et GS est à attribuer à un seul enfant en situation de comptine. Ces

résultats rejoignent donc ceux de Palasis (2013, 2015, 2022) sur deux caractéristiques majeures de l'utilisation du *ne* préverbal chez les enfants de maternelle : la variation individuelle est très importante et l'émergence de *ne* semble conditionnée par la situation de production. Ce dernier point est développé dans la section sur la diglossie (cf. Section IV).

2.3 Perspectives diatopique et diachronique

Dans le cadre du projet ANR-DFG 2012-2015 DADDIPRO³², la perspective acquisitionnelle de l'étude des clitiques nominatifs a été confrontée à une perspective diatopique (dialectes nord-occitans) et diachronique (ancien français).

La Section II.2.1 ci-dessus a proposé une hypothèse de continuité géographique dans le modèle de dispersion des types de cooccurrences dans les dialectes nord-italiens, allant des dialectes de Type 1 le plus à l'est vers les dialectes de Type 4 à la frontière franco-italienne. Nous avons ajouté à ce modèle de dispersion le français enfantin qui présente des caractéristiques des Types 3 et 4 et qui se trouve à quelques kilomètres à l'ouest de la frontière franco-italienne. Dans le projet DADDIPRO, les clitiques nominatifs dans les dialectes nord-occitans ont été soumis aux mêmes tests afin de parfaire nos connaissances sur leur statut morphosyntaxique (Oliviéri et al., 2020). L'étude a montré que les clitiques nord-occitans ne sont pas systématiquement présents avec un verbe fléchi (paradigmes partiels), qu'ils ne doublent jamais un DP indéfini, qu'ils peuvent être pré- ou post-verbales et qu'ils peuvent être séparés du verbe par un marqueur négatif préverbal. Contrairement aux clitiques des dialectes nord-italiens et du français enfantin, les clitiques nord-occitans sont donc analysés comme des arguments syntaxiques, ce qui est congruent avec la description de systèmes linguistiques plus au nord (Picardie et Belgique ; Auger, 2022 ; De Cat, 2005).

Les aspects diachroniques du statut morphosyntaxique des clitiques nominatifs en français ont été examinés à partir de textes depuis l'ancien français

³² « Données et analyses dialectologiques, acquisitionnelles et diachroniques des pronoms sujets en Gallo-Roman », sous la direction de Michèle Oliviéri pour l'Université de Nice et de Georg Kaiser pour l'Université de Konstanz, <https://anr.fr/Projet-ANR-11-FRAL-0007>.

(synthèse dans Olivieri et al., 2020) et à partir de données orales produites par le jeune roi Louis XIII au XVII^e siècle et recueillies par son médecin personnel (Foisil, 1989 ; Palasis, 2015, p. 127-130/226).

Concernant l'ancien français, Olivieri et al. (2020) concluent qu'il existe bien un paradigme de pronoms clitiques (*jo(u), tu, i(l), ele, nos, vos, i(l), eles*) qui est morphologiquement presque identique au paradigme des pronoms toniques, et que les clitiques à ce stade occupent la position syntaxique du sujet. Zumwald Küster (2018) quant à elle mentionne un changement du statut des clitiques vers le XV^e siècle, lorsqu'ils sont passés de formes toniques à formes clitiques. L'élision de *i(l)* est déjà néanmoins attestée dès le XII^e siècle (Pope, 1934) et Chervel (2011) mentionne que la réintroduction de la forme pleine [il] devant consonne s'est faite par le biais de l'enseignement scolaire du français standard. Ce dernier point sera développé dans la section sur la diglossie (cf. Section IV).

Concernant les données enfantines, l'éventuel lien entre les formes de la négation et la présence de clitiques nominatifs a été étudié entre 2;0 et 3;3 chez le jeune roi Louis XIII (Palasis, 2015, p. 127-130/226). Les énoncés négatifs (N = 67) montrent la production de quatre formes différentes dès 2;3, avec une dominance nette pour la forme discontinue (70,1 %), suivie de la forme postverbale (23,9 %) et quelques occurrences de la forme préverbale seule ou accompagnée d'un déterminant négatif (6 %). Par ailleurs, les clitiques nominatifs sont omniprésents dans ce corpus. Par contre, il ne semble pas y avoir de lien dans ce corpus entre l'omission du marqueur négatif *ne* et la présence du clitique puisque tous les énoncés négatifs sauf deux comportent un clitique. Également, à la différence du français enfantin contemporain, les clitiques ne sont pas doublés chez le jeune roi Louis XIII. L'élision du clitique *il* a également été examinée et les données montrent que l'alternance est active (N = 27, 96,3 %) et qu'elle est systématique avec le *il* explétif.

Les données acquisitionnelles du XVII^e siècle montrent donc des différences avec celles du XXI^e siècle : les clitiques nominatifs ne sont pas doublés et leur présence ne semble pas affecter la forme de la négation chez le jeune roi Louis

XIII. Les données de ce corpus indiquent néanmoins l’omniprésence des clitiques nominatifs, l’alternance phonologique de *il* et la variation dans les formes de la négation, qui contrairement aux enfants de maternelle mentionnés dans la Section 2.2 ci-dessus, présentent une augmentation nette de la négation discontinuée au détriment de la négation postverbale entre 2;0 et 3;3 (cf. Palasis, 2015, Tableau 7, p. 128/226).

3. Synthèse

Les recherches rapportées dans les Sections 1 et 2 ont examiné un corpus de phrases déclaratives dans quatre années différentes de maternelle (de 2;3 à 5;11, N = 25 198, cf. Tableau 7), ainsi que les pronoms relatifs et interrogatifs sujets de moyenne section afin de contribuer à l’analyse du phénomène du sujet nul enfantin et au débat sur le statut des clitiques nominatifs en français informel. La Section 3 fournit une synthèse inédite de Palasis (2005, 2010b, 2014, 2015, 2022).

3.1 Clitique nominatif : marqueur préverbal d’accord

Les tests bien connus dans la littérature ont été appliqués à ces données enfantines avec les résultats synthétisés en (18), qui reprennent les caractéristiques mentionnées pour les données adultes en (3) page 44.

- (18) Résultats sur le statut morphosyntaxique du clitique nominatif enfantin :
- i. Les clitiques nominatifs sont omniprésents quel que soit le type de proposition (principale, subordonnée, juxtaposée, coordonnée ; de TPS à GS : 73,4 %-95,7 %-98,8 %-98,9 % ; cf. Tableau 7). L’augmentation du taux est liée à la baisse du taux de sujets nuls (26,6 %-3,8 %-0,7 %-0,2 %). L’omniprésence du clitique est un argument en faveur de son statut morphologique.
 - ii. Les clitiques nominatifs peuvent être doublés par différents types de constituants, avec des ratios de doublement qui suivent la même échelle d’implications que celle proposée pour les dialectes nord-italiens : pronoms toniques > syntagmes nominaux définis > syntagmes nominaux

indéfinis > variables *wh-* (MS : 99,9 %-97,5 %-76,7 %-2,6 % ; Palasis, 2015, p. 123/226). Le doublement par des syntagmes nominaux indéfinis est un argument fort en faveur de l'analyse morphologique des clitiques.

- iii. Les clitiques nominatifs sont quasiment toujours préverbaux (99,97 % ; Total N = 23 761/23 769). La place fixe préverbale est également un argument en faveur du statut morphologique des clitiques.
- iv. L'élision de *il* devant consonne est très productive (par exemple, GS : 96,5 %). L'ajustement phonologique est également un argument en faveur du statut morphologique du clitique.
- v. Le marqueur préverbal *ne*, dont la présence est utilisée comme argument contre l'analyse morphologique, est très rare dans le corpus (1,7 % ; N = 46/2770).

L'ensemble des données apporte des arguments solides en faveur de l'analyse morphologique de la très grande majorité des clitiques nominatifs dans ce corpus. Cela signifie que, lorsqu'un enfant produit *l'escargot i-dort*, *l'escargot* est le sujet syntaxique du verbe, *i-* est un marqueur d'accord préverbal, qui porte possiblement les informations de personne [3], genre [Masculin] et de nombre [Singulier] (mais cf. Palasis, 2010b, p. 241-252 sur le détail de l'émergence des traits de personne, genre, nombre et référentialité en PS), et *dort* est le radical verbal qui ne porte aucune marque d'accord audible. Un parallèle peut être établi avec la structure du syntagme nominal dans lequel l'information grammaticale est également apportée avant l'information lexicale à l'oral, grâce à l'ordre Déterminant-Nom. Ce parallèle illustre alors le rôle fondateur attribué aux morphèmes grammaticaux dans certaines approches récentes de l'acquisition du langage (cf. par exemple Le Normand, 2019).³³

³³ De plus, les travaux de Chevrot, Chabanal & Dugua (2007) et Dugua, Nardy, Liégeois, Chevrot & Chabanal (2017) sur l'acquisition de la liaison en français montrent que la liaison clitique-verbe est acquise avant celle déterminant-nom.

3.2 Français enfantin : langue à sujet nul

L'analyse morphologique du clitique nominatif implique que, lorsqu'aucun pronom tonique ni syntagme nominal ne double le clitique, alors aucun sujet syntaxique n'est phoniquement réalisé. Cette configuration renvoie à celle des langues dites « à sujet nul », telles que l'italien ou les dialectes nord-italiens, dont la configuration avec le clitique marqueur préverbal d'accord est la plus semblable à celle décrite pour le français enfantin.

Dans ce cadre, le phénomène du « sujet nul » enfantin (cf. Section II.1) est donc à renommer « marqueur d'accord préverbal nul » et s'explique par des principes d'économie. Les paradigmes verbaux, qui sont initialement réduits à deux personnes (1sg et 3sg), présentent d'abord un ratio très élevé de radicaux verbaux morphologiquement distinctifs (72,9 % ; par exemple, [sui] « suis » vs [e] « est »), puis, lorsque la personne 2sg émerge, ne nécessitent que deux marqueurs préverbaux pour obtenir trois formes verbales distinctives et donc interprétables (par exemple, [pø] « peux », [typø] « tu peux », [ipø] « i(l) peut »). Il s'agit alors d'un système dans lequel les clitiques préverbaux et les formes verbales apportent l'information flexionnelle de façon complémentaire, non-redondante, comme dans certains dialectes nord-italiens décrits par Pescarini (2022). La comparaison entre ce français enfantin et les dialectes nord-italiens pourrait alors être examinée plus finement en termes de corrélation, complémentarité, et lien de cause à effet entre interprétabilité de la forme verbale et présence du clitique, selon les propositions de Pescarini (2022).

3.3 Différents types d'explétifs

La Section II.2.5 du Chapitre I a abordé la question des explétifs. Leur fonction reste encore à clarifier dans les langues à sujet nul, dans la mesure où ils ne sont pas attendus dans ces systèmes et que les verbes qui appellent un explétif sont des verbes défectifs. Leur présence semble donc inutile. Néanmoins, cette remarque vaut également pour les langues à sujet non-nul et nous avons conclu qu'il pourrait alors simplement s'agir dans les deux types de langues d'un élément

utilisé pour régulariser un paradigme sans aucun apport référentiel, comme son nom l'indique. Néanmoins, les ratios différents d'omissions en fonction du type de verbe (météorologique, modal, etc.), tant dans les dialectes nord-italiens que chez les enfants francophones, et les échelles d'implications similaires dans les deux systèmes (Pescarini, 2014 ; Culbertson & Legendre, 2014 ; Poletto, 2000 ; Pescarini, 2020, p. 127) engagent à davantage de recherches afin d'affiner encore nos connaissances sur le fonctionnement de ces différents types d'explétifs.

Enfin, la Section 2 a également mis en avant que certaines données enfantines n'entraient pas dans le cadre de l'analyse en langue à sujet nul (les DP non-doublés, les clitiques *il* non-élidés devant consonne, les énoncés qui comportent une négation discontinue). Ces cas relativement rares seront examinés et expliqués dans le cadre de l'hypothèse de la diglossie appliquée au français (cf. Section IV). Mais avant cela, tournons-nous vers la question de la structure des questions *wh-* qui représente une autre thématique majeure dans la description de la variation morphosyntaxique du français hexagonal contemporain, tant chez les adultes que chez les enfants.

III. Questions *wh-* : morphosyntaxe et sémantique

À l'exception d'un corpus tout récent (Baunaz & Bonan, 2023), les productions adultes orales contemporaines montrent une majorité de *wh-ex-situ*, que ce soit en français informel en général ou en LAE en particulier (cf. Section II.2.6 du Chapitre I). Les données LAE de Strik & Pérez-Leroux (2011) par exemple renvoient un taux de 61,9 % de questions *wh-ex-situ*. Le défi qui s'ouvre alors aux chercheurs en acquisition est d'expliquer pourquoi les enfants, qui sont en contact avec ce français oral majoritairement *ex-situ*, produisent une majorité de *wh-in-situ* (entre 62,5 % et 90,9 %, cf. Hamann, 2006 et revues de la littérature dans Palasis et al., 2019, p. 136-141/226 et Palasis et al., 2023, p. 201-205/226).

De nouveau, mon approche repose sur une étude approfondie des productions enfantines en vue de chercher l'origine de cette disparité dans le système linguistique des enfants considéré comme ayant une cohérence interne

(cf. Section 1.2 ci-dessus) et non pas dans des explications qui renvoient la prévalence de la position in-situ enfantine à des difficultés computationnelles ou un déficit pragmatique chez l'enfant par rapport à l'adulte (Strik, 2007 ; Jakubowicz, 2011 ; Becker & Gotowski, 2015). De nouveau également, les résultats ont été publiés au fur et à mesure du traitement des trois années de la base de données enfantines et l'hypothèse initiale a été affinée puis confirmée au fil des recherches (Palasis & Faure, 2014 ; Palasis et al., 2019 ; Palasis et al., 2023, cf. Section 1 ci-après). En parallèle, une hypothèse a également été formulée sur la position des items *wh-* en français adulte (Faure & Palasis, 2014 ; Faure & Palasis, 2021, cf. Section 2).

1. Chez l'enfant

Cf. pages 135-160/226 du Volume 2 : Palasis, K., Faure, R. & Lavigne, F. (2019). Explaining variation in *wh*-position in child French: A statistical analysis of new seminaturalistic data. *Language Acquisition*, 26(2), 210-234. <https://doi.org/10.1080/10489223.2018.1513004>

Cf. pages 201-226/226 du Volume 2 : Palasis, K., Faure, R. & Meunier, F. (2023). *Wh*-in-situ in child French: Deictic triggers at the syntax-semantics interface. *Journal of French Language Studies*, 33(3), 273-298. <https://doi.org/10.1017/S0959269523000030>

L'objectif principal des recherches sur l'acquisition des questions *wh-* était de détecter un ou plusieurs facteurs qui expliquent la position des éléments *wh-* chez l'enfant et la prévalence de la position *wh*-in-situ chez l'enfant par rapport à l'adulte. La base de données a permis une extraction fine des questions *wh-* produites par seize enfants pendant leurs trois années de maternelle à partir de l'ensemble des questions *wh-* produites (2;5-5;11, N initial = 1 213).

1.1 Méthode

De nombreux filtres ont été appliqués à l'ensemble initial de données afin de ne conserver que les configurations qui autorisent l'alternance in-situ vs ex-situ. Ainsi, toutes les configurations énumérées en (19) ci-dessous ont été écartées du jeu initial de données.

- (19) Configurations écartées :
- i. Questions enchâssées (n = 129) car elles favorisent la position ex-situ (*tu sais combien ça coûte ?* vs **tu sais ça coûte combien ?*).³⁴
 - ii. Questions avec l'adverbe interrogatif *pourquoi* (N = 111) qui ne se trouve qu'en position ex-situ. De plus, cet item est généralement décrit dans la littérature générative comme étant généré dans une position plus haute et donc différente des autres items interrogatifs (Rizzi, 2001 ; Hamann, 2006 ; Myers & Pellet, 2014).
 - iii. Questions avec le pronom interrogatif *qui* sujet (N = 58) qui apparaît toujours à gauche du verbe en raison de sa fonction syntaxique de sujet.
 - iv. Questions avec une structure clivée (N = 116 ; par exemple, *mais c'est quand que tu montres la maison ?* MAT, 3;11) car l'item *wh-* se trouve toujours à gauche du verbe de la deuxième partie non-clivée et il s'agit d'une structure syntaxique particulière (cf. Oiry, 2011, entre autres).
 - v. Questions qui nécessitent la connaissance du contexte discursif afin d'interpréter l'item *wh-* (configuration connue sous le nom de *D-linking* depuis Pesetsky, 1987) car ce lien avec le discours est parfois considéré comme favorisant la position *wh-in-situ* en français (cf. Cheng & Rooryck, 2000, entre autres). Les items interrogatifs *quel, lequel* et leurs variantes ont donc été écartés (N = 59).
 - vi. Questions avec inversion verbe-sujet (N = 11) car elles excluent la position *wh-in-situ* (par exemple, **voit-il qui ?*).
 - vii. Questions qui présentent une caractéristique du français standard, telle que la non-élision du clitique *il* devant consonne ou la négation discontinue, car le français standard présente davantage de questions *wh-ex-situ* (cf. Section II.2 du Chapitre I sur la dichotomie entre le français standard et le français informel/oral et la Section IV ci-après sur la diglossie chez l'enfant ; N = 1 : *il voit quoi ?* KEL, 5;6).

³⁴ Des travaux récents sur le français adulte informel montrent néanmoins qu'il semble y avoir une évolution vers de plus en plus de *wh-in-situ* dans les propositions enchâssées (cf. Gardner-Chloros & Secova, 2018 ; Poletto & Pollock, 2020). Une comparaison avec les données enfantines sera donc à effectuer.

- viii. Répétitions de questions (autorépétitions et répétitions immédiates de la production d'un autre enfant, N = 10).

Les recherches ont ainsi porté sur un total de 718 questions *wh-* réparties sur trois années (PS = 368 ; MS = 192 ; GS = 158) à partir desquelles nous avons tout d'abord considéré les facteurs Item, Position et Section détaillés en (20) ci-dessous en (i-iii). Néanmoins, dès les premières investigations, il est apparu que le verbe semblait également représenter un facteur important dans la position des *wh-* enfantins (Palasis & Faure, 2014). À la suite des travaux de Coveney (1995) qui avait observé une dichotomie dans les questions *wh-* adultes entre verbes lexicaux et la copule *être* et des travaux de Higgins (1973/1979) et Roy (2013) sur les différents types d'énoncés (non-)prédicationnels, nous avons proposé d'examiner les questions enfantines à la lumière d'un nouveau facteur : une tripartition verbale, rapportée en (20iv) ci-dessous, qui consiste à opposer le présentatif *c'est* (forme figée du verbe *être*), les autres formes du verbe *être* (formes libres de *être*, *c'est*-à-dire *être* accompagné de tout pronom autre que *c'*), et les autres formes verbales (formes libres de verbes lexicaux) (détails dans Palasis et al., 2019, p. 141-144/226).

(20) Facteurs examinés :

- i. Position : in-situ, ex-situ
- ii. Item interrogatif :
 - a. Pronom : *quoi/que* et variantes (*à quoi, qu'est-ce que*, etc.), *qui* objet et variantes (*à qui*, etc.)
 - b. Adverbe : *où, comment, quand, combien*
- iii. Section : PS, MS, GS
- iv. Verbe : présentatif *c'est*, formes libres du verbe *être*, verbes lexicaux

Les 718 questions *wh-* ont été soumises à des tests statistiques (Test Cochran-Mantel-Haenszel, Cochran, 1954 ; Mantel & Haenszel, 1959 ; détails dans Palasis et al., 2019, p. 151/226) afin de déterminer l'effet des facteurs Item, Section et Verbe sur la Position de l'élément *wh-* (PS et MS dans Palasis et al., 2019 ; GS et synthèse dans Palasis et al., 2023).

1.2 Facteur 1 : émergence de « qu'est-ce que »

Les résultats généraux montrent une prévalence de la position *wh*-in-situ pendant les trois années malgré une augmentation constante de la position ex-situ au fil des années (cf. Palasis et al., 2023, p. 213/226, Figure 3 : ex-situ : 32,6 %-40,6 %-44,3 %). Ainsi, les enfants produisent significativement plus de *wh*-in-situ que de ex-situ et, malgré l'augmentation de ces derniers, les trois années ne sont pas significativement différentes les unes par rapport aux autres. Les résultats des items *wh*- pris individuellement indiquent également qu'il n'y a pas de modification significative dans leurs positions pendant les années de maternelle, sauf pour le pronom *quoi* dont le ratio in-situ diminue significativement entre MS et GS (73,5 %-59 %, $p < .05$) au profit de l'émergence du pronom ex-situ *qu'est-ce que*, comme illustré en (21) ci-dessous.

Il est alors intéressant de noter qu'il s'agit de la seule paire d'items interrogatifs à présenter des formes morphologiquement différentes en fonction de la position. Par ailleurs, il existe de nombreux travaux et un débat sur le statut morphosyntaxique de *qu'est-ce que* du point de vue de l'enfant, qui est considéré par certains comme étant le fruit d'une computation syntaxique avec inversion du clitique (Obenauer, 1981 ; Plunkett, 2000) et par d'autres comme un bloc inanalysé (Rooryck, 1994 ; Hulk, 1996 ; Strik, 2008). Un argument fort en faveur de l'hypothèse du bloc inanalysé est l'absence totale d'alternance entre *qu(e) + est-ce que* et d'autres éléments *wh*- en PS et MS (Palasis & Faure, 2014). Il n'en demeure pas moins que produire *qu'est-ce que* est long et articulatoirement plus complexe que *quoi*. Par ailleurs, alors qu'il semble simple d'établir une identité entre les autres paires d'items *wh*- malgré leurs deux positions possibles, il est probablement cognitivement plus complexe d'établir que *quoi* et *qu'est-ce que* représentent deux signifiants différents pour un même signifié, ce qui va à l'encontre notamment du Principe de contraste selon lequel les jeunes enfants présupposent que deux signifiants différents renvoient à deux signifiés différents (Clark, 1987).

(21) Alternance *quoi/qu'est-ce que* :

- a. Je fais quoi ? (MAT, 4;11)
- b. Qu'est-ce qu'il a dit ? (VIC, 5;4)

Les données mettent ainsi en avant une première caractéristique développementale des questions *wh-* à la maternelle : l'augmentation significative de la forme ex-situ *qu'est-ce que* au détriment de *quoi*, qui pourrait représenter l'une des explications au fait souvent remarqué que les enfants produisent plus de pronoms in-situ que les adultes (Zuckerman & Hulk, 2001 ; Strik, 2007 ; Strik & Pérez-Leroux, 2011 ; Becker & Gotowski, 2015). La prévalence des *wh*-in-situ pourrait ainsi être liée à l'utilisation initialement peu fréquente de la forme *qu'est-ce que* chez les jeunes enfants par rapport à la forme in-situ *quoi* plus simple et donc plus fréquente (PS : *qu'est-ce que* = 21,1 %, N = 36/171), suivie de l'augmentation significative entre MS et GS mentionnée ci-dessus (26,5 %-41 %, $p < .05$). Je complète ici cette conclusion formulée dans Palasis et al. (2023, pp. 218-219/226) par une comparaison détaillée des taux des items *wh*-ex-situ dans les données LAE de Strik & Pérez-Leroux (2011) qui renvoyaient un taux de 61,9 % de questions *wh*-ex-situ. Une vérification du détail de ce corpus LAE montre que, parmi ces ex-situ figurent 47,7 % de *wh*-est-ce que (et 14,2 % *wh*-ex-situ SV). Il semblerait alors que l'hypothèse de l'utilisation initialement peu fréquente de *qu'est-ce que* chez les enfants soit une explication possible à la différence des taux d'ex-situ entre enfants et adultes.

1.3 Facteur 2 : prévalence du présentatif « c'est »

Concernant le facteur Verbe, les tests statistiques ont révélé que la forme des verbes (Fixe vs Libre) est une variable discriminante pour la position des items *wh*-pendant les trois années (Palasis et al., 2023, p. 212-215/226). Ainsi, la forme Fixe *c'est* émerge davantage avec des *wh*-in-situ, alors que les formes Libre (verbes *être* et lexicaux) émergent davantage avec des *wh*-ex-situ, comme illustré en (22) ci-dessous. Ces corrélations rendent compte de 74,7 % du corpus (N = 536/718). En revanche, le type de verbe (*être* vs lexical) n'est pas une variable discriminante,

comme initialement envisagé (Palasis & Faure, 2014 à la suite des travaux de Coveney, 1995 sur les adultes).

- (22) Effet de la forme du verbe (Fixe vs Libre) sur la position du *wh-* :
- a. Fixe (*c'est*) et in-situ : *c'est qui Taz ?* (LSN, 5;1)
 - b. Libre (*être*) et ex-situ : *où il est Kelian ?* (MAS, 4;7)
 - c. Libre (lexical) et ex-situ : *combien ça coûte ?* (LAN, 5;2)

Par ailleurs, la variable discriminante de forme (Fixe vs Libre) interagit avec la catégorie de l'item *wh-* (Pronom vs Adverbe). Ainsi, la forme Fixe *c'est* émerge davantage avec des pronoms *wh-in-situ*, alors que les formes Libre (*être* et lexical) favorisent la position *wh-ex-situ* et apparaissent largement avec des adverbes, comme illustré également en (22) ci-dessus.

Il semble que ces résultats représentent alors une deuxième pièce du puzzle concernant l'asymétrie observée entre la position des pronoms, majoritairement in-situ chez les enfants mais pas chez les adultes, et la position des adverbes, majoritairement ex-situ chez les enfants et les adultes (Zuckerman & Hulk, 2001 ; Strik, 2007 ; Strik & Pérez-Leroux, 2011 ; Becker & Gotowski, 2015). En effet, les résultats montrent que les adverbes sont majoritairement présents avec des formes verbales Libre (cf. (22b) et (22c) ci-dessus), ce qui implique qu'ils sont majoritairement ex-situ. *A contrario*, les pronoms émergent davantage avec le présentatif *c'est* (cf. (22a)), ce qui implique une position in-situ (Palasis et al., 2023, p. 215-219/226). L'hypothèse de la forme verbale (Fixe vs Libre) rend ainsi également compte de l'asymétrie entre pronoms et adverbes dans les données enfantines. Ce facteur de la forme verbale (Fixe vs Libre) vient alors compléter le facteur sur l'émergence de *qu'est-ce que* en rendant compte de l'autre partie de l'asymétrie sur la différence entre pronoms et adverbes.

Par ailleurs, le facteur de la Forme verbale permet de proposer un aspect développemental à l'hypothèse. En effet, il est par ailleurs documenté que les jeunes enfants étoffent et diversifient graduellement leur système verbal (Guillaume, 1927 ; Clark, 1978 ; Bassano, Eme & Champaud, 2005 ; Palasis, 2012b,

et cf. la discussion dans la Section II.1 sur le phénomène du sujet nul enfantin). Dans ce cadre, il peut être prédit que le ratio de *wh*-in-situ va diminuer également en lien avec l'ajout de verbes lexicaux dans l'inventaire verbal des enfants qui vont progressivement remplacer la forme *c'est* par des verbes sémantiquement plus étoffés.

1.4 Facteur 3 : *deixis et non-présupposition*

Les facteurs 1 et 2 rendent compte de 74,7 % du corpus (N = 536/718). Palasis et al. (2023, pp. 219-223/226) ont également examiné les 25,3 % de questions *wh*- qui représentent des contre-exemples à l'hypothèse de la Forme verbale (N = 182/718). Dans un premier temps, l'extraction des données a montré que la plupart des exceptions sont des questions *wh*-in-situ avec une forme verbale Libre (90,1 %, N = 164/182, cf. Palasis et al., 2023, Tableau 9, p. 219/226, et exemples en (23) ci-dessous).

- (23) Questions *wh*-in-situ avec une forme verbale Libre :
- a. Et là i(l) dit quoi ? (LAN, 3;11)
 - b. Là y a quoi ? (KEL, 5;7)

Dans un second temps, l'étude fine de ces questions a mis en avant une présence significativement plus élevée de l'adverbe *là* dans les contrevenants in-situ que dans les autres questions *wh*- ($p < .01$, Palasis et al., 2023, p. 219-223/226). Or, l'adverbe *là* partage au moins une caractéristique avec le présentatif *c'est*, dans la mesure où tous les deux sont des éléments déictiques qui ancrent donc la question dans un contexte immédiat « ici et maintenant ». Ces questions *wh*-in-situ avec une forme verbale Libre appartiennent donc finalement au même paradigme que les questions in-situ avec *c'est*, c'est-à-dire qu'elles ne présupposent aucun contexte. Les exemples en (23) sont donc tout aussi non-présuppositionnels que les questions *wh*-in-situ avec *c'est*. J'ajoute ici que l'ensemble de ces questions non-présuppositionnelles (les questions in-situ avec *c'est* et les questions in-situ avec une forme verbale Libre représentent la majorité des questions *wh*- du corpus (62,7 %, N = 450/718).

1.5 Synthèse

Les travaux de Palasis & Faure (2014) et Palasis et al. (2019, 2023) ont confirmé que les enfants produisent davantage de questions *wh*-in-situ que les adultes et que la position in-situ reste dominante pendant les trois années de maternelle malgré une progression constante de la position ex-situ (32,6 %-40,6 %-44,3 %).

L'étude minutieuse de la base de données de maternelle a permis de mettre en avant un effet de la Forme verbale sur la position des items *wh*- (Fixe et in-situ vs Libre et ex-situ, cf. (22) ci-dessus) et deux aspects morphosyntaxiques du système enfantin qui permettent de prédire l'augmentation du ratio de ex-situ : l'augmentation du ratio de la forme ex-situ *qu'est-ce que* au détriment du pronom in-situ *quoi* (Facteur 1, cf. Section 1.2) et l'augmentation du ratio de verbes lexicaux par rapport au présentatif *c'est* (Facteur 2, cf. Section 1.3). Enfin, un facteur à l'interface de la morphosyntaxe et de la sémantique a été mis en avant dans la majorité du corpus (62,7 %, N = 450/718) : la non-présupposition, commune aux questions in-situ avec *c'est* d'une part et avec une forme verbale Libre et l'adverbe *là* d'autre part (Facteur 3, cf. Section 1.4).

Ce troisième facteur se positionne ainsi à l'interface de la morphosyntaxe (positions et formes des items *wh*-) et de la sémantique (sémantique verbale, deixis). Il représente notre lien avec l'hypothèse proposée pour rendre compte de la distribution des *wh*-in-situ et ex-situ non seulement en français enfantin mais également en français adulte informel.

2. Français informel : langue à *wh*-in-situ

En parallèle des travaux menés sur les données enfantines, nos recherches ont également contribué aux nombreux débats sur et autour de la position des items *wh*- en français adulte oral (Faure & Palasis, 2014, 2021 ; cf. Section II.2.6 du Chapitre I). Cette problématique est déroulée ici en tirant le fil de l'interface de la morphosyntaxe et de la sémantique mentionnée dans la section ci-dessus.

La majorité du corpus enfantin (62,7 %) se compose de questions *wh-* qui ne requièrent aucune présupposition particulière quant à d'éventuelles informations contextuelles. Cet aspect non-présuppositionnel est possible en français notamment grâce au présentatif *c'est* qui comporte la copule *être*, forme verbale sémantiquement vide, et le pronom démonstratif élidé *c'*, qui donne un aspect déictique à la forme verbale. Une question posée avec *c'est* est donc une question qui appelle une identification ou une définition sur l'item interrogé dans le cadre énonciatif du « ici et maintenant » (par exemple, *c'est qui Taz ?*). Il s'agit d'une question effective (non-rhétorique), sans présupposition sémantique et sans déplacement syntaxique du *wh-* (puisque les données enfantines ont montré que ce type de question est majoritairement *wh-in-situ*). *A contrario*, les questions avec *wh-ex-situ*, c'est-à-dire avec un déplacement syntaxique de l'item *wh-*, comportent majoritairement une forme verbale Libre, c'est-à-dire forcément plus riche au niveau sémantique que le présentatif *c'est*.

À partir des corrélations observées entre formes verbales et positions *wh-* chez les enfants, l'hypothèse développée est que le déplacement syntaxique de l'élément *wh-* serait déclenché non pas par l'interrogation *per se* (puisque'il existe des interrogations *in-situ*), mais par la sémantique du verbe (Palasis et al., 2019, p. 156/226). En poursuivant cette ligne de raisonnement, les conclusions sont synthétisées en (24).

(24) Conclusions sur le français enfantin :

- a. Ce n'est pas l'interrogation (en termes génératifs, le trait [+*wh*] en C, cf. Section 2.6 du Chapitre I) qui déclenche le déplacement de l'item *wh-* vers la périphérie gauche.
- b. Modalité interrogative (trait [+*wh*] en C) et déplacement de l'item *wh-* peuvent donc être dissociés.
- c. Vu que le système enfantin génère une majorité de questions *wh-in-situ*, le français enfantin peut être décrit comme une langue à *wh-in-situ*, comme le mandarin (en termes génératifs, avec un déplacement « furtif » de l'opérateur parce que le trait [+*wh*] en C est faible).

- d. C'est donc un autre trait qui déclenche le déplacement des *wh*-ex-situ dans la syntaxe du français enfantin et, par extension de l'hypothèse, en français informel, donc également adulte.

3. Chez l'adulte : trait [Exclusion]

Cf. pages 162-200/226 du Volume 2 : Faure, R. & Palasis, K. (2021). Exclusivity! *Wh*-fronting is not optional *wh*-movement in Colloquial French. *Natural Language & Linguistic Theory*, 39(1), 57-95, <https://doi.org/10.1007/s11049-020-09476-w>

L'hypothèse que le français informel est une langue à *wh*-in-situ, formulée sur la base de l'étude de données enfantines, a été mise à l'épreuve des tests traditionnellement appliqués dans la littérature par les linguistes à partir de données du français oral adulte (Faure & Palasis, 2014, 2021, p. 169-178/226).³⁵ Ces tests traditionnels ont mis en avant des comportements différents des *wh*-in-situ et ex-situ et, comme en français enfantin, ont pointé vers un mouvement *wh*-qui pourrait être déclenché par un trait autre que le trait interrogatif. La conclusion pour le français enfantin est donc étendue au français adulte et la conclusion générale pour le français informel est qu'il s'agit d'une langue à *wh*-in-situ. Il est intéressant de noter que Hauser, Aristodemo & Donati (2023) arrivent à la même conclusion pour la langue des signes française.

Comme pour le français enfantin, les contextes qui déclenchent les *wh*-ex-situ ont donc été examinés avec une attention particulière portée à la notion de contraste, présente à la fois chez les enfants et les adultes, comme illustré en (25) avec une série de questions enfantines. Le contraste en (25) provient du contexte : l'enfant interroge sur la position *relative* de trois petites cartes qui sont à placer sur un support commun (Loto des animaux) et, chaque position étant unique, celle-ci exclut les autres possibilités. En d'autres termes, il n'y a qu'une réponse possible à chaque question, ce qui exclut toutes les autres dans ce cas particulier.

³⁵ Cette partie est peu développée ici car les compétences en sémantique et en logique qui ont permis de générer la composition de l'opérateur [Exclusion] dans Faure & Palasis (2021, p. 191-192/226) sont exclusivement celles de Richard Faure.

(25) Série de questions *wh*-ex-situ enfantines (MAS, 2;7) :

- a. Où il est le poisson ?
- b. Où il est le perroquet ?
- c. Où elle est la tortue ?

Globalement, il apparaît que les questions *wh*-ex-situ chez les adultes sont optimales lorsque la réponse à la question est le fruit d'une sélection à partir d'un ensemble, avec exclusion d'au moins une des possibilités. Le trait qui déclenche le déplacement du *wh*-ex-situ est alors appelé [Exclusion] et est formalisé en termes syntaxiques et sémantiques et est illustré en (26) où a. et b. portent tous les deux le trait interrogatif [+wh] faible en tête de phrase et où seul b. connaît le déplacement du *wh*- (pour vérifier l'Opérateur [Excl] au-dessus de IP ; cf. détails de la computation dans Faure & Palasis, 2021, p. 191-193/226).

(26) Trait [Exclusion] : déclencheur des *wh*-ex-situ :

- a. [+wh faible] *Tu as vu* [-Excl] *qui* ? (qui = pas de présupposé sur la réponse = peut être n'importe qui)
- b. [+wh faible] [+Excl] *Qui tu as vu* ~~qui~~ ? (qui = présupposé sur la réponse = qui fait partie d'un ensemble connu du locuteur et de l'interlocuteur, avec l'exclusion d'au moins une possibilité)

Il reste alors à décrire le ou les chaînons manquants qui permettront d'attester de l'émergence et du développement de ce trait [Excl] chez les enfants à partir de l'effet de la forme verbale (Fixe vs Libre) mis en avant chez des enfants de maternelle. Il serait également intéressant de regarder les données des dialectes nord-italiens, qui sont également décrits comme des systèmes mixtes en termes de positions *wh*-in-situ et ex-situ (Baunaz & Bonan, 2023, fn 4 ; Bonan, 2021) ; cf. Section II.2.6 du Chapitre I). Ce parallèle est d'autant plus intéressant que les travaux ci-dessus ont établi un autre fonctionnement parallèle entre français informel et dialectes nord-italiens quant au statut des clitiques nominatifs en tant que marqueurs préverbaux d'accord.

IV. Hypothèse de la diglossie

Le principe de la variation est de présenter au moins deux variantes pour un même phénomène linguistique (cf. Section II.3 du Chapitre I). Ainsi, les sections ci-dessus ont montré qu'en français informel, tant adulte qu'enfantin, un DP est généralement doublé par un clitique nominatif, qu'un clitique nominatif est majoritairement présent même lorsqu'il y a un DP sujet, que le clitique *il* est souvent éliidé devant consonne, que la négation est généralement produite dans une forme simple, postverbale, que les items *wh-* sont in-situ ou ex-situ et que les ratios sont différents entre adultes et enfants, mais qu'aucun de ces phénomènes ne représente 100 % des occurrences. Les résultats rapportés dans les sections précédentes pour les adultes et les enfants sont synthétisés dans le Tableau 11 ci-après (inédit).

Tableau 11 : récapitulatif de caractéristiques du français informel (adultes et enfants)

%	Adultes		Enfants		
	Adulte-Adulte	LAE	PS	MS	GS
Doublément du DP ³⁶	60-74	73-100	98,5	98,2	96,0
Présence du clitique ³⁷	94,1	n.c.	95,7	98,8	98,9
Élision de <i>il / _C</i> ³⁸	79,6	n.c.	98,4	99,1	96,5
Omission de <i>ne</i> ³⁹	81,2-98,9	*16,7-93,8	98,3	97,9	98,8
<i>Wh</i> -in-situ ⁴⁰	10,5-45,9	16,6-51	67,4	59,4	55,7

*contexte scolaire

Le Tableau 11 montre que, pour toutes les caractéristiques du français oral, les enfants présentent quasi-systématiquement des taux plus élevés que les adultes, même par rapport au LAE, mais sans atteindre les 100 %. Cela signifie que les jeunes enfants utilisent très majoritairement le système grammatical du

³⁶ Pourcentage des DP doublés par rapport au nombre total de DP dans chaque corpus. Pour les enfants, les calculs sont effectués à partir des données des lignes b et d du Tableau 5.

³⁷ Pourcentage des énoncés qui comportent un clitique nominatif par rapport à l'ensemble des énoncés déclaratifs dans chaque corpus. Pour les enfants, cf. ligne (a+b) du Tableau 7.

³⁸ Pourcentage des clitics *il* éliidés par rapport à l'ensemble des clitics *il* dans chaque corpus (référentiels et explétifs confondus). Pour les enfants, cf. Tableau 9.

³⁹ Pourcentage des énoncés sans *ne* par rapport à l'ensemble des énoncés verbaux négatifs dans chaque corpus. Pour les enfants, cf. Tableau 10.

⁴⁰ Pour les enfants, cf. Figure 3 de Palasis et al. 2023, p. 213/226.

français oral, mais qu'il existe déjà une petite part de variation dans ce système pour chacune des caractéristiques étudiées.

De plus, les travaux sur les données orales adultes ont montré que certains phénomènes semblent corrélés. En LAE, Culbertson (2010, p. 95) par exemple a mis en avant une variabilité du taux d'omission de *ne* en fonction de la présence du clitique nominatif : 93,7 % d'omissions (N = 4 897/5 229) après un clitique nominatif (avec ou sans DP cooccurrent) et 16,7 % (N = 4/24) après un DP sans clitique nominatif. Culbertson (2010, p. 97) propose alors d'expliquer le taux de rétention de *ne* en présence d'un clitique nominatif en LAE (6,3 %) par le fait que « les locuteurs maintiennent des grammaires qui encodent aussi bien le registre du français standard que celui du français informel dans leur langue [...] » et l'auteure montre qu'un modèle linguistique qui « mélange deux grammaires » rend mieux compte des données qu'un modèle qui ne postule pas la présence de deux grammaires ($p < 0.001$, test du taux de vraisemblance de Lehmann, 1986 ; cf. Section II.3.2 du Chapitre I). Hirschbühler & Labelle (2004, p. 196) ont également émis une hypothèse de « deux grammaires en compétition », une avec *ne* et une sans *ne*, à la suite de l'étude du corpus du jeune roi Louis XIII (cf. Section II.2.3 ci-dessus).

Les Sections 1 et 2 ci-après rapportent mes travaux sur l'hypothèse de la diglossie formulée à partir de l'étude des données des seize enfants de maternelle pour lesquels le français est une L1 (cf. Section 1) et à partir de recherches initiées plus récemment sur deux enfants supplémentaires dans la même classe pour lesquels le français est une L2P (cf. Section I.2 du Chapitre I et Section 2 ci-après).

1. Français L1

- Cf. pages 41-58/226 du Volume 2 : Palasis, K. (2011). Français spontané et français normé : bien plus que deux variétés d'un même français. Implications acquisitionnelles et didactiques. In O. Bertrand & I. Schaffner (Eds.), *Variétés, variations & formes du français* (pp. 55-72). Palaiseau: Les Éditions de l'École Polytechnique, <https://www.researchgate.net/profile/Katerina-Palasis/publications>
- Cf. pages 83-102/226 du Volume 2 : Palasis, K. (2013). The case for diglossia: Describing the emergence of two grammars in the early acquisition of Metropolitan French. *Journal of French Language Studies*, 23(1), Numéro spécial « L'hypothèse d'une diglossie en France », Massot, Benjamin & Paul Rowlett (Eds.), 17-35. <https://doi.org/10.1017/S0959269512000348>

Traditionnellement, dans les grammaires du français ou dans les travaux sur les données adultes, la variation est décrite en prenant comme point de départ la variété standard à partir de laquelle la variété non-standard est alors envisagée (Wagner & Pinchon, 1962 ; Lambrecht, 1981 ; Blanche-Benveniste, 2000 ; Grevisse & Goosse, 2008, entre autres). Ainsi, Lambrecht (1981, p. 6-7) par exemple considère l' « élimination » du clitique *ne*, l' « érosion phonétique » des clitiques, le « remplacement » de *nous* par *on*, etc. Or, le point de vue acquisitionnel oblige à reconsidérer ces descriptions dans la mesure où la relation entre les deux variantes est inversée dans la chronologie du développement linguistique (Palasis, 2011b, p. 44-49/226 ; Palasis, 2013, p. 86-96/226) : par exemple un jeune enfant « n'omet » pas le *ne*, mais démarre son acquisition de la négation avec la forme simple, postverbale pour ensuite étoffer son système et éventuellement « ajouter » une forme négative avec le *ne* préverbal.

Palasis (2011b, pp. 47-49/226) a montré que la forme non-élide du clitique *il* devant consonne émerge bien après les formes qui présentent l'alternance en fonction du contexte phonologique. Ainsi, les quatorze formes non-élidees en PS (cf. Tableau 9 ci-dessus) sont produites par trois enfants en tout et les formes les plus précoces émergent chez LUS et MAI à 3;2 alors que l'alternance est attestée depuis la première séance (à 2;9 et 2;8, respectivement). Un décalage est également observé concernant la négation simple vs discontinue et la production du DP sans clitique nominatif (Palasis, 2013, pp. 91-94/226). Les quinze formes avec *ne* en PS (cf. Tableau 10 ci-dessus) sont produites par un total de quatre enfants (dont deux produisent également le clitique non-élide : LUS et MAI). LUS est la première à produire *ne* à 2;10, suivie par MAI à 3;2. Les émergences de *il* non-élide et de *ne* sont donc concomitantes chez MAI. Il est important de noter qu'elles se trouvent dans une seule et même production et qu'il s'agit d'une chanson (MAI : *il ne chante plus, il ne saute plus, il ne danse plus*). Enfin, les DPs sans clitique associé sont produits par un total de neuf enfants en PS (N = 37, cf. Tableau 7 ci-dessus), dont LUS, qui à elle-seule en produit 46,1 %, mais pas MAI. En revanche, VIC, qui produit une seule occurrence de *ne* en PS à 3;7 produit

également un seul DP sans clitique dans le même énoncé, à l'occasion d'une comptine (VIC : *le facteur n'est pas passé*).

L'examen individuel et chronologique des données de petite section de maternelle met ainsi en lumière quatre caractéristiques importantes de ce français enfantin : (i) la variation non-standard vs standard n'est pas présente chez tous les enfants en fin de petite section, (ii) les variantes standards des différents phénomènes linguistiques observés émergent toujours après les variantes non-standards, (iii) les variantes standards de différents phénomènes émergent souvent concomitamment dans un même énoncé, et (iv) les formes standards sont souvent produites dans des contextes très spécifiques, différents des contextes spontanés habituels : il s'agit notamment de chants, de comptines, et d'activités de narration à partir de livres.

Les caractéristiques non-aléatoires de la variation dans les données de petite section ont ainsi permis d'étayer l'hypothèse de la diglossie appliquée au français à partir de données adultes (cf. Section II.3.2 du Chapitre I) et de confirmer les prédictions développementales formulées en (5) et répétées en (27) ci-dessous.

(27) Hypothèse de la diglossie et prédictions développementales :

- i. La variété « basse » (terminologie de Ferguson, 1959) est maîtrisée par tous les locuteurs de la communauté linguistique, alors que la variété « haute » est utilisée avec des degrés différents de compétence en fonction de l'âge du locuteur (préscolaire ou pas) puis de son niveau de scolarisation. Tous les locuteurs de la communauté linguistique ne sont donc pas forcément diglosses.
- ii. Les deux variétés se développent de façon séquentielle : la variété « basse » est systématiquement acquise avant la variété « haute ». En fonction de l'âge de l'enfant, on devrait donc pouvoir constater des étapes de l'acquisition/apprentissage de la variété « haute » suite à l'acquisition initiale de la variété « basse ».

Dans le cadre des travaux sur les données acquisitionnelles, les deux variétés de français ont été nommées « français spontané » (Grammaire 1) et « français normé » (Grammaire 2) et confirment l’hypothèse forte de Massot (2010, p. 90) qui déclare qu’un locuteur « [...] s’engage pour l’énoncé entier dans l’une de ses deux compétences [...]. » (cf. Palasis, 2011b, pp. 49-51/226). Le Tableau 12 ci-après récapitule les caractéristiques de ces deux grammaires qui, bien qu’apparentées, sont donc considérées comme distinctes (cf. Faure & Palasis, 2021, pp. 165-167/226 pour une synthèse sur la diglossie et Palasis, 2011b, p. 51/226 pour l’ensemble des formes différentes du sujet dans les deux grammaires). Faure & Palasis (2021) avaient laissé la question ouverte quant au statut in-situ ou ex-situ de G2. Dans Palasis et al. (2023), un seul énoncé interrogatif enfantin in-situ comporte une caractéristique de G2 (non-élision du clitique devant consonne : *il voit quoi ?*), ce qui pourrait indiquer que G2 dispose également des deux positions (la position ex-situ étant notamment obligatoire avec l’inversion verbe-clitique), mais les données sont tout à fait insuffisantes pour trancher cette question à partir du corpus de maternelle.

Tableau 12 : récapitulatif des caractéristiques des deux grammaires du français

		G1	G2
Clitique nominatif	Statut	Morphologique	Syntaxique
	Présence	Obligatoire	Facultative
	Clise	Proclise	Pro ou enclise
	Association DP	Doublement	Dislocation
	Ajustement /_C	Oui	Non
	Paramètre	Sujet nul	Sujet non-nul
Négation	Type	Postverbale	Discontinue
Interrogation	Paramètre	<i>Wh</i> -in-situ	?

Je rejoins ainsi les propos de Culbertson (2010) et de Hirschbühler & Labelle (2004) sur l’existence de deux grammaires chez les enfants francophones, et ce peut-être depuis le XVII^e siècle (cf. également Larrivée, 2022). Je rejoins également les propos de Ferguson (1959) sur les modalités d’acquisition ou d’apprentissage de ces deux grammaires : G1 est acquise par tous les enfants dès le plus jeune âge ; G2, par contre, va se développer de façons diverses chez les jeunes enfants en fonction de la présence de G2 ou non dans leur environnement

familial quotidien. Certains locuteurs seront diglosses (2L1), d'autres auront plus de difficultés avec G2, qui dans les cas les plus défavorables ne sera que peu développée à la fin de la scolarité obligatoire (cf. par exemple Jisa, 2004). Les résultats étayent donc l'hypothèse de la diglossie de Ferguson (1959) appliquée au français (Massot & Rowlett, 2013) et mettent en avant que tous les locuteurs du français ne sont pas forcément diglosses (Palasis, 2011b, 2013).⁴¹

L'hétérogénéité des environnements linguistiques en termes d'apports précoces et quotidiens de G2 rejoint notamment la question de l'impact des catégories socio-professionnelles des parents et de leurs niveaux de diplomation (cf. Section II.1.2 du Chapitre I), car LUS et MAI, qui font partie des rares enfants à produire G2 dès la PS, sont toutes les deux filles d'enseignantes. Il peut alors être supposé qu'elles bénéficient d'un environnement à la fois attentif aux différentes variantes de la langue et propice à l'activité de lecture partagée entre adulte et enfant qui est connue pour être bénéfique à la fois pour l'apprentissage du vocabulaire mais également pour l'acquisition de la morphosyntaxe (cf. Mol & Bus, 2011 pour une méta-analyse ; Grolig, 2020).

Par ailleurs et pour terminer, l'utilisation des deux grammaires dans des contextes différents ouvre ces recherches plutôt morphosyntaxiques aux aspects socio-pragmatiques de l'acquisition (cf. par exemple Nardy, Chevrot & Barbu, 2014), les enfants n'utilisant pas leurs deux grammaires de façon indifférenciée. La Section 2 explore ces dimensions socio-pragmatiques en examinant les données de deux enfants supplémentaires dont le français représente non pas une L1 mais une L2P (Palasis, 2022).

2. Français L2P

La variation de la morphosyntaxe du français a fait l'objet des sections précédentes, dans le discours adulte-adulte, en LAE, et dans le discours infantin

⁴¹ De nombreux travaux existent également sur la diglossie en arabe chez les enfants et les adultes (cf. Khamis-Dakwar, Froud & Gordon, 2011 ; Leikin, Ibrahim & Eghbaria, 2014 ; Froud & Khamis-Dakwar, 2018, entre autres).

L1. La variation morphosyntaxique du français est également bien documentée dans le cadre de l'apprentissage du français en tant que L2 par des adultes, grâce à des travaux qui examinent les points développés ci-dessus, notamment le doublement du DP (Nagy, Blondeau & Auger, 2003), l'élision du clitique *il* devant consonne (Howard, Lemée & Regan, 2006), la négation simple vs discontinue (Donaldson, 2017), et les formes interrogatives (Donaldson, 2016). Les études montrent que, généralement, les adultes L2 (re-)connaissent les différentes variantes morphosyntaxiques du français, même lorsqu'elles ne sont pas explicitées en cours, et que ces locuteurs adultes L2 sont ainsi capables d'adapter leur registre à leur interlocuteur, notamment en utilisant les variantes informelles lors de leurs interactions avec des locuteurs natifs (cf. Edmonds, 2020 pour une revue des problématiques et des points de variation). Ces capacités représentent des marqueurs d'une compétence socio-pragmatique élevée en L2 (Taguchi, 2019), sachant que l'existence de deux ou plusieurs variantes pour un même phénomène linguistique est considérée comme une difficulté importante dans l'apprentissage d'une L2 (Sorace, 2003 ; Moyer, 2021). Dans Palasis (2022), j'initie des recherches sur l'acquisition de la variation morphosyntaxique du français oral d'un troisième point de vue encore peu documenté (Taguchi, 2019, p. 9), celui de l'acquisition L2P. Deux facteurs permettent d'utiliser le terme « acquisition » : le jeune âge des enfants (2;7-4;11) et la transmission essentiellement implicite de la variation (environnement quotidien avant l'entrée en école primaire ; cf. Section I.2 du Chapitre I).

En plus des seize enfants L1 suivis pendant leurs trois années de maternelle, les enregistrements ont permis de collecter les données de deux enfants non-francophones sur la même période : ELA, L1 russe (N = 1 613), et RIA, L1 portugais (N = 1 355). Une étude contrastive L1-L2P est ainsi proposée concernant : (i) l'éventuelle existence d'une variation morphosyntaxique chez ELA et RIA sur les formes du sujet, l'ajustement phonologique du clitique *il* au contexte, et les formes de la négation, et (ii) la distribution des variantes en fonction du contexte.

Les Tableaux 13 et 14 ci-dessous rapportent les répartitions des différents types de sujets chez ELA et RIA et sont donc à mettre en regard avec les données des enfants L1 (Tableau 7, page 76). Concernant l'utilisation du DP seul, les données montrent des tendances globalement similaires à celles des enfants L1 : les taux sont très bas, plus précisément pour ELA et RIA, nuls en PS et MS et au-dessus du taux L1 en GS. Sur l'ensemble des catégories, une comparaison statistique des distributions montre que L1 et L2P ne présentent pas de différence significative en PS, mais que les grammaires des deux L2P deviennent significativement différentes entre elles en MS, puis significativement différentes entre elles et par rapport à la distribution L1 en GS (Tests exacts de Fisher, $p < .05$).

Tableau 13 : formes du sujet en maternelle (ELA – L1 russe)

ELA (ru)		PS (2;11-3;7)		MS (4;0-4;6)		GS (4;11-5;6)	
		N	%	N	%	N	%
a	Clitique seul	147	63,9	349	73,3	236	84,3
b	Clitique + DP	77	33,5	127	26,7	39	13,9
c	Sujet nul	6	2,6	0	0,0	1	0,4
d	DP seul	0	0,0	0	0,0	4	1,4
e	Elision (coord)	0	0,0	0	0,0	0	0,0
	TOTAL	230	100,0	476	100,0	280	100,0
a+b	TOTAL	224	97,4	476	100,0	275	98,2

Tableau 14 : formes du sujet en maternelle (RIA – L1 portugais)

RIA (po)		PS (2;7-3;2)		MS (3;7-4;2)		GS (4;7-5;2)	
		N	%	N	%	N	%
a	Clitique seul	113	66,5	186	64,4	137	60,6
b	Clitique + DP	51	30,0	100	34,6	85	37,6
c	Sujet nul	6	3,5	3	1,0	0	0,0
d	DP seul	0	0,0	0	0,0	4	1,8
e	Elision (coord)	0	0,0	0	0,0	0	0,0
	TOTAL	170	100,0	289	100,0	226	100,0
a+b	TOTAL	164	96,5	286	99,0	222	98,2

Une étude détaillée des distributions reste encore à faire, mais l'étude des contextes d'émergence du DP sans clitique chez ELA et RIA montre d'ores et déjà que le contexte est le même que chez les enfants L1 : les enfants sont en situation de lecture partagée, c'est-à-dire qu'ils manipulent des livres et produisent des phrases en faisant semblant de lire, comme illustré en (28) ci-après (RIA : 4;7, L2P et MAX : 4;5, L1). Une seconde similitude avec les enfants L1 est le fait que les

variantes formelles des différents phénomènes linguistiques émergent concomitamment. RIA produit une seule occurrence de négation discontinue sur l'ensemble des trois années et cette variante formelle de la négation émerge avec un DP sans clitique (cf. (28)). ELA par contre ne produit aucune négation discontinue. Enfin, concernant l'ajustement du clitique *il* au contexte, ELA montre une très forte progression des formes pleines de *il* devant consonne en MS et en GS, avec un taux très élevé de 19 % de formes non-éolidées devant consonne en GS (au-dessus du taux L1 le plus élevé : 12,9 %). RIA, quant à elle, ne produit aucune forme pleine devant consonne. L'hypothèse à vérifier est que ces divergences pourraient être liées à leurs systèmes L1, notamment pour ELA car les pronoms du russe ne sont pas clitiques comme en français.

3. Synthèse

Cette étude préliminaire du français L2P chez deux enfants de maternelle non-francophones a mis en avant des caractéristiques similaires avec les enfants L1 à différents niveaux : (i) la présence des mêmes variantes morphosyntaxiques, formelles et informelles, en L1 et en L2P, (ii) la présence dominante des variantes informelles en L1 et en L2P, et (iii) l'émergence progressive de la variante formelle, souvent amorcée par le contexte (notamment l'activité de lecture partagée) ou par la production d'un camarade L1, lui-même en situation de lecture, comme illustré en (28) et indiqué par la ligne « commentaires » (%com) de la base de données.

(28) Contexte d'émergence du français formel (L1 et L2P) : « lecture »

MAX : Sanbarbe s'assoit pour jouer aux cartes.

%com : faisant semblant de lire

RIA : mais Sanbarbe ne veut pas que...

%com : faisant semblant de lire

MAX : Sanbarbe est sur l'eau tout mouillé.

%com : faisant semblant de lire

En termes de processus acquisitionnels de la morphosyntaxe, les similitudes entre L1 et L2P mises en avant par la base de données sont attendues car tous les enfants sont dans la même tranche d'âge et développent ainsi leurs compétences linguistiques dans des contextes essentiellement de transmission implicite avec une modalité orale (cf. Section 1.2 du Chapitre I). Il est donc posé que c'est essentiellement la mémoire non-déclarative (celle des procédures et des généralisations) qui est impliquée, quelle que soit la langue (cf. Tableau 1, page 22). À la lumière de ces premiers résultats qui portent également sur la (re-)connaissance de la variation et de ses contextes distributionnels, il semblerait que les compétences socio-pragmatiques des jeunes enfants soient également développées selon des principes implicites, en L1 et en L2P.

L'hypothèse de la diglossie francophone proposée pour les enfants dont le français est leur L1 (cf. Section IV.1) semble alors pouvoir être étendue au français L2P, et se traduit alors par l'acquisition de deux L2P, avec la même réserve qu'en diglossie L1 : tous les locuteurs n'atteignent pas le même niveau de compétence en G2. Une dernière question à laquelle un début de réponse est apporté ici concerne le fait que ces enfants L2P aient réussi à intégrer des éléments de la variation morphosyntaxique du français avant certains enfants L1. Une réponse se trouve peut-être dans la littérature abondante sur les avantages cognitifs qu'entraîne le plurilinguisme (cf. Siegal et al., 2010 ; Kovács, 2015, entre autres) et dans les travaux naissants sur les interactions entre développement du langage et socialisation à la maternelle (Dai et al., 2022).

Bilan et perspectives

Le Chapitre I de ce texte d'Habilitation à Diriger des Recherches a proposé une synthèse actualisée et commentée de recherches sur les processus qui sous-tendent l'acquisition du langage (cf. Section I) et sur le français hexagonal non-standard (appelé aussi informel ou oral) utilisé par les adultes à l'adresse d'adultes et par les adultes à l'adresse d'enfants (cf. Section II et Tableau 11, p. 102 pour une comparaison des deux). Chaque section a également proposé un panorama des principales approches théoriques qui visent à formaliser les processus acquisitionnels d'une part et la variation morphosyntaxique du français d'autre part (cf. Sections I.3 et II.3). Le Chapitre I a ainsi permis d'explicitier mes choix en termes de cadres théoriques : celui d'un modèle plutôt « GU + statistiques » pour les processus acquisitionnels (cf. Section I.4) et celui de la diglossie pour la variation morphosyntaxique du français (cf. Section II.4).

Les travaux synthétisés dans le Chapitre II représentent alors la rencontre entre certains fondements théoriques de la Grammaire générative et un apport empirique grâce à une base de données enfantines collectée pendant trois ans auprès d'une classe de maternelle en France (cf. Tableau 3, p. 63) et partiellement partagée sur la plateforme CHILDES (Palasis, 2012a).⁴² Les résultats, qui proviennent tous de l'étude détaillée de ces données, par phénomène linguistique, par enfant, par année, etc., permettent alors d'apporter une description fine du développement du système linguistique enfantin pendant les années de maternelle et de proposer des analyses qui mettent en avant les dynamiques endogènes à ce système enfantin et non pas d'éventuelles déficiences de compétence ou de performance, comme souvent évoquées (cf. Section I).

Les travaux sélectionnés pour cette HDR portent sur trois thématiques principales : les formes du sujet, les structures des questions *wh-* et la

⁴² <http://chilides.talkbank.org/access/French/Palasis.html>

formalisation de la variation. Les résultats principaux sont synthétisés en (29) ci-après.

(29) Récapitulatif des résultats principaux :

- i. Le phénomène du « sujet » nul enfantin est décrit comme une absence du clitique nominatif, redondant dans un système enfantin initial dont les formes verbales sont majoritairement morphologiquement distinctives (cf. Section 1 et Tableau 6, p. 70).
- ii. Le clitique nominatif est un marqueur préverbal d'accord (cf. Section 2 et synthèse en (18), p. 87), ce qui entraîne une analyse du français enfantin en tant que langue à « sujet nul », avec de grandes similitudes avec les dialectes nord-italiens.
- iii. La position in-situ ou ex-situ des éléments *wh-* dans le système enfantin est un effet de la forme du verbe (Fixe : in-situ vs Libre : ex-situ, cf. (22) p. 96).
- iv. Le français enfantin est une langue à « *wh-in-situ* », ce qui dissocie la modalité interrogative du déplacement de l'item *wh-* (cf. (24) p. 99).
- v. La sémantique verbale et la deixis dans les questions enfantines représentent des ponts vers une explication longitudinale entre fonctionnements enfantin et adulte, dans lequel le trait [Exclusion] est proposé comme déclencheur du déplacement *wh-ex-situ* (cf. Section III.3).
- vi. La variation morphosyntaxique en français informel (enfantin et adulte, cf. Tableau 12, p. 106) est formulée dans le cadre de la diglossie : tous les locuteurs acquièrent G1 dans l'enfance et le développement de G2 dépend de l'environnement précoce, que le français soit une L1 ou une L2P (cf. Section IV).

À partir de ces résultats, des investigations restent à mener sur l'effet de fréquence générée par un système linguistique (cf. Section II.1.3), les différents types d'explétifs (cf. Section II.2.5), et les nombreuses catégories de questions laissées de côté pour le moment (cf. (19) p. 92). L'étude du développement des clivées, qui représentent une autre structure syntaxique très étudiée (Lambrecht,

2001), commence tout juste dans le cadre d'une collaboration internationale (Lahousse et al., 2023 ; Wensink et al., 2024). La perspective diachronique de la diglossie a été effleurée et mériterait d'être poursuivie, ainsi que celle de l'acquisition L2P avec d'autres phénomènes linguistiques documentés en L2 tels que l'interrogation et les formes du futur (Donaldson, 2016 ; Gudmestad, Edmonds, Donaldson & Carmichael, 2020). Au niveau interlinguistique, les parallèles mentionnés entre le français informel et les dialectes nord-italiens (langues à sujet nul avec marqueur préverbal d'accord et langues qui présentent à la fois des *wh*-in-situ et ex-situ) pourraient être prolongés. Enfin, l'interface entre morphosyntaxe et socio-pragmatique est un domaine émergent dans le cadre de l'acquisition du langage. Ce domaine d'investigation pluridisciplinaire mériterait donc également d'être développé.

Par ailleurs, des recherches collaboratives sur des aspects plus cognitifs du traitement linguistique ont été initiées et ont vocation à être continuées sur l'effet de l'âge sur le traitement des pronoms anaphoriques et cataphoriques (Arslan, Palasis & Meunier, 2020) et sur l'effet de l'espacement des séances d'apprentissage sur la rétention des informations lexicales et grammaticales en L2 (Edmonds, Gerbier, Palasis & Whyte, 2021 ; Whyte, Edmonds, Palasis & Gerbier, 2022).

Enfin, des problématiques qui touchent à la transition développementale entre le français spontané (G1) et le français normé (G2) ont également émergé ces dernières années dans mes recherches, notamment avec la question du traitement graphique des morphèmes silencieux tels que le <-s> du pluriel nominal chez les adultes et chez les enfants à l'école primaire et au collège (Palasis & Meunier, 2017 ; de los Angeles Cortes Arboleda, 2022 ; Lemaire, 2022) et la rédaction de narrations par des élèves dysphasiques (Broc, Palasis & Olive, 2020). Recherche et éducation semblent alors se rejoindre, avec tous les enjeux sociétaux que cela représente. La Section IV.1 a montré que tous les enfants L1 acquièrent G1, mais que tous les enfants n'entrent pas au CP avec le même niveau de G2. (Re-)connaître G1 à l'école primaire en tant que système linguistique à part entière

que tous les enfants utilisent de façon inconsciente initialement pourrait alors représenter un socle commun de connaissances dès la maternelle pour aller vers une transmission explicite de la morphosyntaxe de G2 à l'école élémentaire et au collège.

Références bibliographiques

- Ambridge, B., Rowland, C. F. & Pine, J. M. (2008). Is structure dependence an innate constraint? New experimental evidence from children's complex-question production. *Cognitive Science*, 32(1), 222-255. <https://doi.org/10.1080/03640210701703766>
- Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, 42(2), 239-273. <https://doi.org/10.1017/S030500091400049X>
- Ambridge, B. & Lieven, E. V. M. (2015). A constructivist account of language acquisition. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp. 478-510). Chichester: Wiley & Sons.
- Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, 40(5-6), 509-559. <https://doi.org/10.1177/0142723719869731>
- Ambridge, B. & Wagner, L. (2021). Testable theories of core first language acquisition. *Journal of Child Language*, 48(5), 859-861. <https://doi.org/10.1017/S0305000921000581>
- Aoun, J., Hornstein, N. & Sportiche, D. (1981). Some aspects of wide scope quantification. *Journal of Linguistic Research*, 1(3), 69-95.
- Aristote. ([1994]). *De anima*. Paris: Gallimard.
- Armstrong, N. & Smith, A. (2002). The Influence of Linguistic and Social Factors on the Recent Decline of French *ne*. *Journal of French Language Studies*, 12(1), 23-41. <https://doi.org/10.1017/S0959269502000121>
- Arslan, S., Palasis, K. & Meunier, F. (2020). Electrophysiological differences in older and younger adults' anaphoric but not cataphoric pronoun processing in the absence of age-related behavioural slowdown. *Scientific Reports*, 10(19234). <https://doi.org/10.1038/s41598-020-75550-3>
- Ashby, W. J. (1976). The loss of the negative morpheme *ne* in Parisian French. *Lingua*, 39(1-2), 119-137. [https://doi.org/10.1016/0024-3841\(76\)90060-7](https://doi.org/10.1016/0024-3841(76)90060-7)
- Ashby, W. J. (1977a). Interrogative forms in Parisian French. *Semasia*, 4, 35-52.
- Ashby, W. J. (1977b). *Clitic Inflection in French: A Historical Perspective*. Amsterdam: Rodopi.
- Ashby, W. J. (1980). *Is French Becoming a Verb-Initial Language?* Paper presented at the LSA Summer Meeting.

- Ashby, W. J. (1981). The loss of the negative particle *ne* in French: A syntactic change in progress. *Language*, 57(3), 647-687. <https://doi.org/10.2307/414345>
- Ashby, W. J. (1982). The Drift of French Syntax. *Lingua*, 57(1), 29-46. [https://doi.org/10.1016/0024-3841\(82\)90069-9](https://doi.org/10.1016/0024-3841(82)90069-9)
- Ashby, W. J. (1984). The elision of /l/ in French clitic pronouns and articles. In E. Pulgram (Ed.), *Romanitas: Studies in Romance Linguistics* (pp. 1-16). Ann Arbor: University of Michigan Press.
- Ashby, W. J. (1988). Français du Canada/français de France: divergence et convergence. *The French Review*, 61(5), 693-702. <https://www.jstor.org/stable/393384>
- Ashby, W. J. (2001). Un nouveau regard sur la chute du *ne* en français parlé tourangeau: s'agit-il d'un changement en cours? *Journal of French Language Studies*, 11(1), 1-22. <https://doi.org/10.1017/S0959269501000114>
- Ashby, W. J. (2023). *On Spoken French. An Ashby Reader*. Amsterdam: John Benjamins.
- Auger, J. (1994). *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*. (PhD Dissertation), University of Pennsylvania.
- Auger, J. (1995). On the history of relative clauses in French and some of its dialects. In H. Andersen (Ed.), *Historical Linguistics 1993: Selected Papers from the 11th International Conference on Historical Linguistics* (pp. 19-32). Amsterdam: John Benjamins.
- Auger, J. (2022). Picard subject clitics. An analysis at the interface of syntax, phonology, and prosody. In G. Alboiu & R. King (Eds.), *Points of Convergence in Romance Linguistics. Papers selected from the 48th Linguistic Symposium on Romance Languages (LSRL 48), Toronto, 25-28 April 2018* (pp. 11-34). Dordrecht: John Benjamins.
- Ayoun, D. (2003). *Parameter-setting in Language Acquisition*. London & New York: Continuum Press.
- Bahtchevanova, M. & van Gelderen, E. (2016). The interaction between the French subject and object cycles. In E. van Gelderen (Ed.), *Cyclical Change Continued* (pp. 113-136). Amsterdam: John Benjamins.
- Baker, M. C. (1996). *The Polysynthesis Parameter*. New York & Oxford: Oxford University Press.
- Baker, M. C. (2005). Mapping the Terrain of Language Learning. *Language Learning and Development*, 1(1), 93-129. https://doi.org/10.1207/s15473341lld0101_6

- Barra-Jover, M. (2004). Interrogatives, négatives et évolution des traits formels du verbe en français parlé. *Langue française*, 141, 110-125.
- Barra-Jover, M. (2010). « Le » français ou ce qui arrive lorsqu'un état de choses est observé comme une entité. *Langue française*, 168(4), 3-18.
- Bassano, D. & Mendes-Maillochon, I. (1994). Early grammatical and prosodic marking of utterance modality in French: a longitudinal case study. *Journal of Child Language*, 21(3), 649-675.
<https://doi.org/10.1017/S0305000900009491>
- Bassano, D., Eme, P.-E. & Champaud, C. (2005). A naturalistic study of early lexical development: General processes and inter-individual variations in French children. *First Language*, 25(1), 67-101.
<https://doi.org/10.1177/0142723705049119>
- Bates, A. & Pearl, L. (2019). What do you think that happens? A quantitative and cognitive modeling analysis of linguistic evidence across socioeconomic status for learning syntactic islands. In M. M. Brown & B. Dailey (Eds.), *Proceedings of the 43rd Annual Boston University Conference on Language Development* (Vol. 1, pp. 42-56). Somerville, MA: Cascadilla Press.
- Baude, O. (Ed.) (2006). *Corpus Oraux. Guide des bonnes pratiques 2006*. Orléans & Paris: Presses Universitaires d'Orléans & CNRS Editions.
- Baunaz, L. (2011). *The Grammar of French Quantification*. Dordrecht: Springer.
- Baunaz, L. & Patin, C. (2011). Prosody refers to semantic factors: Evidence from French *wh*- words. In H. Yoo & E. Delais-Roussarie (Eds.), *Actes de Interface Discours et Prosodie 2009* (pp. 96-107):
http://makino.linguist.jussieu.fr/idp09/docs/IDP_actes/Articles/baunaz.pdf.
- Baunaz, L. (2016). French "quantifiers" in questions: Interface strategies. *Acta Linguistica Hungarica*, 63(2), 125-168.
<https://doi.org/10.1556/064.2016.63.2.1>
- Baunaz, L. & Bonan, C. (2023). Activation levels: A fresh perspective on French *wh* in-situ. *Isogloss. Open Journal of Romance Linguistics*, 9(1/7), 1-43.
<https://doi.org/10.5565/rev/isogloss.305>
- Becker, M. & Gotowski, M. (2015). Explaining children's *wh*-in situ questions: Against economy. In E. Grillo & K. Jepson (Eds.), *Proceedings of the 39th annual Boston University Conference on Language Development [BUCLD 39]* (pp. 88-100). Somerville, MA: Cascadilla Press.
- Béguelin, M.-J., Coveney, A. & Guryev, A. (Eds.). (2018). *L'interrogative en français*. Berne: Peter Lang.

- Belletti, A. (2017). Internal Grammar and Children's Grammatical Creativity against Poor Inputs. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.02074>
- Benavides-Varela, S. & Gervain, J. (2017). Learning word order at birth: A NIRS study. *Developmental Cognitive Neuroscience*, 25, 198-208. <https://doi.org/10.1016/j.dcn.2017.03.003>
- Benveniste, E. (1946). Structure des relations de personne dans le verbe *Problèmes de linguistique générale*, 1 (pp. 225-236). Paris: Editions Gallimard, 1966.
- Bergelson, E. & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-month-olds. *Proceedings of the National Academy of Sciences*, 114(49), 12916-12921.
- Berko, J. (1958). The Child's Learning of English Morphology. *Word*, 14(2-3), 150-177. <https://doi.org/10.1080/00437956.1958.11659661>
- Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, 17(2), 89-98. <https://doi.org/10.1016/j.tics.2012.12.002>
- Blanche-Benveniste, C. (1983). L'importance du 'français parlé' pour la description du 'français tout court'. *Recherches sur le français parlé*, 5, 23-45.
- Blanche-Benveniste, C. (1991). *Le français parlé : Etudes grammaticales*. Paris: Editions du Centre National de la Recherche Scientifique.
- Blanche-Benveniste, C. (1994). Quelques caractéristiques grammaticales des "sujets" employés dans le français parlé des conversations. In M. Yaguello (Ed.), *Subjecthood and Subjectivity. Proceedings of the Colloquium "The Status of the Subject in Linguistic Theory"* (pp. 77-107). London: Editions Ophrys.
- Blanche-Benveniste, C. (2000). *Approches de la langue parlée en français*. Paris: Editions Ophrys.
- Blanche-Benveniste, C. (2003a). Les formes grammaticales de réalisation des sujets et leur inégale représentation en français contemporain. In J.-M. Merle (Ed.), *Le sujet* (pp. 73-90). Paris: Editions Ophrys.
- Blanche-Benveniste, C. (2003b). La langue parlée. In M. Yaguello (Ed.), *Le grand livre de la langue française* (pp. 317-344). Paris: Editions du Seuil.
- Blanche-Benveniste, C. & Martin, P. (2010). *Le français : Usages de la langue parlée*. Leuven: Peeters.
- Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

- Boë, L.-J., Sawallis, T. R., Fagot, J., Badin, P., Barbier, G., Captier, G., . . . Schwartz, J.-L. (2019). Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science. *Science Advances*, 5(12), eaaw3916. <https://doi.org/10.1126/sciadv.aaw3916>
- Boeckx, C. & Leivada, E. (2013). Entangled parametric hierarchies: Problems for an overspecified universal grammar. *PLoS ONE*, 8(9), e72357. <https://doi.org/10.1371/journal.pone.0072357>
- Bohn, M. & Frank, M. C. (2019). The pervasive role of pragmatics in early language. *Annual Review of Developmental Psychology*, 1, 223-249. <https://doi.org/10.1146/annurev-devpsych-121318085037>
- Bonan, C. (2021). *Romance Interrogative Syntax. Formal and typological dimensions of variation*. Amsterdam: John Benjamins.
- Bošković, Ž. (2000). Sometimes in [Spec, CP], sometimes in-situ. In R. Martin, D. Michaels & J. Uriagereka (Eds.), *Step by Step: Essays on Minimalism in Honor of Howard Lasnik* (pp. 53-87). Cambridge, MA: MIT Press.
- Brandi, L. & Cordin, P. (1989). Two Italian Dialects and the Null Subject Parameter. In O. Jaeggli & K. Safir (Eds.), *The Null Subject Parameter* (pp. 111-142). Dordrecht: Kluwer Academic Publishers.
- Broc, L., Palasis, K. & Olive, T. (2020). Ecrire des narrations permet aux élèves dysphasiques de mobiliser leurs connaissances syntaxiques. In F. Neveu, B. Harmegnies, L. Hriba, S. Prévost, & A. Steuckardt (Eds.), *SHS Web of Conferences: 7^e Congrès Mondial de Linguistique Française* (Vol. 78, pp. 10003). Montpellier: EDP Sciences.
- Buson, L., Nardy, A., Rousset, I., & Zhang, C. (2022). Pratiques stylistiques en contexte scolaire: usages contrastés du 'ne' de négation chez les adultes et les enfants d'une école maternelle française. In F. Neveu, S. Prévost, A. Steuckardt, G. Bergounioux, & B. Hama (Eds.), *SHS Web of Conferences: Congrès Mondial de Linguistique Française - CMLF 2022* (Vol. 138, pp. 06018). Orléans: EDP Sciences.
- Cabredo Hofherr, P. (2004). Les clitiques sujets du français et le paramètre du sujet nul. *Langue française*, 141, 99-109.
- Camacho, J. A. (2013). *Null Subjects*. Cambridge: Cambridge University Press.
- Canut, E. (2014). Acquisition des constructions syntaxiques complexes chez l'enfant français entre 2 et 6 ans *Congrès Mondial de Linguistique Française - CMLF 2014 - SHS Web of Conferences 8*: EDP Sciences, <https://doi.org/10.1051/shsconf/20140801092>.
- Cappeau, P. (Ed.) (2021). *Une grammaire à l'aune de l'oral ?* Rennes: Presses Universitaires de Rennes.

- Chang, L. (1997). *Wh-in-situ Phenomena in French*. (MA thesis), University of British Columbia.
- Cheng, L. L.-S. & Rooryck, J. (2000). Licensing *Wh-in-situ*. *Syntax*, 3(1), 1-19. <https://doi.org/10.1111/1467-9612.00022>
- Chervel, A. (2011). Histoire de l'orthographe française : réformes, enseignement, niveau. In O. Bertrand & I. Schaffner (Eds.), *Variétés, variations & formes du français* (pp. 195-210). Palaiseau: Les Editions de l'Ecole Polytechnique.
- Chevrot, J.-P., Chabanal, D. & Dugua, C. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage : trois études de cas. *Journal of French Language Studies*, 17(1), 103-128. <https://doi.org/10.1017/S0959269506002663>
- Chifflet, L. (1659). *Essay d'une parfaite grammaire de la langue française*. Anvers: Jacques van Meurs.
- Chomsky, N. (1957). *Syntactic Structures*. La Haye: Mouton ; traduction française *Structures syntaxiques*, 1969, Paris: Editions du Seuil.
- Chomsky, N. (1959). A Review of B.F. Skinner's *Verbal Behavior*. *Language*, 35(1), 26-58 ; reprinted in Lust, Barbara C. & Claire Foley (eds.), 2004, *First Language Acquisition. The Essential Readings*, Oxford: Blackwell, p. 2025-2055.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press ; traduction française *Aspects de la théorie syntaxique*, 1971, Paris: Editions du Seuil.
- Chomsky, N. (1968). *Language and Mind*. New York: Harcourt, Brace & World, Inc.
- Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon Books ; traduction française *Réflexions sur le langage*, 1981, Paris: Flammarion.
- Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures* (traduction française *Théorie du gouvernement et du liage : les conférences de Pise*, 1991, Paris: Editions du Seuil ed.). Dordrecht: Foris.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.
- Chomsky, N. (1998). Some Observations on Economy in Generative Grammar. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, & D. Pesetsky (Eds.), *Is the Best Good Enough? Optimality and Competition in Syntax* (pp. 115-127). Cambridge, MA: Massachusetts Institute of Technology.
- Chomsky, N. (1999). Derivation by Phase *MIT Working Papers in Linguistics*, 18 ; reprinted in Ken Hale: *A Life in Language*, Kenstowicz, Michael (ed.), 2001, Cambridge, MA: The MIT Press, p. 1-52.

- Chomsky, N. (2005). Three Factors in Language Design. *Linguistic Inquiry*, 36(1), 1-22. <https://doi.org/10.1162/0024389052993655>
- Chomsky, N., Gallego, Á. J. & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*, 229-261. <https://doi.org/10.5565/rev/catjl.288>
- Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, 51(1-2), 61-75. <https://doi.org/10.1177/00238309080510010501>
- Clark, E. V. (1978). Strategies for Communicating. *Child Development*, 49(4), 953-959.
- Clark, E. V. (1987). The Principle of Contrast: A Constraint on Language Acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1-33). Hillsdale, NJ: Lawrence Erlbaum.
- Clark, E. V. (2004). Pragmatics and Language Acquisition. In L. R. Horn & G. L. Ward (Eds.), *The Handbook of Pragmatics* (pp. 562-577). Oxford: Blackwell.
- Cochran, W. G. (1954). Some Methods for Strengthening the Common χ^2 Tests. *Biometrics*, 10(4), 417-451.
- Cognola, F. & Casalicchio, J. (Eds.). (2018). *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*. Oxford: Oxford University Press.
- Cohen, M. (1924). Sur les langages successifs de l'enfant. *L'acquisition du langage oral et écrit*, 31 (1993), 33-43.
- Coopmans, C. W., De Hoop, H., Kaushik, K., Hagoort, P., & Martin, A. E. (2022). Hierarchy in language interpretation: evidence from behavioural experiments and computational modelling. *Language, Cognition and Neuroscience*, 37(4), 420-439. <https://doi.org/10.1080/23273798.2021.1980595>
- Cornips, L. (2015). The no man's land between syntax and variationist sociolinguistics: The case of idiolectal variability. In A. Adli, M. Garcia Garcia & G. Kaufmann (Eds.), *Variation in Language: System- and Usage-based Approaches* (pp. 147-172). Berlin: Walter de Gruyter.
- Côté, M.-H. (2001). On the Status of Subject Clitics in Child French. In M. Almgren, A. Barrena, M. Ezeizabarrena, I. Idiazabal, & B. MacWhinney (Eds.), *Proceedings of the 8th Conference of the IASCL* (pp. 1314-1330). Somerville, MA: Cascadilla Press.
- Coveney, A. (1989). Pragmatic constraints on interrogatives in spoken French. *York Papers in Linguistics*, 13, 89-99.

- Coveney, A. (1995). The use of the QU-final interrogative structure in spoken French. *Journal of French Language Studies*, 5(2), 143-171. <https://doi.org/10.1017/S0959269500002738>
- Coveney, A. (1996). *Variability in Spoken French: Interrogation and Negation*. Exeter: Elm Bank; reissued with a Supplement, Bristol: Intellect Books (2002).
- Coveney, A. (1997). L'approche variationniste et la description de la grammaire du français : le cas des interrogatives. *Langue française*, 115, 88-100.
- Coveney, A. (2002). *Variability in Spoken French: A Sociolinguistic Study of Interrogation and Negation*. Bristol: Elm Bank.
- Coveney, A. (2011). A language divided against itself? Diglossia, code-switching and variation in French. In F. Martineau & T. Nadasdi (Eds.), *Le français en contact : Hommages à Raymond Mougéon* (pp. 51-85). Québec: Presses de l'Université de Laval.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87-114. <https://doi.org/10.1017/S0140525X01003922>
- Crain, S. & Nakayama, M. (1987). Structure Dependence in Grammar Formation. *Language*, 63(3), 522-543. <https://doi.org/10.2307/415004>
- Crain, S., Koring, L. & Thornton, R. (2017). Language acquisition from a biolinguistic perspective. *Neuroscience & Biobehavioral Reviews*, 81(B), 120-149. <https://doi.org/10.1016/j.neubiorev.2016.09.004>
- Creissels, D. (1995). *Éléments de syntaxe générale*. Paris: Presses Universitaires de France.
- Crisma, P. (1992). On the Acquisition of Wh-Questions in French. *Geneva Generative Papers*, 0(1-2), 115-122.
- Culbertson, J. (2010). Convergent Evidence for Categorical Change in French: From Subject Clitic to Agreement Marker. *Language*, 86(1), 85-132. <https://www.jstor.org/stable/40666300>
- Culbertson, J. & Legendre, G. (2014). Prefixal agreement and impersonal 'il' in Spoken French: Experimental evidence. *Journal of French Language Studies*, 24(1), 83-105. <https://doi.org/10.1017/S0959269513000380>
- D'Alessandro, R. (2015). Null subjects. In A. Fábregas, J. Mateu & M. T. Putnam (Eds.), *Contemporary Linguistic Parameters: Contemporary Studies in Linguistics* (pp. 201-226). London: Bloomsbury Academic.

- Dabrowska, E. (2000). From formula to schema: The acquisition of English questions. *Cognitive Linguistics*, 11(1-2), 83-102. <https://doi.org/10.1515/cogl.2001.013>
- Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, 178, 222-235. <https://doi.org/10.1016/j.cognition.2018.05.018>
- Dai, S., Bouchet, H., Karsai, M., Chevrot, J.-P., Fleury, E., & Nardy, A. (2022). Longitudinal data collection to follow social network and language development dynamics at preschool. *Scientific Data*, 9(1), 777. <https://doi.org/10.1038/s41597-022-01756-x>
- Dardier, V. (2004). *Pragmatique et pathologies. Comment étudier les troubles de l'usage du langage*. Rosny-sous-Bois: Bréal éditions.
- Darwin, C. (1877). A Biographical Sketch of an Infant. *Mind*, 2, 285-294 ; traduction française "Biographie d'un petit enfant en quelques traits" dans Delefosse, J.-M. Odéric (éd.), 2010, *Sur le langage de l'enfant. Choix de textes de 1876 à 1962*, p. 2081-2090.
- Das, R., Jana, N., Arora, N., & Sengupta, S. (2020). Ultrasound assessment of fetal hearing response to vibroacoustic stimulation. *The Journal of Maternal-Fetal & Neonatal Medicine*, 33(14), 2326-2332. <https://doi.org/10.1080/14767058.2018.1548600>
- De Boysson-Bardies, B. (1996). *Comment la parole vient aux enfants. De la naissance jusqu'à deux ans*. Paris: Editions Odile Jacob.
- De Cat, C. (2002). *French Dislocation*. (PhD thesis), University of York.
- De Cat, C. (2005). French Subject Clitics are not Agreement Markers. *Lingua*, 115(9), 1195-1219. <https://doi.org/10.1016/j.lingua.2004.02.002>
- De Cat, C. (2007). *French Dislocation. Interpretation, Syntax, Acquisition*. Oxford & New York: Oxford University Press.
- de Houwer, A. (2019). Développement et évaluation globale du langage chez le jeune enfant plurilingue : le rôle central de l'environnement linguistique. In S. Kern (Ed.), *Le développement du langage chez le jeune enfant. Théorie, clinique, pratique* (pp. 175-200). Louvain-la-Neuve: De Boeck Supérieur.
- de la Cruz-Pavía, I. & Gervain, J. (2023). Six-month-old infants' perception of structural regularities in speech. *Cognition*, 238, 105526. <https://doi.org/10.1016/j.cognition.2023.105526>
- de los Angeles Cortes Arboleda, S. (2022). *L'apprentissage du morphème silencieux du pluriel des noms en français*. (Master 2 de Psychologie, Parcours

Psychologie du développement, des apprentissages et de l'éducation),
Université Côte d'Azur.

- Dehaene-Lambertz, G. (2017). The human infant brain: A neural architecture able to learn language. *Psychonomic Bulletin & Review*, 24(1), 48-55. <https://doi.org/10.3758/s13423-016-1156-9>
- Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron*, 88(1), 2-19. <https://doi.org/10.1016/j.neuron.2015.09.019>
- Dehaene, S. (2018). *Apprendre ! Les talents du cerveau, le défi des machines*. Paris: Odile Jacob.
- Demuth, K. (1996). Collecting Spontaneous Production Data. In D. McDaniel, C. McKee & H. S. Cairns (Eds.), *Methods for Assessing Children's Syntax* (pp. 3-22). Cambridge, MA: The MIT Press.
- Demuth, K. & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, 35(1), 99-127. <https://doi.org/10.1017/S0305000907008276>
- Déprez, V. & Pierce, A. (1993). Negation and Functional Projections in Early Grammar. *Linguistic Inquiry*, 24(1), 25-67.
- Déprez, V., Syrett, K. & Kawahara, S. (2012). Interfacing information and prosody: French *wh-in-situ* questions. In I. Franco, S. Lusini & A. Saab (Eds.), *Romance Languages and Linguistic Theory 2010. Selected papers from Going Romance Leiden 2010* (pp. 135-154). Amsterdam: John Benjamins.
- Déprez, V., Syrett, K. & Kawahara, S. (2013). The interaction of syntax, prosody, and discourse in licensing French *wh-in-situ* questions. *Lingua*, 124, 4-19. <https://doi.org/10.1016/j.lingua.2012.03.002>
- Di Sciullo, A. M. & Boeckx, C. (Eds.). (2011). *The Bilingual Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford: Oxford University Press.
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1), 158-164. <https://doi.org/10.1038/nn.4186>
- Donaldson, B. (2016). Aspects of interrogative use in near-native French: Form, function, and register. *Linguistic Approaches to Bilingualism*, 6(4), 467-503. <https://doi.org/10.1075/lab.14024.don>

- Donaldson, B. (2017). Negation in near-native French: Variation and sociolinguistic competence. *Language Learning*, 67(1), 141-170. <https://doi.org/10.1111/lang.12201>
- Dryer, M. S. (2013). Position of Interrogative Phrases in Content Questions. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online* (pp. <http://wals.info/chapter/93>). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Dugua, C., Nardy, A., Liégeois, L., Chevrot, J.-P., & Chabanal, D. (2017). L'acquisition des liaisons après les clitiques préverbaux est-elle spécifique? Apport d'une expérimentation à grande échelle. *Journal of French Language Studies*, 27(1), 73-86. <https://doi.org/10.1017/S0959269516000387>
- Dye, C., Kedar, Y. & Lust, B. (2019). From lexical to functional categories: New foundations for the study of language development. *First Language*, 39(1), 9-32. <https://doi.org/10.1177/0142723718809175>
- Edmonds, A. (2020). *L'acquisition d'une langue seconde : éclairages lexical et variationniste et réflexions méthodologiques*. (Habilitation à Diriger des Recherches), Université Paris 8
- Edmonds, A., Gerbier, E., Palasis, K., & Whyte, S. (2021). Understanding the distributed practice effect and its relevance for the teaching and learning of L2 vocabulary. *Lexis*, 18: *Lexical learning and teaching*. <https://doi.org/10.4000/lexis.5652>
- Eisenbeiss, S. (2010). Production methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 11-34). Amsterdam: John Benjamins.
- Ervin, S. (1964). Imitation and Structural Change in Children's language. In E. H. Lenneberg (Ed.), *New Directions in the Study of Language* (pp. 163-189). Cambridge, MA: The MIT Press.
- Eubank, L. & Gregg, K. R. (1999). Critical Periods and (Second) Language Acquisition: Divide et Impera. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 65-99). Mahwah, NJ: Lawrence Erlbaum.
- Ezeizabarrena, M.-J. (2003). Null Subjects and Optional Infinitives in Basque. In N. Müller (Ed.), *(In)vulnerable Domains in Multilingualism* (pp. 83-106). Amsterdam: John Benjamins.
- Faure, R. & Palasis, K. (2014). *French wh-questions: Crossing diglossia and topicality*. Paper presented at the 44th Linguistic Symposium on Romance Languages (LSRL), London, Ontario.

- Faure, R. & Palasis, K. (2021). Exclusivity! *Wh*-fronting is not optional *wh*-movement in Colloquial French. *Natural Language & Linguistic Theory*, 39(1), 57-95. <https://doi.org/10.1007/s11049-020-09476-w>
- Ferdinand, R. A. (1996). *The Development of Functional Categories: The Acquisition of the Subject in French*. La Hague: Holland Academic Graphics.
- Ferguson, C. A. (1959). Diglossia. *Word*, 15(2), 325-340. <https://doi.org/10.1080/00437956.1959.11659702>
- Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist*, 66(6), 103-114.
- Finn, A. S., Kalra, P. B., Goetz, C., Leonard, J. A., Sheridan, M. A., & Gabrieli, J. D. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology*, 142, 212-220. <https://doi.org/10.2307/415004>
- Finn, A. S., Kharitonova, M., Holtby, N., & Sheridan, M. A. (2019). Prefrontal and hippocampal structure predict statistical learning ability in early childhood. *Journal of Cognitive Neuroscience*, 31(1), 126-137. https://doi.org/10.1162/jocn_a_01342
- Foisil, M. (Ed.) (1989). *Journal de Jean Héroard Médecin de Louis XIII*. Paris: Fayard.
- Fonseca-Greber, B. (2000). *The change from pronoun to clitic to prefix and the rise of null subjects in spoken Swiss French*. (PhD dissertation), University of Arizona, Arizona.
- Fonseca-Greber, B. & Waugh, L. R. (2003). The Subject Clitics of Conversational European French: Morphologization, Grammatical Change, Semantic Change, and Change in Progress. In R. Núñez-Cedeño, L. López & R. Cameron (Eds.), *A Romance Perspective on Language Knowledge and Use* (pp. 99-117). Amsterdam: John Benjamins.
- Fonseca-Greber, B. (2013). La morphologisation de *qui*. *Journal of French Language Studies*, 23(3), 401-421. <http://doi.org/10.1017/S0959269512000270>
- Fonseca-Greber, B. (2018). A relative verb for Spoken French: The missing link. *Lingua*, 210-211, 1-29. <https://doi.org/10.1016/j.lingua.2018.03.007>
- Forest, T. A., Schlichting, M. L., Duncan, K. D., & Finn, A. S. (2023). Changes in statistical learning across development. *Nature Reviews Psychology*, 2(4), 205-219. <https://doi.org/10.1038/s44159-023-00157-0>
- Franck, J., Millotte, S., Posada, A., & Rizzi, L. (2013). Abstract knowledge of word order by 19 months: An eye-tracking study. *Applied Psycholinguistics*, 34(2), 323-336. <https://doi.org/10.1017/S0142716411000713>

- Friedemann, M.-A. (1993). The Underlying Position of External Arguments in French: A Study in Adult and Child Grammar. *Language Acquisition*, 3(3), 209-255. https://doi.org/10.1207/s15327817la0303_2
- Friedemann, M.-A. & Rizzi, L. (2000). The Acquisition of Syntax: Introduction. In M.-A. Friedemann & L. Rizzi (Eds.), *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics* (pp. 1-25). London: Longman.
- Friederici, A. D., Mueller, J. L. & Oberecker, R. (2011). Precursors to natural grammar learning: preliminary evidence from 4-month-old infants. *PLoS ONE*, 6(3), e17920. <https://doi.org/10.1371/journal.pone.0017920>
- Friedmann, N., Belletti, A. & Rizzi, L. (2021). Growing Trees: The acquisition of the left periphery. *Glossa: A Journal of General Linguistics*, 6(1), 131. <https://doi.org/10.16995/glossa.5877>
- Froud, K. & Khamis-Dakwar, R. (2018). Neurophysiological investigations in studies of Arabic linguistics: The case of Arabic diglossia. In E. Benmamoun & R. Bassiouney (Eds.), *The Routledge Handbook of Arabic Linguistics* (pp. 249-262). New York: Routledge.
- Gadet, F. (1996). Niveaux de langue et variation intrinsèque. *Palimpsestes*, 10, 17-40.
- Gadet, F. (2003). *La variation sociale en français*. Paris: Editions Ophrys.
- Gadet, F. (2007). La variation de tous les français. *Linx*, 57, 155-164.
- Gardner-Chloros, P. & Secova, M. (2018). Grammatical change in Paris French: in situ question words in embedded contexts. *Journal of French Language Studies*, 28(2), 181-207. <https://doi.org/10.1017/S0959269518000091>
- Gelderen, E. v. (2008). Where did Late Merge go? Grammaticalization as Feature Economy. *Studia Linguistica*, 62(3), 287-300.
- Gelderen, E. v. (2009). Feature economy in the Linguistic Cycle. In P. Crisma & G. Longobardi (Eds.), *Historical Syntax and Linguistic Theory* (pp. 93-109). Oxford: Oxford University Press.
- Gerken, L., Quam, C. & Goffman, L. (2019). Adults Fail to Learn a Type of Linguistic Pattern that is Readily Learned by Infants. *Language Learning and Development*, 15(4), 279-294. <https://doi.org/10.1080/15475441.2019.1617149>
- Gervain, J. (2015). Plasticity in early language acquisition: the effects of prenatal and early childhood experience. *Current Opinion in Neurobiology*, 35, 13-20. <https://doi.org/10.1016/j.conb.2015.05.004>

- Geveler, J. & Müller, N. (2015). *Wh-fronting and Wh-in-situ in the Acquisition of French: Really Variants?* In P. Guijarro Fuentes, K. Schmitz & N. Müller (Eds.), *The Acquisition of French in Multilingual Contexts* (pp. 43-65). Bristol: Multilingual Matters Ltd.
- Gillet, P., Hommet, C. & Billard, C. (Eds.). (2000). *Neuropsychologie de l'enfant : une introduction*. Marseille: Solal.
- Goldberg, A. E. (1995). *Constructions : A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Goodluck, H. & Kazanina, N. (2020). Fragments along the way: Minimalism as an account of some stages in first language acquisition. *Frontiers in Psychology*, 11, 584. <https://doi.org/10.3389/fpsyg.2020.00584>
- Gopnik, A. & Bonawitz, E. (2015). Bayesian models of child development. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2), 75-86. <https://doi.org/10.1002/wcs.1330>
- Gordishevsky, G. & Schaeffer, J. (2002). On Null Subjects in Child Russian. In Y. Otsu (Ed.), *The Proceedings of the Third Tokyo Conference on Psycholinguistics* (pp. 115-137). Tokyo: Hituzi Publishing Company.
- Goucha, T., Zaccarella, E. & Friederici, A. D. (2017). A revival of Homo loquens as a builder of labeled structures: Neurocognitive considerations. *Neuroscience & Biobehavioral Reviews*, 81(B), 213-224. <https://doi.org/10.1016/j.neubiorev.2017.01.036>
- Grevisse, M. & Goosse, A. (2008). *Le bon usage*. Bruxelles: De Boeck & Duculot (14ème édition).
- Grolig, L. (2020). Shared storybook reading and oral language development: A bioecological perspective. *Frontiers in Psychology*, 11, 1818. <https://doi.org/10.3389/fpsyg.2020.01818>
- Gudmestad, A., Edmonds, A., Donaldson, B., & Carmichael, K. (2020). Near-native sociolinguistic competence in French: Evidence from variable future-time expression. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée*, 23(1), 169-191. <https://doi.org/10.37213/cjal.2020.28845>
- Guillaume, P. (1927). Le développement des éléments formels dans le langage de l'enfant. *Journal de Psychologie*, 24, 203-229.
- Gurgand, L., Lamarque, L., Havron, N., Bernard, J. Y., Ramus, F., & Peyre, H. (2023). The influence of sibship composition on language development at 2 years of age in the ELFE birth cohort study. *Developmental Science*, 26(4), e13356. <https://doi.org/10.1111/desc.13356>

- Guryev, A. & Larrivé, P. (2021). Routines discursives comme contextes d'émergence de l'interrogative partielle in situ en diachronie. *Langue française*, 4, 75-90.
- Guryev, A., Dekhissi, L. & Bonan, C. (2024). L'interrogative *in situ* : aspects formels, pragmatiques et variationnels. *Journal of French Language Studies, First View*. <https://doi.org/10.1017/S0959269524000097>
- Hamann, C. & Plunkett, K. (1998). Subjectless Sentences in Child Danish. *Cognition*, 69(1), 35-72. [https://doi.org/10.1016/S0010-0277\(98\)00059-6](https://doi.org/10.1016/S0010-0277(98)00059-6)
- Hamann, C. (2006). Speculations about Early Syntax: The Production of Wh-Questions by Normally Developing French Children and French Children with SLI. *Catalan Journal of Linguistics*, 5, 143-189.
- Hamlaoui, F. (2011). On the role of phonology and discourse in Francilian French *wh*-questions. *Journal of Linguistics*, 47(1), 129-162. <https://doi.org/10.1017/S0022226710000198>
- Hansen, A. B. & Malderez, I. (2004). Le *ne* de négation en région parisienne : une étude en temps réel. *Langage et société*, 107(1), 5-30.
- Hart, B. & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: Paul H Brookes Publishing.
- Hart, B. & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American educator*, 27(1), 4-9.
- Hartshorne, J. K., Tenenbaum, J. B. & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263-277. <https://doi.org/10.1016/j.cognition.2018.04.007>
- Hauser, C., Aristodemo, V. & Donati, C. (2023). A subject advantage in covert dependencies: The case of *wh*-question comprehension in French Sign Language. *Syntax*, 26(3), 280-310. <https://doi.org/10.1111/synt.12257>
- Heap, D. & Roberge, Y. (2001). Cliticisation et théorie syntaxique, 1971-2001. *Revue québécoise de linguistique*, 30(1), 63-90. <https://doi.org/10.7202/000514ar>
- Heap, D., Oliviéri, M. & Palasis, K. (2017). Clitic Pronouns. In A. Dufter & E. Stark (Eds.), *Manual of Romance Morphosyntax and Syntax* (Vol. 17, pp. 183-229). Berlin: De Gruyter.
- Higgins, F. R. (1973/1979). *The pseudo-cleft construction in English*. (PhD dissertation ; published by Garland), Massachusetts Institute of Technology, Cambridge, MA.

- Hilton, H. (2022). *Enseigner les langues avec l'apport des sciences cognitives*. Vanves: Hachette Education.
- Hirschbühler, P. & Labelle, M. (1994). Changes in verb position in French negative infinitival clauses. *Language Variation and Change*, 6(2), 149-178.
- Hirschbühler, P. & Labelle, M. (2004). Enclisis and proclisis of object pronouns at the turn of the 17th century: The speech of the future Louis XIIIth. In R. S. Gess & D. Arteaga (Eds.), *Historical Romance Linguistics: Retrospective and perspectives* (pp. 187-208). Amsterdam: John Benjamins.
- Howard, M., Lemée, I. & Regan, V. (2006). The L2 acquisition of a phonological variable: The case of /l/deletion in French. *Journal of French Language Studies*, 16(1), 1-24. <https://doi.org/10.1017/S0959269506002298>
- Hulk, A. (1991). Les pronoms clitiques sujets et la théorie linguistique *Actes du XVIIIe Congrès International de Linguistique et de Philologie Romanes* (pp. 504-514). Tübingen: Niemeyer.
- Hulk, A. (1996). The Syntax of Wh-Questions in Child French. *Amsterdam Series in Child Language Development*, 5, 129-172.
- Hunkeler, H. (2005). *Aspects of the evolution of the interactions mother-child: case study of two dizygotic twin children between 15 and 26 months*. Université de Rouen. Unpublished manuscript.
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45(3), 337-374. [https://doi.org/10.1016/S0010-0285\(02\)00500-5](https://doi.org/10.1016/S0010-0285(02)00500-5)
- Hyams, N. M. (1986). *Language Acquisition and the Theory of Parameters*. Dordrecht: D. Reidel Publishing Company.
- Iglesias, O. & Larrivée, P. (2014). Une approche idiolectale de la chute de *ne* en français contemporain *Actes du Congrès Mondial de Linguistique Française - CMLF 2014*: SHS Web of Conferences.
- Jaber, A., Donati, C. & Geraci, C. (2022). On the properties of null subjects in sign languages: The case of French Sign Language (LSF). *The Linguistic Review*, 39(4), 655-686. <https://doi.org/10.1515/tlr-2022-2100>
- Jakubowicz, C. (1996). Crosslinguistic Investigation. In D. McDaniel, C. McKee & H. S. Cairns (Eds.), *Methods for Assessing Children's Syntax* (pp. 257-285). Cambridge, MA: The MIT Press.
- Jakubowicz, C. & Rigaut, C. (1997). L'acquisition des clitiques nominatifs en français. In A. Zribi-Hertz (Ed.), *Les pronoms : morphologie, syntaxe et typologie* (pp. 57-99). Saint-Denis: Presses Universitaires de Vincennes.

- Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua*, 121(3), 339-351. <https://doi.org/10.1016/j.lingua.2010.10.006>
- Jisa, H. (2004). Growing into Academic French. In R. A. Berman (Ed.), *Language Development across Childhood and Adolescence* (pp. 135-161). Amsterdam: John Benjamins.
- Johnson, J. S. & Newport, E. L. (1989). Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology*, 21(1), 60-99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Jung, Y., Walther, D. B. & Finn, A. S. (2021). Children automatically abstract categorical regularities during statistical learning. *Developmental Science*, 24(5), e13072. <https://doi.org/10.1111/desc.13072>
- Kaiser, G. A. (1994). More about INFL-lection and Agreement: The Acquisition of Clitic Pronouns in French. In J. M. Meisel (Ed.), *Bilingual First Language Acquisition: French and German Grammatical Development* (Vol. 7, pp. 131-159). Amsterdam & Philadelphia: John Benjamins.
- Kayne, R. S. (1975). *French Syntax. The Transformational Cycle* (traduction française *Syntaxe du français. Le cycle transformationnel*, 1977, Paris: Seuil ed.). Cambridge, MA: The MIT Press.
- Kayne, R. S. (2000). A Note on Clitic Doubling in French. In R. S. Kayne (Ed.), *Parameters and Universals* (pp. 163-186). Oxford & New York: Oxford University Press.
- Kedar, Y., Casasola, M., Lust, B., & Parmet, Y. (2017). Little words, big impact: Determiners begin to bootstrap reference by 12 months. *Language Learning and Development*, 13(3), 317-334. <https://doi.org/10.1080/15475441.2017.1283229>
- Kern, S. (Ed.) (2019). *Le développement du langage chez le jeune enfant. Théorie, clinique, pratique*. Paris: De Boeck Supérieur.
- Khamis-Dakwar, R., Froud, K. & Gordon, P. (2011). Acquiring diglossia: Mutual influences of formal and colloquial Arabic on children's grammaticality judgments. *Journal of Child Language*, 39(1), 61-89. <http://doi.org/10.1017/S0305000910000784>
- Khomsi, A. (1982). Langue maternelle et langage adressé à l'enfant. *Langue française*, 54, 93-107.
- Kovács, A. M. (2015). Cognitive adaptations induced by a multi-language input in early development. *Current Opinion in Neurobiology*, 35, 80-86. <https://doi.org/10.1016/j.conb.2015.07.003>

- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843. <https://doi.org/10.1038/nrn1533>
- Labelle, M. (1988). *Prédication et mouvement : le développement de la construction relative chez les enfants francophones*. (Thèse de Doctorat), Université d'Ottawa.
- Labov, W. (1963). The Social Motivation of a Sound change. *Word*, 19(3), 273-309. <https://doi.org/10.1080/00437956.1963.11659799>
- Lahousse, K. (2022). Parametric variation in Romance synchronic V2 through the lens of French. *Isogloss*, 8(3, Special issue *Residual Verb Second in Romance*, Cruschina, S. Fabregas, A. & Ch. Meklenborg Nilsen, eds.), 1-30. <https://doi.org/10.5565/rev/isogloss.206>
- Lahousse, K., Jourdain, M., Thonissen, K., & Palasis, K. (2023). *The acquisition of the syntax-information structure interface: On the emergence of c'est 'it is' in child French L1*. Paper presented at the 56th Annual meeting of the Societas Linguistica Europaea, 29th August-1st September 2023, Athens.
- Lambrecht, K. (1981). *Topic, Antitopic and Verb Agreement in Non-standard French*. Amsterdam: John Benjamins.
- Lambrecht, K. (1984). A Pragmatic Constraint on Lexical Subjects in Spoken French. *Chicago Linguistic Society*, 20, 239-256.
- Lambrecht, K. (2001). A framework for the analysis of clefts constructions. *Linguistics*, 39(3), 463-516. <https://doi.org/10.1515/ling.2001.021>
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar*. Stanford: Stanford University Press.
- Langacker, R. W. (1988). A Usage-Based Model. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (pp. 127-161). Amsterdam: John Benjamins.
- Larrivée, P. (2019). Historical pragmatics, explicit activation and *wh in situ* in French. In I. Feldhausen, M. Elsig, I. Kuchenbrandt, & M. Neuhaus (Eds.), *Romance Languages and Linguistic Theory 15. Selected papers from 'Going Romance' 30, Frankfurt* (pp. 113-131). Dordrecht: John Benjamins.
- Larrivée, P. & Guryev, A. (2021). Variantes formelles de l'interrogation. Présentation. *Langue française*, 4, 9-24.
- Larrivée, P. (2022). Is Medieval French diglossic? New evidence on remnant V2 and register. *Isogloss*, 8(2)(8), 1-16. <https://doi.org/10.5565/rev/isogloss.114>
- Le Normand, M.-T. (2019). Productive use of syntactic categories in typical young French children. *First Language*, 39(1), 45-60. <https://doi.org/https://doi.org/10.1177/0142723718778920>

- Le Normand, M.-T. & Thai-Van, H. (2022). The role of Function Words to build syntactic knowledge in French-speaking children. *Scientific Reports*, 12(544), 1-15. <https://doi.org/10.1038/s41598-021-04536-6>
- Legate, J. A. & Yang, C. (2002). Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review*, 19(1-2), 151-162. <https://doi.org/10.1515/ling.2001.021>
- Legendre, G., Barrière, I., Goyet, L., & Nazzi, T. (2010). Comprehension of Infrequent Subject-Verb Agreement Forms: Evidence from French-learning Children. *Child Development*, 81(6), 1859-1875. <https://doi.org/10.1111/j.1467-8624.2010.01515.x>
- Legendre, G., Culbertson, J., Barrière, I., Nazzi, T., & Goyet, L. (2010). Experimental and Empirical Evidence for the Status and Acquisition of Subject Clitics and Agreement Marking in Adult and Child Spoken French. In V. Torrens, L. Escobar, A. Gavarro, & J. Gutiérrez (Eds.), *Movement and Clitics* (pp. 333-360). Newcastle: Cambridge Scholars Publishing.
- Lehmann, E. (1986). *Testing Statistical Hypotheses*. New York: Wiley.
- Leikin, M., Ibrahim, R. & Eghbaria, H. (2014). The influence of diglossia in Arabic on narrative ability: evidence from analysis of the linguistic and narrative structure of discourse among pre-school children. *Reading and Writing*, 27(4), 733-747.
- Leivada, E. (2020). Mid-level generalizations of generative linguistics: Experimental robustness, cognitive underpinnings and the interdisciplinarity paradox. *Zeitschrift für Sprachwissenschaft*, 39(3), 357-374. <https://doi.org/10.1515/zfs-2020-2021>
- Leivada, E. & Murphy, E. (2022). A demonstration of the uncomputability of parametric models of language acquisition and a biologically plausible alternative. *Language Development Research*, 2(1), 105-138. <https://doi.org/10.34842/2022-585>
- Lemaire, S. (2022). *Trajectoire développementale de l'apprentissage du pluriel sur les noms communs*. (Master 2 de Psychologie, Parcours Psychologie du développement, des apprentissages et de l'éducation), Université Côte d'Azur.
- Levine, D., Pace, A., Luo, R., Hirsh-Pasek, K., Michnick Golinkoff, R., de Villiers, J., . . . Sweig Wilson, M. (2020). Evaluating socioeconomic gaps in preschoolers' vocabulary, syntax and language process skills with the Quick Interactive Language Screener (QUILS). *Early Childhood Research Quarterly*, 50, 114-128. <https://doi.org/10.1016/j.ecresq.2018.11.006>

- Liang, Y. (2023). *Quantitative Syntax, Formal Syntax and Information Theory: bridging gaps by studying French variation*. (Thèse de doctorat), Université Paris Cité.
- Lidz, J. & Gagliardi, A. (2015). How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annual Review of Linguistics*, 1, 333-353. <https://doi.org/10.1146/annurev-linguist-030514-125236>
- Lightbown, P. (1977). *Consistency and Variation in the Acquisition of French*. (PhD Dissertation), Columbia University, New York.
- Lightfoot, D. W. (2017). Discovering new variable properties without parameters. *Linguistic Analysis*, 41(3), 409-444.
- Locke, J. (1690). *Essai sur l'entendement humain, Livres I et II*. Paris: Vrin ; traduit par Jean-Michel Vienne, 2001.
- Longobardi, G. (2018). Principles, parameters, and schemata: A radically underspecified UG. *Linguistic Analysis*, 41(3-4), 517-558. <https://eprints.whiterose.ac.uk/31977/>
- Luo, R., Masek, L. R., Alper, R. M., & Hirsh-Pasek, K. (2022). Maternal question use and child language outcomes: The moderating role of children's vocabulary skills and socioeconomic status. *Early Childhood Research Quarterly*, 59, 109-120. <https://doi.org/10.1016/j.ecresq.2021.11.007>
- MacWhinney, B. (2000a). *The CHILDES Project. Tools for Analyzing Talk: Transcription Format and Programs* (Vol. 1, 3rd edition). Mahwah, NJ: Lawrence Erlbaum Associates, or at <http://childes.talkbank.org/>.
- MacWhinney, B. (2000b). *The CHILDES Project. Tools for Analyzing Talk: The Database* (Vol. 3rd edn, vol. 2). Mahwah, NJ: Lawrence Erlbaum Associates, or at <http://childes.talkbank.org/>.
- Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology*, 19(23), 1994-1997. <https://doi.org/10.1016/j.cub.2009.09.064>
- Mantel, N. & Haenszel, W. (1959). Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *Journal of the National Cancer Institute*, 22(4), 719-748.
- Manzini, M. R. & Savoia, L. M. (2005). *I dialetti italiani e romanci: morfosintassi generativa*. Torino: Edizioni dell'Orso.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, 283(5398), 77-80. <https://doi.org/10.1126/science.283.5398.77>

- Marino, C. & Gervain, J. (2019). The impact of generative linguistics on psychology: Language acquisition, a paradigm example. *Acta Linguistica Academica*, 66(3), 371-396.
- Martin, F., Nie, Y., Alexiadou, A., & Guasti, M. T. (2022). Wearing Causation on its Sleeve: Overt cause in Child French Causatives. In Y. Gong & F. Kpogo (Eds.), *BUCLD 46: Proceedings of the 46th annual Boston University Conference on Language Development* (Vol. 2, pp. 497-510). Somerville, MA: Cascadilla Press.
- Martineau, F. & Mougeon, R. (2003). A Sociolinguistic Study of the Origins of *ne* Deletion in European and Quebec French. *Language*, 79(1), 118-152. <https://www.jstor.org/stable/4489387>
- Massot, B. (2008). *Français et diglossie. Décrire la situation linguistique française contemporaine comme une diglossie : arguments morphosyntaxiques*. (Thèse), Université Paris 8.
- Massot, B. (2010). Le patron diglossique de variation grammaticale en français. *Langue française*, 168(4), 87-106.
- Massot, B. & Rowlett, P. (2013). L'hypothèse d'une diglossie en France. *Journal of French Language Studies*, 23(1). <https://doi.org/10.1017/S0959269512000336>
- Mathieu, E. (2004). The mapping of form and interpretation: the case of optional WH-movement in French. *Lingua*, 114(9-10), 1090-1132. <https://doi.org/10.1016/j.lingua.2003.07.002>
- Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft*, 28(1), 5-34.
- Menks, W. M., Ekerdt, C., Lemhöfer, K., Kidd, E., Fernández, G., McQueen, J. M., & Janzen, G. (2024). Developmental changes in brain activation during novel grammar learning in 8-25-year-olds. *Developmental Cognitive Neuroscience*, 66, 101347. <https://doi.org/10.1016/j.dcn.2024.101347>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.
- Miller, P. H. (1992). *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland (PhD Thesis, 1991, Utrecht University).
- Miller, P. H. & Monachesi, P. (2003). Les pronoms clitiques dans les langues romanes. In D. Godard (Ed.), *Langues romanes. Problèmes de la phrase simple* (pp. 67-123). Paris: CNRS Editions.

- Mol, S. E. & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267-296.
- Morin, Y.-C. (1979). La morphophonologie des pronoms clitiques en français populaire. *Cahier de linguistique*, 9, 1-36.
- Moyer, A. (2021). *The Gifted Language Learner: A Case of Nature or Nurture?* Cambridge: Cambridge University Press.
- Munaro, N., Poletto, C. & Pollock, J.-Y. (2001). Eppure si muove! On Comparing French and Bellunese Wh- Movement. In P. Pica & J. Rooryck (Eds.), *Linguistic Variation Yearbook* (Vol. 1, pp. 167-180). Amsterdam: John Benjamins.
- Munaro, N. & Pollock, J.-Y. (2005). Qu'est-ce-que (*qu*)-est-ce que ? A Case Study in Comparative Romance Interrogative Syntax. In G. Cinque & R. S. Kayne (Eds.), *The Oxford Handbook of Comparative Syntax* (pp. 542-606). New York: Oxford University Press.
- Myers, L. L. (2007). *WH-Interrogatives in Spoken French: a Corpus-Based Analysis of their Form and Function*. (Dissertation), University of Texas.
- Myers, L. L. & Pellet, S. (2014). *Pourquoi* in spoken French. In S. Katz Bourns & L. L. Myers (Eds.), *Perspectives on Linguistic Structure and Context* (pp. 157-182). Amsterdam: John Benjamins.
- Nadasdi, T. (1995). *Variation morphosyntaxique et langue minoritaire: Le cas du français ontarien*. (Doctoral dissertation), University of Toronto, Toronto.
- Nagy, N., Blondeau, H. & Auger, J. (2003). Second language acquisition and “real” French: An investigation of subject doubling in the French of Montreal Anglophones. *Language Variation and Change*, 15(1), 73-103.
- Nallet, C. & Gervain, J. (2021). Neurodevelopmental Preparedness for Language in the Neonatal Brain. *Annual Review of Developmental Psychology*, 3(1), 41-58. <https://doi.org/10.1146/annurev-devpsych-050620-025732>
- Nardy, A., Chevrot, J.-P. & Barbu, S. (2014). Sociolinguistic convergence and social interactions within a group of preschoolers: A longitudinal study. *Language Variation and Change*, 26(3), 273-301.
- Negro, I. & Genelot, S. (2009). Les prédicteurs en grande section maternelle de la réussite en lecture en fin de première année d'école élémentaire : l'impact du nom des lettres. *Bulletin de Psychologie*, 3(501), 291-306.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L'Année psychologique*, 101, 447-462.

- Newport, E. L., Gleitman, H. & Gleitman, L. R. (1977). Mother, I'd rather do it myself: some effects and non-effects of maternal speech style. In C. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109-149). Cambridge: Cambridge University Press.
- Obenauer, H.-G. (1981). Le principe des catégories vides et la syntaxe des interrogatives complexes. *Langue française*, 52(1), 100-118.
- Obenauer, H.-G. (1994). *Aspects de la syntaxe A-barre. Effets d'intervention et mouvements des quantifieurs*. (Thèse de Doctorat d'Etat), Université de Paris VIII.
- Oiry, M. (2011). Acquisition of Long-Distance Questions in French. *Language Acquisition*, 18(3), 205-208. <https://doi.org/10.1080/10489223.2011.580680>
- Oliviéri, M., Kaiser, G. A., Palasis, K., Zimmermann, M., & Faure, R. (2020). Quand la dialectologie, la diachronie et l'acquisition se parlent : Étude comparative des pronoms sujets en occitan et en français. In J. Sibille (Ed.), *La microvariation syntaxique dans les langues romanes de France. Actes du colloque Symila, Toulouse, 11 et 12 juin 2015* (pp. 81-98). Limoges: Lambert-Lucas, <https://www.researchgate.net/profile/Katerina-Palasis/publications>.
- Organisation internationale de la Francophonie. (2022). *La langue française dans le monde*. Paris: Gallimard.
- Palasis, K. (2005). *Problèmes d'acquisition et le Paramètre du Sujet Nul*. (Mémoire de Master), Université de Nice Sophia Antipolis.
- Palasis, K. & Oliviéri, M. (2007). Beyond Deficiency. Reconsidering the Null Subject Phenomenon. In S. Baauw, J. Van Kampen & M. Pinto (Eds.), *The Acquisition of Romance Languages. Selected Papers from The Romance Turn II 2006* (Vol. LOT Occasional Series 8, pp. 127-143). Utrecht: LOT Publications, <https://dspace.library.uu.nl/handle/1874/296778>.
- Palasis, K. (2009). *Syntaxe générative et acquisition : le sujet dans le développement du système linguistique du jeune enfant*. (Thèse), Université Nice Sophia Antipolis.
- Palasis, K. (2010a). Introducing New French Child Data: Thoughts on their Gathering and Coding. *Corpus*, 9 'La syntaxe de corpus', 33-51, <https://journals.openedition.org/corpus/1801>.
- Palasis, K. (2010b). *Syntaxe générative et acquisition : le sujet dans le développement du système linguistique du jeune enfant*. Villeneuve d'Ascq: ANRT Diffusion.
- Palasis, K. (2011a). *Pour une meilleure représentativité des données orales en grammaire générative : problèmes et solutions à partir d'un nouveau corpus enfantin français*. Paper presented at the VIII^e Colloque des Jeunes Chercheurs

Praxiling : Corpus, données, modèles : approches qualitatives et quantitatives, Montpellier.

- Palasis, K. (2011b). Français spontané et français normé : bien plus que deux variétés d'un même français. Implications acquisitionnelles et didactiques. In O. Bertrand & I. Schaffner (Eds.), *Variétés, Variations & Formes du français* (pp. 55-72). Palaiseau: Les Éditions de l'École Polytechnique, <https://www.researchgate.net/profile/Katerina-Palasis/publications>.
- Palasis, K. (2012a). *Romance French Child Corpus 1 & 2 (2;5-4;11)*. Retrieved from: <http://childes.talkbank.org/access/French/Palasis.html>
- Palasis, K. (2012b). Economy Principles within Child Speech: When the Nominative Clitic does not Surface. In S. Ferré, P. Prévost, L. Tuller, & R. Zebib (Eds.), *Selected Proceedings of the Romance Turn IV Workshop on the Acquisition of Romance Languages* (pp. 261-284). Newcastle: Cambridge Scholars Publishing, <https://www.researchgate.net/profile/Katerina-Palasis/publications>.
- Palasis, K. (2013). The case for diglossia: Describing the emergence of two grammars in the early acquisition of Metropolitan French. *Journal of French Language Studies*, 23(1), Special Issue 'L'hypothèse d'une diglossie en France', Massot, Benjamin & Paul Rowlett (eds.), 17-35. <https://doi.org/10.1017/S0959269512000348>
- Palasis, K. (2014). Étude morphosyntaxique des pronoms sujets à partir d'un corpus de maternelle. In R. Mir-Samii (Ed.), *Du sujet et de son absence dans les langues* (pp. 259-270). Le Mans: Université du Maine, <http://doi.org/10.13140/2.1.1586.9440>.
- Palasis, K. & Faure, R. (2014). *Grammatical and interpretive constraints on the development of wh-questions in L1 French*. Paper presented at the Romance Turn VI Conference, September 18-20, Universitat de les Illes Balears, Palma de Mallorca.
- Palasis, K. (2015). Subject clitics and preverbal negation in European French: Variation, acquisition, diatopy and diachrony. *Lingua*, 161, 125-143. <http://dx.doi.org/10.1016/j.lingua.2014.11.012>
- Palasis, K. & Meunier, F. (2017). *Coût cognitif des morphèmes silencieux*. Paper presented at the 58ème Congrès annuel de la Société Française de Psychologie, Symposium 9 : Processus cognitifs et structure morphologique des mots, Nice, <https://www.researchgate.net/profile/Katerina-Palasis/publications>.
- Palasis, K., Faure, R. & Lavigne, F. (2019). Explaining variation in *wh*-position in child French: A statistical analysis of new seminaturalistic data. *Language Acquisition*, 26(2), 210-234. <https://doi.org/10.1080/10489223.2018.1513004>

- Palasis, K. (2022). *La compétence socio-pragmatique en français L2 entre 3 et 6 ans : Etude de la variabilité morphosyntaxique chez deux enfants non-natifs dans leurs interactions avec des enfants natifs à la maternelle*. Paper presented at the Pragl2 : L'acquisition de la compétence pragmatique en L2, Nice, France, <https://www.researchgate.net/profile/Katerina-Palasis/publications>.
- Palasis, K., Faure, R. & Meunier, F. (2023). *Wh-in-situ in child French: Deictic triggers at the syntax-semantics interface*. *Journal of French Language Studies*, 33(3), 273-298. <https://doi.org/10.1017/S0959269523000030>
- Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.
- Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins.
- Pearl, L. & Goldwater, S. (2016). Statistical learning, inductive bias, and Bayesian inference in language acquisition. In J. Lidz, W. Snyder & J. Pater (Eds.), *The Oxford Handbook of Developmental Linguistics* (pp. 664-695). Oxford: Oxford University Press.
- Pearl, L. (2021). Theory and predictions for the development of morphology and syntax: A Universal Grammar + statistics approach. *Journal of Child Language*, 48(5), 907-936. <https://doi.org/10.1017/S0305000920000665>
- Pelucchi, B., Hay, J. F. & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80(3), 674-685.
- Perfors, A., Tenenbaum, J. B. & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, 118(3), 306-338. <https://doi.org/10.1016/j.cognition.2010.11.001>
- Pescarini, D. (2014). La distribuzione dei clitici soggetto espletivi : tipologia e prospettive parametriche. *L'Italia dialettale*, 75, 229-246.
- Pescarini, D. (2020). *La microvariation syntaxique dans les langues romanes. Un modèle paramétrique*. (Habilitation à Diriger des Recherches), Université Côte d'Azur.
- Pescarini, D. (2022). Microvariation in agreement inflection: Subject clitics vs inflection. *Word Structure*, 15(3), 358-379.
- Pesetsky, D. (1987). Wh-in-Situ: Movement and Unselective Binding. In E. Reuland & A. Ter Meulen (Eds.), *The Representation of (In)definiteness* (pp. 98-129). Cambridge, MA: MIT Press.
- Pierce, A. (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars* (Vol. 14). Dordrecht: Kluwer Academic Publishers.

- Pine, J. M. (1994). The language of primary caregivers. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition* (pp. 15-37). Cambridge: Cambridge University Press.
- Pine, J. M. & Lieven, E. (1997). Slot and frame patterns and the development of the determiner category. *Applied Linguistics*, 18(2), 123-138. <https://doi.org/10.1017/S0142716400009930>
- Platon. ([1999]). *Menon*. Paris: Le Livre de Poche ; traduction de Jacques Cazeaux.
- Plunkett, B. (2000). The Emergence of Periphrastic Questions in Child French. In M. Perkins & S. Howard (Eds.), *New Directions in Language Development and Disorders* (pp. 105-117). New York: Springer.
- Plunkett, B. (2002). Null subjects in child French interrogatives: A view from the York Corpus. In C. D. Pusch & W. Raible (Eds.), *Romance corpus linguistics: Corpora and spoken language* (pp. 441-452). Tübingen: Narr.
- Pohl, J. (1965). Observations sur les formes d'interrogation dans la langue parlée et dans la langue écrite non-littéraire. *Actes du Xe Congrès International de Linguistique et de Philologie Romanes*, 2, 501-512.
- Poletto, C. (1993). *La sintassi del soggetto nei dialetti italiani settentrionali*. Padova: Unipress.
- Poletto, C. (2000). *The Higher Functional Field: Evidence from Northern Italian Dialects*. New York & Oxford: Oxford University Press.
- Poletto, C. & Pollock, J.-Y. (2005). On *wh*-clitics, *wh*-doubling and apparent *wh*-*in-situ* in French and some north eastern Italian dialects. *Recherches linguistiques de Vincennes*, 33, 135-156. <https://doi.org/10.4000/rlv.1317>
- Poletto, C. & Pollock, J.-Y. (2020). Remnant Movement and Smuggling in Some Romance Interrogative Clauses. In A. Belletti & C. Collins (Eds.), *Smuggling in Syntax*. Oxford: Oxford University Press.
- Pope, M. K. (1934). *From Latin to Modern French*. Manchester: Manchester University Press.
- Prévost, P. (2009). *The Acquisition of French: the Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism, and L2 Acquisition*. Amsterdam: John Benjamins.
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crèpin, G. (1988). Fetal hearing. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 29(3), 191-212.
- Quillard, V. (2000). *Interroger en français parlé : Etudes syntaxique, pragmatique et sociolinguistique*. (Thèse), Université de Tours.

- Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax: the Nature of Early Child Grammars*. Oxford: Blackwell.
- Radford, A. (2000). Children in Search of Perfection: Towards a Minimalist Model of Acquisition. *Essex Research Reports in Linguistics*, 34, <http://privatewww.essex.ac.uk/~radford/PapersPublications/perfection.htm>.
- Rasetti, L. (1996). Null Subjects and Root Infinitives in the Child Grammar of French. *Geneva Generative Papers*, 4(2), 120-132.
- Reinhart, T. (1998). *Wh*-in-situ in the framework of the minimalist program. *Natural Language Semantics*, 6(1), 29-56. <https://doi.org/10.1023/A:1008240014550>
- Rindermann, H. & Baumeister, A. E. E. (2015). Parents' SES vs. parental educational behavior and children's development: A reanalysis of the Hart and Risley study. *Learning and individual differences*, 37, 133-138. <https://doi.org/10.1016/j.lindif.2014.12.005>
- Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
- Rizzi, L. (1986a). Null Objects in Italian and the Theory of *pro*. *Linguistic Inquiry*, 17, 501-557.
- Rizzi, L. (1986b). On the Status of Subject Clitics in Romance. In O. Jaeggli & C. Silva-Corvalan (Eds.), *Studies in Romance Linguistics* (Vol. 24, pp. 391-419). Dordrecht: Foris.
- Rizzi, L. (1991/1996). Residual verb second and the *Wh* Criterion. In A. Belletti & L. Rizzi (Eds.), *Parameters and Functional Heads* (pp. 63-90). Oxford: Oxford University Press.
- Rizzi, L. (1992). Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives. *Geneva Generative Papers*, 0(2), 102-114.
- Rizzi, L. (1997). A Parametric Approach to Comparative Syntax: Properties of the Pronominal System. In L. Haegeman (Ed.), *The New Comparative Syntax* (pp. 268-285). Harlow: Addison Wesley Longman.
- Rizzi, L. (2001). On the position "Int(errogative)" in the left periphery of the clause. In G. Cinque & G. Salvi (Eds.), *Current Studies in Italian Syntax: Essays offered to Lorenzo Renzi* (pp. 287-296). Oxford: Elsevier.
- Rizzi, L. (2006). On the form of chains: Criterial positions and ECP effects. In L. Cheng & N. Corver (Eds.), *Wh-Movement: Moving On* (pp. 97-133). Cambridge, MA: The MIT Press.
- Roberge, Y. (1990). *The Syntactic Recoverability of Null Arguments*. Kingston & Montreal: McGill-Queen's University Press.

- Rooryck, J. (1994). On two types of underspecification: Towards a feature theory shared by syntax and phonology. *Probus*, 6(2-3), 207-233. <https://doi.org/10.1515/prbs.1994.6.2-3.207>
- Rossi, M. (1999). *L'intonation, le système du français : description et modélisation*. Paris: Editions Ophrys.
- Roy, I. (2013). *Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface*. Oxford: Oxford University Press.
- Ruwet, N. (1986). Note sur les verbes météorologiques. *Revue québécoise de linguistique*, 15(2), 43-55. <https://doi.org/10.7202/602560ar>
- Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, 274, 1926-1928. <https://www.jstor.org/stable/2891705>
- Sanchez-Alonso, S. & Aslin, R. N. (2022). Towards a model of language neurobiology in early development. *Brain and Language*, 224, 105047. <https://doi.org/10.1016/j.bandl.2021.105047>
- Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *The Journal of the Acoustical Society of America*, 119(6), EL61-EL67.
- Shi, R., Legrand, C. & Brandenberger, A. (2020). Toddlers track hierarchical structure dependence. *Language Acquisition*, 27(4), 397-409. <https://doi.org/10.1080/10489223.2020.1776010>
- Shlonsky, U. (2012). Notes on wh in situ in French. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, & C. Poletto (Eds.), *Functional Heads: The Cartography of Syntactic Structures* (Vol. 7, pp. 242-252). New York: Oxford University Press.
- Siegal, M., Surian, L., Matsuo, A., Geraci, A., Iozzi, L., Okumura, Y., & Itakura, S. (2010). Bilingualism accentuates children's conversational understanding. *PLoS ONE*, 5(2), e9004. <https://doi.org/10.1371/journal.pone.0009004>
- Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Slabakova, R. (2006). Is there a critical period for semantics? *Second Language Research*, 22(3), 302-338. <https://doi.org/10.1191/0267658306sr270oa>
- Smith, M. E. (1926). An Investigation of the Development of the Sentence and the Extent of Vocabulary in Young Children. *University of Iowa Studies in Child Welfare*, III(5) (traduction française "Chapitre II : Le développement de la phrase" dans Delefosse, J.-M. Odéric (éd.), 2010, *Sur le langage de l'enfant. Choix de textes de 1876 à 1962*), 193-209.
- Snow, C. (1972). Mother's speech to children learning language. *Child Development*, 43, 549-565.

- Sorace, A. (2003). Near-nativeness. In C. Doughty & M. H. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 130-151). Oxford: Blackwell.
- Sportiche, D. (1999). Subject Clitics in French and Romance, Complex Inversion and Clitic Doubling. In K. B. Johnson & I. Roberts (Eds.), *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli* (pp. 189-222). Dordrecht: Kluwer.
- Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4(3), 232-243. <https://doi.org/10.1162/jocn.1992.4.3.232>
- Stark, E. (2014). Réanalyses dans la graphie : 'l'écrit spontané' dans les SMS et le statut des pronoms clitiques du français contemporain. *Langages*, 4, 131-148.
- Stark, E. & Binder, L. (2021). L'inversion du sujet clitique en français oral: ultime apannage des interrogatives? *Langue française*, 4, 25-40.
- Strik, N. (2007). L'acquisition des phrases interrogatives chez les enfants francophones. *Psychologie française*, 52(1), 27-39. <https://doi.org/10.1016/j.psfr.2006.07.003>
- Strik, N. (2008). *Syntaxe et acquisition des phrases interrogatives en français et en néerlandais: Une étude contrastive*. (Thèse de doctorat), Université de Paris 8.
- Strik, N. & Pérez-Leroux, A. T. (2011). Jij doe wat girafe?: *Wh*-movement and inversion in Dutch-French bilingual children. *Linguistic Approaches to Bilingualism*, 1(2), 175-205.
- Suppes, P., Smith, R. & Léveillé, M. (1973). The French Syntax of a Child's Noun Phrases. *Archives de Psychologie*, 42, 207-269.
- Svenonius, P. (2002). Introduction. In P. Svenonius (Ed.), *Subjects, Expletives, and the EPP* (pp. 3-27). New York: Oxford University Press.
- Taguchi, N. (Ed.) (2019). *The Routledge Handbook of Second Language Acquisition and Pragmatics*. New York: Routledge.
- Theakston, A. L. & Lieven, E. (2017). Multiunit sequences in first language acquisition. *Topics in Cognitive Science*, 9(3), 588-603. <https://doi.org/10.1111/tops.12268>
- Tomasello, M. (1992a). The social bases of language acquisition. *Social Development*, 1(1), 67-87.
- Tomasello, M. (1992b). *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.

- Tomasello, M. (2000). The Item-based Nature of Children's Early Syntactic Development. *Trends in Cognitive Sciences*, 4(4), 156-162. [https://doi.org/10.1016/S1364-6613\(00\)01462-5](https://doi.org/10.1016/S1364-6613(00)01462-5)
- Tomasello, M. (2003). *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge, MA & London: Harvard University Press.
- Tomasello, M. (2006). Acquiring Linguistic Constructions. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology: Cognition, Perception and Language (Volume 2)* (pp. 255-298). Hoboken, NJ: John Wiley & Sons.
- Tóth-Fáber, E., Nemeth, D. & Janacsek, K. (2023). Lifespan developmental invariance in memory consolidation: evidence from procedural memory. *PNAS Nexus*, 2(3), pgad037. <https://doi.org/10.1093/pnasnexus/pgad037>
- Travis, L. D. (1984). *Parameters and effects of word order variation*. (Ph.D. thesis), MIT.
- Turker, S., Seither-Preisler, A. & Reiterer, S. M. (2021). Examining individual differences in language learning: A neurocognitive model of language aptitude. *Neurobiology of Language*, 2(3), 389-415. https://doi.org/10.1162/nol_a_00042
- Ullman, M. T. (2016). The Declarative/Procedural Model: A Neurobiological Model of Language Learning, Knowledge, and Use. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 953-968). London: Academic Press.
- Ullman, M. T. (2020). The Declarative/Procedural Model: A Neurobiological Motivated Theory of First and Second Language. In B. VanPatten, G. D. Keating & S. Wulff (Eds.), *Theories in Second Language Acquisition* (pp. 128-161). London: Routledge.
- van der Kant, A., Männel, C., Paul, M., Friederici, A. D., Höhle, B., & Wartenburger, I. (2020). Linguistic and non-linguistic non-adjacent dependency learning in early development. *Developmental Cognitive Neuroscience*, 45, 100819. <https://doi.org/10.1016/j.dcn.2020.100819>
- Veneziano, E. (2000). Interaction, conversation et acquisition du langage dans les trois premières années. In M. Kail & M. Fayol (Eds.), *L'acquisition du langage. Le langage en émergence. De la naissance à 3 ans* (pp. 231-265). Paris: Presses Universitaires de France.
- Vercelli, D. & Piattelli-Palmarini, M. (2009). Language in an Epigenetic Framework. In M. Piattelli-Palmarini, J. Uriagereka & P. Salaburu (Eds.), *Of Minds & Language: A dialogue with Noam Chomsky in the Basque Country* (pp. 97-107). Oxford: Oxford University Press.
- Vergnaud, J.-R. & Zubizarreta, M. L. (2005). The representation of focus and its implications: towards an alternative account of some 'intervention effects'. In

- J. Koster, H. Van der Hulst & H. Van Riemsdijk (Eds.), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk* (pp. 641-660). Berlin & New York: Mouton de Gruyter.
- Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective Effects of Age of Acquisition on Morphological Priming: Evidence for a Sensitive Period. *Language Acquisition*, 25(3), 315-326. <https://doi.org/10.1080/10489223.2017.1346104>
- Villeneuve, A.-J. & Auger, J. (2013). 'chtileu qu'i m'freumereu m'bouque i n'est point coér au monne': Grammatical variation and diglossia in Picardie. *Journal of French Language Studies*, 23(1), 109-133. <https://doi.org/10.1017/S0959269512000385>
- Wagner, R. & Pinchon, J. (1962). *Grammaire du français*. Paris: Hachette.
- Walter, H. (1988). *Le français dans tous les sens*. Paris: Éditions Robert Laffont.
- Waterfall, H. R., Sandbank, B., Onnis, L., & Edelman, S. (2010). An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language*, 37(3), 671-703. <https://doi.org/10.1017/S0305000910000024>
- Watson, J. B. (1924). *Behaviorism*. London: Kegan, Paul, Trench, Trubner & Co.
- Wei, X., Adamson, H., Schwendemann, M., Goucha, T., Friederici, A. D., & Anwender, A. (2023). Native language differences in the structural connectome of the human brain. *NeuroImage*, 270, 119955. <https://doi.org/10.1016/j.neuroimage.2023.119955>
- Wensink, T., Lahousse, K., de Cat, C., Palasis, K., & Busson, B. (2024). *The Syntactic Development of Cleft Sentences in L1 French*. Paper presented at the 57th Annual Meeting of the Societas Linguistica Europaea (SLE 2024), Helsinki, 21-24/08/2024.
- Werker, J. F. (1994). Cross-language speech perception: Developmental change does not involve loss. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 95-120). Cambridge, MA: The MIT Press.
- Werker, J. F. & Hensch, T. K. (2015). Critical Periods in Speech Perception: New Directions. *Annual Review of Psychology*, 66, 173-196. <https://doi.org/10.1146/annurev-psych-010814-015104>
- Whyte, S., Edmonds, A., Palasis, K., & Gerbier, E. (2022). ESPACE L2: Exploring spacing effects in explicit and implicit online learning of L2 English. In B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouëсны, & J. Whelpton (Eds.), *Intelligent CALL, granular systems, and*

learner data: Short papers from EUROCALL 2022 (pp. 386-391): Research-publishing.net, <https://doi.org/10.14705/rpnet.2022.61.9782383720157>.

Xu, Q., Chodorow, M. & Valian, V. (2023). How infants' utterances grow: A probabilistic account of early language development. *Cognition*, 230, 105275. <https://doi.org/10.1016/j.cognition.2022.105275>

Yang, C. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*, 8(10), 451-456. <https://doi.org/10.1016/j.tics.2004.08.006>

Yang, C. (2012). Computational models of syntactic acquisition. *WIREs Cognitive Science*, 3, 205-213. <https://doi.org/10.1002/wcs.1154>

Yang, C., Crain, S., Berwick, R. C., Chomsky, N., & Bolhuis, J. J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience and Biobehavioral Reviews*, 81(B), 103-119. <https://doi.org/10.1016/j.neubiorev.2016.12.023>

Yedetore, A., Linzen, T., Frank, R., & McCoy, R. T. (2023). How poor is the stimulus? Evaluating hierarchical generalization in neural networks trained on child-directed speech. *arXiv preprint arXiv:2301.11462*. <https://arxiv.org/pdf/2301.11462.pdf>

Zanuttini, R. (1997). *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford: Oxford University Press.

Zribi-Hertz, A. (1994). La syntaxe des clitiqes nominatifs en français standard et en français avancé. *Travaux de Linguistique et de Philologie*, 32, 131-148.

Zribi-Hertz, A. (2011). Pour un modèle diglossique de description du français : quelques implications théoriques, didactiques et méthodologiques. *Journal of French Language Studies*, 21(2), 231-256. <https://doi.org/10.1017/S0959269510000323>

Zuckerman, S. & Hulk, A. (2001). Acquiring optionality in French wh-questions: An experimental study. *Revue québécoise de linguistique*, 30(2), 71-97. <https://doi.org/10.7202/000520ar>

Zumwald Küster, G. (2018). *Est-ce que* et ses concurrents. In M.-J. Béguelin, A. Coveney & A. Guryev (Eds.), *L'interrogative en français* (pp. 95-118). Berne: Peter Lang.