



**HAL**  
open science

# L'acquisition d'un paradigme : éclairage multidimensionnel sur la mise en place des déterminants chez quatre enfants entre 1;6 et 3;5

Marine Le Mené Guigourès

► **To cite this version:**

Marine Le Mené Guigourès. L'acquisition d'un paradigme : éclairage multidimensionnel sur la mise en place des déterminants chez quatre enfants entre 1;6 et 3;5. Linguistique. Université Sorbonne Paris Cité, 2017. Français. NNT : 2017USPCA031 . tel-01719263

**HAL Id: tel-01719263**

**<https://theses.hal.science/tel-01719263>**

Submitted on 28 Feb 2018

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

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

**COMUE SORBONNE PARIS CITÉ**  
**UNIVERSITÉ SORBONNE NOUVELLE - PARIS 3**  
ED 268 - LANGAGE ET LANGUES : DESCRIPTION, THEORISATION,  
TRANSMISSION  
EA 7345 - CLESTHIA

Thèse présentée à l'Université Sorbonne Nouvelle - Paris 3  
pour l'obtention du titre de Docteur  
Discipline : Sciences du langage

Marine Le Mené Guigourès

**L'ACQUISITION D'UN PARADIGME**  
**ÉCLAIRAGE MULTIDIMENSIONNEL**  
**SUR LA MISE EN PLACE DES DÉTERMINANTS**  
**CHEZ QUATRE ENFANTS ENTRE 1;6 ET 3;5**

Thèse dirigée par  
Madame la Professeure Anne Salazar Orvig  
(Université Sorbonne Nouvelle - Paris 3)

Soutenue le 11 février 2017

Jury :

Madame Sophie Kern, Chargée de recherche CNRS, Laboratoire Dynamique du Langage (UMR 5596), Université Lyon 2, Rapporteur

Madame Elena Lieven, Professeure, Université de Manchester

Madame Aliyah Morgenstern, Professeure, Université Sorbonne Nouvelle-Paris 3

Monsieur Christophe Parisse, Chargé de recherche INSERM, Laboratoire Modyco (UMR 7114), Université Paris-Ouest Nanterre, Rapporteur

Madame Anne Salazar Orvig, Professeure, Université Sorbonne Nouvelle-Paris 3, Directrice de thèse

Madame Edy Veneziano, Professeure émérite, Université Paris Descartes

**COMUE SORBONNE PARIS CITÉ**  
**UNIVERSITÉ SORBONNE NOUVELLE - PARIS 3**  
ED 268 - LANGAGE ET LANGUES : DESCRIPTION, THEORISATION,  
TRANSMISSION  
EA 7345 - CLESTHIA

Thèse présentée à l'Université Sorbonne Nouvelle - Paris 3  
pour l'obtention du titre de Docteur  
Discipline : Sciences du langage

Marine Le Mené Guigourès

**L'ACQUISITION D'UN PARADIGME**  
**ÉCLAIRAGE MULTIDIMENSIONNEL**  
**SUR LA MISE EN PLACE DES DÉTERMINANTS**  
**CHEZ QUATRE ENFANTS ENTRE 1;6 ET 3;5**

Thèse dirigée par  
Madame la Professeure Anne Salazar Orvig  
(Université Sorbonne Nouvelle - Paris 3)

Soutenue le 11 février 2017

Jury :

Madame Sophie Kern, Chargée de recherche HDR CNRS, Laboratoire Dynamique du Langage (UMR 5596), Université Lyon 2, Rapporteur

Madame Elena Lieven, Professeure, Université de Manchester

Madame Aliyah Morgenstern, Professeure, Université Sorbonne Nouvelle-Paris 3

Monsieur Christophe Parisse, Chargé de recherche INSERM, Laboratoire Modyco (UMR 7114), Université Paris-Ouest Nanterre, Rapporteur

Madame Anne Salazar Orvig, Professeure, Université Sorbonne Nouvelle-Paris 3, Directrice de thèse

Madame Edy Veneziano, Professeure émérite, Université Paris Descartes



# RÉSUMÉ

## L'ACQUISITION D'UN PARADIGME ÉCLAIRAGE MULTIDIMENSIONNEL SUR LA MISE EN PLACE DES DÉTERMINANTS CHEZ QUATRE ENFANTS ENTRE 1;6 ET 3;5

De par leur présence à l'interface entre plusieurs domaines linguistiques (phonologie, morpho-syntaxe, sémantique ou pragmatique), les déterminants ont depuis longtemps suscité l'intérêt des chercheurs en acquisition. On trouve ainsi de nombreux travaux portant notamment sur les contraintes phono-prosodiques, lexico-syntaxiques, sémantiques ou pragmatiques à l'œuvre dans le processus d'acquisition de ces formes, mais peu d'entre eux ont cherché à proposer une analyse combinée de ces facteurs, en tenant compte à la fois des morphèmes adultes et des formes de transition (omissions et fillers). Nous avons donc souhaité, dans le cadre de ce travail, décrire l'émergence des déterminants et de leurs proto-formes en essayant de saisir, dans une perspective multidimensionnelle intégrant à la fois des facteurs distributionnels et fonctionnels, quels facteurs influencent le plus ce processus, et à quelles périodes.

Cette thèse repose sur l'analyse d'un corpus longitudinal de quatre enfants, âgés de 1;6 pour le plus jeune lors de la première séance sélectionnée à 3;5 pour le plus âgé lors de la dernière. Nous avons repéré au sein de ce corpus tous les noms produits par les enfants ainsi que les formes les précédant et chaque syntagme a fait l'objet d'une analyse phono-prosodique, lexico-syntaxique, sémantique, pragmatique et discursive.

Les résultats montrent d'une part que tous les facteurs distributionnels peuvent - à des degrés différents - influencer la production des formes pré-nominales, et d'autre part, que les usages que les enfants font des formes témoignent d'une sensibilité précoce à certains contrastes pragmatico-discursifs. Toutefois, aucun des facteurs pris séparément n'a permis de rendre compte de l'intégralité des productions des enfants. L'analyse conjointe de ces facteurs distributionnels et fonctionnels a confirmé les tendances dégagées au préalable et a mis en évidence l'influence d'un ensemble de facteurs (et en particulier, phono-prosodique, syntaxique et pragmatique) plus que d'un facteur unique ou prédominant tout au long du processus d'acquisition des déterminants.

*Mots clés : acquisition du langage, acquisition des déterminants, fillers, phono-prosodie, cadres lexico-syntaxiques, pragmatique, discours*

# ABSTRACT

## ACQUIRING A PARADIGM MULTIDIMENSIONAL PERSPECTIVE ON THE SETTING UP OF DETERMINERS IN A CORPUS OF FOUR CHILDREN FROM 1;6 TO 3;5

Given their presence at the crossroads of different linguistic fields (phonology, morphosyntax, semantics or pragmatics), determiners have for a long time aroused the interest of language acquisition researchers. We thus find several studies dealing with the phono-prosodic, lexical and syntactic, semantic or pragmatic constraints involved in the acquisition of determiners. But only few of them have provided a combined analysis of all these factors, considering both adult morphemes and transitional forms (omissions and filler syllables). Our study aims to describe the emergence of filler syllables and determiners and understand - in a multidimensional perspective examining simultaneously distributional and functional factors - which factors may affect the acquisition process, and over which periods of time.

Our study is based on a corpus of four children, aged from 1;6 for the youngest in the first selected session to 3;5 for the eldest in the last session. We first identified all the noun phrases produced in the data and each occurrence was then analysed according to phono-prosodic, lexical, syntactic, semantic, pragmatic and discursive factors.

Our results show that each distributional factor may influence, at different degrees, the production of pre-nominal forms and that children's use of forms shows early sensitivity to pragmatic factors. However, none of the factors taken separately have accounted for all of the children's productions. An additional combined analysis of both distributional and functional factors has supported the initial trends and highlighted the influence of a series of factors (and in particular, phono-prosodic, syntactic and pragmatic), more than a single and predominant one throughout the acquisition process of determiners.

*Keywords: language acquisition, acquisition of determiners, filler syllables, prosodic phonology, lexical and syntactic frames, pragmatics, discourse*

## REMERCIEMENTS

*Ça y est, la fin est proche et voici venu le moment de dire merci à tou(te)s celles et ceux qui m'ont aidée, soutenue, encouragée, supportée le temps de cette thèse et qui ont fait de ces dernières années un circuit si agréable à emprunter.*

*La règle veut que l'on débute une page de remerciements par ceux que l'on adresse à son/sa directeur(trice) de thèse. Il arrive probablement que ce(tte) dernier(e) ne soit pas toujours le/la plus concerné(e) et que la raison plus que le cœur nous invite à respecter la tradition, mais pour ma part, c'est avec un plaisir immense que j'adresse mes premiers remerciements à ma directrice, Anne Salazar Orvig, qui a minutieusement contribué, à chaque étape, à la préparation et à la rédaction de ce travail. Anne, tu as accompli bien plus que ce que tes fonctions te demandaient. Tu m'as suivie, formée, épaulée avec une telle assiduité que les « profonde gratitude », « plus vifs remerciements » et autres formules consacrées ne suffisaient pas à te dire combien je te suis reconnaissante. Je te remercie de m'avoir fait confiance en acceptant de suivre cette thèse (il fallait à l'époque être très optimiste ou un peu déraisonnable !), de m'avoir initiée à la recherche, seule mais aussi et surtout en équipe, de m'avoir donné l'opportunité et l'envie de découvrir et redécouvrir le métier d'enseignante, et plus globalement, merci pour les éclairages nombreux que tu m'as apportés sur la vie universitaire. C'est en très grande partie grâce à toi que j'ai eu l'impression d'avoir un parcours de thésarde si peu monotone et solitaire et j'espère que tu sais tout ce que je te dois...*

*Je tiens également à remercier très chaleureusement Aliyah Morgenstern, pour avoir accepté de lire cette thèse et d'en présider le jury, mais aussi pour toute la sympathie qu'elle m'a témoignée ces dernières années, en me laissant participer à son séminaire Multimod et à son projet sur les dîners familiaux. Aliyah, grâce à toi j'ai rencontré Camille, Stéphanie et Yann, avec qui j'ai noué de très belles amitiés, alors j'espère que tu ne te sépareras jamais de cette envie inégalable de rassembler les gens. Tu sais mieux que personne rendre les temps de travail conviviaux et toutes les réunions 'dîners familiaux' sont pour moi d'excellents souvenirs. Et puis dernière chose, merci de m'avoir donné la chance d'écouter du jazz à l'arrière de la voiture de Sandro, ça restera à jamais un moment fondateur de mon parcours !*

*J'adresse par ailleurs mes sincères remerciements à Madame Edy Veneziano pour sa participation à mon jury de thèse et pour ses précieux commentaires lors de nos dernières rencontres. Je tiens également à exprimer toute ma gratitude à Madame Sophie Kern, Madame Elena Lieven et Monsieur Christophe Parisse pour avoir accepté de faire partie de ce jury et de prendre sur leur temps que j'imagine déjà bien rempli pour lire cette thèse. Votre présence au sein de mon jury a rendu les derniers moments de rédaction plus encourageants et stimulants.*

*Ce travail n'aurait jamais pu être mené à bien sans la participation des enfants et des parents qui en constituent le corpus. Un immense merci donc à Antonin et son frère, Solal, Lubin et sa sœur, Léonie, Naël et son frère, Tristan, de m'avoir acceptée comme visiteuse régulière et partenaire de jeu, mais aussi à Elise et Thomas, Séverine et Antoine ainsi qu'à Viviane et Sylvain pour m'avoir sans discuter donné accès à leur intimité familiale. J'ai appris beaucoup de la vie de parent auprès de vous et mes temps de recueil figurent parmi les meilleurs de ma thèse. Ce corpus n'aurait pas non plus été recueilli sans l'aide de Marie Leroy-Collombel qui a su convaincre son réseau étampois de me laisser m'installer avec ma caméra, alors merci Marie de ta générosité et de ta bienveillance !*

*Je me dois bien sûr de remercier également mon université, mon école doctorale et son directeur, mon équipe de recherche et sa directrice, le ministère de l'enseignement supérieur et de la recherche, l'Agence Nationale de la Recherche ainsi que l'ex-IDEX et mes parents pour le financement de cette longue période doctorale et de mes nombreuses missions à l'étranger. Si cette recherche a pu être menée, c'est aussi grâce à votre soutien matériel et je vous en suis extrêmement reconnaissante.*

*Sur le GR du doctorat j'ai eu la chance de rencontrer d'autres personnes fantastiques avec qui j'ai partagé bien plus que les problématiques de la thèse. Merci à Anna, Aurélia, Camille, Cricri, Elise, Fuchs, Greg, Janina, Juju, Laetitia, Marie, Nao, Pauline, Salma, Stefano, Steph et Stéphane pour (dans le désordre) les soirées olives-rouge, les cidres à Beaubourg et tout le reste, les accras rue de Charonne, les cafés à l'ILPGA et à Censier, les nuits blanches à Fribourg et j'en passe... On le devinera aisément, ma thèse c'est aussi ça, c'est vous. Une mention spéciale 'formation' à Christine et Julien qui, sans relâche, ont répondu à mes questions d'enseignante et chercheuse en herbe. Qu'aurais-je fait sans vous ? Et à Naomi pour la patience dont elle a su faire preuve face à mes questions de phono, et pour les beaux projets des Pilgrim Sisters à venir ! L'avenir est dans l'interface, toi-même tu sais !*

*Un grand merci aussi aux collègues de l'ILPGA - Adeline, Bertrand, Charlotte, Emre, Maximilien et Tiphonie - que j'ai eu l'occasion de (re)découvrir avant, pendant et en dehors du projet COMET et qui m'ont chaleureusement encouragée.*

*Je tiens également à remercier celles et ceux qui sont déjà cités par ailleurs mais qui ont en plus - et presque de leur plein gré - accepté de relire cette thèse pour en corriger les fautes d'orthographe, les placements hasardeux de virgules et les lourdeurs. Merci beaucoup à Emilie, Julien, et à ma mère. Toutes les erreurs qui figurent encore dans ce manuscrit peuvent donc leur être imputées.*

*Cette fin de thèse c'est aussi l'occasion de dire merci à mes ami-e-s, ma famille d'ami-e-s, qui ne m'ont pas toujours aidée à trouver le droit chemin de la soutenance mais qui m'ont apporté tellement de magie que je ne peux que les en remercier. Alors merci à Bouquette, Dorothee, Emilie, Faby, Fanny, Floflo, Gillen, Juju G., Kokoff, le Belge, le Furet, Ludi, Marine, Max, Nono, Onur et Shabnam. La liste est grande mais certainement pas plus grande que l'amitié que j'ai pour vous.*

*Je tiens enfin à remercier ma famille, et particulièrement ma grand-mère, pour l'amour qu'elle me porte, et pour ses pensées et encouragements quotidiens, ainsi que ma mère et Denis, pour m'avoir depuis longtemps déjà laissé faire des choix d'études et de carrière risqués, sans jamais les remettre en question. Sans votre confiance et votre soutien inébranlables, je n'aurais jamais entamé cette thèse et l'aurais encore moins terminée. Je mesure quotidiennement ma chance de vous avoir comme parents.*



# SOMMAIRE

RÉSUMÉ.....	5
ABSTRACT .....	6
REMERCIEMENTS .....	7
SOMMAIRE.....	9
INTRODUCTION .....	11
<b><u>PARTIE I - APPROCHES THÉORIQUES.....</u></b>	<b>17</b>
<b>CHAPITRE I - DES APPROCHES VARIEES SUR LE DEVELOPPEMENT DU LANGAGE DE L'ENFANT .....</b>	<b>19</b>
1. L'ACQUISITION DU LANGAGE : UNE FACULTE INNEE ?.....	22
2. L'ACQUISITION DU LANGAGE : UNE CONSTRUCTION PROGRESSIVE .....	33
3. DES PERSPECTIVES DIVERGENTES SUR L'ACQUISITION DES CATEGORIES SYNTAXIQUES .....	53
4. PERSPECTIVE ADOPTEE POUR APPROCHER LES DONNEES .....	58
<b>CHAPITRE II - L'EMERGENCE DE LA CATEGORIE 'DETERMINANT' CHEZ L'ENFANT .....</b>	<b>65</b>
1. GENERALITES SUR LE DETERMINANT CHEZ L'ADULTE.....	68
2. UN PARADIGME EN CONSTRUCTION : L'ETUDE DES FILLERS.....	80
3. LE TRAITEMENT DES FILLERS ET DES DETERMINANTS DANS LA LITTERATURE, REFLET DE LA DIVERSITE DES APPROCHES METHODOLOGIQUES ET THEORIQUES .....	94
4. ENJEUX D'UNE ETUDE SUR L'ACQUISITION DU PARADIGME DES DETERMINANTS .....	124
<b><u>PARTIE II - QUESTIONS DE RECHERCHE ET METHODOLOGIE.....</u></b>	<b>131</b>
<b>CHAPITRE III - QUESTIONS DE RECHERCHE.....</b>	<b>133</b>
1. DESCRIPTION DES FORMES.....	136
2. CONTRAINTES DISTRIBUTIONNELLES.....	137
3. INFLUENCES DES FACTEURS FONCTIONNELS.....	138
4. MISE EN PERSPECTIVE DES FACTEURS DISTRIBUTIONNELS ET FONCTIONNELS.....	139
<b>CHAPITRE IV - CORPUS ET METHODOLOGIE.....</b>	<b>141</b>
1. QUESTIONNEMENTS PREALABLES A LA CONSTITUTION DU CORPUS.....	143
2. LE RECUEIL DE CORPUS A L'HEURE DU PARTAGE DE DONNEES .....	146
3. PRESENTATION DES ENFANTS .....	149
4. LES SITUATIONS D'ENREGISTREMENT .....	154
5. LA PLACE DE L'OBSERVATEUR .....	155
6. LA TRANSCRIPTION DU CORPUS.....	163
7. L'ANALYSE DU CORPUS.....	178
<b><u>PARTIE III - ANALYSES ET RESULTATS.....</u></b>	<b>185</b>
<b>CHAPITRE V - DISTRIBUTION ET DESCRIPTION DES FORMES PRE-NOMINALES .....</b>	<b>187</b>
1. IDENTIFICATION DES POSITIONS DITES 'PRE-NOMINALES' ET CATEGORISATION DES FORMES.....	189
2. DEVELOPPEMENT DES FORMES PRE-NOMINALES CHEZ LES QUATRE ENFANTS DU CORPUS .....	196
3. BILAN COMPARATIF SUR L'EVOLUTION DE LA POSITION PRE-NOMINALE DANS LE CORPUS.....	207
<b>CHAPITRE VI - INFLUENCE DES FACTEURS DISTRIBUTIONNELS SUR LA MISE EN PLACE DES DETERMINANTS.....</b>	<b>223</b>
1. INFLUENCE DE LA STRUCTURE SYLLABIQUE.....	226
2. INFLUENCE DU STATUT SEMANTIQUE DU NOM ASSOCIE A LA FORME PRE-NOMINALE .....	239
3. IMPACT DU CONTEXTE LEXICAL ET SYNTAXIQUE.....	259

## SOMMAIRE

<b>CHAPITRE VII - INFLUENCE DES FACTEURS FONCTIONNELS SUR LA MISE EN PLACE DES DETERMINANTS.....</b>	<b>299</b>
1. INFLUENCE DES CARACTERISTIQUES PRAGMATICO-DISCURSIVES DES REFERENTS .....	301
2. LE ROLE DES REPRISES DU NOM DANS L'ACQUISITION DES FORMES.....	339
<b>CHAPITRE VIII - OBSERVATION CROISEE DES INFLUENCES DISTRIBUTIONNELLES ET FONCTIONNELLES.....</b>	<b>369</b>
1. TRAITEMENT STATISTIQUE DES DONNEES .....	373
2. PREVALENCE -CONSTANTE OU PERIODIQUE- D'UN FACTEUR OU COMPLEMENTARITE DES FACTEURS ?376	
<b>CONCLUSION .....</b>	<b>389</b>
1. SYNTHESE GENERALE .....	389
2. UN TABLEAU A COMPLETER .....	392
<b>BIBLIOGRAPHIE .....</b>	<b>399</b>
<b><u>PARTIE IV - ANNEXES .....</u></b>	<b><u>415</u></b>
<b>FORMULAIRE D'AUTORISATION PARENTALE .....</b>	<b>417</b>
<b>CONVENTIONS DE TRANSCRIPTION.....</b>	<b>419</b>
<b>DÉCOUPAGE DES TOURS DE PAROLE EN ÉNONCÉS .....</b>	<b>429</b>
<b>TABLE DES MATIÈRES.....</b>	<b>433</b>

# INTRODUCTION

En français, le déterminant est, avec le nom, un des deux éléments constitutifs du syntagme nominal, dont la présence est requise de façon quasiment systématique. On relève dans les grammaires modernes (*i.a.* Riegel *et al.*, 2004) deux ensembles de déterminants, les définis et les indéfinis, se partageant chacun en sous-ensembles et pouvant faire l'objet de variations, en fonction du genre et du nombre des noms auxquels ils sont associés.

Pour l'enfant, l'acquisition du déterminant implique donc que ces différentes formes soient maîtrisées et que leur production devant les noms soit généralisée, mais aussi et surtout qu'elles soient produites de façon appropriée, attendue, dans le discours (Karmiloff-Smith, 1979). Les déterminants étant porteurs des caractéristiques sémantico-référentielles des noms auxquels ils sont associés (Charolles, 2002), l'acquisition du paradigme passe également par la maîtrise des associations entre les formes et les fonctions portées par ces formes.

L'usage du déterminant n'est donc pas seulement le reflet d'une compétence grammaticale formelle mais aussi une voie d'entrée, une fenêtre, sur les ressources dont dispose l'enfant - en dialogue avec l'autre - pour comprendre et s'approprier la dimension fonctionnelle de la langue. Au même titre que l'émergence des formes devant les verbes peut nous renseigner non seulement sur le développement grammatical de l'enfant mais aussi sur certains aspects psychologiques, cognitifs ou sociaux de son développement (Caët, 2013), la mise en place des déterminants devant les noms nous informe sur les compétences pragmatiques, et plus largement interactionnelles, de l'enfant.

Cette dimension fonctionnelle a d'ailleurs été largement étudiée dans la littérature, de même que la dimension plus formelle de la mise en place du paradigme. On trouve ainsi un grand nombre

de travaux qui ont cherché à décrire et situer les premières formes de déterminants produites par les enfants (*i.a.* Bassano, 2010), et d'autres qui se sont attachés à expliquer leur présence/absence ainsi que les types de déterminants produits par des facteurs que l'on appellera distributionnels - phono-prosodiques (Demuth, 1996), lexico-syntaxiques (Pine et Lieven, 1997) et sémantiques (Bassano, 2005) - et/ou fonctionnels, et liés donc à des facteurs interactionnels et plus spécifiquement, pragmatico-discursifs (Rozendaal, 2008 ; Salazar Orvig *et al.*, 2013).

Certains auteurs, comme Demuth (1996 ; Demuth et Tremblay, 2008), ont par exemple argumenté en faveur d'une influence du rythme de la langue ou du nombre de syllabes du nom sur la production des formes pré-nominales. D'autres auteurs, comme Pine et Lieven (1997 ; Lieven *et al.*, 1997), ont pour leur part cherché à montrer que les déterminants ne constituent pas d'emblée une catégorie grammaticale conforme à celle observée chez l'adulte mais qu'ils sont d'abord produits associés à des cadres lexicaux et/ou syntaxiques spécifiques. Pour ces auteurs, la production des premiers définis et indéfinis serait donc déterminée par les contextes lexico-syntaxiques dans lesquels ils sont produits. On trouve aussi dans la littérature des études sur l'impact du sémantisme des noms sur la production des premiers déterminants. D'après Bassano notamment (2005, 2008, 2010), l'enfant produirait davantage de déterminants devant les noms référant à des entités inanimées que devant ceux référant à des entités animées. Enfin, de nombreux auteurs (Karmiloff-Smith, 1979 ; Rozendaal, 2008 ; Salazar Orvig *et al.*, 2013) ayant cherché à évaluer la maîtrise par l'enfant des associations entre formes et fonctions ont observé les contextes pragmatiques dans lesquels les premiers déterminants sont produits, et bien que controversés, les résultats témoignent - partiellement au moins - de la sensibilité de l'enfant à certaines caractéristiques pragmatiques des référents (comme leur caractère spécifique/non spécifique) lors de la production des premiers déterminants. Cette fois encore, le contexte - ici, pragmatique - pourrait expliquer la distribution des déterminants.

De nombreuses études (Peters et Menn, 1993 ; Kilani-Schoch et Dressler, 2000 ; Veneziano et Sinclair, 2000) ont par ailleurs montré que l'enfant n'acquiert pas d'emblée les morphèmes adultes mais produit, en concurrence avec les formes attendues, des omissions et des fillers. Les fillers sont des syllabes de remplissage, le plus souvent constituées d'une seule voyelle ou d'une association 'consonne + voyelle', que l'enfant peut produire durant une période, devant les mots lexicaux (proto-noms et proto-verbos). Si ces éléments peuvent être interprétés comme des

précurseurs de déterminants, de pronoms, de prépositions ou encore d’auxiliaires, ils ne peuvent toutefois pas être interprétés comme des morphèmes de la langue adulte, parce qu’ils n’en partagent pas toutes les caractéristiques. Les fillers ont été largement étudiés, afin de comprendre leur statut dans les premières productions de l’enfant mais aussi leur évolution au cours du développement langagier. Il a notamment été proposé dans la littérature que leur production puisse être liée à des facteurs phono-prosodiques (Demuth, 2001 ; Veneziano et Sinclair, 2000 ; Veneziano, 2014b) ou sémantiques (Bassano, 2005), et quelques auteurs se sont également penchés sur leur éventuelle dimension pragmatique (notamment, Salazar Orvig et Morgenstern, 2002).

Ces travaux sur les déterminants et les fillers nous apportent un éclairage considérable sur l’acquisition du paradigme mais ne proposent pas - à l’exception de quelques-uns - d’observation combinant l’analyse de plusieurs facteurs susceptibles d’influencer simultanément, ou de façon successive, la production des formes pré-nominales et ont au contraire le plus souvent privilégié l’analyse d’un seul de ces facteurs. Par ailleurs, toutes les études sur les déterminants ne se sont pas intéressées aux formes de transition et ont donc parfois exclu fillers et omissions des analyses, réduisant ainsi de façon notable les traces à notre disposition pour comprendre au mieux le processus d’acquisition des déterminants.

L’objectif principal de cette thèse est donc de mettre en perspective certaines des dimensions qui pourraient, d’après la littérature, conditionner la production des formes pré-nominales (dimensions phono-prosodique, lexico-syntaxique, sémantique et pragmatico-discursive), afin de saisir d’une part si toutes jouent un rôle dans le processus d’acquisition des déterminants chez les enfants de notre corpus, et d’autre part, à quelles périodes du développement langagier elles agissent.

Afin de répondre à notre objectif, nous avons fait le choix d’une étude longitudinale, nous permettant d’appréhender le développement presque intégral du paradigme, des premières formes produites devant le nom à la maîtrise des morphèmes adultes. Notre corpus est constitué des données de quatre enfants. Nous en avons suivi trois (Antonin, Lubin et Naël), et avons par ailleurs utilisé les données d’un quatrième enfant (Adrien), préalablement recueillies par N. Yamaguchi dans le cadre de sa thèse (Yamaguchi, 2012), et disponible sur la base de données CHILDES. L’enfant le plus jeune a 1;6 an au début des séances que nous avons sélectionnées et

## INTRODUCTION

le plus âgé à la fin a 3;5 ans. Tous les enregistrements ont été effectués au domicile des enfants dans des activités librement choisies par les parents et le plus souvent, les enfants ont été filmés en interaction avec leurs parents et les membres de leur fratrie.

Après avoir été transcrites, les données ont été observées à la lumière de deux axes d'analyse principaux : d'une part, la mise en place et la distribution des formes pré-nominales et d'autre part, le rôle joué par certains facteurs distributionnels et fonctionnels sur l'émergence de ces formes. La première partie de nos analyses nous permettra de pouvoir dessiner une vue d'ensemble du développement de la position pré-nominale chez chaque enfant et suite à la seconde partie de nos analyses, basée sur l'observation des contextes phono-prosodiques, lexico-syntaxiques, sémantiques et discursifs dans lesquels les noms sont produits, nous essaierons de définir quels facteurs entrent en jeu - et à quel moment - dans la construction progressive du déterminant.

Nos analyses sont ancrées dans une approche de l'acquisition du langage basée sur l'usage (*i.a.* Tomasello, 2003, 2009). Nous considérons en effet que l'enfant n'est pas doté d'un dispositif inné et spécifique au langage, mais qu'à l'inverse, il apprend à parler en s'appuyant sur les données fournies par son environnement, et particulièrement sur ses échanges avec un autre plus expérimenté. A ce titre, nous envisageons l'acquisition de la catégorie 'déterminant' par l'enfant comme une acquisition progressive, construite grâce aux différents usages de déterminants qui jalonnent l'expérience que l'enfant fait de sa langue. Dans cette perspective, nous rendrons compte de l'émergence de la catégorie en considérant non seulement la période à laquelle les déterminants de forme adulte deviennent les productions dominantes mais aussi la période pendant laquelle les formes de transition sont majoritaires. Nous mettrons par ailleurs en perspective, aussi souvent que possible, le développement des formes et les productions conjointes de l'enfant et de son interlocuteur, afin d'appréhender le mieux possible le rôle du discours et plus globalement, de l'interaction, dans le développement de la position pré-nominale.

De façon à restituer le plus fidèlement possible la démarche de recherche adoptée pour cette étude, nous avons organisé cette thèse en trois parties. Dans une première partie, comprenant deux chapitres (*Chapitre I - Des approches variées sur le développement du langage de l'enfant* et *Chapitre II - L'émergence de la catégorie 'déterminant' chez l'enfant*), nous ferons dans un premier temps, un état des

lieux des différentes perspectives théoriques sur l'acquisition du langage, avant d'exposer les choix théoriques faits dans le cadre de ce travail, et dans un second temps, une revue de la littérature consacrée aux déterminants et aux fillers. Au sein de la seconde partie, nous ferons au chapitre III (*Questions de recherche*), l'inventaire des questions soulevées dans ce travail, et au chapitre IV (*Corpus et méthodologie*), nous introduirons le lecteur au corpus recueilli, aux problématiques posées par le recueil et la transcription, ainsi qu'à quelques éléments relatifs à notre méthodologie d'analyse (sur laquelle nous reviendrons de façon plus exhaustive dans les chapitres suivants). La troisième et dernière partie de la thèse sera consacrée aux résultats de notre analyse et à la discussion de ces résultats. Au sein du chapitre V (*Distribution et description des formes pré-nominales*), nous ferons la description des formes produites en position pré-nominale par chacun des enfants du corpus ; au chapitre VI (*Influence des facteurs distributionnels sur la mise en place des déterminants*), nous présenterons les résultats des analyses phono-prosodiques, sémantiques et lexico-syntaxiques et au chapitre VII (*Influence des facteurs fonctionnels*), les résultats des analyses portant sur les caractéristiques pragmatico-discursives des référents ainsi que sur les reprises de l'enfant. Au chapitre VIII (*Observation croisée des influences distributionnelles et fonctionnelles*), nous mettrons en perspective les différents facteurs considérés dans les chapitres précédents afin d'en mesurer le poids respectif. Enfin, en conclusion, nous ferons la synthèse générale des principaux résultats et apports de ce travail ainsi que l'inventaire des nombreuses perspectives qui en découlent.





# PARTIE I

## *Approches théoriques*

---

- CHAPITRE I** Des approches variées sur le développement du langage de l'enfant
- CHAPITRE II** L'émergence de la catégorie 'déterminant' chez l'enfant



# CHAPITRE I

*Des approches variées sur le développement  
du langage de l'enfant*

---



**B**ien que l'intérêt pour le développement du langage de l'enfant soit apparu très tôt - on peut notamment citer les travaux d'Itard (1801), de Taine (1876) ou encore de Darwin (1877) - les premières études systématiques du langage de l'enfant ne datent que du début du XX<sup>ème</sup> siècle et la discipline n'est véritablement apparue que dans les années 1970. En quelques décennies seulement, le domaine de l'acquisition du langage a vu émerger un nombre considérable de travaux et de théories distinctes.

Les controverses théoriques ont marqué et continuent de marquer la discipline. À la suite de Hirsh-Pasek et Golinkoff (1996) ou Kail (2000), nous considérerons le positionnement des chercheurs du domaine comme s'articulant autour de plusieurs questions fondamentales :

- le langage est-il inné ou lié à des apprentissages ?
- quelle est la nature de l'input qui permettrait le développement langagier de l'enfant ?
- l'enfant dispose-t-il de capacités initiales ? Si tel est le cas, sont-elles d'emblée linguistiques ou d'ordre cognitif ou social ?
- les mécanismes en jeu dans le développement du langage de l'enfant sont-ils spécifiques au langage ou identiques aux mécanismes à l'œuvre pour tous les autres domaines du développement ?

Les réponses apportées à ces questions permettent d'opposer deux types de théories : les théories qu'Hirsh-Pasek et Golinkoff appellent 'inside-out', comprenant des approches considérant notamment le langage comme inné et les théories que ces auteures nomment 'outside-in', pour lesquelles à l'inverse le langage est issu des apprentissages de l'enfant. On trouve dans chacun de ces deux grands axes théoriques une multitude d'approches et de travaux différents sur le développement du langage. Nous reviendrons au sein de ce chapitre sur cette diversité théorique

et sur les débats actuels dans le domaine de l'acquisition du langage. A la suite de cette présentation, nous définirons le cadre théorique dans lequel s'inscrit ce travail.

Nous commencerons notre état des lieux des différentes perspectives théoriques par l'approche innéiste de l'acquisition du langage. Nous présenterons dans cette première partie du chapitre les travaux menés sur le développement du langage de l'enfant dans le cadre de la Grammaire Universelle et s'interrogeant sur l'émergence des catégories linguistiques ainsi que sur les mécanismes permettant le développement du langage.

## **1. L'acquisition du langage : une faculté innée ?**

L'approche innéiste générativiste du langage proposée par Chomsky dans les années 1960 s'inscrit en réaction à la conception béhavioriste de l'apprentissage, représentée à l'époque par Bloomfield ou Skinner. A l'inverse des comportementalistes qui considèrent le langage comme l'apprentissage d'une somme d'habitudes verbales (notamment, Skinner, 1957), les générativistes envisagent le langage comme une faculté qui ne s'apprend pas et qui est présente dans le matériel génétique de l'enfant dès sa naissance. Cette grammaire universelle et innée n'impliquerait aucune participation de l'interaction ni du modèle adulte dans le développement langagier de l'enfant. Cet argument est étayé, entre autres, par l'idée que l'environnement linguistique de l'enfant n'est pas suffisamment riche pour que ce dernier puisse s'en servir comme modèle.

Nous reviendrons dans les paragraphes suivants sur les principaux arguments à l'origine du courant générativiste ainsi que sur l'évolution du dispositif d'acquisition du langage défendu par ses théoriciens.

### **1.1 Créativité de l'usage du langage et récursivité**

Les travaux de Chomsky sur la grammaire générative ont pour beaucoup été élaborés autour de la 'dimension créative de l'usage du langage' ('creative aspect of language use'). Cet aspect est notamment mis en évidence dès l'introduction de son ouvrage de 1964 dont nous avons reproduit un extrait ci-dessous :

« The central fact to which any significant linguistic theory must address itself is this: a mature speaker can produce a new sentence of his language on the appropriate occasion, and other speakers can understand it immediately, though it is equally new to them. Most of our linguistic experience, both as speakers and hearers, is with new

sentences; once we have mastered a language, the class of sentences with which we can operate fluently and without difficulty or hesitation is so vast that for all practical purposes (and, obviously, for all theoretical purposes), we may regard it as indefinite » (Chomsky, 1964 : 7).

Pour Chomsky, l'être humain a la capacité de créer un nombre infini de phrases, en combinant les mots de sa langue à partir d'un ensemble fini de phonèmes. Cette faculté permet également à tout locuteur d'une langue de comprendre les productions de ses interlocuteurs, quand bien même leurs énoncés n'auraient jamais été entendus par ce locuteur au préalable. L'usage du langage serait donc en lui-même une activité créatrice et cette créativité distinguerait le langage de l'espèce humaine des moyens de communication des autres espèces animales (Chomsky 2006[1968]). A la suite des philosophes cartésiens, Chomsky considère que la spécificité de l'Homme n'est pas tant d'avoir un instrument de communication, ce que partagent certains animaux, mais bien plutôt un outil créatif pour exprimer librement sa pensée et produire des réponses appropriées dans des situations nouvelles (Chomsky, 2002[1966] ; Voss, 1973). Pour Chomsky, il s'agit là du phénomène le plus important à considérer dans le cadre de la théorisation linguistique. A la suite de la définition de la dimension créative du langage qu'il propose en 1964, il ajoute : « it is clear that a theory of language that neglects this "creative" aspect is of only marginal interest » (Chomsky, 1964 : 8).

Au-delà de l'aspect créatif de l'usage du langage, les générativistes soulignent également une autre des spécificités du langage humain : sa récursivité. Les propositions produites par un locuteur peuvent être enchâssées, par coordination ou subordination, rendant ainsi le langage humain particulièrement complexe. Cette notion d'abord étudiée par les travaux en logique et sur les langages formels renvoie dans la terminologie chomskyenne à l'idée que la syntaxe est productive, qu'il n'est pas possible de fixer une limite à la longueur des phrases, et cela quelle que soit la langue étudiée (Egré, 2011). Pour Chomsky, c'est ce trait caractéristique du langage humain qui rend possible la créativité dans l'usage de la parole.

Ces deux spécificités du langage humain sont à la source des travaux de Chomsky : c'est pour résoudre ce 'mystère' de la créativité du langage et pour éclairer la créativité humaine dans son ensemble que ses recherches en linguistique ont été menées. McGilvray, dans la seconde édition de *Cartesian Linguistics* écrit : « Chomsky assumes [...] that the science of language can illuminate

human creativity, and that close observation of human creativity can tell us something about the nature of language and the structure of our minds » (2002[1966] : 11).

## **1.2 Opposition compétence-performance**

S'il peut sembler fastidieux de dissocier au fil des ouvrages de Chomsky la notion de récursivité de celle de créativité, c'est parce que les deux sont intrinsèquement liées. La première renvoie à la créativité de la faculté de langage et la seconde à la créativité dans l'usage que font les locuteurs du langage. Les règles permettant la créativité de parole seraient issues de la maîtrise de la récursivité syntaxique. La récursivité serait à rattacher à la compétence des locuteurs et celle de créativité (dans l'usage) à leur performance.

Cette opposition compétence-performance, introduite par Chomsky (1965) distingue ce qui est de l'ordre de la connaissance qu'a un locuteur-auditeur ('speaker-hearer') de sa langue, du savoir linguistique qu'il possède (compétence) et ce qui est de l'ordre de l'usage réel que fait un locuteur de sa langue en situation d'énonciation, de la mise en œuvre de son savoir linguistique en communication, aussi bien en production qu'en réception (performance). Pour les générativistes, la compétence est une propriété ou une fonction mentale, à laquelle il ne serait possible d'accéder à travers la performance linguistique du locuteur que dans une situation de communication idéalisée, qui serait préservée des difficultés d'attention et de mémoire, des distractions, ou bien encore des erreurs des locuteurs, qui peuvent communément affecter la performance linguistique. Contrairement à la compétence qui représenterait la capacité mentale des locuteurs à comprendre et produire un nombre infini d'énoncés, la performance ne représenterait qu'un échantillon très faible des énoncés pouvant être produits dans une langue donnée. La compétence linguistique implique la connaissance des règles grammaticales et inclut au-delà du domaine de la syntaxe des composantes phonétiques, phonologiques, sémantiques et morphologiques. Nous y reviendrons ultérieurement (section 1.6) lors de la présentation des théories du bootstrapping.

Selon Chomsky, la théorie de la grammaire qui doit être ciblée par le linguiste est celle de la compétence, et donc de la grammaire interne du locuteur, et non de la performance, puisque les productions verbales effectives des sujets ne sont pas le véritable reflet de leur compétence linguistique. Cet extrait, tiré de l'introduction de l'ouvrage de Chomsky publié en 1965, en



témoigne : « a grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence » (1965 : 4).

### **1.3 La notion de 'pauvreté du stimulus'**

S'il n'y est pas fait référence explicitement dans les travaux de Chomsky avant 1965, un des arguments fondamentaux des générativistes pour justifier la nature innée du langage est celui de la 'pauvreté du stimulus' (Pullum et Scholz, 2002 ; Thomas, 2002). Les données linguistiques à disposition de l'enfant seraient insuffisantes pour permettre son développement langagier et expliquer la créativité et la complexité des énoncés produits par les locuteurs, enfants et adultes. Les énoncés entendus ou adressés à l'enfant sont agrammaticaux ou incomplets, et ne fournissent donc pas de bons 'modèles', comme l'explique Chomsky dans une note de son ouvrage de 1965 : « It also seems apparent that much of the actual speech observed consists of fragments and deviant expressions of a variety of sorts » (Chomsky, 1965 : 201).

Il relève par ailleurs dans cette même note qu'aucune information métalinguistique n'est offerte aux enfants ; les adultes n'enseignent pas la langue à apprendre aux enfants : « It seems clear that many children acquire first or second languages quite successfully even though no special care is taken to teach them and no special attention is given to their progress » (Chomsky, 1965 : 200-201). L'enfant ne pourrait pas, selon la perspective générativiste, procéder au même travail que le linguiste puisqu'il n'aurait pas à sa disposition dans le langage qui lui est adressé, les exemples grammaticaux et agrammaticaux fondamentaux pour comprendre les principes de la syntaxe de sa langue (Chomsky, 2002[1957] ; Fortis, 2008 ; Pinker, 1996[1984]). Ce matériau ne fournirait pas de façon suffisamment transparente les règles qui sous-tendent le fonctionnement des catégories grammaticales.

### **1.4 Les trois composantes de la grammaire générative**

En s'appuyant sur cette idée que l'environnement linguistique des locuteurs ne peut pas expliquer à lui seul la créativité de leurs énoncés et leur niveau de compétence linguistique, Chomsky postule l'existence d'un système de règles - une grammaire générative - qui constituerait la capacité linguistique innée des locuteurs :

« On the basis of a limited experience with the data of speech, each normal human has developed for himself a thorough competence in his native language. This competence can be represented, to an as yet undetermined extent, as a system of rules that we can call the grammar of his language » (Chomsky, 1965 : 8-9).

La grammaire innée des locuteurs leur permettrait de pallier les insuffisances de la langue à laquelle ils sont exposés et notamment de reconnaître des énoncés acceptables ou inacceptables, ambigus, ou encore syntaxiquement corrects mais sémantiquement incorrects.

Dans la théorie Standard, trois composantes du langage sont prises en compte : les composantes syntaxique, sémantique et phonologique. Contrairement aux premiers travaux de Chomsky qui ne se concentraient que sur la dimension syntaxique, cette théorie convoque donc également l'étude du sens et du son (Dubois-Charlier et Vautherin, 2008). La composante syntaxique se partage entre règles de base, à l'origine de ce que Chomsky appelle la structure profonde, et règles transformationnelles qui transposent la structure profonde en une structure dite de surface ou structure superficielle. La structure profonde comprend les règles fondamentales du fonctionnement de la langue que sont les règles syntagmatiques, et la structure de surface permet quant à elle de transformer la syntaxe de la structure profonde en une autre syntaxe, adaptée à la langue du locuteur. La composante sémantique permet de fournir au locuteur les informations nécessaires à l'interprétation de la signification de la structure profonde et la composante phonologique, d'y associer les sons à produire dans la langue. D'après Schane, « les composantes sémantique et phonologique ont donc une nature purement interprétative, puisque leur fonction est seulement d'assigner, respectivement, un sens et une prononciation aux phrases fournies par la composante syntaxique » (1967 : 3).

### **1.5 Du 'Modèle Standard' au 'Programme Minimaliste' : évolution de la perspective générativiste sur l'acquisition du langage**

La théorie décrite par Chomsky dans son ouvrage de 1965 - théorie du 'Modèle Standard' - est modifiée au fil des recherches des générativistes. La perspective sur la Grammaire Universelle définie dans cet ouvrage ainsi que l'approche développée sur l'acquisition du langage sont ensuite reprises et partiellement modifiées dans les travaux ultérieurs. Une revue exhaustive de la littérature impliquerait un retour sur chacune des étapes de l'évolution de la grammaire générative, mais notre point étant ici de faire la synthèse de la perspective générativiste sur

L'acquisition du langage, nous concentrerons notre discussion sur les travaux fondateurs pour les recherches sur le langage de l'enfant. Nous reviendrons brièvement dans les paragraphes qui suivent sur le 'dispositif d'acquisition du langage' ('Language Acquisition Device') proposé par Chomsky en 1965, sur la 'Théorie des Principes et Paramètres' (Chomsky, 1981) et sur les travaux plus récents des générativistes issus du 'Programme Minimaliste' (Chomsky, 1993).

Toute la théorie de la Grammaire Universelle repose sur l'idée que les structures de la langue sont innées. Comme le signalent Schaff et Brendel (1973), cette conception de l'acquisition « n'est pas fortuite, mais bel et bien nécessaire dans le système conceptuel de la grammaire générative. Sans elle en effet, il est impossible de fonder la thèse sur l'universalité de la grammaire [...] » (1973 : 13). Parallèlement au développement de la théorie du Modèle Standard (1965), Chomsky introduit une approche de l'acquisition du langage, fondée sur l'idée que l'enfant disposerait d'un dispositif spécifique à l'acquisition du langage. Ce dispositif - accessible à l'enfant dès le début de son acquisition de la langue - serait spécifique à l'Homme et fonctionnerait distinctement des autres domaines de la cognition. Pour Chomsky, le Language Acquisition Device (LAD) pré-programme l'enfant à l'acquisition d'une langue en lui fournissant les universaux linguistiques qui constituent la Grammaire Universelle. Ces universaux innés expliqueraient qu'il soit possible pour un individu d'apprendre une langue et de comprendre et produire des énoncés encore jamais entendus :

« Such universal conditions are not learned ; rather, they provide the organizing principles that make language learning possible, that must exist if data are to lead to knowledge. By attributing such principles to the mind, as an innate property, it becomes possible to account for the quite obvious fact that the speaker of a language knows a great deal that he has not learned » (Chomsky, 2002[1966] : 94).

Chomsky distingue deux types d'universaux linguistiques innés : les universaux formels et les universaux substantiels. Les premiers permettent à l'enfant de saisir les contraintes qui régissent le fonctionnement des langues naturelles, et les seconds, les catégories des éléments de la langue, ainsi que leurs fonctions et les relations qu'ils peuvent entretenir. Ces derniers concernent des « généralisations portant sur les unités phonologiques, morphologiques, syntaxiques ou sémantiques supposées être les éléments constitutifs de n'importe quelle langue humaine » (Egré, 2011 : 47).

Le LAD justifierait également la rapidité d'acquisition du langage et le fait que le rythme d'acquisition du langage soit comparable d'un enfant à l'autre, et cela indépendamment de l'environnement linguistique. Par ailleurs, si l'exposition à la langue est nécessaire pour activer les mécanismes en jeu dans le dispositif d'acquisition du langage, cette théorie attribue un rôle négligeable au contexte d'acquisition.

Comme le signale Ingram (1989), le Modèle Standard et les travaux de Chomsky commencent dans les années 1960 à ne plus seulement influencer les théories en syntaxe mais également les travaux en acquisition. Le modèle proposé par Chomsky est repris dans le but de décrire et expliquer le développement grammatical précoce de l'enfant. McNeill (1966, 1970) contribue notamment à l'enrichissement de la théorie du Language Acquisition Device initialement proposée par Chomsky.

Dans la lignée des travaux de ce dernier, McNeill propose une théorie complète sur l'acquisition de la syntaxe chez l'enfant. Fondée sur l'idée que l'enfant a à sa disposition dans son bagage génétique une grammaire universelle, la théorie proposée par McNeill établit qu'à la base de cette grammaire se trouve le socle des fonctions grammaticales (sujet, prédicat, objet, modifieur) qui sous-tendraient la structure de toutes les langues. L'enfant serait capable de maîtriser ces relations grammaticales élémentaires à la fin de la période des premières combinaisons.

McNeill se positionne par rapport à la grammaire à pivots de Braine (1963). Il conçoit que dans la première partie de la période des énoncés à deux termes, la grammaire de l'enfant soit semblable à la grammaire à pivots, bien qu'à l'inverse de ce qui est proposé par Braine, les mots pivots ne se définiraient pas par leur fréquence et leur position fixe mais par leur place dans la hiérarchie universelle des fonctions grammaticales. La fin de la période des énoncés à deux termes est marquée pour McNeill par la présence chez l'enfant d'une grammaire de base dotée des fonctions syntaxiques rudimentaires universelles et qui émerge grâce à l'exposition de l'enfant au langage adulte. Cette théorie fait donc intervenir une composante développementale dans la mise en place de la syntaxe (Ingram, 1989), qui n'était pas présente dans les développements antérieurs sur le LAD.

La théorie de McNeill est ensuite reprise et développée par Pinker (1996)[1984]. Comme McNeill, Pinker considère que l'enfant possède certaines des catégories syntaxiques adultes dès la naissance et que celles-ci émergent progressivement dans le langage de l'enfant. Il propose en

plus de ce qui est avancé dans la théorie de McNeill une hypothèse précise sur les étapes par lesquelles passe l'enfant dans le développement de sa grammaire (Ingram, 1989) et attribue une part importante à la sémantique dans la définition des catégories grammaticales. Nous reviendrons ultérieurement au travail de Pinker (au point 1.6.1) lorsque nous nous intéresserons au mécanisme du bootstrapping sémantique.

Un autre modèle - celui des 'Principes et Paramètres' - est proposé par Chomsky en 1981, pour d'une part légitimer le principe de la Grammaire Universelle malgré les variations inter-langues et d'autre part, proposer un modèle plus homogène après « la multiplication de règles et de transformations parfois contradictoires » (Kail, 2000 : 11) ayant fait suite à la théorie du Modèle Standard. Dans cette conception de la Grammaire Universelle, les générativistes considèrent toujours le langage comme une faculté innée qui dépend seulement d'une exposition minimale à la langue qui entoure l'enfant. L'acquisition est perçue comme tout aussi 'instantanée' que dans la théorie du LAD (Chomsky, 1988) mais les règles avancées pour expliquer l'acquisition du langage diffèrent de celles antérieurement proposées.

Dans le cadre de cette théorie, la Grammaire Universelle serait constituée d'une structure de base, de règles fondant le système linguistique - les principes - et de règles - les paramètres - statuant sur les potentielles variations ou écarts d'une langue par rapport aux principes de base. La théorie des Principes et Paramètres accorde au travers du principe de projection une place importante au lexique qui, d'après Chomsky, aiderait l'enfant à maîtriser la syntaxe de sa langue. Le principe de projection permettrait en effet à l'enfant d'associer les classes grammaticales des éléments lexicaux et leur fonctionnement syntaxique, en tant que constituants à l'intérieur d'une phrase. Les noms et les verbes reconnus par l'enfant se 'projetteraient' par exemple en syntagmes nominaux et syntagmes verbaux (Kail, 2000). Les paramètres permettraient par ailleurs d'adapter les principes à la langue spécifique en acquisition, ceux-ci s'ajustant en fonction de l'exposition de l'enfant à l'input de son environnement langagier. Ce phénomène a été invoqué à de nombreuses reprises pour expliquer par exemple que malgré la notion de Grammaire Universelle, un locuteur puisse appréhender les variations d'une langue à l'autre quant à l'usage obligatoire ou non des pronoms sujets (Hyams, 1986).

Suite aux travaux sur la théorie des Principes et Paramètres, Chomsky élabore à partir de 1993 la théorie du 'Programme Minimaliste'. Les principes fondamentaux de ce nouveau modèle restent

les mêmes que ceux proposés dans les versions antérieures de la théorie générativiste : les règles qui organisent les langues sont universelles et la faculté de langage est une propriété spécifique de l'homme, à dissocier de ses autres facultés. L'objectif est également semblable ; il s'agit de continuer à expliquer les contraintes d'une langue particulière dans le cadre d'un système invariant et universel (Bogny Yapo, 2007).

Dans la continuité des travaux précédents, le Programme Minimaliste développe en les modulant les arguments avancés dans la théorie des Principes et Paramètres. L'enfant posséderait une structure intrinsèque à la faculté de langage qui favoriserait son acquisition langagière. Cette structure serait constituée de paramètres qui se mettraient en place entre 0 et 6 ans. A sa naissance, l'enfant ne serait donc pas prédisposé à l'acquisition d'une langue particulière, seule la mise en place rapide des paramètres le conditionnerait à l'acquisition de la langue particulière de son environnement. A la différence de la théorie précédente, le Programme Minimaliste stipule que les principes et les paramètres sont en nombre restreint dans la grammaire de l'enfant, sans quoi la période d'acquisition de la langue serait beaucoup plus longue. La variation inter-langues prend une place bien moins importante dans le processus d'acquisition dans ce nouveau cadre minimaliste, le nombre de paramètres lui étant dédiés est donc limité, ce qui explique qu'ils soient acquis rapidement. En comparaison de la théorie précédente, Chomsky simplifie par ailleurs considérablement la liste des opérations présentées dans la théorie des Principes et Paramètres ; les différents modules de la grammaire générative établis antérieurement et à acquérir par l'enfant sont bien moins complexes et nombreux.

## **1.6 Les théories du « bootstrapping »**

D'autres chercheurs travaillant dans le cadre de la grammaire générative se sont davantage focalisés sur les mécanismes d'apprentissage auxquels l'enfant peut avoir recours pour acquérir les propriétés de sa langue. Ces théories dites du 'bootstrapping' partagent avec les théories chomskyennes l'idée que l'enfant possède un dispositif spécifique et inné d'acquisition du langage mais cherchent davantage à en expliquer le contenu et l'évolution (Kail, 2000). Au-delà des différences entre les nombreux types de bootstrapping proposés dans la littérature, on pourra noter que ces théories « ont en commun l'idée que l'enfant peut utiliser des indices présents dans l'input langagier ou des connaissances déjà établies dans un domaine linguistique pour acquérir

d'avantage de savoir linguistique dans le même domaine ou au sein d'un autre domaine » (Höhle, 2009 : 363, *notre traduction*). Pour la plupart des travaux s'intéressant aux mécanismes de bootstrapping, la part du rôle joué par l'input dans le développement du langage est plus importante que dans les développements proposés par Chomsky.

### 1.6.1 Le bootstrapping sémantique

Ancré dans la théorie des Principes et Paramètres (Chomsky, 1981), le concept de bootstrapping est d'abord introduit par Pinker en 1984. Dans son ouvrage *Language learnability and language development*, il propose en effet l'idée que l'enfant s'appuierait sur les représentations sémantiques qu'il se fait des énoncés de l'input pour attribuer des étiquettes grammaticales aux mots de sa langue.

Ces représentations sémantiques - rendues possibles par le contexte des énoncés produits dans l'input de l'enfant et par sa capacité innée à comprendre le sens des mots - seraient reliées par l'enfant à des éléments syntaxiques présents dans sa grammaire. Cette capacité à faire correspondre sémantique et syntaxe permettrait à l'enfant d'acquérir d'une part les caractéristiques morphosyntaxiques propres à chaque catégorie grammaticale et d'autre part, les relations syntaxiques appropriées dans sa langue (Höhle, 2009).

D'autres approches sont nées suite à celles de Pinker, dont deux d'entre elles également intégrées au cadre général de la Grammaire Universelle : le mécanisme de bootstrapping syntaxique initialement proposé par Gleitman (1990) et le mécanisme de bootstrapping prosodique proposé dans sa première version par Gleitman et Wanner (1982).

### 1.6.2 Le bootstrapping syntaxique

L'approche de Gleitman (1990) établit que les enfants déduisent une partie de la signification des mots qui leur sont inconnus à partir des cadres syntaxiques dans lesquels apparaissent ces mots nouveaux. C'est donc, comme pour Pinker (1996[1984]), une association de la sémantique et de la syntaxe qui aide l'enfant à approfondir ses compétences linguistiques mais dans l'approche de Gleitman, c'est la syntaxe qui fournit une passerelle vers l'acquisition de la sémantique et non l'inverse.

Les expériences menées sur les verbes laissent apparaître que l'enfant s'appuierait sur le contexte syntaxique de l'énoncé pour accéder au sens des verbes. Ainsi, pour reprendre l'exemple de Höhle (2009), les verbes produits dans des contextes transitifs seraient analysés par l'enfant comme des verbes causatifs, par opposition aux verbes produits dans des contextes intransitifs.

S'il est difficile d'établir que les connaissances syntaxiques de l'enfant lui permettent de déduire la signification précise des verbes (Pinker, 1996[1984]), elles lui offriraient *a minima* l'accès à la catégorie sémantique d'appartenance des verbes produits dans son input.

### 1.6.3 Le bootstrapping prosodique et phonologique

D'autres recherches, comme celles de Gleitman et Wanner (1982) ou plus tard, Morgan et Demuth (1996), ont établi que la prosodie pourrait elle aussi jouer un rôle fondamental dans la mise en place du langage chez l'enfant, et plus particulièrement dans la découverte des unités syntaxiques. Certains indices prosodiques tels que le rythme, l'accentuation ou l'intonation contribueraient à aider l'enfant à segmenter la parole entendue, à acquérir les règles d'organisation des mots de sa langue et à en catégoriser syntaxiquement les différents éléments. Les informations prosodiques et phonologiques tirées de l'input permettraient à l'enfant d'accéder à une information syntaxique rudimentaire qui favoriserait par la suite des compétences inductives plus complexes. D'autres indices - sémantiques ou syntaxiques par exemple - seraient alors accessibles pour l'enfant. La part à attribuer à la prosodie et à la phonologie deviendrait donc moins importante au fil du développement langagier de l'enfant mais « at the root, however, it is the signal that conveys words, phrases, sentences, indeed grammars to each new generations » (Morgan et Demuth, 1996 : 20). Nous reviendrons aux travaux de Demuth et au rôle de la prosodie dans le développement du langage au sein du chapitre suivant, lorsque nous traiterons la question de l'émergence des fillers et des déterminants devant le nom.

On notera par ailleurs que seules quelques approches situées dans le cadre de la grammaire générative sont listées ici mais que d'autres variantes relatives à la sémantique, la syntaxe ou la prosodie, ou bien encore liées à la distribution des éléments de la langue dans la phrase et à la typologie des langues ont été avancées dans de très nombreux travaux depuis la parution de ceux fondateurs de Pinker. Nous reviendrons notamment au point 3.1 de ce chapitre sur la



perspective adoptée par les chercheurs générativistes pour traiter la question de l'acquisition des catégories syntaxiques.

## **2. L'acquisition du langage : une construction progressive**

Les différentes approches que nous venons de présenter sont fondées sur l'idée que l'enfant possède une connaissance innée des structures linguistiques et n'envisagent donc pas que le langage puisse faire l'objet d'un apprentissage. Cette perspective a largement été remise en cause par d'autres approches théoriques. Toutes ne s'accordent pas sur les processus à l'œuvre dans le développement linguistique, en ce qui concerne notamment le rôle des mécanismes cognitifs, la part de l'enfant ou de l'interaction dans le développement langagier mais elles sont néanmoins unies autour d'une même notion : le langage n'est pas une découverte pour l'enfant mais bien une construction ; d'après Hirsh-Pasek et Golinkoff (1996), les théories dites 'outside-in' ont en commun l'idée que l'enfant « works to learn language structure ; it does not come for free » (1996 : 19).

Cette seconde partie du chapitre s'attachera à faire la synthèse de ces différentes approches et à décrire ce qui fonde leur unité aussi bien que leur diversité. Le point suivant (2.1) sera consacré aux recherches faisant une place importante, voire majoritaire, aux mécanismes cognitifs en jeu dans l'acquisition du langage, et nous nous concentrerons au point 2.2 sur les études ayant davantage (ou par ailleurs) mis en avant le rôle des déterminants sociaux dans le développement langagier.

### **2.1 Autonomie de l'enfant dans le processus d'acquisition langagière**

L'ensemble des théories qui partagent l'idée que le langage n'est pas inné peuvent être organisées autour d'un continuum aux extrémités duquel figureraient le rôle des capacités cognitives d'une part, et le rôle de l'environnement et de l'interaction d'autre part.

Parmi les travaux qui privilégient les facteurs cognitifs aux facteurs interactionnels, on trouve ceux de l'école constructiviste de Genève, incarnée par Piaget ou Sinclair. Pour les constructivistes piagétiens, développements cognitif et langagier sont étroitement liés et l'acquisition du langage est nécessairement conditionnée par un certain degré de développement

cognitif ; autrement dit, la mise en place du langage repose sur un socle de capacités cognitives de base, considéré comme un prérequis indispensable à toute acquisition langagière.

Ce socle de base qui se met en place entre la naissance de l'enfant et ses deux ans est ce que Piaget (1990[1937]) appelle le stade de l'intelligence sensori-motrice. Pendant cette période, l'enfant fait l'expérience du réel qui l'entoure, et cette expérience lui permet d'accéder au stade ultérieur (dit 'pré-opératoire'), à la fonction symbolique ou de représentation et au langage. Pour Piaget, le langage, reposant sur une faculté d'abstraction, est d'abord structuré par l'action et la découverte du concret, comme le signale Veneziano (2000) :

« la capacité de construire des 'catégories' d'objets et d'actions, des relations entre ceux-ci, et l'acquisition de la notion concomitante d'objet permanent, acquises pendant la période sensorimotrice, fournissent des bases pour construire des relations de signification stables et pour saisir la nature de la relation de signification qui lie les signifiants langagiers à leurs signifiés » (2000 : 233).

Contrairement aux théories innéistes, le langage n'est donc pas, dans une perspective piagétienne, une faculté qui s'acquiert de façon innée mais qui repose sur des structures cognitives déjà présentes chez l'enfant lorsqu'il aborde l'acquisition de sa langue.

L'école constructiviste piagétienne n'est pas la seule à mettre en avant l'importance des prérequis cognitifs. On retrouve aussi cette idée dans les travaux de Braine par exemple (*i.a.* 1971, 1992). D'après lui, l'enfant dispose d'habiletés innées mais celles-ci seraient d'ordre cognitif et non syntaxique. Braine (1992) reprend ainsi la théorie syntaxique développée par Pinker (1996[1984]) en substituant l'hypothèse d'une grammaire innée par celle d'une capacité cognitive innée et propose par exemple que l'enfant puisse s'appuyer sur ces connaissances cognitives pour faire la distinction entre prédicat et argument.

Comme Braine, et contrairement aux générativistes, Slobin (*i.a.* 1973, 2004) appréhende les fonctionnements cognitif et langagier comme fondamentalement liés, et considère par ailleurs que l'acquisition du langage est régie par des principes cognitifs universels. A partir de l'observation de données enfantines, il dégage des principes opératoires qui, d'après lui, permettraient à l'enfant d'extraire et segmenter le langage entendu. Il argumente par exemple en faveur d'une acquisition plus rapide des unités généralement produites en fin d'énoncé, segment auquel l'enfant serait plus sensible.

Influencés à la fois par la perspective fonctionnaliste telle qu'elle a été pensée par Slobin, et par les apports de la psychologie cognitive, Bates et MacWhinney (*i.a.* 1989) développent un modèle dit 'de compétition' pour décrire la mise en place de la grammaire chez l'enfant. Contrairement à des modèles antérieurs, celui-ci n'est pas construit sur la base des catégories abstraites de la langue mais sur un système de correspondance entre formes et fonctions, autrement dit, entre grammaire et signification. Dans la lignée des travaux de Slobin, le modèle de compétition conçoit le langage comme une faculté qui doit faire l'objet d'un apprentissage et qui repose sur des mécanismes cognitifs de base, généraux et universels. Ce modèle d'acquisition, élaboré suite à de nombreuses observations et comparaisons inter-langues, est en effet fondé sur l'idée que dans un environnement linguistique stimulant, l'enfant est capable, grâce à des processus cognitifs généraux, de faire le lien entre des signes intrinsèquement vides de sens et des référents, et ainsi, de pouvoir inférer le sens des mots, ou plus largement, des énoncés.

Notre objectif n'étant pas ici de rendre compte de façon exhaustive de la manière dont a été traitée la question de la cognition dans la littérature acquisitionniste, nous limiterons donc notre exposé à ces quelques illustrations. Signalons toutefois que si les études citées nous semblent fondatrices, elles ne témoignent que partiellement de la diversité des approches ayant argumenté en faveur de la primauté des capacités cognitives sur les capacités langagières, et de la dépendance entre développement cognitif et langagier. Nous verrons par ailleurs dans les pages suivantes que les approches qui mettent en avant le caractère essentiel de l'interaction pour le développement langagier ne négligent pas pour autant le rôle des mécanismes cognitifs dans le processus d'acquisition du langage.

Outre la place importante consacrée aux mécanismes cognitifs, les travaux que nous venons d'examiner ont également en commun le fait de considérer que l'enfant est actif et relativement autonome dans le processus d'acquisition langagière, et que le rôle joué par l'interaction, bien que non négligeable, n'est pas central.

A la lecture des travaux que nous venons de citer et qui s'inscrivent dans des perspectives théoriques variées (constructivisme piagétien, fonctionnalisme), on peut à première vue considérer - dès lors que leurs auteurs n'envisagent pas que l'enfant possède une grammaire universelle et innée mais des mécanismes cognitifs généraux garantissant un apprentissage réussi de la langue - qu'ils se distinguent nettement des approches générativistes. Ces approches

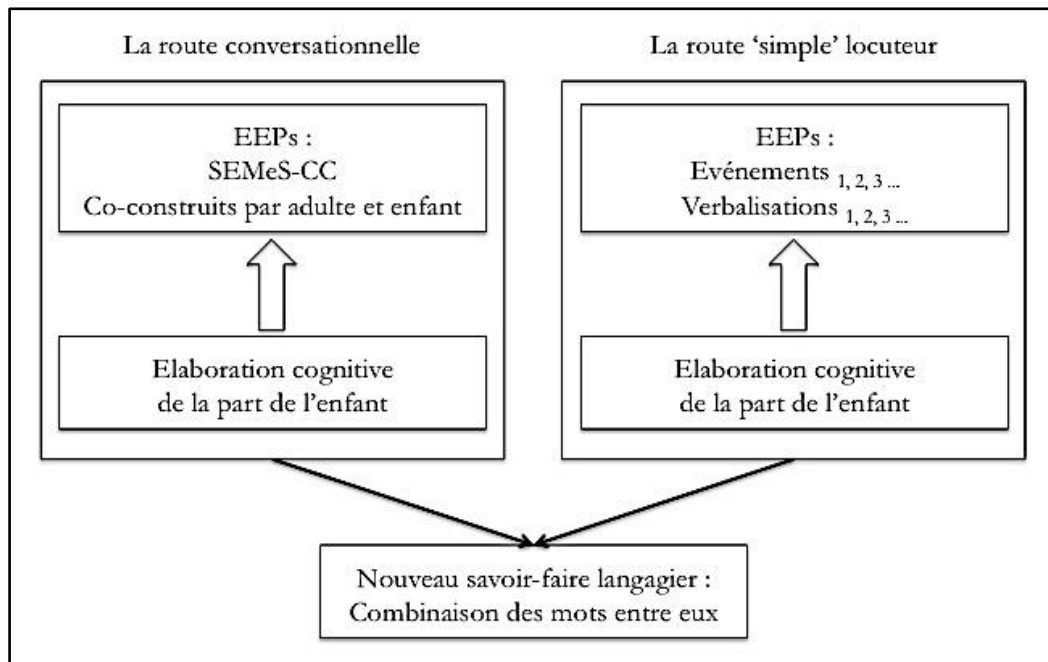
partagent toutefois avec la conception chomskyenne l'idée que la contribution de l'entourage de l'enfant et de l'interaction est minimale.

Dans la perspective piagétienne initiale, l'acquisition du langage n'est certes pas perçue comme déconnectée du rapport de l'enfant à son environnement linguistique, mais c'est avant tout le rôle de l'enfant, qui agit sur son environnement et qui, plus particulièrement, traite de façon autonome les énoncés qu'il entend et en assimile les règles, qui est privilégié. Le langage se construirait donc avant les interactions sociales, et non au cœur de ces interactions, comme cela est envisagé dans le cadre des études interactionnistes, que nous aborderons ultérieurement.

Toutefois, parmi des études plus récentes conduites dans un cadre théorique initial piagétien, la place de l'interaction et des processus conversationnels est réexaminée à la lumière de l'interactionnisme vygotkien. Dans les travaux de Veneziano (*i.a.* 1999, 2000) par exemple, on retrouve l'idée que les savoir-faire langagiers de l'enfant reposent sur une combinaison entre élaboration cognitive et développement autonome de nouveaux savoir-faire langagiers, basés sur des savoir-faire antérieurs, mais on constate par ailleurs que pour l'auteure, la tendance interprétative des adultes - de la mère notamment - et les échanges co-construits entre adulte et enfant, peuvent jouer un rôle très favorable au développement du langage.

Dans le cadre d'un travail portant notamment sur la transition entre énoncés à un terme et énoncés à deux termes, Veneziano (1999) propose ainsi que le passage entre ces deux stades puisse être déclenché par des processus d'acquisition issus de deux 'routes' distinctes, considérées comme « alternative[s] et/ou complémentaire[s] » (1999 : 218) : la 'route conversationnelle' et la 'route du *simple* locuteur'. Dans le premier cas, l'enfant peut construire des suites d'énoncés à un mot en succession en collaboration avec l'adulte (SEMeS-CC) ; autrement dit, l'adulte et le fonctionnement conversationnel peuvent soutenir le développement de SEMeS, dont le degré de complexité ne serait peut-être pas atteint par l'enfant sans l'étayage de l'adulte, à ce stade de son développement langagier. *A contrario*, dans le cas de la 'route du *simple* locuteur', l'enfant utilise ses propres productions comme un input, pour accéder à de nouvelles acquisitions ; pour l'auteure, « les progrès dans le répertoire lexical et dans la manière de l'utiliser peuvent, en eux-mêmes, ouvrir la voie à la combinaison des mots » (*ibid.* : 219). Ces deux 'routes' ont en commun pour Veneziano d'impliquer un traitement cognitif de la part de l'enfant. La schématisation du modèle proposé dans le cadre de cette étude est reproduite ci-dessous.

Schéma 1 - Les processus d'acquisition : route conversationnelle et route du 'simple' locuteur (Veneziano, 1999 : 219)



*EEPs : événements énonciatifs*

*SEMeS-CC : suites d'énoncés à un mot en succession - soutenues par le cadre conversationnel*

Dans le cadre des approches proposées par Slobin et Braine auxquelles nous faisons référence plus haut, l'enfant doit être entouré de langage pour pouvoir s'en saisir et le traiter. On peut donc penser qu'elles ne négligent pas le rôle de l'environnement social et de l'input mais elles n'accordent pas pour autant à l'interaction un statut de pilier primordial à l'apprentissage. Le langage adressé à l'enfant est seulement envisagé comme un matériau que l'enfant traite pour s'approprier sa langue, sans pour autant que l'interaction avec l'autre et le soutien de ses interlocuteurs ne soient traités comme des conditions nécessaires au développement cognitif et langagier.

Enfin, dans le cadre du modèle proposé par Bates et MacWhinney, le rôle de l'environnement dans lequel l'enfant évolue n'est pas non plus ignoré, comme en témoigne cet extrait d'un chapitre de MacWhinney (2001) : « If a young child is reared in isolation, language acquisition will be difficult. If the child is given strong scaffolding and support, acquisition will be facilitated » (2002 : 17). On relève ici que le développement du langage n'est donc pas, pour l'auteur, dissocié du soutien et de l'étaiyage que l'enfant peut se voir offrir par ses interlocuteurs. Il convient néanmoins d'ajouter que si pour un enfant élevé sans contact avec un environnement

linguistique, l'acquisition du langage serait 'difficile', elle n'en est que 'facilitée' par un environnement stimulant, et non pleinement conditionnée par cet environnement et par la qualité des échanges auxquels l'enfant prend part. La question de l'influence du contexte est davantage traitée au sein du modèle de compétition en termes de traitement de l'input et de fréquence de production des mots et des structures. Plus ils sont fréquents dans l'input, plus l'enfant les acquiert rapidement, et *a contrario*, moins ils sont produits et plus l'enfant a de chances de les assimiler tardivement. Appliqué à une réflexion sur l'acquisition d'une langue seconde, tout l'enjeu du modèle est, selon MacWhinney (2001), de « comprendre comment l'input linguistique peut être structuré pour maximiser un apprentissage efficace. Quels aspects phonologiques, syntaxiques, sémantiques et morphologiques de l'input sont utilisés par l'apprenant pour 'décoder' la langue nouvelle » (2002 : 1). Dans le cas où l'apprenant n'aurait pas accès à un environnement social riche et au sein duquel est parlée la langue en cours d'acquisition, il pourrait par ailleurs créer son propre système de 'soutien autonome' ('autosupport'), au moyen de stratégies telles que l'écoute de la radio ou le visionnage de la télévision ou de films. Nous ne sommes pas en mesure de discuter ici du rapport qu'entretient un apprenant à l'input de sa langue seconde ni de la pertinence didactique des stratégies proposées ; toutefois, il nous semble intéressant de mentionner ici ce travail pour ce qu'il révèle de la prise en compte de l'interaction, dont l'intérêt prédominant serait d'apporter des modèles linguistiques sur lesquels l'enfant autonome, peut s'appuyer, plus que des échanges au travers desquels l'enfant co-construit, avec son interlocuteur, ses compétences langagières.

D'autres études ont davantage mis l'accent sur l'idée que l'enfant apprend à parler avant tout pour communiquer avec l'autre et que ceux qui interagissent avec lui jouent un rôle essentiel dans son processus d'acquisition. Au point suivant, nous passerons en revue quelques-uns des aspects majeurs développés dans le cadre de ces travaux.

## **2.2 Apprendre à parler : un processus social**

Nous présenterons, dans la première partie de cette section 2.3 certains des travaux ayant placé la motivation sociale de l'enfant et de son acquisition langagière au cœur de leur réflexion, en abordant en particulier la question de l'intention communicative du jeune enfant et des différentes fonctions qui peuvent être attribuées à ses premières productions. Dans une seconde

partie, nous passerons en revue certaines des thématiques largement renseignées dans la littérature interactionniste, en abordant dans un premier temps la question des caractéristiques globales de l'interaction adulte-enfant considérées comme nécessaires au développement langagier (situations ritualisées, attention conjointe), et dans un second temps, nous discuterons de façon spécifique des caractéristiques linguistiques de cette interaction (nature du langage adressé à l'enfant et conduites verbales étayantes). Enfin, au sein de la section 2.3.3, nous traiterons de la dépendance entre les productions langagières de l'enfant et l'input entendu ou directement adressé, ou plus globalement, du rapport entre les usages langagiers de l'enfant et son expérience communicative. Nous ferons appel dans cette section aux études menées dans le cadre de la théorie usage-based.

### **2.2.1 Enjeux communicationnels de l'acquisition du langage**

Nous l'avons signalé plus tôt, l'ensemble des travaux qu'il est possible de rassembler sous l'étiquette 'non innéistes' ne sont pas comparables en tous points. Parmi les divergences figurent notamment les questions de la motivation sociale de l'enfant qui apprend à parler et celle des ressources dont il dispose, grâce au soutien de ses partenaires conversationnels, pour accéder aux différentes fonctions du langage.

Si certains auteurs comme Piaget ont d'abord mis en avant la continuité entre développement cognitif et développement langagier, nous verrons dans les pages qui suivent que d'autres ont insisté davantage sur la continuité et la réciprocité entre intentions communicatives et acquisition du langage. De même, bien qu'il soit déjà question dans la perspective fonctionnaliste de Slobin, ou plus tard, de Bates ou MacWhinney, de la dimension fonctionnelle du langage, on constate que le rapport entre les formes et leurs fonctions est surtout rendu accessible à l'enfant grâce à des habiletés cognitives, et dans une moindre mesure, grâce à son activité communicative.

A l'inverse, dans des études menées en pragmatique développementale ou dans un cadre interactionniste, l'enfant est un être social dont les premières acquisitions sont profondément conditionnées pour et par la communication avec l'autre. Dans ce cadre, le développement langagier n'est pas abordé en termes d'acquisition de la langue en elle-même et pour elle-même, et donc exclusivement au travers des formes, mais bien plutôt en termes de maîtrise des combinaisons entre les formes et les fonctions portées par ces formes. Par ailleurs, la place

attribuée aux facteurs cognitifs est donc secondaire, et celles des motivations communicationnelles de l'enfant et du soutien de son entourage sont posées comme prioritaires.

Parmi les travaux qui ont contribué à mettre au premier plan les intentions communicatives et les fonctions des premiers énoncés du jeune enfant, on trouve le travail fondateur de Halliday (notamment, 1973), basé sur l'observation de l'acquisition du langage de son fils, Nigel. D'après Halliday, le point de départ pour l'enfant n'est pas linguistique, ou grammatical, mais social ; autrement dit, l'intérêt de l'apprentissage langagier ne repose pas sur la maîtrise de la langue mais sur ce qu'elle offre comme potentialités communicatives. L'enfant n'apprend pas à parler pour apprendre à parler, et ainsi seulement maîtriser les phonèmes, les lexèmes ou les structures syntaxiques de la langue, mais aussi et surtout pour interagir avec autrui, pour avoir les moyens de communiquer. Dans cette perspective, il s'est intéressé aux premiers actes de communication de Nigel et a pu établir que durant une première phase (de 9 mois à 16 mois et demi), les productions verbales et non verbales de son fils couvrent six fonctions distinctes que nous synthétisons dans le tableau ci-dessous.

**Tableau 1 - Fonctions couvertes par le langage du jeune enfant telles qu'elles ont été proposées par Halliday (1973)**

Fonction <i>instrumentale</i> (ou <i>I want</i> function)	Volonté d'obtenir objet ou aide
Fonction <i>régulatrice</i> (ou <i>you do this</i> function)	Requête d'action adressée à une personne spécifique
Fonction <i>interactionnelle</i> (ou <i>me and you</i> function)	Volonté d'attention partagée avec un membre de son entourage (le plus souvent, la mère)
Fonction <i>personnelle</i> (ou <i>here I come</i> function)	Expression d'un état personnel (sentiments, plaisir, ...)
Fonction <i>heuristique</i> (ou <i>tell me why</i> function)	Observations et questions adressées à l'interlocuteur (notamment, dénominations et demandes de dénomination)
Fonction <i>imaginative</i> (ou <i>let's pretend</i> function)	Jeux de faire-semblant

Notons que se développe dans une seconde phase (de 16 mois et demi à 22 mois et demi) une septième fonction, la fonction informative : le langage est alors utilisé par l'enfant pour transmettre de l'information à ses interlocuteurs. En parallèle, les fonctions instrumentale, régulatrice et interactionnelle se combinent et donnent lieu à la fonction 'pragmatique', et de même, la combinaison des fonctions personnelle, heuristique et imaginative permet le développement de la fonction 'mathématique'. Dans le premier cas, l'enfant agit sur le monde



extérieur au travers de ses productions, et dans le second, le langage est utilisé comme un moyen d'apprentissage. Si elle n'est pas observée chez Nigel, Halliday signale toutefois qu'il existerait une troisième phase, que Bernicot (1998) synthétise en écrivant qu'elle « correspond à l'entrée dans le système adulte où selon Halliday le nombre de fonctions est illimitée » (1998 : 15).

Les travaux de Halliday ont ensuite donné lieu à un certain nombre d'études qui ont également cherché à décrire le langage de l'enfant au travers des différentes fonctions qu'il peut couvrir. Sur la base des actes de langage tels qu'ils ont été initialement définis par Austin (1969) et Searle (1982), ou plus tard, Searle et Vanderveken (1985), des auteures comme Bernicot ou Marcos ont proposé des classifications et des analyses des actes de langage chez l'enfant, en prenant en considération aussi bien les productions verbales des enfants que leurs gestes et productions vocales non linguistiques. A titre d'exemple, Bernicot (1992) s'est particulièrement intéressée aux actes 'directifs' (ou demandes) de l'enfant - définis par Ryckebusch et Marcos (2000) comme des « actions, gestes, productions vocales ou [des] combinaisons entre ces différents moyens d'expression [sollicitant] une action sur un objet ou une action de l'adulte » (2000 : 137) - et Marcos et Bernicot (1997) aux actes dits 'assertifs', qui « désignent ou caractérisent un lieu, objet, action ou personne » (*ibid.*).

Ces travaux ont en commun avec ceux de Halliday le fait de considérer que les intentions communicatives de l'enfant se manifestent très tôt et qu'elles sont présentes avant même que ne soient maîtrisées les formes linguistiques. Les intentions communicatives sont donc premières et la correspondance entre les formes linguistiques standardisées, conventionnelles, et leurs fonctions est acquise ensuite. Si toutes ces études se focalisent particulièrement sur les intentions communicatives des enfants, les fonctions du langage ou les actes de langage, thèmes consacrés dans le courant de la pragmatique développementale, elles s'intéressent néanmoins aussi à l'influence de l'interaction, au langage adressé à l'enfant ou à l'étayage linguistique offert par les partenaires de l'enfant et présentent donc des points de convergence avec les thématiques développées dans la perspective interactionniste telle qu'elle a été pensée par Vygotski et plus tard, Bruner. Les contributions majeures de ces deux auteurs, et plus globalement, de la théorie interactionniste, seront développées au point suivant (2.3.2).

### **2.2.2 Un processus soutenu par l'interlocuteur**

Aux fondements de l'interactionnisme on trouve l'idée que le développement et le fonctionnement de l'être humain sont inscrits dans la culture d'une communauté : l'individu n'est pas isolé et ne se développe donc pas indépendamment du contexte dans lequel il évolue et de ses interactions avec le monde et les autres. Pour Vygotski par exemple (1997[1936]), le développement de l'enfant ne résulte pas d'une simple maturation du cerveau mais bien au contraire des activités de l'enfant, inscrites dans un contexte culturellement situé. L'enfant est en interaction constante avec les membres de son entourage et ces interactions sont à la source de son appropriation des codes de sa communauté, et notamment du langage. Il ne s'agit pas pour autant de nier l'existence de capacités biologiques et cognitives innées et en particulier, de prédispositions pour le langage, mais de considérer que dès le début de sa vie, l'enfant est tourné vers l'autre.

Au sein de ces interactions, l'adulte est considéré comme l'interface entre la culture, les savoirs communs, les codes sociétaux et l'enfant ; le développement humain ne peut ainsi être perçu pour les interactionnistes que « comme un processus d'assistance, de collaboration entre enfant et adulte, l'adulte agissant comme médiateur de la culture » (Bruner, 1983 : 8). Les codes linguistiques faisant partie intégrante de la culture, l'adulte est donc aussi médiateur de la construction langagière de l'enfant, au travers notamment de situations d'interaction routinières, ritualisées, au sein desquelles l'adulte soutient l'activité de l'enfant, la lui facilite et le guide. Dans les deux points suivants, nous aborderons de façon plus précise la façon dont sont décrites dans la littérature interactionniste, d'une part l'activité de communication entre l'enfant et l'adulte, et d'autre part, les conduites langagières des partenaires de l'enfant<sup>1</sup>.

#### *2.3.2.1 Des contextes interactionnels favorables au développement langagier*

Dans l'ouvrage de référence « Le développement de l'enfant : savoir faire, savoir dire » (1983), Bruner liste quatre capacités dont l'enfant serait, selon lui, doté de façon innée :

---

<sup>1</sup> Signalons que les études qui seront citées dans les points qui suivent se sont concentrées sur un modèle occidental d'interaction dyadique et de langage adressé à l'enfant. Un grand nombre d'études (*i.a.* Schieffelin et Ochs, 1986) ont suggéré que dans d'autres contextes culturels, les enfants apprennent à parler dans des contextes variés d'interactions polyadiques.

- 1- *Disponibilité orientée vers les moyens et les fins* : l'enfant serait d'emblée orienté vers la recherche de moyens pour parvenir à un but, à des fins qui puissent être satisfaisantes pour lui ;
- 2 - *Transactionnalité* : l'enfant est un être social, et ses activités sont dès sa naissance tournées vers l'autre et vers la communication ;
- 3 - *Systematisation* : dans les premiers mois de sa vie, l'enfant tend à systématiser son action, à prendre part à des activités définies, répétées, et dont le nombre est limité ;
- 4 - *Abstraction* : à la suite de Piaget, Bruner considère que l'enfant organise le monde qui l'entoure avec des règles abstraites (notamment, gestion de l'espace et du temps).

Si Bruner considère que certains aspects du développement de l'enfant sont dus à des capacités innées, il remet néanmoins en question l'existence d'un Dispositif d'Acquisition du Langage (Language Acquisition Device, LAD), tel qu'il a été proposé par Chomsky (*cf.* point 1.5 de ce chapitre). L'acquisition du langage ne serait rendue possible pour Bruner que par l'aide de l'adulte et au travers de scénarios routiniers, ritualisés. Ces deux composantes - routines et tutelle de l'adulte - constituent ce qu'il appelle le 'Système de Support à l'Acquisition du Langage' (Language Acquisition Support System, LASS).

C'est en analysant deux suivis longitudinaux d'enfants en interaction avec leur mère - Richard et Jonathan, suivis respectivement de 5 mois à 2 ans et de 3 mois à un 1 an et demi, environ deux fois par mois - que Bruner a caractérisé les premières activités des enfants et les scénarios dans lesquels mère et enfant sont impliqués. Il s'intéresse aux jeux, mis en place dans les deux dyades (notamment, jeux de 'coucou', de cache-cache, d'échanges d'objets) et constate que ces activités partagées entre adulte et enfant sont organisées comme des scénarios standardisés au sein desquels les rôles sont d'abord délimités et deviennent ensuite réversibles, lorsque l'enfant est suffisamment familiarisé avec l'exercice.

Ces scénarios très ritualisés et routiniers font sens pour l'enfant qui connaît la situation et lui permettent d'acquérir l'usage conventionnel des symboles linguistiques :

« L'acquisition du langage commence avant que l'enfant n'émette son premier discours lexico-grammatical. Elle commence quand la mère et l'enfant créent un format d'interaction prédictible qui peut servir de microcosme pour la communication et la constitution d'une réalité partagée. Les transactions qui se produisent dans de tels formats constituent 'l'input' à partir duquel l'enfant peut alors maîtriser la grammaire, comment référer et signifier, et comment réaliser ses intentions en communiquant » (1983 : 129).

Dans la mesure où l'enfant peut anticiper sur ce que va faire l'adulte, ce type d'interaction très ritualisée lui permet de développer sa prévisibilité et ainsi, d'interpréter plus facilement les nouveautés (ou 'variations') introduites par l'adulte d'une part, et d'appréhender les présupposés et les aspects implicites de la conversation d'autre part, lorsque certaines des étapes du scénario ne sont pas suivies par la mère. Bruner montre donc comment les activités ludiques, en ce qu'elles sont naturellement structurées et conventionnalisées, permettent à la fois à l'enfant d'accéder au matériau linguistique fourni par l'adulte et à la production d'actes langagiers de plus en plus flexibles et adaptables.

Sur la base du travail de Bruner et de Wittgenstein (notamment, 1953), Tomasello s'est lui aussi intéressé aux caractéristiques des interactions adulte-enfant susceptibles de favoriser chez l'enfant l'apprentissage de l'usage conventionnel des symboles linguistiques. Il a également constaté que les dyades mère-enfant s'engagent dans des activités ritualisées, et a par ailleurs relevé au sein de ces activités l'existence d'épisodes dits 'd'attention conjointe', qui témoigneraient de la compréhension chez l'enfant de l'intentionnalité de son interlocuteur :

« A partir de 9 mois, ils [les enfants] s'engagent en effet dans un certain nombre de comportements que l'on qualifie d'attention conjointe, semblant indiquer qu'ils commencent à comprendre que les autres sont, comme eux-mêmes, des êtres intentionnels : on peut suivre, diriger ou partager les relations qu'ils entretiennent avec des entités appartenant au monde extérieur » (2004 : 62).

Dans ces épisodes d'attention conjointe, enfant et adulte portent leur attention sur un même objet (ou plus généralement, un même référent) et communiquent ensemble à propos de ce centre d'intérêt commun. Pour Tomasello, ce partage d'attention joue un rôle essentiel dans le développement du langage en facilitant l'entrée de l'enfant dans l'intersubjectivité : l'enfant accède, au travers de ces épisodes, aux intentions communicatives de son partenaire et apprend à comprendre la façon dont l'adulte utilise les symboles linguistiques.

Plusieurs études se sont focalisées sur l'influence de l'attention conjointe sur le développement du langage de l'enfant. Tomasello et Todd (1983) ont ainsi montré que les enfants qui passent le plus de temps dans des épisodes d'attention conjointe avec leur mère entre l'âge de 12 et 18 mois possèdent un vocabulaire plus étendu à l'âge de 18 mois. De même, Baldwin (1991) a cherché à faire apprendre de nouveaux mots à de jeunes enfants (de 19 mois) dans deux situations distinctes : dans la première, l'adulte offrait de nouveaux mots en suivant ce que retenait

l'attention de l'enfant, et dans la seconde, l'offre était faite quand l'enfant ne regardait pas l'objet, l'enfant devant donc, *a posteriori*, focaliser son attention sur l'objet que l'adulte avait dénommé. Les résultats ont montré que dans le premier cas, lorsque l'attention est partagée entre les participants, l'apprentissage des nouveaux éléments lexicaux était favorisé.

Les études qui se sont intéressées de près aux interactions entre enfant et adulte nous permettent donc de constater que certaines des caractéristiques de cette situation de communication - caractère ritualisé de l'échange et disposition des participants à l'attention conjointe - soutiennent le développement langagier de l'enfant. Nous verrons au point suivant que d'autres études, qui se sont cette fois focalisées à un niveau plus fin sur les productions verbales des dyades mère-enfant, ont également montré que l'interlocuteur soutient - par le biais du langage qu'il adresse à l'enfant - le processus d'acquisition du langage, et que pour sa part, l'enfant co-construit ses productions langagières en s'appuyant sur les offres de l'adulte.

### 2.3.2.1 Spécificités linguistiques de l'interaction adulte-enfant

En réponse aux arguments des comportementalistes considérant que le langage est acquis par simple imitation, et aux générativistes qui ont avancé que l'input fourni par l'adulte était trop imparfait et complexe pour que l'enfant apprenne à parler par le biais de l'interaction (*cf.* section 1.3 de ce chapitre), un grand nombre de travaux ont été menés pour connaître la nature du langage que les adultes - et en particulier, les mères occidentales - adressent aux enfants. Depuis les travaux fondateurs de Ferguson (1964, 1977), on sait que spontanément, les mères modifient, ajustent, leurs productions linguistiques de façon à les rendre plus accessibles aux enfants. Le langage adressé à l'enfant (LAE) fait l'objet de modifications sur différents niveaux langagiers et est adapté en fonction de l'âge de l'enfant et de ses capacités langagières. Certaines caractéristiques du LAE sont propres aux premières interactions entre adulte et enfant, et d'autres interviennent quand l'enfant commence à parler.

De nombreux travaux ont notamment montré que sur le plan de la phonologie et de la prosodie, le LAE présente des différences notables avec le langage adressé à l'adulte (entre autres, Ferguson, 1964 ; de Boysson-Bardies, 1996). Les résultats de ces études soulignent que le registre prosodique du LAE est plus haut que celui du langage adressé à l'adulte, et qu'il varie en fonction de l'âge de l'enfant : plus l'enfant est jeune, plus le registre est haut. Il a également été observé

que les mères tendent à exagérer les contours intonatifs, ainsi qu'à accentuer et allonger les mots à contenu sémantique. Le rythme de leurs productions est aussi plus lent et augmente progressivement au cours du développement de l'enfant. D'autres modifications, telles que des pauses plus longues en fin d'énoncés ou une plus grande intelligibilité du LAE, ont été mises en relief dans les études précédemment mentionnées, et là encore, il apparaît que ces caractéristiques évoluent selon l'âge de l'enfant, en diminuant progressivement lorsque l'âge et le développement langagier de l'enfant augmentent.

Dans le cadre d'une synthèse sur les caractéristiques du LAE, Rondal (1978) note que le lexique utilisé par les mères est moins complexe, moins diversifié et moins abstrait que celui du langage adressé à l'adulte. Le contenu sémantique du discours des mères est simplifié lui aussi et la nature des relations sémantiques qui sont exprimées dans leurs énoncés varie en fonction du niveau de développement linguistique des enfants (Glanzer et Todd, 1975).

Si on s'intéresse aux caractéristiques syntaxiques du LAE, on observe en particulier que celui-ci présente peu d'éléments non grammaticaux, comme le signale Rondal (1978) : « le langage maternel est très généralement bien formé grammaticalement ; souvent mieux formé, à ce point de vue, semble-t-il, que la majorité des échanges verbaux entre adultes familiers où l'ellipse, les interruptions de débit et autres modifications formelles en cours d'énoncé contribuent à rendre moins évidente la structure syntaxique du discours » (1978 : 76). Outre le caractère grammaticalement bien formé des énoncés des mères, on peut également constater un faible taux d'interruptions, d'hésitations ou de 'faux départs' (*ibid.*). Snow (1972) a par ailleurs montré que les énoncés des mères qui s'adressent aux enfants sont en moyenne plus courts et moins complexes. Il apparaît également que les énoncés produits sont majoritairement des énoncés interrogatifs ou déclaratifs, et rarement des impératifs (De Weck, 2010), mais que cette répartition évolue selon l'âge de l'enfant (Newport, Gleitman et Gleitman, 1977).

Enfin, les études qui se sont intéressées aux spécificités pragmatiques du LAE (*i.a.* Rondal, 1983) ont pu relever d'une part que celui-ci contient plus d'énoncés directifs (questions, demandes d'information, ordres, demandes de clarification ou de confirmation), et d'autre part, que sur le plan conversationnel, la mère réagit aux productions de l'enfant et lui apporte du feedback (entre autres, approbations ou désapprobations, répétitions et reformulations, expansions).

Le LAE a donc été largement observé sous l'angle de ses caractéristiques, mais une autre dimension a également fait l'objet de nombreux travaux, celle de ses effets sur le langage de l'enfant. Les recherches sur les corrélations entre certains traits formels du langage maternel et le niveau de développement langagier de l'enfant ont donné lieu à des résultats assez controversés. En revanche, les études qui se sont penchées sur la dimension conversationnelle du LAE ont fait le constat d'une relation positive entre certaines conduites de la mère, ainsi que la fréquence de ces conduites, et la longueur ou la complexité des énoncés enfantins.

A titre d'exemple, dans le travail de Veneziano (1999) précédemment mentionné (*cf.* point 2.2 de ce chapitre), l'expansion de l'adulte permet - lorsqu'elle suit l'intention communicative de l'enfant - de faciliter la production de suites d'énoncés à un mot en succession, et peut ainsi être considérée, parmi d'autres facteurs, comme un élément favorable à l'accès de l'enfant au stade des énoncés à deux termes. Il a également été montré (*cf.* notamment Clark et Chouinard, 2000 ; Leroy-Collombel, 2009) que les reformulations de l'adulte, outre le fait qu'elles permettent de demander confirmation, de vérifier ce que l'enfant a voulu dire, soutiennent le processus d'acquisition en apportant la forme conventionnelle de l'énoncé. Grâce aux reformulations, l'enfant peut saisir le contraste entre l'énoncé qu'il a produit et l'énoncé conventionnel, et ultérieurement, modifier explicitement sa production ou simplement, la prendre en compte tacitement. Il a été observé dans la littérature que l'adulte peut produire des reformulations pour tous les niveaux linguistiques (phonologique, morphologique, syntaxique), qu'il en produit davantage lorsqu'il interagit avec un très jeune enfant et que le taux de reformulations diminue à mesure que la fréquence d'énoncés erronés décroît.

L'enfant peut donc bénéficier de l'aide de l'adulte à un niveau global, en s'appuyant sur la structure générale de l'activité proposée par l'adulte (*cf.* point 2.3.1), mais il peut aussi profiter d'un soutien local, portant spécifiquement sur ses productions langagières. L'adulte peut appuyer le développement langagier en amont, par des demandes, des sollicitations, et également en aval, au travers de réactions d'accord, d'approbation, mais aussi de reprises, de reformulations - permettant des 'corrections' implicites - ou d'expansions.

Ces constats concordent avec les travaux de Vygotski ou Bruner que nous évoquons plus haut et qui ont tous deux avancé que le processus de développement de l'enfant, et notamment langagier, repose sur la médiation de l'adulte. Grâce à cette relation asymétrique où l'autre - qu'il

s'agisse d'un adulte ou d'un enfant plus âgé - est plus expert, l'enfant peut avancer dans son développement en faisant davantage qu'il ne ferait s'il était seul face à une tâche. Au sein de cette relation de tutelle, l'adulte peut notamment aider l'enfant par le biais de ce que Bruner appelle 'l'étaillage', qui consiste pour lui « à prendre en mains ceux des éléments de la tâche qui excèdent initialement les capacités du débutant, lui permettant ainsi de concentrer ses efforts sur les seuls éléments qui demeurent dans son domaine de compétence et de les mener à terme » (1983 : 263).

Cette idée, au cœur de la théorie interactionniste, est déjà présente dans les travaux de Vygotski (1997[1936]), pour qui l'ensemble des tâches que l'enfant n'arrive pas à faire seul mais qu'il parvient à faire avec la médiation d'un adulte détermine ce qu'il appelle la 'zone proximale de développement' de l'enfant (ZPD). Reprise ultérieurement par Bruner, cette zone est définie comme « la distance entre le niveau de développement actuel tel qu'on peut le déterminer à travers la façon dont l'enfant résout des problèmes seul et le niveau de développement potentiel tel qu'on peut le déterminer à travers la façon dont l'enfant résout des problèmes lorsqu'il est assisté par l'adulte ou collabore avec d'autres enfants plus avancés » (1983 : 287). La ZPD doit donc être distinguée du niveau actuel de développement de l'enfant - correspondant à ce que celui-ci est capable de faire seul - et les interventions de l'adulte doivent de préférence se situer dans cette zone s'il veut contribuer au développement de l'enfant.

En effet, l'étaillage n'est efficace qu'à condition de cibler des apprentissages qui puissent être accessibles pour l'enfant, et donc, dans les limites de ce que son développement cognitif et/ou langagier lui permettent. Si la tâche proposée est trop complexe, l'enfant ne pourra pas tirer profit de l'étaillage de l'adulte. A l'inverse, si elle porte exclusivement sur des compétences déjà acquises, le développement de l'enfant ne sera pas non plus encouragé. L'adulte doit donc adapter son étaillage en fonction des acquis de l'enfant et à chaque nouvelle étape, restreindre l'aide qu'il lui apportait à l'étape précédente, afin de lui permettre à terme de réaliser la tâche de façon autonome.

Le rapide tour d'horizon que nous venons de faire des principaux concepts fondateurs de la théorie interactionniste nous a permis d'entrevoir le rôle central qui y est attribué au partenaire de l'enfant, aussi bien dans le cadre global des activités entre l'adulte et l'enfant que dans celui plus spécifique des échanges verbaux au sein de la dyade. Nous consacrerons le point suivant à une revue des travaux - menés principalement dans une perspective basée sur l'usage - qui se sont



intéressés à la façon dont l'enfant puise dans son expérience communicative pour apprendre à parler, et en particulier, à la question de la dépendance entre input et usages langagiers de l'enfant. Nous verrons ainsi que pour certains auteurs (*i.a.* Tomasello, 2000b), le développement du langage consiste en une acquisition fragmentée des formes et des structures syntaxiques fréquemment produites dans l'input de l'enfant.

### 2.2.3 Des usages langagiers dépendants de l'expérience communicative

Nous avons vu que parmi les approches qui considèrent l'acquisition du langage comme un phénomène qui se construit au cœur de l'interaction et des échanges avec l'autre, certains travaux se sont particulièrement intéressés aux caractéristiques étayantes des interactions adulte-enfant (*cf.* sections précédentes). D'autres ont davantage cherché à expliquer comment l'enfant acquiert ses premiers mots et ses premières constructions syntaxiques. Les études menées dans une perspective basée sur l'usage ont notamment mis en relation la nature et la fréquence des items lexicaux et syntaxiques produits par l'adulte et les productions de l'enfant, des premiers énoncés à un terme à la production systématique d'énoncés de plusieurs termes conformes à la grammaire adulte.

A l'inverse du courant générativiste, le courant usage-based envisage la construction de la grammaire comme un produit de l'utilisation du langage. Pour Tomasello, l'enfant n'apprend pas explicitement, il agit, interagit, et apprend dans l'interaction. L'acquisition des mots aussi bien que de la grammaire serait basée sur l'usage : la fonction communicative des mots serait appréhendée en discours et les structures de la langue seraient extraites des modèles ('patterns') d'énoncés à plusieurs termes produits dans la langue adulte.

Nous l'avons déjà signalé (*cf.* point 2.3.2.1), l'acquisition du lexique s'expliquerait, dans une perspective usage-based, par la capacité de l'enfant à interpréter l'intention communicative de son partenaire au travers d'épisodes d'attention conjointe. En comprenant l'intention de son interlocuteur, l'enfant peut se saisir des moyens linguistiques pour catégoriser les objets. Le stock lexical de l'enfant dépendrait donc du temps préalablement passé dans des scènes d'attention conjointe (*cf.* étude de Tomasello et Todd, 1983 déjà citée ou Tomasello, Mannle et Kruger, 1986). L'apprentissage du lexique reposant sur cette compréhension *in situ* de l'intentionnalité de

l'adulte, il serait dans un premier temps fortement ancré dans la situation de communication, avant de pouvoir être transposé à d'autres contextes.

De nombreux travaux menés dans une perspective basée sur l'usage se sont par ailleurs intéressés à l'acquisition de la syntaxe et ont notamment signalé que l'acquisition des constructions grammaticales semble très clairement corrélée aux énoncés entendus par l'enfant. Notons que la notion de 'construction', à la base de la théorie usage-based de l'acquisition du langage est définie par Tomasello comme « a unit of language that comprises multiple linguistic elements used together for a relatively coherent communicative function, with sub-functions, being performed by the elements as well » (Tomasello, 2003 : 100) et que cette notion rassemble des structures linguistiques variables et qui ne possèdent pas toutes le même niveau de complexité et/ou d'abstraction. Dans une étude de 2003, Cameron-Faulkner *et al.* ont ainsi fait l'analyse du langage que 12 mères anglophones adressaient à leur enfant de 2 ans et ont montré entre autres, que plus de la moitié des énoncés des mères commençait par une des 52 expressions référencées comme les plus fréquentes par les auteurs (et le plus souvent composées de deux mots ou morphèmes). Par ailleurs, environ 45% de ces énoncés commencent par un des 17 mots les plus couramment utilisés, parmi lesquels on trouve par exemple 'what', 'that', 'it' ou 'you'. A propos de cette étude, Tomasello remarque donc que « many, indeed the majority, of the utterances children hear are grounded in highly repetitive item-based frames that they experience dozens, in some cases hundreds of times everyday » (2003 : 112). Cameron-Faulkner *et al.* soulignent également que les productions des enfants reflètent largement celles des mères, dans la mesure où les enfants produisent en grande partie les mêmes expressions que celles relevées dans le langage des mères, et dans certains cas, dans des proportions largement comparables.

Au cours de ses interactions avec son environnement, l'enfant saisisrait donc des éléments lexico-syntaxiques récurrents chez l'adulte, qu'il pourrait ensuite réutiliser dans ses propres productions. Dans un premier temps, ces productions se limitent à des 'holophrases', autrement dit, des constructions dont les éléments, s'ils sont considérés de façon distincte, ne sont pas transposés à d'autres contextes. Certains des éléments du langage adulte sont repris par l'enfant mais de façon figée (exemples : 'I-wanna-do-it', 'lemme-see', 'where-the-bottle'). Pour Wittek et Tomasello (2005) par exemple, le marquage syntaxique précoce est à envisager comme circonscrit à certains verbes, avec lesquels les constructions sont apprises une par une. L'enfant peut exprimer - par

des éléments morphologiques ou par l'ordre des mots - les rôles des participants dans une situation donnée mais ne généraliserait pas ces marquages à des situations similaires sans que cela n'ait été mentionné antérieurement dans le discours de l'adulte. L'enfant construirait donc des îlots proto-syntaxiques liés à chaque verbe, et non transférables d'un verbe à l'autre.

Dans les étapes ultérieures, l'enfant devient capable d'analyser ces figements en éléments distincts, réutilisables dans d'autres énoncés et d'autres contextes linguistiques. Un élément  $x$  d'abord associé de façon spécifique à un élément  $y$  peut alors être transposé à une autre construction  $x+z$ . L'enfant apprend progressivement à étendre l'usage qu'il fait de ces unités. La portée de ces usages est d'abord limitée à quelques combinaisons avant que ne soient créées de véritables constructions productives et abstraites. Signalons toutefois que la mise en place de la généralisation et de l'abstraction peut être relevée à des moments très variés d'une construction à l'autre, selon la nature du langage entendu par l'enfant :

« Children construct these abstractions only gradually and in piecemeal fashion, with some categories and constructions appearing much before others that are of a similar type from an adult perspective - due quite often to differences in the language that individual children hear » (Tomasello, 2006 : 258).

L'étude de Cameron-Faulkner *et al.* (2003) que nous venons de citer met en avant la correspondance entre la forme des structures récurrentes chez l'adulte et la forme des productions de l'enfant. Pour Tomasello, il existerait par ailleurs un autre type de correspondance entre les énoncés de l'adulte et ceux de l'enfant, reposant cette fois sur la fonction des énoncés. Bien que produites comme un tout figé, les premières constructions de l'enfant seraient porteuses d'une intention communicative spécifique, semblable à celle des productions de l'adulte. Au même titre que d'autres approches précédemment citées (*cf.* point 2.3.2.1), l'approche usage-based considère que l'enfant apprend à utiliser le langage au travers d'actes de communication, plus que par (et pour) l'apprentissage strict de mots isolés que l'enfant combine ensuite entre eux au moyen de règles linguistiques. L'apprentissage des mots et de la grammaire correspondrait donc à l'extraction des formes présentes dans le discours de l'adulte, mais aussi et surtout à l'extraction de la valeur communicative de chacun de ses composants. Dans cette perspective, l'acquisition du langage repose sur sa dimension symbolique et l'acquisition de la grammaire n'en est que dérivée. La compétence linguistique adulte est alors conçue comme « a structured inventory of meaningful linguistic constructions - including both the more regular and

the more idiomatic structures in a given language (and all structures in between) » (Tomasello, 2006 : 258).

Rappelons enfin que dans la théorie usage-based, l'acquisition du langage repose sur des processus cognitifs généraux qui permettent à l'enfant qui entend des énoncés individuels de (re)construire ensuite les catégories abstraites de la langue. Tomasello (*i.a.* 2009) regroupe ces processus en deux catégories : les éléments socio-cognitifs liés à ce qu'il appelle 'l'intention-reading', propres à l'espèce humaine et donnant accès à l'enfant aux symboles linguistiques et à la dimension fonctionnelle du langage (notamment, disposition à l'attention conjointe et compréhension des intentions communicatives de l'autre), et les éléments liés au 'pattern-finding', communs cette fois à tous les primates et concentrant les compétences cognitives impliquées dans le processus d'abstraction (entre autres, catégorisation, construction de schémas, capacité d'analogie).

Pour les travaux qui considèrent l'acquisition du langage comme un produit de l'usage, l'enfant développe donc ses compétences langagières grâce à des processus cognitifs mais aussi grâce au langage qui lui est adressé. L'acquisition du lexique est envisagée comme dépendante des épisodes d'attention conjointe entre adulte et enfant, et l'acquisition de la syntaxe - ou plus précisément, l'acquisition des constructions lexico-syntaxiques - s'appuierait sur les structures présentes dans l'input. La perspective usage-based considère donc que l'acquisition des formes et des fonctions du langage est ancrée dans l'expérience communicative de l'enfant, mais se démarque de l'approche interactionniste, dans la mesure où elle se focalise davantage sur l'impact de la fréquence et de la distribution des formes dans le langage adressé à l'enfant que sur le rôle du dialogue et des conduites langagières de l'adulte.

Pour conclure l'inventaire que nous avons cherché à établir dans cette seconde partie du chapitre, nous pouvons donc dire que les hypothèses et conclusions innéistes ont été largement remises en question, par un grand nombre de travaux ayant mis en avant la dimension progressive de l'acquisition du langage. Cet état des lieux nous a également permis d'observer l'hétérogénéité des approches référencées comme 'non innéistes'. Nous aurons pu remarquer entre autres, que celles-ci se distribuent selon la place qu'elles accordent aux facteurs cognitifs et à l'interaction dans le processus d'acquisition du langage. Pour certaines, le langage s'acquiert progressivement et à condition que l'enfant soit inséré dans un environnement linguistique, mais l'apprentissage

langagier repose essentiellement sur des mécanismes cognitifs (*cf.* perspective constructiviste piagétienne ou fonctionnaliste non interactionniste) ; pour d'autres, l'acquisition du langage repose avant tout sur les échanges de l'enfant avec son entourage et celui-ci est conçu comme le pilier central du développement de l'enfant (*cf.* approche interactionniste) ; et pour d'autres encore, l'enfant apprendrait à parler à la fois grâce à des processus cognitifs et à l'interaction, mais davantage pour ce que celle-ci représente - au travers de l'input - pour l'apport des formes et du sens que pour le caractère indispensable du dialogue et des conduites étayantes de l'interlocuteur (*cf.* approche usage-based).

Notre travail s'intéressant à l'acquisition de la catégorie 'déterminant', nous nous pencherons dans la troisième partie de ce chapitre sur le traitement de l'acquisition des catégories syntaxiques dans la littérature. Nous limiterons notre exposé à deux approches au sein desquelles cette question a été la plus documentée : l'approche innéiste et l'approche usage-based.

### **3. Des perspectives divergentes sur l'acquisition des catégories syntaxiques**

Les deux perspectives majeures sur l'acquisition du langage que nous venons de présenter, considérant pour l'une que le langage est acquis de façon innée chez l'enfant, et pour l'autre que cette acquisition est progressive, conduisent nécessairement à ce que soient développées dans la littérature des conclusions divergentes sur l'acquisition des catégories syntaxiques, et notamment des morphèmes grammaticaux (déterminants, pronoms, auxiliaires, flexions, etc). Le cadre théorique du travail des chercheurs, qu'il soit ancré par exemple dans une conception générativiste ou basée sur l'usage, se reflète dans les études consacrées spécifiquement à l'acquisition des éléments qui constituent la grammaire de l'enfant, et de façon complémentaire, ces études viennent étayer le modèle global d'acquisition du langage soutenu par les auteurs.

Nous aborderons dans les deux sections qui suivent certains des travaux ayant traité cette question de l'acquisition des catégories fonctionnelles chez l'enfant, et conduits d'une part dans une perspective innéiste (section 3.1), et d'autre part, dans une perspective basée sur l'usage (section 3.2).

### 3.1 Perspective innéiste

Les différents travaux menés dans le cadre de la grammaire générative se distribuent en deux hypothèses majeures, qui se distinguent l'une de l'autre par le fait de considérer qu'il existe ou non une continuité entre les catégories abstraites de la grammaire adulte (lexicales et grammaticales ou fonctionnelles) et celles de l'enfant<sup>2</sup>.

Dans la première hypothèse, considérant qu'il y a continuité entre la grammaire de l'adulte et la grammaire de l'enfant, il n'y aurait pas de véritable différence entre les catégories syntaxiques de l'adulte et celles du langage de l'enfant. Qu'il s'agisse des catégories lexicales ou grammaticales, elles seraient acquises d'emblée (et donc présentes dans la grammaire de l'enfant dès le début de l'acquisition) et les omissions d'éléments grammaticaux seraient liées à des questions de performance ou de sous-spécification de certaines catégories.

On trouve parmi les travaux de chercheurs adeptes de cette hypothèse ceux de Bloom (1990) sur la production d'énoncés sans sujet en anglais, ceux de Demuth (1992) sur l'acquisition des catégories fonctionnelles en Sesotho, ou encore ceux de Hoekstra et Hyams (1995) sur les infinitifs indépendants en néerlandais et Valian (1986, 2009) sur l'acquisition de la catégorie 'déterminant' en anglais (nous reviendrons aux travaux de Valian ultérieurement dans cette thèse, cf. section 3.2 du deuxième chapitre). Dans le cadre de ces études, les auteurs considèrent que le développement de la syntaxe est indépendant du développement des autres domaines du langage (contrairement aux hypothèses établies dans les théories du bootstrapping présentées précédemment), et que la structure qui constitue le noyau de la Grammaire Universelle est présente dès le début du processus d'acquisition.

Dans l'étude de Bloom (1990), les énoncés contenant un sujet seraient construits avec des locutions verbales significativement plus courtes que ceux au sein desquels le sujet est omis, ce qui permet à l'auteur d'avancer que l'enfant omet le sujet lorsque l'énoncé est trop long. Les différences entre la grammaire adulte et celle de l'enfant s'expliqueraient donc pour l'auteur par des limitations cognitives. Dans le cas du travail de Demuth (1992), les catégories fonctionnelles

---

<sup>2</sup> Nous reprenons ici en la simplifiant la typologie théorique établie par Granfeldt (2003) dans le cadre de sa thèse. Notons que si cette présentation nous semble pertinente pour dresser un inventaire rapide des travaux menés dans une perspective générativiste, nous nous devons de signaler, à la suite de l'auteur, qu'elle simplifie quelque peu le tableau - plus complexe - des recherches qui ont été menées sur l'acquisition des catégories fonctionnelles.

seraient présentes dans le langage de l'enfant, et certaines seraient, de façon précoce, pleinement utilisées, tandis que d'autres ne seraient pas réalisées ou phonétiquement sous-spécifiées pendant une certaine période. Ces conclusions sont partagées par Hoekstra et Hyams (1995), qui considèrent également que certains traits seraient sous-spécifiés dans les débuts de l'acquisition, et que dans le cas précis de leur étude, une fois le trait approprié acquis (le trait 'Nombre'), le nombre d'énoncés contenant des infinitifs indépendants comme seuls éléments verbaux diminue. Bien que cette hypothèse considère que l'enfant a accès aux mêmes catégories que l'adulte, il peut donc être envisagé que l'enfant ne fasse pas d'emblée l'exploitation complète de la grammaire à sa disposition.

Si nous ne faisons ici mention que d'une seule hypothèse de continuité, il est à noter toutefois que les chercheurs ayant adhéré à cette hypothèse n'ont pas tous défini le même degré de continuité entre la grammaire de l'enfant et celle de l'adulte. A titre d'exemple, les travaux de Bloom ou de Valian que nous venons de citer revendiquent l'existence d'une continuité forte entre grammaire enfantine et grammaire adulte, tandis que Demuth ou Hoekstra et Hyams considèrent plutôt pour leur part que cette continuité entre les deux est faible.

Dans la seconde hypothèse, dite de 'discontinuité', représentée par des chercheurs comme Radford (notamment, 1990), il n'y aurait pas de continuité entre la grammaire adulte et celle de l'enfant, le premier système de catégories grammaticales de l'enfant étant construit sur des principes distincts de celui de l'adulte. Dans une première étape du développement langagier, dite pré-catégorielle, les catégories lexicales et fonctionnelles seraient absentes du système de l'enfant (pendant la période de production des énoncés à un terme) ; dans une seconde étape - au stade des énoncés à deux termes - apparaîtraient les catégories lexicales et dans un dernier temps, les catégories fonctionnelles (au stade des énoncés à plus de deux termes). D'après l'auteur, le développement du langage serait conditionné par ces différentes étapes génétiquement déterminées de 'maturation' de la Grammaire Universelle. On retrouve cette notion de maturation dans un travail préalable de Borer et Wexler (1987), partisans quant à eux d'une continuité faible entre grammaire de l'enfant et grammaire cible adulte.

Les différences entre les productions de l'adulte et celles de l'enfant s'expliqueraient ainsi par l'absence des catégories fonctionnelles dans la grammaire précoce de l'enfant. Radford (1990)

précise par ailleurs que les formes observées dans les productions de l'enfant de moins de deux ans dans des positions de morphèmes fonctionnels ne seraient que des 'imposteurs'.

### **3.2 Perspective basée sur l'usage**

Nous avons vu plus haut (*cf.* section 2.3.3) que l'approche usage-based considère que pour la plupart des enfants, les premières productions sont composées d'éléments isolés ('item based') et que le développement langagier consiste en une acquisition parcellaire des formes et des structures syntaxiques de la langue. Sur la base de ce qu'il entend, analyse et stocke, l'enfant construit graduellement les structures abstraites de la langue adulte et peut, après une période marquée d'abord par la production d'holophrases, puis par celle de schémas lexicaux et syntaxiques spécifiques, généraliser les contextes dans lesquels il utilise les différentes unités de sa langue.

Parallèlement aux constructions linguistiques, les catégories syntaxiques deviennent elles aussi de plus en plus abstraites. En s'appuyant sur les exemples fournis dans son input et sur sa compréhension de la fonction communicative des unités produites dans les énoncés de ses interlocuteurs, l'enfant peut constituer un ensemble de catégories de mots remplissant la même fonction. Contrairement à la perspective générativiste, l'approche usage-based n'envisage donc pas que l'enfant possède dès sa naissance une connaissance et une compréhension des catégories syntaxiques, mais bien plutôt que ces catégories se construisent petit à petit, grâce au langage que l'enfant entend et qui lui est adressé. D'après Tomasello (2000a), l'hypothèse générativiste d'une continuité entre la grammaire de l'adulte et la grammaire de l'enfant n'a d'ailleurs jamais été confirmée par une base empirique solide, et à l'épreuve des données (expérimentales et naturelles), cette hypothèse d'un système inné de catégories syntaxiques ne peut être que contestée, au profit d'une alternative constructiviste, plus à même de rendre compte de façon réaliste de l'acquisition du langage et de la grammaire chez l'enfant.

Dans une revue de littérature récente sur l'acquisition des catégories syntaxiques, Ambridge (sous presse) soutient lui aussi l'idée que les approches innéistes n'expliquent que de façon lacunaire l'émergence des catégories syntaxiques. D'après lui, considérer que ces catégories sont présentes d'emblée chez l'enfant peut, en apparence, résoudre le problème de leur acquisition mais ne fait en réalité que remplacer une question épineuse par une autre : comment les enfants identifient-ils



les occurrences de ces catégories dans leur input ? Il ajoute par ailleurs une autre dimension au débat opposant générativistes et constructivistes en proposant que si les catégories syntaxiques ne sont en effet pas innées, elles ne correspondraient pas non plus aux catégories induites de l'input que décrivent les chercheurs non innéistes, et il serait même inutile de considérer l'existence chez les locuteurs de catégories syntaxiques abstraites.

Selon Ambridge, les explications apportées par les travaux constructivistes - et reposant pour la plupart sur la capacité de l'enfant à faire l'analyse des régularités distributionnelles des unités de la langue - envisagent que les catégories absentes dans les premiers temps du développement langagier sont *in fine*, dans la grammaire adulte, les mêmes catégories syntaxiques traditionnelles que celles qui sont présupposées dans la littérature innéiste. Or, ces catégories n'existeraient pas pour lui en tant que telles dans la grammaire des locuteurs. Ceux-ci produisent et comprennent des énoncés qu'ils n'ont encore jamais entendus, en généralisant les exemplaires de mots, syntagmes ou énoncés qu'ils ont stockés (de même que leur signification), mais « ne forment en aucun cas une abstraction isolée correspondant à une catégorie telle que le nom ou le verbe » (sous presse, *notre traduction*). Si, ainsi formulée, l'idée que les catégories syntaxiques n'existent pas semble particulièrement innovante dans le champ des études constructivistes, Ambridge signale néanmoins qu'il se peut que pour certains auteurs, le recours à des étiquettes comme 'nom' et 'verbe' reflète davantage un besoin de raccourci commode qu'un véritable engagement théorique. Il relève par ailleurs que chez Tomasello (2003), qui considère que l'enfant ne fait pas une analyse exclusivement basée sur la distribution des formes mais également sur leur fonction ('functionally based distributional analysis'), la question des catégories paradigmatiques a déjà été discutée, en termes de regroupements fonctionnels plus que d'adéquation à une catégorisation linguistique traditionnelle, comme en témoignent les lignes qui suivent :

« These categories [catégories paradigmatiques telles que 'nom' et 'verbe'] are formed through a process of functionally based distributional analysis in which concrete linguistic items (such as words or phrases) that serve the same communicative functions in utterances and constructions over time are grouped together into a category » (2003 : 301).

Si Ambridge a récemment consacré un travail spécifique à l'acquisition des catégories syntaxiques, cette question est déjà partiellement traitée dans le chapitre de Ambridge et Lieven sur le constructivisme, paru en 2015, et conformément à ce que nous venons de voir, les auteurs

ne considèrent pas là non plus qu'il puisse être utile de regrouper sous l'étiquette d'une même catégorie tous les éléments dont les propriétés sont semblables :

« there is no need for learners to construct either conventional Universal Grammar style grammatical categories such as [NOUN] and [DETERMINER], or even finer-grained language-specific, input-based categories such as [ENGLISH MASS NOUN] or [ENGLISH TRANSITIVE VERB]. The reason is that, provided that the child stores the semantic, phonological (etc.) properties of every word and every construction slot, which she will have to do anyway, she already knows which words can be used in any given construction : the words whose properties overlap sufficiently with those of the construction. There is nothing to be gained by grouping together the words that can act as fillers for the same slot, and calling them a “category” » (2015 : 483).

Nous pouvons donc déduire de ce rapide état des lieux que pour la plupart des chercheurs ancrés dans un cadre théorique constructiviste et basé sur l'usage, l'acquisition des catégories syntaxiques est graduelle, non innée, et rendue possible par l'apport de l'input. Toutefois, l'existence même de regroupements en catégories syntaxiques traditionnelles que l'enfant induirait de son environnement langagier est remise en question dans de récents travaux.

Pour clore ce premier chapitre, nous consacrerons la section qui suit (section 4) à la présentation de quelques remarques concernant d'une part l'ancrage théorique global de cette thèse, et d'autre part, à la perspective qui sera spécifiquement adoptée pour traiter d'une question récurrente quand on s'intéresse à l'acquisition du langage, celle de la variation. Nous présenterons donc d'abord brièvement l'approche théorique que nous avons retenue pour cette étude du développement de la catégorie 'déterminant' ; et nous verrons ensuite (aux points 4.2 et 4.3) qu'au vu de la spécificité de la langue et du contexte culturel étudiés, seule la variation individuelle - des enfants et de leurs interlocuteurs - pourra être véritablement prise en compte.

## **4. Perspective adoptée pour approcher les données**

### **4.1 Approche théorique retenue pour l'étude du développement de la catégorie 'déterminant'**

A la lecture de nos questions de recherche (chapitre III) et de notre méthodologie d'analyse (chapitre IV), le lecteur pourra constater que nous avons fait le choix dans cette thèse de mettre en perspective des approches méthodologiques et des questionnements de chercheurs issus de

courants théoriques variés ; toutefois, l'organisation de ces questionnements et analyses repose sur une approche spécifique du processus d'acquisition du langage chez l'enfant.

Contrairement à la perspective générativiste présentée dans la première partie de ce chapitre, nous considérons, à la suite des travaux fondateurs de constructivistes comme Piaget (notamment, 1966[1923]) ou Sinclair (*i.a.* 1967), que la faculté de langage n'est pas innée, qu'elle se développe, se construit progressivement, en s'appuyant sur des mécanismes cognitifs de base. Si nous estimons que l'enfant a un rôle considérable à jouer dans ce processus, nous ne négligeons pas, comme a pu le faire l'école de Genève dans ses premiers travaux, le rôle joué par la dimension sociale du langage. Dans la lignée des travaux de psychologues et de linguistes comme Bruner (*i.a.* 1987), Vygotski (notamment, 1997[1936]) ou plus récemment, Tomasello (*i.a.* 2004), nous envisageons l'acquisition du langage comme un processus social, au sein duquel l'enfant apprend à parler non pas pour apprendre des sons, des mots, ou des structures syntaxiques, mais bien pour pouvoir communiquer, interagir avec l'autre. Dans cette perspective, l'adulte, ou plus généralement, l'autre plus expert, est le médiateur des acquisitions de l'enfant, il est le pilier grâce auquel l'enfant s'approprie la langue de son environnement.

A ce titre, l'acquisition de la catégorie 'déterminant' n'est pas considérée dans ce travail comme innée mais graduelle, et jalonnée d'étapes de transition avant que les formes adultes ne soient pleinement maîtrisées, tant sur le plan formel que fonctionnel. Comme d'autres morphèmes grammaticaux, le déterminant n'est pas, d'après nous, la manifestation d'une simple familiarisation de l'enfant avec la grammaire de sa langue, ou l'indice de sa progression dans la maîtrise d'un module abstrait et exclusivement formel de la langue. Au contraire, au-delà de l'éclairage qu'il nous apporte sur le développement de la dimension formelle du langage, le déterminant - ainsi que ses précurseurs - nous informent sur sa dimension sociale, sur la façon dont l'enfant construit sa grammaire en prenant en considération les échanges auxquels il prend part et en s'appuyant sur le langage qui lui est adressé. Dans cette perspective, nous n'appréhenderons pas le développement du déterminant comme un parcours que l'enfant emprunte seul et basé uniquement sur des facteurs formels, mais comme une co-construction entre l'enfant et l'adulte, ou l'ainé, avec qui il interagit, et dépendant également de facteurs interactionnels. Ainsi, notre étude sera ancrée dans un cadre non seulement constructiviste, mais aussi basé sur l'usage que l'enfant et son entourage font des déterminants. Nous garderons

toutefois à l'esprit que si le corpus sur lequel s'appuie notre analyse nous permet d'entrevoir les usages des participants, ils ne restent qu'un reflet partiel de leurs pratiques langagières quotidiennes, et des formats d'interaction dans lesquels les enfants sont quotidiennement impliqués, avec un échantillon beaucoup plus vaste d'interlocuteurs, qui conditionnent au même titre que les échanges avec les interlocuteurs que nous avons filmés, les usages que les enfants font de la langue.

## **4.2 Une étude linguistiquement et culturellement située**

Nous savons, d'après des études comparatives telles que celle de Choi et Gopnik (1995), que la structure de la langue peut influencer la nature des premiers mots. Dans le cadre de ces travaux, les auteurs ont par exemple comparé le développement du premier lexique chez des enfants en cours d'acquisition de l'anglais - langue pour laquelle la structure canonique est Sujet-Verbe-Objet - et des enfants en cours d'acquisition du coréen - qui, contrairement à l'anglais, est une langue avec une structure de type Sujet-Objet-Verbe. Ils ont montré que de par leur position plus saillante en coréen qu'en anglais, les verbes sont présents en plus grande proportion au sein des premiers mots des enfants coréens qu'au sein du premier lexique des enfants anglophones. De nombreux autres travaux sur le développement du lexique ou de la grammaire ont fait depuis des constats semblables ; et l'acquisition des déterminants n'échappant pas à cette question de la variation inter-langues, nous verrons au chapitre II que, comme le lexique, elle peut faire l'objet de variations importantes d'une langue à l'autre. Or, nous avons fait le choix dans ce travail de ne nous focaliser que sur l'émergence du déterminant en français. De fait, si nous pourrions avoir recours à des illustrations de travaux menés sur d'autres langues, dont la structure peut être plus ou moins éloignée de celle du français, nous ne dépasserons pas dans la discussion de nos résultats les frontières imposées par la langue de nos locuteurs, le français.

Il nous semble par ailleurs important de préciser que notre travail est ancré dans une perspective culturelle spécifique, que l'on peut de façon quelque peu simpliste qualifier 'd'occidentale'. Dans les familles que nous avons suivies, les parents considèrent que l'enfant est un interlocuteur à part entière, et cela même durant les premières étapes de son acquisition langagière, ils s'engagent avec lui dans des interactions individuelles et s'ajustent à son niveau de compétence. Or, on sait depuis notamment les travaux de Schieffelin et Ochs (1986) ou Rabain-Jamin (1998) que des

facteurs culturels, comme le format d'interaction et la nature du langage adressé à l'enfant, peuvent influencer l'entrée de l'enfant dans le langage et entraîner une chronologie différente du développement de ses compétences communicationnelles. Dès lors, nous appréhenderons notre étude et le rôle attribué à l'interaction dans ce qu'ils ont de culturellement situés.

### 4.3 Une place faite à la variation individuelle

Si au vu de notre étude il nous est difficile de pouvoir véritablement prendre en compte la variation liée à la diversité des langues et des formats de socialisation langagière, nous accorderons cependant une attention particulière à la variation individuelle.

Un grand nombre d'études, parmi lesquelles celles de Bates *et al.* (1995) sur le lexique ou d'autres consacrées spécifiquement à l'acquisition des déterminants (*i.a.* Bassano et Eme, 2001) et sur lesquelles nous reviendrons ultérieurement, ont montré une grande diversité dans les rythmes d'acquisition des enfants (et particulièrement à partir de 12-13 mois dans l'étude de Bates *et al.*).

Nous savons par ailleurs qu'outre les différences de rythme d'acquisition, les enfants peuvent acquérir la langue par des biais distincts, autrement dit, avec des styles ou des stratégies qui varient d'un enfant à l'autre. Les recherches sur la variation individuelle ont notamment mis en évidence que le plus couramment, l'enfant entre dans le langage par une stratégie dite 'analytique', en réduisant ses productions à une seule unité, parfois figée, et combine ensuite ces mots isolés en séquences de deux mots puis construit peu à peu des énoncés conformes à la cible adulte, comprenant des mots grammaticaux. A l'inverse, pour les enfants dont la stratégie est dite 'globale', le développement ne va pas d'une production de mots isolés à une complexification mais suit plutôt une progression allant d'un ensemble d'éléments non analysés à la combinaison significative des unités de la langue (Bates *et al.*, 1988 ; Nelson, 1981). Jisa et Richaud (1994) signalent toutefois qu'« on ne constate pas chez l'enfant l'utilisation exclusive d'une stratégie globale ou analytique, aux dépens de l'autre, mais plutôt la dominance de l'une ou de l'autre » (1994 : 10).

On trouve d'autres différences de style entre les enfants, en ce qui concerne notamment la nature des premiers mots produits, lors de l'acquisition d'une même langue (Nelson, 1973). Certains enfants font leurs premiers pas dans le langage à travers des noms communs qui désignent le plus souvent des objets de leur environnement, et pour d'autres, la liste des premiers mots est

plus hétérogène. Les noms restent majoritaires, mais s'ajoutent aux noms des termes à valeur plus sociale ou expressive, que l'on pourrait associer à d'autres catégories grammaticales et des structures figées, produites comme un tout.

Notre objectif ici n'étant pas de faire l'état des lieux complet des travaux sur le phénomène de variation individuelle, nous nous en tiendrons à ces quelques exemples illustrant bien la nécessité de considérer chaque enfant comme potentiellement doté d'un rythme d'acquisition et de stratégies individuelles, et parfois non comparables au développement prototypique observé dans la littérature. Si ces différences ont un impact sur la quantité et la nature du premier lexique produit par l'enfant, elles peuvent également influencer le développement des morphèmes grammaticaux. L'analyse de nos données tendra donc autant que possible à ne pas généraliser nos constats et à prendre en compte cette diversité, de rythme comme de style.

Signalons enfin que les enfants ne sont pas les seuls à faire l'objet de variation individuelle. Snow (1977) a par exemple montré qu'il existe des styles de discours variables d'une mère à l'autre, certaines ayant pour objectif de solliciter l'enfant et de l'impliquer dans l'échange en cours, et d'autres ayant davantage pour but de contrôler, de réguler son activité. Tomasello et Farrar (1986) ont également montré que certaines mères tendent à suivre l'intérêt de leur enfant lorsqu'elles échangent avec lui, en parlant d'objets déjà au centre de son attention, et d'autres privilégient l'apport d'éléments nouveaux, en parlant d'objets n'ayant pas encore attiré l'attention de l'enfant. Les auteurs ont pu constater que les enfants dont les mères ont un profil semblable au premier, et qui donc suivent l'attention de l'enfant, possèdent un lexique plus vaste à 21 mois que les enfants dont les mères déplacent le centre d'attention.

Ces résultats montrent bien que les différences de style maternel peuvent influencer le développement du langage de l'enfant ; ils font donc partie des facteurs à considérer lors de l'analyse des données, et en particulier, lors de l'observation de résultats hétérogènes, voire contradictoires, d'un enfant à l'autre.



Ce premier chapitre nous a permis de revenir sur les deux approches théoriques majeures qui s'opposent dans le champ de l'acquisition du langage, considérant pour l'une que la mise en place

du langage chez l'enfant est une faculté innée, et pour l'autre qu'au contraire, elle se construit progressivement. Nous nous sommes par ailleurs arrêtée brièvement sur la perspective adoptée par ces deux approches sur l'acquisition spécifique des catégories syntaxiques, et nous avons clos ce chapitre par quelques remarques concernant le cadre théorique dans lequel cette étude est ancrée et le traitement qui sera fait de la variation individuelle. Dans le chapitre qui suit, nous traiterons spécifiquement des travaux portant sur la catégorie 'déterminant', en abordant dans un premier temps les caractéristiques de la cible adulte, et dans un second temps, le développement du paradigme chez l'enfant.





## CHAPITRE II

*L'émergence de la catégorie  
'déterminant' chez l'enfant*

---



**L**e chapitre précédent - consacré aux différentes approches théoriques qui constituent le domaine de l'acquisition du langage - nous a permis de constater que l'origine et la nature des catégories syntaxiques dans le langage précoce de l'enfant sont au cœur des recherches et des débats théoriques. En effet, l'acquisition des premiers morphèmes - tant grammaticaux que lexicaux - a fait et continue de faire l'objet de nombreuses études, portant sur la chronologie et le développement de chacune de ces catégories, dans une langue donnée ou dans une perspective inter-langues.

En français, comme dans d'autres langues, les chercheurs se sont pour beaucoup concentrés sur l'acquisition des contraintes grammaticales relatives à l'acquisition des deux catégories centrales de la langue, le nom et le verbe, processus auquel il peut être fait référence dans la littérature sous l'appellation de « grammaticalisation ». Bassano précise avoir recours à ce terme pour renvoyer « aux processus par lesquels les jeunes enfants mettent en place les contraintes grammaticales du système de la langue qu'ils acquièrent, autrement dit, les contraintes morphologiques et syntaxiques » (2008 : 17). Ce deuxième chapitre portera sur la grammaticalisation du nom chez l'enfant francophone, au travers de l'acquisition du paradigme du déterminant.

Une première partie du chapitre sera consacrée au déterminant chez l'adulte du point de vue morphologique et sémantico-référentiel. La deuxième section nous permettra de passer en revue les différents travaux ayant proposé une description des précurseurs des déterminants dans les productions enfantines. Dans un troisième temps, nous dresserons un panorama des différentes hypothèses, à la fois divergentes et complémentaires, proposées dans la littérature pour décrire et expliquer l'acquisition des formes pré-nominales (proto-formes et déterminants de forme adulte).

Enfin, nous exposerons dans une quatrième et dernière partie du chapitre les enjeux actuels d'une étude sur l'acquisition du paradigme des déterminants.

## 1. Généralités sur le déterminant chez l'adulte

### 1.1 La complexité de la cible à atteindre

En français, le nom doit dans la plupart des cas être précédé d'un déterminant « pour constituer un groupe nominal bien formé dans la phrase de base » (Riegel *et al.* 2004 : 151). En effet, à l'exception de certains contextes - noms propres (a), prédicats (b), noms abstraits (c), noms en position sujet exprimant le pluriel ou des substances (d) et noms inclus dans certains syntagmes prépositionnels (e) (Müller et Kupisch, 2007) -, le nom se distingue principalement des autres catégories grammaticales par le déterminant qui l'accompagne et qui lui est systématiquement antéposé.

#### Exemples 1 - Noms produits non précédés d'un déterminant (Müller et Kupisch 2007 : 46)

- (a) Marie est venue
- (b) Marie est chanteuse
- (c) Avoir peur, avoir froid
- (d) A cette fête, on servait moules et crevettes jusque tard dans la nuit. (Roodenburg, 2004)
- (e) Je vais en train, café sans sucre

L'acquisition du déterminant par l'enfant se manifeste donc en partie par sa production quasi-systématique devant les noms et par la maîtrise des contextes dans lesquels sa production n'est pas attendue.

Nous pouvons par ailleurs noter que le déterminant en français est morphologiquement complexe. Sa catégorie se décline en un paradigme relativement hétérogène composé de dix sous-ensembles de déterminants (sur lesquels nous reviendrons au point suivant) que les grammaires modernes (*i.a.* Riegel *et al.*, 1994) organisent en deux grandes classes : les déterminants définis et les déterminants indéfinis (envisagés tous deux ici dans leur acception la plus large). Au sein de chacune des catégories qui composent ces deux classes, les déterminants varient selon le genre et le nombre des noms auxquels ils sont associés. Si la catégorie 'déterminant' peut être considérée comme difficile à saisir pour l'enfant, c'est donc aussi parce qu'elle est représentée en français par une grande variété de sous-ensembles et de formes.

Ces différentes catégories de déterminants sont mobilisées en français en fonction des caractéristiques sémantico-référentielles des noms, qui varient notamment selon les intentions communicatives des locuteurs et les représentations qu'ils peuvent avoir de la connaissance de leurs interlocuteurs. Le déterminant possède en effet un impact sur la façon dont est traitée la référence portée par le syntagme nominal en discours et d'un déterminant à l'autre, la valeur référentielle du syntagme ne sera pas la même. Les grammaires modernes considèrent par exemple que les déterminants définis possèdent en commun le fait de présenter le référent du nom comme identifiable par l'interlocuteur, par le biais du contexte linguistique, du contexte situationnel ou des connaissances partagées ou supposées partagées par le locuteur et l'interlocuteur. A l'inverse des définis, les indéfinis précèdent le nom dans les cas où le référent du nom est quelconque, non identifiable pour l'interlocuteur, que l'information est nouvelle, ou encore lorsque le locuteur ne juge pas utile de donner l'information. Nous reviendrons sur ces caractéristiques référentielles et discursives du déterminant au sein de la section 1.3 de ce chapitre.

L'acquisition du déterminant par l'enfant passe donc par l'acquisition de sa valeur grammaticale et de la multiplicité de ses formes, mais également par la maîtrise de ses fonctions sémantico-discursives. Le déterminant est une composante plurifonctionnelle du langage adulte et il ne peut être considéré comme pleinement acquis chez l'enfant que lorsque toutes ces fonctions sont assimilées.

## 1.2 Caractéristiques formelles du déterminant chez l'adulte

Nous venons de l'évoquer, le déterminant est considéré au sein du système grammatical français comme un des deux constituants obligatoires (dans la plupart des contextes) du syntagme nominal de forme canonique : syntagme nominal → déterminant + nom. C'est la présence du déterminant qui permet de donner son statut de syntagme nominal au nom et par là même de pouvoir lui attribuer un statut syntaxique dans l'énoncé.

Le déterminant porte les marques de genre et de nombre du nom auquel il est associé. A l'exception de certains cas tels que 'certains/certaines', la marque de genre sur le déterminant disparaît lorsque le nom est au pluriel. Quant au nombre, il est toujours marqué à l'écrit (sauf pour certains quantifieurs invariables) mais peut ne pas être différencié à l'oral. Notons

également que dans le cas des possessifs, le déterminant marque aussi le lien du nom à la personne.

On trouve dans la grammaire traditionnelle<sup>1</sup> deux grandes sous-classes d'éléments susceptibles de précéder le nom pour former avec lui un syntagme nominal : les articles (définis, indéfinis et partitifs) et les adjectifs déterminatifs (possessifs, démonstratifs, indéfinis, numéraux, interrogatifs, exclamatifs et relatifs). Les articles et les adjectifs déterminatifs occupant la même position devant le nom, la grammaire moderne a opéré un regroupement catégoriel afin de ne plus parler que d'une classe unique de mots appelés 'déterminants'. De fait, certains des éléments inclus dans la catégorie des adjectifs déterminatifs, tels que 'même' ou 'quelconque', ne répondant plus aux caractéristiques de la classe des déterminants se sont vus rattachés à d'autres catégories.

Si l'on s'en tient à la répartition proposée dans les grammaires modernes, on trouve donc deux classes de déterminants, les déterminants définis et indéfinis. Les définis, qui - nous le rappelons - partagent, du point de vue des grammairiens, le fait de présenter le référent comme une entité identifiable, sont répartis en trois sous-ensembles :

- les articles définis (*le, la, les* et formes élidées devant voyelle ou 'h' muet) ;
- les déterminants démonstratifs (*ce, cette, ces*) ;
- les déterminants possessifs (*mon, ton, son, etc.*).

Les indéfinis, qui réfèrent quant à eux à des individus quelconques de la classe, sont organisés en sept sous-catégories :

- les articles indéfinis (*un, une, des* et formes élidées devant voyelle ou 'h' muet) ;
- les articles partitifs (*du, de le, de la*) ;
- les déterminants numéraux cardinaux (*un, deux, trois, mille, etc.*) ;
- les déterminants dits indéfinis ou quantifieurs exprimant une quantité nulle, la singularité, la pluralité ou encore la totalité (*aucun(e), nul(le), chaque, certain(e)s, différent(e)s, divers(e)s, quelque(s), peu de, plusieurs, tout(e), tou(te)s*) ;
- les déterminants interrogatifs (*quel(le), quel(le)s*) ;
- les déterminants exclamatifs : (*quel(le), quel(le)s*) ;

---

<sup>1</sup> Cf. notamment Mitterand (1980) pour un commentaire sur la nomenclature grammaticale destinée à l'enseignement du français dans le Second Degré (datée de juillet 1975).

- les déterminants relatifs (*lequel, laquelle, lequel(le)s*).

Si le déterminant assume un rôle important d'un point de vue syntaxique, en permettant au nom de devenir un syntagme nominal et ainsi, d'être porteur d'une fonction dans l'énoncé, notons qu'il joue également un rôle sémantico-référentiel considérable en français. Il est généralement dit du déterminant qu'il sert à 'actualiser' le nom, il lui donne un sens en discours en transformant une entité abstraite en un élément référentiel. En français, dans la mesure où le nom en lui-même ne désigne rien à l'exception d'un concept, il ne peut pas, seul, faire acte de référence. Ce sont alors les déterminants qui, ancrés dans un acte d'énonciation, lui font « correspondre un certain secteur de la réalité (ils font donc passer du sens au référent) » (Ducrot et Schaeffer, 1995 : 371). Si les deux classes de déterminants, définis et indéfinis, ainsi que leurs sous-ensembles servent tous à actualiser le nom, nous verrons au point suivant qu'ils portent des valeurs sémantico-référentielles différentes.

### **1.3 Caractéristiques sémantico-référentielles du syntagme nominal chez l'adulte**

Nous venons de voir au point précédent qu'en français, le nom (devant lequel est attendu un déterminant) peut, s'il est produit seul, désigner un concept mais ne peut pas référer à une entité existante dans les contextes où il n'est pas actualisé par un déterminant. De façon complémentaire, le déterminant quel qu'il soit ne peut pas renvoyer à des éléments à contenu sémantique et donc être référentiellement autonome, mais il porte la caractérisation de la référence du nom.

Ainsi, le syntagme nominal qu'il soit minimal (déterminant et nom) ou étendu (déterminant, nom et modifieur(s)) fait partie - au même titre que les noms propres et les pronoms - de la catégorie des expressions référentielles, que Charolles définit comme « toutes les expressions à même de faire allusion à une entité extralinguistique (concrète, abstraite, générique, spécifique, etc.) existant dans un monde quelconque (2002 : 247).

Nous traiterons dans cette section des caractéristiques sémantico-référentielles des déterminants définis et indéfinis, et par extension, des syntagmes nominaux ou expressions nominales auxquels ils sont associés. Dès lors que notre intérêt dans la troisième partie de cette thèse se portera surtout sur les premiers contrastes dans les usages que font les enfants des formes de transition

et des formes adultes, et au sein de la catégorie des formes adultes, des contrastes entre articles définis et indéfinis, nous nous focaliserons uniquement, dans les deux sections qui suivent, sur les traits sémantico-référentiels des syntagmes au sein desquels sont produits des articles définis (1.3.1) et indéfinis (1.3.2)<sup>2</sup>.

### 1.3.1 Traits sémantico-référentiels des syntagmes définis

Comme l'ensemble des déterminants définis, l'article défini implique, dans son emploi spécifique, que l'interlocuteur puisse identifier, repérer, le référent associé au syntagme nominal qu'il introduit. D'après Riegel *et al.*, « l'article défini présuppose l'existence et l'unicité : il n'y a pas d'autre(s) référent(s) accessible(s) qui vérifie(nt) la description de la réalité désignée par le GN » (1994 : 54). L'interlocuteur peut ensuite s'appuyer sur le contexte linguistique (le syntagme nominal défini peut par exemple, dans le contexte antérieur, avoir été introduit par un syntagme nominal indéfini, les deux entretenant alors une relation anaphorique), sur le contexte situationnel (et notamment, sur l'accessibilité visuelle du référent), ou bien encore sur les connaissances que cet interlocuteur partage avec le locuteur, pour identifier l'entité à laquelle il est fait référence au moyen du syntagme nominal défini.

D'après Charolles (2002), les syntagmes définis se partagent entre descriptions complètes et incomplètes. Les premières ne concernent que les expressions qui ne renvoient en elles-mêmes qu'à un seul référent (exemple : 'la voiture immatriculée 3309WB22'), ou à un seul groupe de référents (exemple : 'les membres actuels de l'académie Goncourt'). A l'inverse, les syntagmes définis incomplets (exemple : 'la voiture') peuvent dans l'absolu faire référence à une infinité d'entités ou de groupes d'entités particuliers. Les descriptions définies complètes restreignent la référence de telle manière qu'elles ne sont plus valables que pour un seul élément ou groupe d'éléments, et c'est ce qui leur confère leur autonomie référentielle. Aucun contexte n'est donc requis pour comprendre à quel élément le locuteur fait référence, à la différence des descriptions définies incomplètes. C'est donc notamment grâce à un contexte référentiel restreint et à la

---

<sup>2</sup> Nous faisons référence ici aux articles définis et indéfinis, par opposition à la macro-catégorie des déterminants définis et indéfinis, mais nous parlerons ultérieurement (dans la suite de ce chapitre et dans les suivants) de définis et d'indéfinis ou de déterminants définis et indéfinis pour faire mention de ces deux catégories d'articles.



coopérativité des interlocuteurs qu'un syntagme défini incomplet peut être interprété comme référant à un élément singulier.

Pour reprendre les termes de Charolles, nous pouvons ainsi dire que :

« les SN définis signalent qu'une entité ou un ensemble d'entités est la seule ou sont les seules d'un certain type. Cette indication (présupposition) a besoin d'être justifiée contextuellement. Soit le SN défini comporte en lui-même des adjoints qui spécifient les circonstances dans lesquelles cette unicité est garantie (descriptions définies complètes), soit le SN ne fournit aucune information de nature à satisfaire cette exigence (descriptions définies incomplètes) » (2002 : 103).

Outre son emploi spécifique - lorsqu'il réfère donc à des occurrences particulières - l'article défini peut aussi être employé avec une valeur générique, et ainsi référer à des entités générales. Dans son emploi générique, le défini renvoie soit à une classe (exemple : 'les chiens sont fidèles'<sup>3</sup>), soit un à un type (exemple : 'le chien est l'ami de l'homme'<sup>4</sup>). Le défini peut par ailleurs, au-delà de ses emplois dans des syntagmes nominaux référant à des éléments spécifiques (concrets et comptables), et dans des syntagmes référant à des génériques, désigner des entités indénombrables (exemples : 'la farine', 'la poussière'), abstraites (exemples : 'l'oisiveté', 'le courage'), ou renvoyant à des événements (exemples : 'la crise', 'la pollution').

### 1.3.2 Traits sémantico-référentiels des syntagmes indéfinis

A l'inverse des déterminants définis qui constituent un ensemble relativement homogène, les déterminants que l'on rassemble traditionnellement sous l'étiquette 'indéfinis' sont beaucoup plus diversifiés, sur les plans syntaxique comme sémantique, et sont d'ailleurs rassemblées dans cette catégorie des unités dont la nature peut varier d'une grammaire à l'autre. On note néanmoins dans le classement que nous avons retenu (*cf.* section 1.2) que toutes les formes des déterminants indéfinis désignent (si aucune autre précision n'est apportée au sein de l'énoncé) des individus quelconques d'une classe, sans qu'il soit possible d'identifier le référent de façon univoque. Dans cette perspective, l'article indéfini est d'après Riegel *et al.* (1994), « à la fois le plus typique et le plus répandu des déterminants qui marquent la référence à un ou plusieurs éléments quelconques » (1994 : 159). Nous verrons toutefois que s'ils renvoient à des entités quelconques, les syntagmes indéfinis peuvent également être aussi spécifiés que les syntagmes définis.

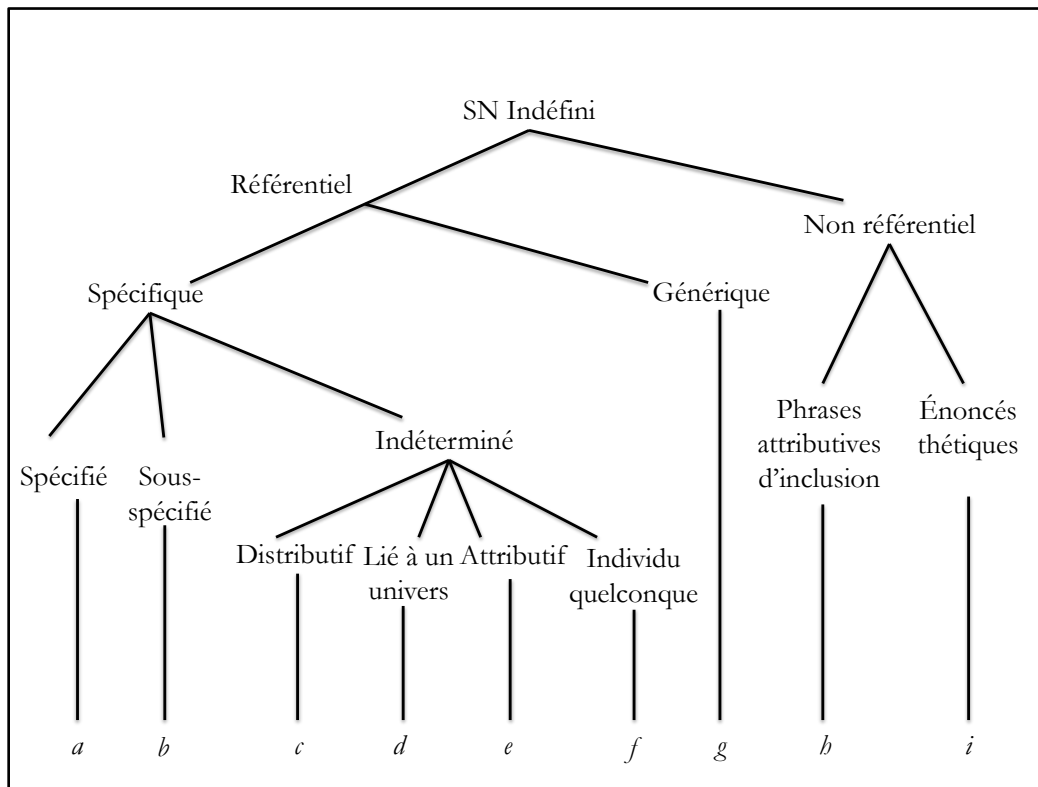
---

<sup>3</sup> Exemple tiré de Charolles (2002 : 248).

<sup>4</sup> *Ibid.*

Si l'on s'en tient à la classification proposée par Charolles (2002) pour l'indéfini singulier (que nous reproduisons ci-dessous, cf. schéma 1), on relève tout d'abord que si une part des usages des syntagmes nominaux indéfinis renvoie à des usages référentiels, une autre part est utilisée dans le cadre d'énoncés non référentiels, au sein de ce que Charolles appelle des 'phrases attributives d'inclusion' et des 'énoncés thétiques'. Dans le premier cas (cf. exemple h), la fonction portée par le syntagme indéfini est prédicative : le syntagme nominal indéfini ne renvoie pas dans ce contexte à un élément de la classe désignée par le nom, mais il lui attribue les propriétés de cette classe. Dans le deuxième cas (cf. exemple i), Charolles considère là aussi qu'aucun objet n'est visé et que le locuteur formule un jugement, dans le cadre d'une construction impersonnelle entièrement prédicative.

**Schéma 1 - Schéma synthétique des usages du syntagme indéfini (Charolles, 2002 : 182)**



- a. Je lis un article de Robert Jackson sur les indéfinis paru en 1972 dans le numéro 19 de la Revue de Philosophie Expérimentale.  
 b. Je lis un article.  
 c. Paul a offert une montre à ses neveux.  
 d. J'ai rêvé que je dansais le tango avec une infirmière.  
 e. Trace un trait.

f. *Un roman policier, on le lit en trois heures* (D. Van de Velde)

g. *Un roman policier, ça se lit en trois heures.*

b. *C'est un disque* (R. Martin)

i. *Il manque une ceinture.* (à propos d'un imperméable)

Parmi les usages référentiels, on trouve d'une part, à l'opposé des usages non référentiels, les syntagmes indéfinis renvoyant à des entités spécifiques - qu'elles soient spécifiées, sous-spécifiées ou indéterminées - et d'autre part, les usages référentiels génériques.

Dans les contextes où la référence est spécifique, la description faite avec le syntagme nominal indéfini peut donc référer à une entité particulière (*cf.* exemple b) mais ne pas être suffisamment précise pour être repérable par l'interlocuteur quel que soit le contexte ; la description est alors dite 'sous spécifiée'. A l'inverse, dans l'exemple 'Je lis un article de Robert Jackson sur les indéfinis paru en 1972 dans le numéro 19 de la Revue de Philosophie Expérimentale' (exemple a), les précisions données sur l'article le rendent repérable, la description est donc dite 'spécifiée' et l'usage préférerait probablement l'emploi d'un article défini.

De la même façon que les expressions nominales définies, les expressions nominales indéfinies possèdent une tête lexicale qui fournit des éléments descriptifs et catégoriels sur les unités auxquelles ils renvoient et leur permettent ainsi des descriptions aussi spécifiées que celles faites avec des syntagmes définis ; cependant, leur mode de donation de la référence est très différent de celui des définis. Les informations fournies dans l'exemple b rendent en effet l'article mentionné identifiable mais elles n'en font pas l'élément unique qu'il serait si le défini était utilisé, dans la mesure où c'est le fait qu'il soit lu qui lui confère une forme d'unicité, et non la description qui en est faite, aussi précise soit-elle. Il serait toujours possible d'argumenter que malgré son caractère spécifique, cette description n'est pas suffisante pour référer à un seul et unique article.

Bien qu'ils ne renvoient pas nécessairement, comme les définis, à un référent unique, les syntagmes indéfinis peuvent avoir la même autonomie référentielle que les descriptions définies complètes, dans la mesure où ils ne nécessitent pas de contexte préalable pour être identifiables par l'interlocuteur, et c'est d'ailleurs pour cette raison qu'ils sont employés de façon prioritaire pour introduire de nouveaux référents dans le discours. Les descriptions indéfinies sont cependant dépendantes des énoncés dans lesquels ils sont produits. D'après Ducrot, « ce n'est

pas le SN indéfini qui opère un acte de référence, bien plutôt [...] c'est l'énoncé pris en bloc qui fournit les informations nécessaires pour que la référence puisse être opérée ensuite » (1972 : 156) ; c'est donc l'énoncé entier et non le syntagme nominal qui établirait la référence.

On trouve par ailleurs, parmi les cas de référence spécifique (par opposition à la référence générique) les emplois que Charolles considère comme 'indéterminés'. Ceux-ci se déclinent en quatre sous-catégories : les syntagmes nominaux indéfinis renvoyant à une référence dite 'distributive' ou à un référent 'lié à un univers', les usages dits 'attributifs' et enfin, les indéfinis référant à un 'individu quelconque'. Dans les contextes où l'indéfini est employé avec une valeur indéterminée, on ne parle pas d'une chose qui est la seule d'un certain type, mais on parle d'une classe et on en extrait éventuellement un exemplaire correspondant au nom donné. C'est le cas en *c* ('Paul a offert une montre à ses neveux') où l'on extrait un exemplaire quelconque de la classe 'montre', et où par ailleurs l'interprétation est considérée comme 'distributive' (et non 'collective') dans la mesure où on envisagera - de par l'usage personnel qui est généralement fait de cet objet - que chacun des neveux a reçu une montre différente.

Dans l'exemple *d* ('J'ai rêvé que je dansais le tango avec une infirmière'), le locuteur fait référence à une personne qui au sein de la classe 'infirmière' répond à la spécification 'danser le tango' ; on extrait donc là aussi un élément d'une classe et celui-ci ne peut exister que dans le rêve dont il est question ou renvoyer à des éléments de la réalité. Ces référents sont dits 'liés à un monde' parce qu'ils sont introduits par des verbes ou certains circonstants donnant naissance à des univers fictifs.

Dans l'exemple 'Trace un trait' (exemple *e*), le locuteur demande à son interlocuteur de dessiner un élément qui soit conforme aux critères de sa description, et ne peut être spécifique que dans l'idée qu'il s'en fait mais pas au moment où l'énoncé est produit puisque l'élément n'existe pas. Cet usage de l'indéfini est dit 'attributif', l'objet de discours est indéterminé et n'existe - avant qu'il soit actualisé - qu'au travers de ses propriétés et dans les représentations de celui qui produit l'énoncé.

Dans le cas de l'exemple *f* ('Un roman policier, on le lit en trois heures'), le roman policier est spécifique mais quelconque, indéterminé, puisqu'aucune prédication ne permet de considérer un exemplaire particulier de cette classe ; cet emploi de l'indéfini ressemble donc aux usages attributifs mais le référent est plus abstrait et ne dépend pas des représentations du locuteur.

Enfin, comme les articles définis, l'article indéfini singulier s'emploie également avec une valeur générique, proche des usages spécifiques indéterminés que nous venons d'évoquer. Dans l'exemple *g* ('Un roman policier, ça se lit en trois heures'), l'élément quelconque étant repris par le démonstratif 'ça', il devient alors un exemplaire prototypique, représentatif de la classe introduite par le nom. Le sens générique de l'article indéfini est néanmoins conditionné, lorsque le syntagme indéfini est sujet, par le fait que le prédicat puisse être recevable pour toute la classe (Riegel *et al.*, 1994).

### 1.3.3 Le modèle adulte de combinaisons forme-fonction

Les caractéristiques des définis et des indéfinis, listées ci-dessus, nous permettent donc d'établir d'une part que dans un contexte de référence spécifique, lorsque le référent ciblé est unique, qu'il a déjà été mentionné ou est accessible dans le contexte (ou bien les deux à la fois), et/ou qu'il est connu de l'interlocuteur, alors on pourrait s'attendre à ce que soit produit un article défini. Lorsqu'à l'inverse le référent est quelconque (et par la suite spécifié ou non par la prédication), qu'il n'a pas fait l'objet d'un repérage préalable (autrement dit, qu'il est nouveau dans le discours) et n'est pas connu de l'interlocuteur, on pourrait plutôt s'attendre à la production d'un article indéfini. Dès lors que le référent est nouveau, l'indéfini pourrait également être adéquat même si l'interlocuteur partage la connaissance du référent avec le locuteur. D'autre part, dans un contexte de référence générique, on pourrait s'attendre à observer le plus fréquemment des articles définis singulier et pluriel ainsi que des articles indéfinis singulier<sup>5</sup>.

On trouve dans la littérature sur l'acquisition de la référence nominale chez l'enfant des travaux qui ont cherché à établir une grille des combinaisons possibles entre formes morphosyntaxiques et facteurs pragmatiques dans la langue adulte, afin de pourvoir décrire les usages des enfants et les comparer à ceux de leurs interlocuteurs adultes. On peut par exemple citer le travail de Rozendaal et Baker (2008) qui se sont intéressées - à partir d'un corpus d'enfants anglophones, francophones et néerlandophones de 2;0 à 3;3 - à l'influence des facteurs pragmatico-discursifs

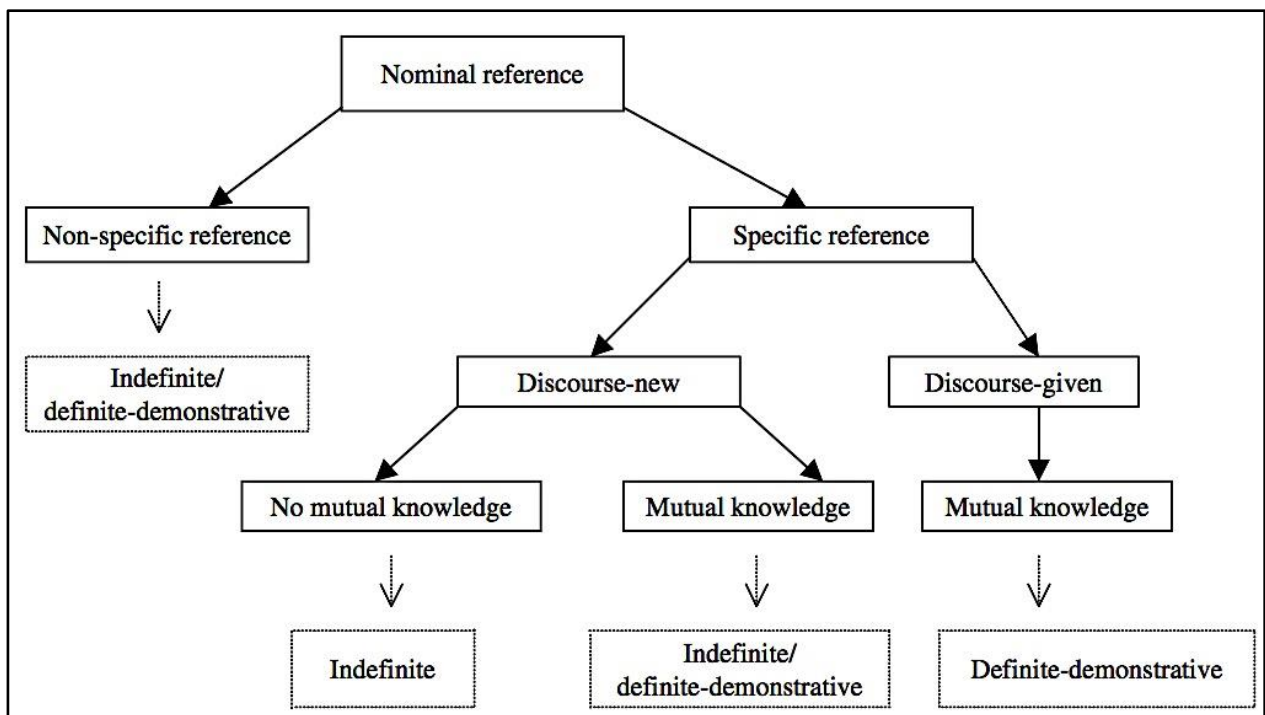
---

<sup>5</sup> Riegel *et al.* notent néanmoins que l'emploi des articles défini et indéfini dans des contextes génériques n'est pas exclusif ; d'autres déterminants comme le démonstratif ou l'indéfini 'plusieurs' peuvent être utilisés avec une valeur générique dans un énoncé dont le sens global est générique (exemple : 'Les hommes nous en veulent. Ah ! Cette pauvre femme qui a toujours tous les torts (dit par une femme)', Plusieurs hommes seront toujours plus forts qu'un seul' (Riegel *et al.*, 1994 : 153).

sur les usages qui pouvaient être faits par les enfants des déterminants définis et indéfinis, et ont dans ce cadre proposé une grille des usages adultes majoritaires.

D'après les auteures, en français comme en anglais ou en néerlandais, le type de déterminant produit dépend du caractère spécifique/non spécifique du référent<sup>6</sup>, de la familiarité du référent pour le locuteur et/ou l'interlocuteur et enfin, du statut discursif du référent (déjà mentionné dans le discours ou non). La figure ci-dessous (schéma 2) présente l'emploi possible des définis et des indéfinis (ainsi que des démonstratifs) en fonction de ces différents facteurs pragmatiques.

**Schéma 2 - Combinaisons des facteurs pragmatiques et des formes morphosyntaxiques pour la référence nominale chez l'adulte (Rozendaal et Baker, 2008 : 776)**



Selon Rozendaal et Baker (*op. cit.*), la première distinction à faire en ce qui concerne les usages des déterminants porte sur le caractère spécifique, autrement dit particulier, ou non spécifique du référent. Bien que l'indéfini soit de loin le plus utilisé dans le contexte de référence non spécifique, les auteures notent néanmoins, à la suite de Lyons (1999), que certains définis et

<sup>6</sup> Pour les auteures, la distinction entre spécifique et non spécifique est ici basée sur une notion pragmatique et non sémantique : dans le cas de la référence spécifique, le contexte discursif indique que le locuteur a à l'esprit une entité particulière, à l'inverse des cas de référence non spécifique. Le modèle de référence est donc construit du point de vue du locuteur, pour qui l'entité est ou n'est pas spécifique (ou particulière) à l'instant où il y fait référence.

démonstratifs peuvent également être produits, de même que certains noms indénombrables produits sans déterminant. Dans leur article de 2008, les auteures ne proposent pas d'exemple de référence non spécifique marquée par une forme autre qu'un syntagme nominal indéfini, mais dans le cadre de sa thèse, Rozendaal (2008) précise que si l'association entre référence non spécifique et déterminant indéfini est en effet la plus observée, tous les types de combinaisons 'déterminant + nom', ainsi que les noms nus indénombrables et les noms propres peuvent être utilisés dans des contextes de référence non spécifique, comme en attestent, d'après l'auteure, les exemples en anglais ci-dessous.

**Exemples 2 - Référence non spécifique marquée par un déterminant défini, un déterminant numéral et un nom propre (Rozendaal, 2008 : 47)<sup>7</sup>**

- a. The new computer that we are going to buy must be much faster than our old one.
- b. Tom plans to bring up three children on his own – but first he needs to find a woman to produce them for him.
- c. Debbie's still waiting for Mr Right – I wonder when she'll finally see there's no such thing and settle with Bob.

Dans le cas de la référence spécifique, le référent peut être connu du locuteur mais nouveau dans le discours et le choix du déterminant repose alors sur le degré de connaissance partagée des interlocuteurs. L'indéfini est attendu dans les cas où la connaissance du référent n'est pas partagée et à l'inverse, lorsqu'elle l'est, le défini ou le démonstratif sont attendus. Lorsque le référent a déjà été mentionné, il est le plus souvent repris au moyen de pronoms, mais dans le cas où un nom est utilisé, ce dernier est associé à un défini ou un démonstratif.

Qu'il s'agisse de travaux portant initialement sur la référence dans la langue adulte ou dans les premières productions de l'enfant, on peut conclure des observations faites dans ce point 1.3 qu'il existe des associations préférentielles entre articles définis et indéfinis et contextes sémantico-référentiels. Certaines combinaisons peuvent donc être considérées comme 'appropriées' et d'autres comme 'inappropriées', ou du moins non 'optimales' (Rozendaal, 2008). Nous garderons toutefois à l'esprit que lorsque l'on s'intéresse à des interactions naturelles, il n'existe pas une seule forme qui remplisse une fonction particulière et une seule fonction remplie par une forme spécifique. Par ailleurs, lorsque nous ferons l'analyse des usages des déterminants dans les productions des enfants de notre corpus, nous nous reporterons, si cela est nécessaire,

---

<sup>7</sup> Rozendaal tire les exemples *b* et *c* de Lyons (1999 : 170).

aux observations faites dans la littérature (notamment par Rozendaal, 2008) sur les combinaisons forme-fonction dans le langage adressé à l'enfant, afin d'éviter de traiter comme inadéquates et propres à l'enfant des associations présentes dans l'input adulte, et de conclure trop hâtivement à un manque de compétences pragmatiques chez l'enfant.

Après ce rapide préalable sur les caractéristiques formelles et sémantico-référentielles du déterminant chez l'adulte, nous consacrerons la suite de ce chapitre à une revue de la littérature sur l'acquisition du déterminant chez l'enfant, en commençant par un état des lieux spécifique sur les études consacrées aux fillers (section 2), qui sera suivi d'un inventaire des différentes perspectives adoptées dans la littérature pour traiter la question de la mise en place du paradigme des déterminants (section 3).

## **2. Un paradigme en construction : l'étude des fillers**

### **2.1 Remarques préliminaires sur les fillers**

Si les études sur la construction du paradigme du déterminant ont le plus souvent traité cette question en termes de présence ou d'absence du déterminant, un certain nombre d'entre elles ont néanmoins porté sur les formes de transition produites par les enfants au cours de la mise en place du paradigme, et avant la stabilisation des formes adultes. La littérature sur ces formes de transition s'accorde pour reconnaître leur contribution significative à la connaissance des processus de grammaticalisation du nom et de l'usage qui en est fait par les enfants mais les débats autour de leur statut et de leur développement sont nombreux.

Dès lors que l'on s'intéresse aux nombreuses études et approches qui ont traité la question des fillers, on observe que la terminologie employée par les chercheurs pour y faire référence est variable. On peut signaler notamment qu'en 1937, Grégoire relève dans son corpus la présence de « proto-marqueurs syntaxiques », Bloom (1970) y fait également référence plus tard sous l'étiquette « placeholders », et Dore *et al.* (1976) les évoquent dans leur travail sous la mention « pre-syntactic devices ». Depuis les années 1990 et l'article princeps de Peters et Menn (1993), c'est le terme « fillers » (ou « filler syllables » dans la littérature anglophone) qui est le plus communément utilisé pour référer à ces formes. On note cependant que cette appellation est discutée, en particulier dans les travaux de Veneziano (entre autres, 2003) ou de Bertin (2011) qui



lui préfèrent respectivement les termes « éléments additionnels préfixés » (« Prefixed Additional Elements ») et « formes de transition » ou « formes transitoires ». Veneziano explique faire ce choix terminologique pour des raisons de neutralité, et à la suite de ce premier constat, Bertin cherche elle aussi à s'approcher de la façon la plus neutre possible du phénomène décrit en évitant « l'image d'un mécanisme où un réservoir d'informations serait rempli de propriétés adéquates » (2011 : 236). La dénomination « fillers » évoquant l'idée d'un vide à remplir par l'enfant - qu'il soit prosodique ou syntaxique - et préétabli, préconçu pour les morphèmes grammaticaux ultérieurement produits, il est en effet envisageable que cette appellation puisse être trop chargée théoriquement pour être représentative de tous les travaux sur le développement de la morphologie chez l'enfant.

Nous ferons néanmoins le choix d'avoir recours au terme « fillers » - en alternance avec l'appellation « formes de transition » - dans le cadre de cette thèse. Signalons cependant que nous ne l'utiliserons pas dans la perspective d'une adéquation avec l'*a priori* théorique qu'il peut convoquer mais bien seulement parce qu'il est le plus fréquemment utilisé dans la littérature et nous semble par conséquent être le plus opportun pour favoriser une continuité entre les travaux précédents sur ces formes et notre présente étude. Bien que Veneziano préfère dans ses travaux le recours au terme 'éléments additionnels préfixés', elle signale à ce propos dans l'article consacré aux fillers de l'*Encyclopedia of Language Development* que « le terme *fillers* est le plus établi et est maintenant utilisé sans qu'y soit associé de façon préconçue un cadre interprétatif spécifique » (2014a : 215-216, *notre traduction*).

La question terminologique n'est pas la seule manifestation de la diversité des points de vue autour de la question des fillers. Ces différentes perspectives sont également reflétées dans les définitions qui sont données des formes de transition. La définition prototypique des fillers, partagée par la plupart des études, fait état de voyelles ou de syllabes 'consonne-voyelle' précédant les mots lexicaux (ou plus exactement les proto-mots lexicaux), non conformes à des mots de la langue adulte et qui figurent dans le langage du tout jeune enfant à des « places » où l'on pourrait trouver produits quelques mois plus tard des déterminants, des auxiliaires, mais aussi des prépositions ou des pronoms. On peut citer pour exemples en français des cas de fillers pré-nominaux relevés par Bassano : [se a ʃa] '*c'est FILLER chat*', [mɛtʁ e ʃosɔ̃] '*mettre FILLER chausson*' (2008 : 33), ou de fillers pré-nominal et pré-verbal relevés par Veneziano : [agy] '*FILLER*

*grue*, [ə'pik] *FILLER pique*' (2014a : 215). Si la définition générale mentionnée plus haut semble faire consensus, on note cependant très vite des variations concernant le statut attribué à ces formes au sein du parcours d'acquisition des mots grammaticaux et de leur devenir dans le langage de l'enfant. Ces variations - qui peuvent être liées au cadre théorique adopté par les chercheurs et/ou aux facteurs considérés comme prépondérants dans la mise en place des morphèmes grammaticaux - feront l'objet d'un développement spécifique au sein de ce chapitre (section 3). Avant d'aborder cette question des différentes perspectives adoptées dans la littérature pour décrire et expliquer la production des formes de transition, nous reviendrons dans les pages suivantes sur quelques observations générales au sujet des fillers.

Signalons enfin que cette thèse ayant pour objectifs de décrire et de comprendre comment se développe le paradigme du déterminant chez l'enfant, le point de vue qui y sera adopté pour traiter des fillers sera le plus souvent restreint à la catégorie des fillers pré-nominaux. Le phénomène des fillers nous intéressera en effet davantage pour ce qu'il représente et apporte comme éclairage dans l'évolution de la catégorie 'déterminant' qu'en tant que catégorie transitionnelle 'fillers' à part entière. Nous n'aurons donc pas pour ambition, ni dans la revue de littérature qui suit ni dans le traitement ultérieur des données, de couvrir tout le champ des études et des questions soulevées par l'étude des fillers - soumises à une grande diversité que l'on considère les fillers pré-verbaux ou bien pré-nominaux - mais seulement celles concernées par les fillers produits devant les noms. Les sections qui suivent (sections 2.2 à 2.5) s'intéresseront néanmoins aux fillers dans leur ensemble mais les deux parties suivantes du chapitre (sections 3 et 4) se focaliseront quant à elles sur le développement du paradigme du déterminant et donc sur les fillers pré-nominaux.

## **2.2 Un phénomène situé dans le temps du développement**

### **2.2.1 Transition entre énoncés à un terme et énoncés à deux termes**

Les différents travaux menés sur les fillers ou sur les formes adultes de morphèmes pré-nominaux ou pré-verbaux s'accordent pour situer les premières productions de ces formes autour de la période de transition des énoncés à un terme aux énoncés à deux termes. Bien que cette période apparaisse le plus souvent chez l'enfant un peu avant la deuxième année, c'est en termes d'étapes dans le développement linguistique et non d'âge que l'apparition des fillers est

discutée et cela pour des raisons évidentes de variabilité inter-individuelle (sur laquelle nous reviendrons au point 2.4).

Si l'on constate le plus souvent que les premières occurrences des fillers apparaissent au moment de la transition entre énoncés à un et à deux termes, c'est, d'après la littérature, parce qu'il existe une relation de dépendance forte entre production des premières formes pré-nominales et capacité de combinaison de l'enfant (Veneziano et Sinclair, 2000). L'apparition des fillers fait suite à une première période de production et de stabilisation des énoncés à un terme et plus précisément, à la période du développement linguistique pendant laquelle l'enfant produit ce qui est communément appelé des Suites d'Énoncés à un Mot en Succession (SEMeS) en français et Successive Single-Word Utterances (SSWUs) en anglais (*i.a.*, Veneziano, 1997a). Durant cette période, l'enfant produit dans des tours de parole différents des énoncés à un mot qui une fois assemblés forment un enchaînement sémantiquement cohérent et préparent les premières productions d'énoncés à deux termes (*cf.* également 'constructions verticales' décrites par Scollon, 1979). Ainsi, dans les données de Veneziano et Sinclair (2000), l'augmentation soudaine de la part des fillers à 1;7 fait suite aux premières apparitions de SEMeS trois mois plus tôt, et cette hausse de la proportion des fillers précède ensuite l'augmentation du nombre d'énoncés combinant deux termes. Enfin, lorsque les énoncés à deux termes et plus deviennent majoritaires dans les productions de l'enfant, la part des fillers diminue et les formes adultes deviennent plus fréquentes. Les fillers seraient donc liés au développement de la combinatoire, et leur apparition favorisée par le passage des énoncés à un terme aux énoncés à deux termes.

### **2.2.2 Masse lexicale critique et développement des formes pré-nominales**

Nous pouvons par ailleurs noter qu'à la suite des travaux de Bates et de ses collègues (notamment, Marchman et Bates, 1994), Bassano souligne que « conformément à l'hypothèse de la « masse critique », l'accumulation d'un certain stock lexical est nécessaire pour que la grammaticalisation d'une classe de mots s'enclenche » (2005 : 10). Appliquée à la grammaticalisation du nom, cela signifierait qu'un stock minimal de noms est requis pour qu'apparaissent les fillers et les premiers déterminants dans les productions de l'enfant. Dans son corpus (corpus de Pauline), Bassano (*ibid.*) relève par exemple que l'accroissement significatif du nombre de noms produits a lieu autour de 25 mois et que l'explosion dans la grammaticalisation

des noms se produit entre 27 et 28 mois, ce qui confirmerait, d'après elle, l'hypothèse proposée par Marchman et Bates (1994). Le développement des fillers et des déterminants ne serait donc là encore pas conditionné par l'âge de l'enfant mais plutôt largement soumis à des questions de variation individuelle, et plus spécifiquement, lié au moment auquel l'explosion lexicale se produit chez chaque enfant.

### 2.2.3 Concomitance entre fillers, omissions et formes adultes

Un autre élément lié à l'apparition des fillers dans le langage de l'enfant fait consensus dans la littérature : celui de la concomitance entre les fillers et les formes adultes de morphèmes grammaticaux. Pendant la mise en place du déterminant coexistent le plus souvent des formes de transition - fillers mais aussi absences de formes - et des formes adultes. Le développement des positions pré-lexicales n'est pas graduel et l'enfant tâtonne, en produisant en alternance des formes attendues et des formes non attendues. L'alternance se fait d'abord majoritairement entre absences de formes et fillers, et par la suite, majoritairement entre fillers et formes adultes. À titre d'exemple, Bassano et Eme (2001) observent qu'en moyenne, les dix enfants de leur corpus produisent à 20 mois 50% de noms qui ne sont précédés d'aucune forme, 15% d'entre eux sont précédés par des fillers et 12% par des déterminants. Ces observations sont partagées par López Ornat (2003) qui a travaillé sur des données transversales d'enfants hispanophones âgés de 1;7 en moyenne et constaté qu'après une période de production de noms sans déterminant, la plupart des enfants de son corpus produisent en alternance des noms nus ou précédés de fillers ainsi que des noms précédés de déterminants de forme adulte. Dans une étude longitudinale de 2008, Salazar Orvig *et al.* montrent également (*cf.* tableau 1 ci-dessous) qu'à tous les âges observés dans leur corpus (de 1;8 à 3;0), l'enfant produit devant les noms des définis et/ou des indéfinis ainsi que d'autres formes adultes mais également des fillers et des omissions<sup>8</sup>. On note par exemple que la part des déterminants représente 56,15% des formes pré-nominales de 2;4 à 2;6 et devient donc majoritaire mais en alternance, les enfants produisent encore 18,46% de fillers et 25,38% de noms seuls.

---

<sup>8</sup> Pour d'une part faciliter la rédaction mais aussi parce que dans la suite de ce travail nous considérerons les omissions comme des formes à part entière qui contrastent de façon significative avec les fillers et les déterminants, nous aurons recours à l'expression 'production de formes' pour couvrir aussi bien les formes véritablement produites (fillers et déterminants) que les absences de formes (omissions).

**Tableau 1 - Distribution des marques en position pré(proto)-nominale (d'après Salazar Orvig *et al.*, 2008 : 64)**

Âge	Définis	Indéfinis	Autres	Fillers	Omissions
1;8-1;9	1,96%	0,00%	0,00%	31,37%	66,67%
1;10-2;0	23,33%	6,45%	0,50%	26,55%	43,18%
2;1-2;3	20,59%	14,50%	5,38%	26,37%	33,16%
2;4-2;6	36,92%	10,77%	8,46%	18,46%	25,38%
2;7-2;9	25,93%	36,11%	10,65%	4,17%	23,15%
2;10-3;0	31,08%	51,39%	7,17%	1,20%	9,16%

Cette alternance est globalement régulière puisque les formes de transition tendent à diminuer et la part des formes adultes à augmenter, mais le développement n'est pas linéaire pour autant dans la mesure où la part des formes de transition, qu'il s'agisse des fillers ou des omissions, peut augmenter après avoir diminué. Ces formes peuvent aussi réapparaître à un stade avancé du développement langagier.

Ce phénomène de non-linéarité a notamment été observé par Kilani-Schoch et Dressler (2000) qui ont constaté que les formes de transition connaissent ce que les auteurs appellent des 'phases de régression'. Dans le corpus de Sophie, sur lequel porte l'analyse, la diminution des fillers n'entraîne pas systématiquement une augmentation des formes adultes, en position pré-nominale comme pré-verbale. La part des fillers pré-nominaux diminue considérablement à partir de 2;2 mais elle s'accompagne d'une augmentation temporaire des omissions. De la même façon, lorsque la part des clitiques sujets augmente, de nombreux verbes restent produits sans sujet, et cela malgré la disparition des fillers. Les auteurs relèvent par ailleurs qu'après avoir nettement diminué, la proportion de fillers augmente entre 2;4.1 et 2;4.12 avant de diminuer à nouveau jusqu'à leur disparition. Yamaguchi *et al.* (en préparation) ont fait des constats similaires dans le corpus longitudinal de Madeleine qui produit environ 30% de fillers entre 1;11 et 2;1 (29,57% à 1;11 et 27,84% à 2;1) et en produit 39,71% quelques mois plus tard à 2;2.

### 2.3 La construction progressive d'une catégorie de transition

De nombreux chercheurs ont étudié le phénomène des fillers pour ce qu'ils apportent comme éclairage à la grammaticalisation des lexèmes, en les opposant aux omissions d'une part et aux formes adultes d'autre part, mais sans nécessairement prendre en considération la complexité de cette catégorie. Or, bien qu'ils ne soient pas conformes aux exigences requises pour constituer

une catégorie grammaticale adulte, les fillers n'en sont pas moins les représentants d'une catégorie certes éphémère mais qui fait sens dans le langage de l'enfant. Dans cette perspective, un certain nombre de chercheurs (Kilani-Schoch et Dressler, 2000 ; Peters et Menn, 1993 ; Peters, 2001 ; Veneziano et Sinclair, 1997 ; Veneziano, 2014b) se sont ainsi intéressés aux fillers dans une approche développementale, et ont argumenté en faveur d'une évolution de leur statut, allant de pair avec une évolution globale des compétences langagières de l'enfant.

### **2.3.1 Du pré-morphologique au (quasi-)morphologique**

On constate en effet que, dès les premiers travaux consacrés spécifiquement aux fillers, certains chercheurs se sont attachés non seulement à situer le moment de leur apparition et à les décrire à un temps donné mais aussi à décrire et expliquer leur évolution dans le temps. Peters et Menn (1993) ainsi que Veneziano et Sinclair (1997) ont ainsi montré qu'il n'existe pas un seul type de fillers produits dans des positions syntaxiques appropriées et d'emblée grammaticaux mais plutôt différents types de fillers - ou de périodes - d'abord phono-prosodiques avant de devenir grammaticaux. Le modèle proposé par ces auteures conçoit le développement des fillers en trois phases : pré-morphologique, proto-morphologique et morphologique ou quasi-morphologique.

Pendant la période pré-morphologique, les fillers sont avant tout vocaliques, rarement consonantiques, et ne semblent pas être différenciés par l'enfant en fonction de la position qu'ils occupent dans l'énoncé. Les fillers auraient à ce stade des propriétés purement phono-prosodiques. L'enfant qui commence à appréhender les régularités sonores et prosodiques de la langue utiliserait les fillers pour préserver la structure syllabique et/ou rythmique de sa langue et pour allonger ses énoncés. Ces premières formes ne seraient pas à considérer comme distinctes des éléments lexicaux avec lesquels ils sont produits mais plutôt comme des extensions phonologiques de ces éléments. Nous reviendrons sur cette phase pré-morphologique au sein de la section 3.1, consacrée à la place de la phonologie dans la littérature sur l'acquisition des déterminants.

Lors de la deuxième phase, dite proto-morphologique, les fillers continuent à être produits avec une fonction rythmique mais commencent par ailleurs à partager certaines caractéristiques avec les morphèmes grammaticaux adultes, tant du point de vue de la distribution que de la phonologie. On commence, en effet, à pouvoir identifier les classes des éléments ciblés dans la

langue adulte (déterminants, pronoms, auxiliaires et prépositions) et au-delà de la position, à observer des correspondances phonologiques entre les éléments de la même classe. Peters (2001) signale par exemple que les enfants anglophones tendent à produire un coup de glotte à l'initiale des proto-déterminants et à produire des nasales lorsque les fillers se substituent aux auxiliaires. On relève également à ce stade la présence de fillers de plus d'une syllabe qui semblent reproduire de façon figée certains clusters de la langue adulte. Ainsi, en anglais 'umma' pourrait être l'amalgame de 'I'm + gonna', ou 'didja' celui de 'did + you' (Peters, 2001). Ces amalgames ne seraient analysés par l'enfant comme des segments distincts que dans un stade ultérieur.

D'après Peters (2001), la forme phonologique des fillers produits dans une dernière étape, dite morphologique, correspond à celle des formes adultes de façon suffisamment claire pour qu'ils puissent être identifiés sans difficulté. Leur distribution concorde le plus souvent avec la distribution des cibles adultes et ils semblent porter les mêmes fonctions que ces derniers. A ce stade, les fillers se distribueraient en catégories distinctes et seraient organisés de façon systématique. Pour Veneziano (2014b), la dernière étape dite quasi-morphologique est marquée par la coexistence de fillers proto-morphologiques et de morphèmes grammaticaux conformes à la cible adulte, pouvant parfois précéder le même élément lexical. A ce stade, ni les fillers ni les morphèmes conformes à la cible ne sont produits dans tous les contextes où ils sont obligatoires puisque l'on continue à relever la présence de certaines omissions.

Nous aurons l'occasion de signaler au sein de la section 2.4 que la production des fillers peut être soumise à des questions de variation individuelle ; il est donc très probable que tous les enfants ne passent pas par ces trois étapes décrites dans la littérature. Néanmoins, pour certains enfants au moins, les fillers représentent une étape de transition avant la maîtrise des morphèmes grammaticaux, à envisager non pas comme une étape uniforme mais comme la déclinaison de plusieurs sous-étapes témoignant du caractère évolutif et complexe des fillers.

### **2.3.2 Continuité dans le processus de 'grammatisation' des fillers**

Kilani-Schoch (1998), et plus tard Kilani-Schoch et Dressler (2000), ont eux aussi cherché à décrire et expliquer le processus de développement des fillers, que les auteurs désignent comme le processus de 'grammatisation' des fillers. Les auteurs se sont appuyés pour leurs études sur l'analyse des fillers produits dans le corpus de Sophie (auquel nous avons déjà fait référence au

point 2.2.3), enfant francophone suivie de 1;6 à 3;8 ans. Pour ces deux études, ont été référencés comme 'fillers' les éléments additionnels qui occupent les positions pré-verbales et pré-nominales, mais aussi ce que Kilani-Schoch (1998) appelle les 'fillers lexicalisés', c'est-à-dire les voyelles produites à l'initiale de mots tronqués, et dont la présence ne peut pas être expliquée par le système phonologique de l'enfant. C'est le cas par exemple de [əpa] (lapin) ou [ɛka] (chocolat), produits respectivement à 1;6.24 et 1;7.15 et [ətalɔ̃] (pantalon) à 1;11.19<sup>9</sup>. A l'inverse, lorsque les voyelles produites peuvent résulter de substitutions phonologiques régulières dans le langage de l'enfant, elles n'ont pas été caractérisées comme des fillers.

A la suite de Peters et Menn (1993) pour l'anglais ou Veneziano et Sinclair (1997) pour le français, Kilani-Schoch et Dressler (2000) ont argumenté en faveur d'un développement allant du prosodique au morphologique. Les premiers fillers seraient d'abord prosodiques avant d'être grammaticaux, et ils se rapprocheraient progressivement des formes adultes, sur les plans phonétique et fonctionnel, avant d'être remplacés par les morphèmes grammaticaux adultes. Kilani-Schoch (1998) note par exemple que l'enfant passe du pré-morphologique au proto-morphologique lorsque les fillers peuvent être observés dans des positions plus diversifiées telles que les positions à fonction prépositionnelle ou celles caractéristiques d'une structure clivée, ou bien encore lorsqu'ils se rapprochent phonétiquement de cibles adultes identifiables. Outre leurs fonctions prosodiques, les fillers acquièrent progressivement une fonction grammaticale, et d'après l'auteure, « du simple segment vocalique motivé phonologiquement et prosodiquement au morphème grammatical il n'y a pas de rupture mais un développement continu bien que non linéaire » (1998 : 171).

Cette continuité couvrirait donc d'une part l'idée que l'enfant acquiert la maîtrise des morphèmes non pas d'emblée mais au travers d'étapes successives et intermédiaires, pendant lesquelles cette maîtrise n'est que partielle, et d'autre part, l'interaction et la dépendance entre facteurs phonologiques et morphologiques. D'après le modèle constructiviste de développement de la morphologie (modèle dit de 'Morphologie Naturelle') proposé par Dressler et Kilani-Schoch (notamment, 2000), la grammaire de l'enfant est organisée autour de 'modules' qui ne sont pas acquis de façon innée mais construits par l'enfant au travers d'étapes pré- et proto-

---

<sup>9</sup> Notons que ces formes (éléments dits 'non additionnels') n'ont pas été prises en compte comme des fillers chez Veneziano et Sinclair (notamment, 2000).



morphologiques. Au cours de ces étapes, phonologie, prosodie et morphologie constituent un ensemble indifférencié, avant de se scinder au stade morphologique en modules distincts. Dans cette perspective, la phonologie est première et conduit à la mise en place de la morphologie. Le développement des fillers serait donc, d'après Kilani-Schoch et Dressler, parfaitement en adéquation avec ce modèle puisque, dans un premier temps, leur fonction se partagerait entre phonologie et morphologie, et qu'en dernier lieu, lorsque les fillers commencent à correspondre rigoureusement aux morphèmes de la langue adulte (qui ensuite les remplacent), ils constituent la preuve de l'émergence de la morphologie comme module à part entière.

## **2.4 Un phénomène soumis à la variation individuelle**

Nous l'avons signalé plus tôt, les premiers fillers peuvent être produits à des âges différents d'un enfant à l'autre, ce qui en fait comme tout phénomène linguistique un phénomène soumis à la variation individuelle. Mais au-delà de ces questions de rythme dans le développement langagier, un certain nombre d'études montrent que si la majorité des enfants semble avoir recours aux fillers, tous les enfants n'en produisent pas à la même échelle - et éventuellement pas du tout - et durant la même période.

### **2.4.1 Des styles individuels dans la proportion des fillers produits**

Les travaux qui se sont intéressés aux fillers ont montré qu'il existait un certain nombre de tendances générales dans les productions des enfants. Bassano et Eme (2001) qui se sont focalisées sur la variabilité inter-individuelle dans la grammaticalisation des noms en français ont par exemple constaté que tous les enfants de leur corpus (dix enfants) semblaient avoir recours aux fillers comme « procédure d'amorçage de la grammaticalisation du nom » (2001 : 4) puisque à 20 mois, tous produisaient des fillers devant certains noms. Les données des enfants étant compilées nous n'avons pas la possibilité de rendre compte précisément des différences dans les temps d'apparition et dans la fréquence de production des fillers chez chacun des enfants, mais les auteures signalent qu'elles ont trouvé « une certaine variabilité de niveau - donc de rythme - à 20 mois surtout » (2001 : 4). Les dix enfants du corpus ne semblent donc pas présenter de différences de style dans la façon dont le déterminant se met en place mais présentent *a contrario* des différences dans le rythme d'acquisition.

Outre des variations dans le moment d'apparition des fillers, on note dans la littérature que leur proportion peut elle aussi dépendre de profils individuels. La fréquence de production des fillers n'est en effet pas la même chez tous les enfants. Peters signale par exemple dans sa revue des travaux sur les fillers publiée en 2001 que Minh et Seth (enfants dont les productions ont notamment été détaillées dans ses travaux de 1983 ou 1995) produisent des fillers dans des proportions importantes alors qu'à l'inverse, le fils de Menn, Daniel, ne produit pas de fillers syllabiques. Seules quelques occurrences de fillers consonantiques sont relevées dans son corpus (Peters et Menn, 1993). Peters (2001) note également que dans les corpus de Bloom (1970), Eric produit des fillers beaucoup plus fréquemment que Kathryn ou Gia.

Bien qu'ils soient produits dans des proportions différentes, il semblerait que la plupart des enfants passe par une période de production de ces formes de transition. Certains auteurs ont néanmoins relevé que tous les enfants n'en produiraient pas, ou du moins pas pendant les périodes concernées par le recueil de corpus. Dans une étude de 2003 portant sur les variations grammaticales chez des enfants espagnols de 1;4 à 1;10, López Ornat signale par exemple que deux des enfants sur les vingt-trois de l'étude ne produisent pas de fillers. Simonsen (2001) relève également, dans une étude sur la production de fillers en norvégien, qu'un des deux enfants de son corpus, Tomas, ne produit pas du tout de fillers. Dans la mesure où ces enfants acquièrent *in fine* la maîtrise du paradigme des déterminants, au même titre que les enfants passant par une période de transition, López Ornat (*op. cit.*) suggère qu'ils accèderaient à la représentation des syntagmes nominaux canoniques 'déterminant + nom' en ayant recours à d'autres mécanismes de traitement.

De façon plus générale, Peters (1977, 1983) remarque que les enfants expressifs produisent plus de fillers que les enfants référentiels<sup>10</sup>. Pour Demuth (2001) ou Pepinsky *et al.* (2001), ce phénomène pourrait s'expliquer par le fait que les premiers entrent dans le langage au niveau du mot, voire de l'énoncé, alors que les seconds y accèdent d'abord au niveau du pied<sup>11</sup>. Les degrés

---

<sup>10</sup> Pour rappel, chez les enfants dits 'référentiels' comme 'expressifs', les noms communs sont majoritaires dans les premières productions de l'enfant mais chez les enfants expressifs, les premiers mots sont plus diversifiés, contiennent une moindre proportion de noms et des expressions figées telles que 'come-here' ou 'thank-you' (Nelson, 1973).

<sup>11</sup> Le 'pied' est généralement défini dans la littérature (*cf.* par exemple, Selkirk, 1984 ; Nespor et Vogel 1986 ; Demuth, 1996) comme le niveau de la hiérarchie prosodique situé entre le mot phonologique et la syllabe.

supérieurs de la hiérarchie prosodique - sur laquelle nous reviendrons à la section 3.1.2 - entraîneraient une production plus importante de fillers.

La littérature fait également état d'une évolution variable de ces formes d'un enfant à l'autre. Veneziano (2014a) signale que les fillers peuvent être produits chez certains enfants dans de faibles proportions et augmenter lentement jusqu'à une diminution finale ; à l'inverse, chez d'autres enfants la proportion des fillers peut augmenter soudainement, se stabiliser et diminuer jusqu'à leur disparition. On note par ailleurs dans certaines études (Caët, 2013, pour les fillers pré-verbaux ; Yamaguchi *et al.*, en préparation, pour les fillers pré-nominaux) que tous les enfants qui produisent des fillers ne le font pas durant la même période. Qu'il s'agisse des fillers pré-nominaux ou pré-verbaux, la durée de leur production peut varier d'un enfant à l'autre et s'étendre de quelques mois à quelques années.

Lorsqu'ils sont présents dans le langage de l'enfant, les fillers pourront donc apparaître plus ou moins précocement, être produits dans des proportions variables et se distribuer différemment dans le temps.

#### **2.4.2 Un rapport individuel aux contraintes en jeu dans la production des fillers**

Nous verrons dans la deuxième partie de ce chapitre que les outils utilisés dans la littérature pour décrire la présence ou l'absence de fillers et leur évolution sont nombreux et que cette variété s'explique par le statut qui est attribué aux fillers, par l'importance donnée à certaines contraintes en jeu dans l'acquisition du langage - qu'elles soient prosodiques, morphologiques, sémantiques, syntaxiques, ou pragmatiques - ou par le cadre théorique général adopté par les chercheurs.

La plupart des études sur le sujet n'ayant pas proposé d'analyse combinée des facteurs et, au contraire, privilégié l'observation de facteurs spécifiques, il est difficile de pouvoir établir si d'un enfant à l'autre, l'influence de ces contraintes peut être variable. Veneziano (2014a) suggère néanmoins que les enfants ne seraient peut-être pas tous sensibles de la même façon aux mêmes facteurs et pourraient produire des fillers pour remplir des objectifs différents. On pourrait par exemple imaginer que chez certains, les fillers puissent être davantage produits pour remplir une fonction prosodique et chez d'autres pour répondre à des contraintes pragmatico-discursives.

Veneziano (2014a) relève par ailleurs que la fonction portée par les fillers pourrait évoluer au fil du développement langagier de l'enfant, autrement dit, qu'un facteur puisse être prédominant au sein d'une période donnée et qu'un autre facteur puisse devenir prépondérant dans une période ultérieure.

## **2.5 Un phénomène soumis à la variation inter-langues**

La production de fillers a été relevée dans des travaux sur l'allemand, l'anglais, le danois, l'espagnol, le français, le gallois, le grec, l'hébreu, l'italien, le japonais, le néerlandais, le norvégien, le portugais, le sesotho et le suédois (Veneziano, 2014a). Il semblerait donc que si leur présence n'est pas attestée chez tous les enfants, elle l'est pour toutes les langues étudiées pour lesquelles la présence de morphèmes grammaticaux pré-lexicaux est attendue dans la langue adulte. On relève néanmoins dans la littérature que le moment d'apparition des fillers et leur proportion peut varier selon la langue étudiée.

### **2.5.1 Des différences dans le moment d'émergence des fillers et des déterminants**

Lleo et Demuth (1999), qui se sont intéressées aux contraintes prosodiques liées à l'acquisition des déterminants, ont comparé l'âge d'apparition des fillers et des déterminants en espagnol et en allemand. L'analyse de leur corpus (trois enfants hispanophones et quatre enfants germanophones âgés entre 1;4 et 2;3) témoigne d'une apparition plus précoce des formes pré-nominales en espagnol qu'en allemand. Dans les données étudiées, la production des formes pré-nominales en allemand s'accélère entre 1;10 et 2;1 alors que l'accélération est observée en espagnol dès l'âge de 1;4. Cette différence dans le rythme d'acquisition s'expliquerait pour elles par le statut prosodique de ces formes dans la langue cible et notamment, par une prosodification plus complexe du déterminant en allemand qu'en espagnol. En espagnol, les déterminants sont en règle générale des syllabes non accentuées qui précèdent les mots lexicaux et forment avec eux un ensemble prosodique. D'après les auteures, la combinaison d'un déterminant et d'un mot lexical de deux syllabes est en espagnol prosodiquement équivalente à un mot de trois syllabes. A l'inverse, en allemand, le déterminant peut posséder deux statuts prosodiques différents : il peut former à lui seul un pied accentuel lorsqu'il est plein ou être prosodiquement associé au mot précédent lorsqu'il est contracté.

Dans la conclusion de leur étude, les auteures généralisent les constats faits sur l'espagnol et l'allemand et postulent que les déterminants apparaîtraient plus tôt dans les langues romanes que dans les langues germaniques. Des constats similaires ont d'ailleurs été effectués par Belletti et Guasti (2015) dans une étude sur l'acquisition des déterminants en italien et en néerlandais. La prosodification plus complexe des articles en néerlandais expliquerait, d'après elles, le taux plus important d'omissions en néerlandais qu'en italien. Lleo et Demuth envisagent également la possibilité de prédire le moment d'apparition des morphèmes grammaticaux dans d'autres langues, mais à condition d'avoir connaissance des contraintes prosodiques à l'œuvre dans la production des mots constitués d'un seul morphème et de savoir comment sont prosodifiés les mots fonctionnels dans la langue cible. Elles soulignent par ailleurs que leurs résultats viennent contredire l'idée d'une universalité dans le développement des mots grammaticaux préalablement développée dans les travaux de chercheurs générativistes (notamment, Radford, 1990).

### **2.5.2 Influence des propriétés de la langue cible sur la proportion des fillers produits**

Outre des variations dans le moment d'apparition des fillers, certains auteurs avancent également que certaines langues semblent favoriser, plus que d'autres, la production des fillers. Dans l'étude citée à la section précédente, Pepinsky *et al.* (2001) relèvent par exemple que la langue en acquisition ainsi que l'input et la fréquence des formats prosodiques entendus par l'enfant semblent influencer la proportion de fillers produits. Les auteurs suggèrent notamment que les enfants apprenant des langues contenant des mots phonologiques plus complexes et plus fréquents, comme en espagnol, se focalisent sur des niveaux supérieurs de la hiérarchie prosodique et produisent plus de fillers que les enfants apprenant des langues comme l'anglais, pour lesquelles la majorité des mots de la langue sont composés d'un pied monosyllabique.

Veneziano (2014a) signale également que si la présence de fillers est attestée dans un grand nombre de langues - y compris dans des langues sans articles comme le japonais ou agglutinantes comme le turc - il semblerait que certaines d'entre elles, comme le français ou l'anglais, impliquent une proportion plus importante de ces formes que d'autres langues telles que le japonais ou le gallois. Néanmoins, si les résultats de ces études peuvent laisser à penser que la langue produite par l'enfant a un impact non négligeable sur la production des formes de transition, Veneziano (*ibid.*) souligne que le phénomène étant largement soumis à la variation

individuelle, d'autres études sur la variation inter-linguistique seraient requises pour venir étayer ces conclusions. En effet, les recherches sur l'émergence des fillers et des déterminants se sont pour beaucoup focalisées sur l'analyse d'une seule langue ou sur la comparaison de deux langues, et sur des corpus relativement restreints ; une étude à plus grande échelle, comparant davantage de langues au sein d'un vaste corpus permettrait de faire état avec plus de certitude de l'influence de la langue sur la proportion de fillers produits par les enfants.

### **3. Le traitement des fillers et des déterminants dans la littérature, reflet de la diversité des approches méthodologiques et théoriques**

Nous avons pu constater au sein de la deuxième section de ce chapitre que les travaux qui se sont intéressés aux formes de transition sont nombreux, qu'il s'agisse d'études se focalisant exclusivement sur les fillers ou sur ces éléments pour ce qu'ils apportent à la connaissance du processus de construction des catégories 'déterminant' et 'pronom'. Dans cette troisième section, nous traiterons spécifiquement des travaux portant sur les fillers pré-nominaux et les déterminants, dans des approches phono-prosodique, lexico-syntaxique et pragmatico-discursive.

#### **3.1 Perspective phonologique sur le développement des déterminants**

Dans le domaine de la phono-prosodie, les études sur le développement des fillers et des déterminants sont nombreuses mais d'une étude à l'autre, les objectifs principaux restent largement comparables : d'une part, apporter un éclairage autre que morphosyntaxique à la question du développement des déterminants et d'autre part, déterminer grâce aux outils prosodiques et phonologiques si les fillers sont des éléments purement phonologiques, proto-morphologiques ou pleinement morphologiques, au même titre que les formes adultes. Nous présenterons dans les deux sections qui suivent les réponses apportées à ces questions dans la littérature.

##### **3.1.1 La place de la phonologie dans le débat sur le statut des fillers**

Dans une étude de 2005, Tremblay passe en revue les trois hypothèses qui ont été proposées dans des travaux antérieurs pour expliquer la présence des fillers dans le langage enfantin et cherche à observer si les fillers sont seulement des éléments phonologiques qui reflèteraient la sensibilité de l'enfant aux propriétés phono-prosodiques de la langue, s'ils sont proto-

morphologiques et reflèteraient alors la connaissance de l'enfant de certaines catégories fonctionnelles ou enfin, purement morphologiques mais phonologiquement différents de la forme cible adulte en raison d'une incapacité phonétique de l'enfant à les produire.

Certains travaux (notamment, pour l'anglais : Peters et Menn, 1993 ; pour le français : Veneziano et Sinclair, 1997 ; Kilani-Schoch, 1998) ont en effet proposé que dans une première période du développement langagier, les fillers seraient purement phonologiques (ou pré-morphologiques), et ne seraient donc que la manifestation de la sensibilité de l'enfant aux propriétés segmentales et rythmiques de la langue en cours d'acquisition. Les fillers seraient dans un premier stade le reflet d'une correspondance davantage phonologique que morphosyntaxique avec les caractéristiques de l'input. Veneziano et Sinclair (1997) ont indiqué par exemple qu'en français - langue dans laquelle la syllabe finale d'un mot lexical est la plus saillante du point de vue prosodique - l'inventaire des voyelles produites dans les premiers éléments préfixés de leur corpus était le même que celui des voyelles produites devant la position finale, et faisant partie du même mot ou d'un morphème grammatical précédant l'élément lexical. Plus tard, l'analyse du corpus montre que l'enfant reprend dans les fillers les voyelles des morphèmes grammaticaux précédant le plus souvent les mots lexicaux cibles. C'est, d'après Peters (2001), ce type de contraste qui permet au chercheur de différencier la phase pré-morphologique de la phase proto-morphologique.

Tremblay signale, dans son étude, que dans les recherches précédemment mentionnées, la méthode d'analyse employée ne permet pas de définir clairement si au-delà des aspects phonologiques, les premiers fillers ne sont pas également associés à des propriétés morphosyntaxiques. C'est pourquoi elle propose d'observer à quelle fréquence les fillers sont produits dans des contextes où ils répondent aux critères phono-prosodiques de la langue mais pas à ses critères syntaxiques. De par leur incompatibilité (le plus fréquemment) avec la présence de mots fonctionnels, quatre positions sont choisies par l'auteure pour l'analyse : devant les noms propres et les adverbes monosyllabiques, entre les adjectifs et les noms communs monosyllabiques ou entre les noms communs et les adjectifs monosyllabiques, et enfin, entre les quantifieurs et les adverbes monosyllabiques. L'analyse du corpus - celui de Max, enfant francophone observé de 1;9 à 3;2 mais dont seules les douze premières séances de 1;9 à 2;3 ont été conservées pour cette étude - montre que le nombre de fillers produits dans ces positions est nul ou négligeable. Pour l'auteure, ces résultats impliquent que Max, dès les premières étapes de

son développement langagier, serait sensible aux positions dans lesquelles la production d'un mot fonctionnel peut être attendue ou non. Dans la mesure où l'auteure ne relève pas de fillers dans toutes les positions pouvant être phono-prosodiquement attendues en français et non attendues du point de vue syntaxique mais bien seulement dans des positions qui font sens syntaxiquement, les fillers ne seraient donc pas seulement des remplisseurs phonologiques mais répondraient aux contraintes syntaxiques de la langue.

Les résultats de l'analyse de Tremblay ne valident pas davantage le scénario préalablement envisagé par Bottari *et al.* (1993/94) qui voudrait que les fillers soient des marqueurs syntaxiques que l'enfant ne peut pas produire en raison de lacunes au niveau segmental ou phono-prosodique. En ce qui concerne l'argument segmental, Tremblay montre que le système phonologique de l'enfant lui permettrait de produire les phonèmes des formes adultes de déterminants. L'asymétrie observée entre les fillers et les premières syllabes des mots bisyllabiques - qui ne devrait pas être attendue dans le cas d'un simple déficit phonologique - remet en question l'hypothèse d'une acquisition tardive de la représentation phonologique des déterminants chez les enfants, en raison de leur caractère non accentué.

Les résultats de l'analyse segmentale de Tremblay sont confirmés par les récents travaux de Yamaguchi et Rialland (2012, 2013) qui ont observé le parcours d'acquisition des consonnes dans deux corpus longitudinaux d'enfants francophones et montré que les phonèmes consonantiques étaient acquis plus tardivement dans les mots grammaticaux que dans les mots lexicaux. En ce qui concerne spécifiquement les formes pré-nominales, le décalage de plusieurs mois dans le moment d'acquisition des consonnes attendues dans les déterminants laisse donc là encore à penser que si les fillers ne sont pas produits conformément à la cible adulte, ce n'est pas en raison d'un déficit phonologique. Yamaguchi *et al.* (en préparation) montrent par ailleurs que les réalisations des consonnes dans les mots grammaticaux, et donc dans les fillers, ne sont pas phonologiquement proches des consonnes cibles dans les morphèmes grammaticaux adultes. D'après les auteures, ces résultats suggèrent que dans une première période au moins, les fillers consonne-voyelle ne sont ni des morphèmes adultes, ni des approximations des formes adultes mais des entités qui partagent certaines caractéristiques phonologiques propres à la catégorie. Ces régularités phonologiques pourraient être interprétées comme les premières manifestations de la construction d'un proto-paradigme des formes pré-nominales et pré-verbales.



A mi-chemin entre les deux scénarios précédents - phonologique et morphologique - Tremblay propose quant à elle que dès les premiers stades de développement langagier, les fillers puissent avoir une valeur morphosyntaxique et qu'ils ne seraient pas produits de façon aléatoire mais bien de façon systématique et dans des positions syntaxiquement contraintes. D'après l'auteure, les fillers pourraient donc, par leur nature proto-morphologique, nous éclairer sur la façon dont les jeunes enfants parviennent à assimiler les critères syntaxiques de la langue, à attribuer une représentation phonologique à ces morphèmes grammaticaux et à les associer avec un ensemble donné de traits sémantiques (personne, nombre, genre).

Nous avons pu voir précédemment (section 2.3) que cette perspective est partagée dans un certain nombre d'études mais à l'inverse des conclusions de Tremblay, toutes ne concluent pas que les fillers sont d'emblée, dès leur apparition dans le langage de l'enfant, empreints d'une valeur morphosyntaxique, et considèrent par ailleurs que leur statut n'est pas figé mais au contraire qu'il évolue à mesure que le degré de développement langagier de l'enfant augmente. Il est du reste possible d'envisager que les conclusions de Tremblay (2005) puissent être liées à l'âge et au stade de développement langagier de l'enfant. Rappelons que cette étude est fondée sur l'analyse d'une période de six mois chez un seul enfant de 1;9 à 2;3. L'auteure précise que l'analyse n'est pas poursuivie au-delà de 2;3 (MLU : 3,3) parce que le déterminant est acquis et que les fillers ont disparu des productions de l'enfant mais le recueil ayant débuté à 1;9 (MLU : 1,1), nous pouvons nous interroger sur le statut des formes produites avant cette date.

Cette étude ne clôt donc pas le débat sur le statut à attribuer aux fillers, sujet restant encore l'objet de discussions et de controverses, mais elle nous semble particulièrement intéressante dans sa façon de convoquer des arguments et des outils issus de la phonologie pour discuter des conclusions précédemment dressées dans la littérature et statuer sur la valeur morphologique ou non des fillers. L'apport de la phonologie pour la connaissance des fillers, mais aussi des déterminants, ne se limite pas aux seuls travaux que nous venons de mentionner. D'autres auteurs s'y sont intéressés mais en se focalisant cette fois particulièrement sur l'influence de la structure et de la hiérarchie prosodiques de la langue en cours d'acquisition. Une synthèse de ces travaux est présentée dans la section qui suit (section 3.1.2).

### 3.1.2 Structure prosodique de la langue en cours d'acquisition

Certaines recherches sur l'acquisition de la phonologie et en particulier sur l'acquisition phonologique des déterminants (notamment, Demuth, 2001 ; Demuth et Tremblay, 2008 ; Lleó et Demuth, 1999) ont montré que les premiers mots de l'enfant tendent à se conformer à la structure prosodique la plus fréquente dans la langue à laquelle ils sont exposés, tant en termes de rythme que de structure syllabique. Les premiers travaux sur des langues à accent comme l'anglais ou l'allemand (*i.a.* Demuth, 1996 ; Gerken, 1994, 1996) reposent sur l'hypothèse que les enfants apprenant une langue dont la structure est trochaïque - dont l'accent est mis le plus souvent sur la syllabe initiale du mot - auront tendance à omettre les syllabes produites hors de ce cadre, comme par exemple, la syllabe initiale dans les mots de trois syllabes lorsque celle-ci n'est pas accentuée. Appliquée à l'acquisition des déterminants, cette hypothèse implique que dans des langues aux structures trochaïques, la production des déterminants, formes non accentuées, serait contraire au schéma phono-prosodique préférentiel de la langue ; ces morphèmes seraient donc particulièrement sujets à des omissions dans les premières productions de l'enfant. Et en effet, les résultats de ces études montrent que les enfants âgés de moins de deux ans tendent à faire l'omission des syllabes faibles - et notamment des déterminants - lorsqu'elles sont produites devant des syllabes accentuées.

Comme le signale Bassano (2010), en français, la structure rythmique repose davantage sur la syllabe que sur l'accent - qui n'existe pas ou peu - et lorsque certains segments sont accentués, il s'agit de façon préférentielle de la dernière syllabe des mots et non de la première. La structure prosodique est donc en français iambique et non trochaïque. La sensibilité des enfants à cette structure pourrait, d'après de nombreux auteurs (Bassano, 2010 ; Tremblay et Demuth, 2007 ; Veneziano et Sinclair, 2000), se manifester par un effet de la longueur du nom sur la production des formes pré-nominales. Tremblay et Demuth (2007) font l'hypothèse qu'en français, les déterminants seraient, dans une première étape du développement, produits davantage devant des noms monosyllabiques que multisyllabiques. Une réalisation privilégiée des déterminants devant des noms monosyllabiques traduirait la préférence des enfants pour une structure iambique et binaire, majoritaire en français. Les résultats de la recherche de Tremblay et Demuth (*ibid.*) - menée sur deux enfants, suivis de 1;5 à 2;5, montrent que les premiers déterminants sont en effet produits le plus souvent devant les mots d'une seule syllabe. Les résultats montrent par

ailleurs une asymétrie significative entre les segments 'déterminant + nom de deux syllabes' et 'déterminant + nom de trois syllabes', ce qui suggère pour les auteures que les enfants passeraient par une étape intermédiaire avant la maîtrise de la production des déterminants devant des noms de trois syllabes. Malgré des différences individuelles dans le taux de production des déterminants, les deux enfants suivent un parcours développemental similaire : les déterminants sont d'abord produits devant des noms monosyllabiques, puis bisyllabiques et enfin trisyllabiques.

Ces résultats sont en accord avec les travaux précédents de Veneziano et Sinclair (2000) qui ont également constaté dans le cadre d'un suivi longitudinal d'une enfant âgée de 1;3 à 2;2, que les fillers étaient davantage produits devant des noms monosyllabiques que multisyllabiques, mais aussi avec ceux de Tremblay (2006) qui observe - là aussi dans le cadre d'un suivi longitudinal - que les déterminants sont produits plus tôt et en plus grand nombre devant des noms monosyllabiques. Ces conclusions sont partagées par Hilaire, Régol et Jisa (2002) qui montrent que chez des enfants au développement langagier tardif, et équipés d'implants cochléaires, les premiers déterminants sont produits majoritairement devant des noms monosyllabiques. Ces tendances sont également confirmées par Bassano (2010) qui compare les données d'enfants de trois groupes d'âge - 20, 30 et 39 mois - et montre qu'à 20 mois, 50% des noms monosyllabiques sont produits précédés de fillers et de déterminants, contre 16% seulement des noms multisyllabiques ; et à 30 et 39 mois, bien que l'écart entre les noms monosyllabiques et multisyllabiques se resserre, les noms monosyllabiques restent les plus fréquemment précédés de formes pré-nominales.

Cette préférence pour la structure syllabique majoritaire de la langue a également été signalée dans d'autres langues. Certains travaux menés sur le néerlandais (Fikkert, 1994) ou l'anglais (Smith, 1973) montrent que des mots cibles en CVC peuvent être produits par l'enfant suivis d'une voyelle, soit CVCV, ce qui leur permettrait, d'après Demuth (2001), d'éviter la production d'une coda ou de répondre à la structure syllabique préférentielle de la langue. De la même façon, en espagnol, beaucoup de mots sont tronqués et réduits à deux syllabes (Demuth, 2000).

Ces nombreux constats témoignent d'une sensibilité de l'enfant aux préférences prosodiques, et notamment syllabiques, de la langue, mais aussi, d'après Demuth (*ibid.* 2000, 2001), Tremblay et Demuth (2007) ou Pepinsky *et al.* (2001), de l'influence de la hiérarchie prosodique (Selkirk,

1984 ; Nespors et Vogel, 1986) sur les premières productions de l'enfant. Pour ces auteurs, les énoncés de la langue sont organisés en différents niveaux d'unités prosodiques (cf. schéma 3 ci-dessous), allant de la more ('Mora')<sup>12</sup> à l'énoncé phonologique ('Phonological Utterance').

**Schéma 3 - La hiérarchie prosodique (Pepinsky *et al.*, 2001 : 582)**

Utt - Phonological Utterance	I think Sue likes the dolly
IP - Intonational Phrase	Sue likes the dolly
PP - Phonological Phrase	likes the dolly
PW - Phonological Word	the dolly
Ft - Foot	dolly
$\sigma$ - Syllable	doll
$\mu$ - Mora	do

Ainsi, en français, le syntagme nominal constitué d'un déterminant et d'un nom est inclus dans un syntagme phonologique ('Phonological Phrase'), le déterminant étant prosodifié directement au niveau du syntagme phonologique et le nom au niveau du mot phonologique ('Phonological Word'). Le mot phonologique se décomposerait en pieds ('Foot'), constitué de deux syllabes. Dans l'hypothèse proposée par les auteures, les formes produites par l'enfant révèlent une prosodification du déterminant différente de celle de l'adulte. Au lieu d'être prosodifiée directement au niveau du syntagme phonologique, la forme pré-nominale<sup>13</sup> est prosodifiée au niveau du pied binaire. Il y a donc coïncidence chez l'enfant entre syntagme phonologique et mot phonologique, composé préférentiellement d'un pied iambique. Cette licence prosodique du déterminant par le pied binaire expliquerait donc que l'enfant produise préférentiellement des fillers ou des déterminants devant les noms monosyllabiques et omette ces formes devant les

<sup>12</sup> La 'more' est définie dans la littérature (cf. par exemple, Selkirk, 1984 ; Nespors et Vogel 1986 ; Demuth, 1996) comme une unité phonologique plus fine que la syllabe, et figurant donc au dernier niveau de la hiérarchie prosodique.

<sup>13</sup> On remarque que dans cette hypothèse, la distinction entre filler et déterminant n'est pas pertinente, il s'agit pour les auteurs d'expliquer la présence ou l'absence d'une forme, mais pas le type spécifique de forme.

noms bisyllabiques. L'acquisition des niveaux supérieurs de la hiérarchie prosodique expliquerait que progressivement, les enfants puissent produire des formes pré-nominales devant des mots contenant plus d'une syllabe et cessent de les omettre devant les mots bisyllabiques.

Nous l'avons vu plus haut (section 2.4.1), d'après Demuth (2001) ou Pepinsky *et al.* (2001), beaucoup de différences individuelles dans les premières productions de l'enfant pourraient s'expliquer au moyen de ces différents niveaux de la hiérarchie prosodique et en particulier par le niveau par lequel l'enfant entrerait dans le langage. Pour beaucoup d'enfants, le mot est l'unité prosodique préférentielle lors des premières acquisitions mais d'autres se concentrent à un niveau plus élevé sur le syntagme phonologique ('Phonological Phrase'), l'énoncé phonologique ('Phonological Utterance') ou le groupe intonatif ('Intonational Phrase').

Ainsi, d'après Demuth (2001), une approche prosodique des premières productions de l'enfant apporterait donc :

« un cadre pour comprendre non seulement la variation individuelle observée dans l'usage des premiers fillers et leur disparition au fil du temps, mais également un moyen de comprendre pourquoi certaines syllabes et morphèmes sont omis du langage précoce de l'enfant : les deux phénomènes sont gouvernés par les contraintes prosodiques en jeu dans la grammaire de l'enfant à un temps donné » (2001 : 248, *notre traduction*).

## 3.2 Influence du contexte lexical et syntaxique

### 3.2.1 Propriétés lexico-sémantiques du nom

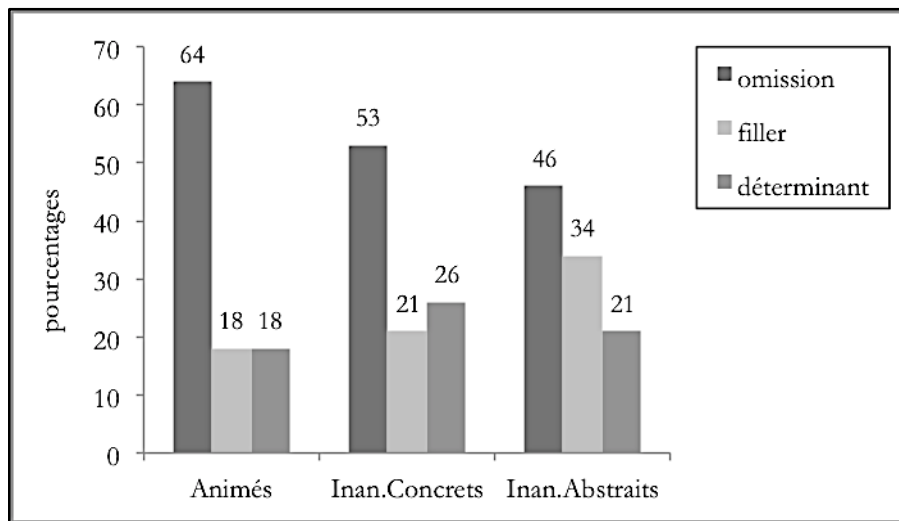
Dans plusieurs de ses travaux, Bassano (2000 ; 2005) fait non seulement l'hypothèse d'une dépendance entre développement lexical et développement grammatical en ce qui concerne la mise en place des déterminants (*cf.* point 2.2.2 de ce chapitre), mais également l'hypothèse d'une dépendance entre propriété lexico-sémantique<sup>14</sup> des noms et degré de grammaticalisation, qu'elle appelle « amorçage lexico-sémantique de la grammaticalisation » (Bassano, 2000). D'après elle, « certains types sémantiques de noms pourraient être plus propices que d'autres à former le substrat des processus de grammaticalisation » (Bassano, 2005). Elle a donc distingué les noms renvoyant à des entités animées de ceux renvoyant à des entités inanimées, et précisé la catégorie

---

<sup>14</sup> Dans le souci d'alléger la lecture, nous aurons recours aux expressions « statut sémantique du nom » ou « propriété sémantique du nom » pour référer au statut sémantique ou à la propriété sémantique des entités auxquelles renvoient les noms.

'inanimés' en traitant les inanimés abstraits et concrets de façon distincte. Dans le corpus analysé (corpus de Pauline), l'enfant omet plus fréquemment le déterminant avec les noms animés (omission à 64%) et à l'inverse, la proportion de déterminants produits est la plus forte avec les inanimés concrets (production d'un déterminant à 26%). Il semblerait donc pour Bassano que la catégorie 'inanimés concrets' serait la catégorie privilégiée pour l'acquisition de la contrainte d'utilisation du déterminant. En ce qui concerne les fillers, on constate en observant le graphique ci-dessous (graphique 1) qu'ils sont, à l'inverse des déterminants, proportionnellement plus produits devant des noms désignant des entités inanimées abstraites que des entités inanimées concrètes. Ces chiffres ne sont cependant pas discutés par Bassano dans le cadre de cette publication.

**Graphique 1 - Procédures d'emploi des noms dans les sous-classes sémantiques de noms (corpus de Pauline) (d'après Bassano, 2005 : 10)**



Dans le cadre de sa thèse, Nashawati (2010) a testé elle aussi l'influence des propriétés lexico-sémantiques des noms sur la production des formes pré-nominales, chez trois enfants francophones de 1;11 à 3;0 ans. Elle a pu constater, comme Bassano, que les déterminants apparaissent préférentiellement devant des noms désignant des entités inanimées mais à l'inverse, elle n'a pas observé d'association particulière entre omissions ou fillers et type spécifique de nom. Les omissions, comme les fillers, précèdent dans son corpus des noms renvoyant aussi bien à des entités animées qu'inanimées.

### 3.2.2 Contrastes dans l'usage des formes en position pré-nominale

D'autres études se sont intéressées à l'influence des lexèmes associés aux formes pré-nominales, mais en se focalisant cette fois sur la question de la combinaison entre déterminant et nom et non plus sur celle de la propriété lexico-sémantique du nom, et en écartant de leur réflexion les fillers et les omissions pour ne prendre en compte que les formes adultes de déterminants.

En 1986, Valian avance, dans son étude sur l'acquisition de six catégories syntaxiques chez l'enfant - déterminant, adjectif, nom, syntagme nominal, préposition, syntagme prépositionnel - que les enfants sont sensibles dès leur plus jeune âge aux propriétés abstraites de la langue de leur environnement et maîtrisent les catégories grammaticales de la langue adulte très précocement.

Dans le cas spécifique des déterminants, Valian considère que la catégorie est acquise si le déterminant apparaît dans le cadre de syntagmes nominaux, devant l'adjectif ou le nom, qu'il n'est pas produit seul et n'est pas combiné avec d'autres déterminants (à l'exception de certains quantifieurs). L'analyse - qui s'appuie sur un corpus de six enfants de 2;0 à 2;5 ans (MLU entre 2,93 et 4,14) - montre d'une part que la performance des enfants dans l'usage qu'ils font des déterminants est comparable et d'autre part, qu'ils ne font pas d'erreur dans l'ordre des mots (le déterminant est produit devant le nom ou l'adjectif) et ne font pas de double emploi du déterminant qui serait grammaticalement incorrect dans la langue adulte. Seules quelques occurrences de déterminants sont produites seules et Valian relève qu'il s'agit soit de numéraux ou de quantifieurs que l'on pourrait trouver produits seuls dans la langue adulte soit de réalisations dues à des hésitations (répétitions, tâtonnements ou faux départs). Toutes les conditions d'emploi du déterminant étant respectées par les enfants, l'auteure conclut donc à une maîtrise de la catégorie 'déterminant'.

Dix ans après cette étude, Pine et Martindale (1996), puis Pine et Lieven (1997), remettent en question les conclusions de Valian et particulièrement, l'idée d'une catégorie 'déterminant' abstraite donnée d'emblée à l'enfant. En se focalisant sur l'opposition définis / indéfinis, Pine et Lieven ont cherché à observer si les premières productions de l'enfant étaient marquées par un taux important d'overlap<sup>15</sup> (ou 'recouvrement contextuel'). Pour ce faire, ils ont sélectionné au

---

<sup>15</sup> L'overlap est un outil utilisé dans la littérature pour mesurer la capacité de l'enfant à transférer certains morphèmes grammaticaux - déterminants mais aussi pronoms - d'un contexte lexical ou syntaxique à un autre.

sein d'un corpus de onze enfants anglophones (âgés de 1 à 3 ans) toutes les occurrences de syntagmes nominaux définis et indéfinis et observé si dans les cas où les mêmes noms étaient produits à plusieurs reprises (chez le même enfant), la forme pré-nominale était identique ou variable d'une occurrence à l'autre. Les résultats montrent un taux très faible de variation dans les cas où le nom est répété, au contraire il semblerait que les enfants tendent à employer les noms avec des formes pré-nominales spécifiques. D'après les auteurs, cet usage figé qui est fait des articles 'the' et 'a' témoigne de la capacité limitée des enfants à transférer les déterminants d'un contexte lexical à l'autre.

Ces résultats sont en accord avec l'approche théorique des auteurs qui considèrent - rappelons le (*cf.* point 3.2 du chapitre précédent) - que les catégories linguistiques sont construites graduellement par l'enfant, sur la base de son expérience linguistique avec son environnement et de ses capacités d'apprentissage et de traitement de l'information.

Signalons par ailleurs que ces résultats concordent également avec ceux de l'étude de López Ornat (2003) déjà mentionnée aux points 2.2 et 2.4. Dans son corpus de 23 enfants hispanophones âgés de 1;4 à 1;10, l'auteure a pu observer que pour un même nom, la même forme pré-nominale est utilisée dans 87% des cas, ce qui suggérerait là aussi que les enfants tendraient dans un premier temps à associer les noms à des déterminants spécifiques.

Suite au travail de Pine et Lieven (1997), Valian *et al.* publient en 2009 les résultats d'une étude sur un corpus de 21 enfants anglophones de 1;10 à 2;8 en interaction avec leur mère. Six mesures d'overlap différentes ont été testées sur le corpus, dont l'alternance entre article indéfini 'a' et article défini 'the' précédemment mesurée par Pine et Lieven. Les auteurs étendent par ailleurs leur analyse à tous les types de déterminants et à la comparaison des séquences 'déterminant + nom' produites chez les enfants et chez les mères. Les résultats montrent d'une part que les erreurs dans les productions de déterminants représentent moins de 1% des occurrences chez les enfants et d'autre part que lorsque les noms sont produits au moins deux fois, ils le sont pour plus de la moitié avec des déterminants différents. Lorsqu'ils sont produits au moins six fois, le déterminant est différent dans 80% des cas. Les auteurs en concluent donc que plus les opportunités de variation sont nombreuses, plus le taux d'overlap est important. Au regard de cette étude, les résultats antérieurs de Pine et Lieven (*ibid.*) faisant état d'un taux de recouvrement contextuel très faible étaient probablement dus à un biais d'échantillonnage.



Pine *et al.* (2013) reviennent sur les critiques méthodologiques adressées par Valian *et al.* en 2009 et dressent une liste de trois critères à respecter pour mesurer le taux d'overlap dans les corpus. Ils proposent de restreindre l'analyse aux noms qui peuvent être précédés à la fois de 'the' et 'a(n)', qui sont utilisés au moins deux fois par un locuteur donné et enfin, aux noms qui sont produits à la fois par l'enfant et l'adulte. Les auteurs constatent que lorsque ces critères sont respectés, les enfants présentent un taux d'overlap (31%) significativement moins important que les adultes (47%). D'après Ambridge et Lieven (2015), ces résultats vont à l'encontre des travaux considérant que les catégories 'déterminant' et 'nom', ainsi que la connaissance sur la façon dont ils se combinent, sont présents chez l'enfant dès la naissance.

Dans le cadre d'un travail sur les facteurs distributionnels et pragmatico-discursifs en jeu dans l'acquisition des définis et des indéfinis, Salazar *et al.* (2013) ont également soumis leur corpus - combinant les données transversales et longitudinales de 24 enfants de 1;10 à 3;0 ans - à une mesure du taux d'overlap en position pré-nominale. A la différence des études précédemment exposées, les auteurs ont pris en compte dans leur analyse les déterminants produits par les enfants mais aussi les formes de transition - omissions et fillers.

A condition de considérer ce contexte plus général de fluctuation au sein de la position pré-nominale, incluant formes adultes et formes de transition, les données de Salazar Orvig *et al.* montrent d'une part que globalement, les phénomènes de fluctuation - entre plusieurs formes de fillers, entre fillers et omissions ou encore entre fillers, omissions et déterminants - sont plus fréquents que les répétitions du même syntagme. Les résultats montrent d'autre part que le taux de contraste entre plusieurs formes augmente avec le degré de développement linguistique ; plus le MLU est important, plus les contrastes en position pré-nominale sont fréquents. Les données laissent cependant à voir qu'à l'exception d'une partie du corpus<sup>16</sup>, ces contrastes se manifestent rarement au sein de la même séance, mais bien plutôt d'une séance à l'autre. Sur les quatre groupes MLU observés, seuls 3,33 à 25% des noms sont concernés par un contraste au sein de la même session, et le taux de répétition du même nom, qu'il soit précédé d'un déterminant, d'un filler ou d'une omission reste relativement stable quelque soit le groupe MLU. Les auteurs relèvent également que pour tous les groupes, le degré de variation individuelle est conséquent. A

---

<sup>16</sup> La partie du corpus concernée est constituée d'un suivi longitudinal d'ateliers de langage en classe de maternelle, systématiquement construits autour du même personnage.

titre d'exemple, sept enfants du corpus ne présentent pas d'overlap du tout, et cela toutes sessions confondues.

Ces auteurs se sont par ailleurs focalisés sur le contraste spécifique entre définis et indéfinis et ont montré que lorsque ces déterminants sont produits fréquemment, particulièrement au sein des groupes MLU les plus avancés (3 et 4), les contrastes avec le même nom sont très rares puisqu'ils ne représentent pas plus de 5% du total des noms. Les résultats montrent là encore une grande variabilité individuelle ; certains enfants ne produisent pas du tout de contraste entre les deux types de déterminants, et d'autres en produisent mais pas pour toutes les séances. Enfin, le contraste entre définis et indéfinis tend à augmenter entre le premier groupe MLU et le troisième mais diminue ensuite radicalement entre le troisième et le dernier, lorsque le MLU devient supérieur à 3. D'après les auteurs, les données ne suggèrent donc pas une association entre développement linguistique et déploiement du paradigme des déterminants, et l'usage des formes s'expliquerait davantage par les contextes discursifs dans lesquels celles-ci sont produites. Nous y reviendrons à la section 3.3 de ce chapitre.

### **3.2.3 Rôle du cadre lexico-syntaxique**

Nous avons pu voir dans la section précédente que le faible taux d'overlap constaté par Pine et Lieven dans leur étude (1997) les amenait à conclure à un usage relativement figé des déterminants et lié à des lexèmes spécifiques. Dans cette étude, les auteurs constatent en effet que les déterminants ne sont pas produits dans des cadres lexico-syntaxiques très diversifiés ; à l'inverse, ceux-ci tendent à être produits dans un nombre de contextes ou *patterns* très limité. En moyenne, 56% des énoncés produits par les enfants et comprenant des articles définis ou indéfinis impliquent un des trois patterns les plus fréquemment produits au sein de chaque corpus. Chez neuf enfants sur onze, ces patterns sont produits dans plus de la moitié des énoncés dans lesquels on trouve des définis ou des indéfinis.

**Tableau 2 - Pourcentage de production des déterminants dans les trois patterns les plus fréquents (d'après Pine et Lieven, 1997 : 132)**

Sujets	Patterns dans lesquels sont produits les déterminants	Productions de déts. comptabilisées (%)
Anna	<i>a + X ; in the + X ; in a + X</i>	60.0
Alan	<i>a + X ; on the + X ; the + X or in the + X</i>	56.0
Charles	<i>a + X ; in the + X ; want a + X</i>	52.2
Helen	<i>that's a + X ; a + X ; like a + X</i>	62.5
Irwin	<i>a + X<sup>a</sup></i>	61.5
Karen	<i>a + X ; want a + X ; that's a + X or in the + X or on the + X</i>	35.1
Laura	<i>a + X ; in the + X ; where's the + X</i>	52.1
Leonard	<i>a + X ; there's a + X or that's a + X or the + X</i>	61.5
Lionel	<i>a + X ; want a + X<sup>b</sup></i>	54.5
Margaret	<i>a + X ; get a + X ; want a + X</i>	41.1
Molly	<i>a + X ; where's the + X ; there's the + X</i>	79.2
Mean		56.0

<sup>a</sup> Seul cadre avec plus d'une occurrence

<sup>b</sup> Seulement deux cadres avec plus d'une occurrence

Ces contextes de production - présentés dans le tableau que nous avons reproduit ci-dessus - peuvent se limiter à des syntagmes nominaux tels que '*a* ou *the + X*' ou à des syntagmes prépositionnels ('*in/on a/the + X*') mais aussi constituer des contextes syntaxiques plus étendus tels que '*get a + X*', '*want + X*' ou '*that's a + X*'.

D'après les auteurs, ces observations suggèrent donc que les déterminants sont d'abord produits dans le cadre d'associations spécifiques avec certains éléments lexico-syntaxiques, avant que s'étende progressivement la gamme de cadres dans lesquels les déterminants apparaissent.

Dans ce contexte, les auteurs signalent que le recours aux outils de la grammaire à pivots proposée par Braine (1963) semble plus pertinent pour analyser les premières combinaisons que l'utilisation des catégories syntaxiques adultes, qui ne sont pas encore acquises par l'enfant. Ces conclusions concordent avec un certain nombre d'études constructivistes sur le développement grammatical précoce de l'enfant (notamment, Tomasello, 1992).

Nous avons pu voir dans les pages précédentes qu'un grand nombre des études de la littérature cherchant à expliquer la mise en place du déterminant reposent sur l'analyse des contraintes distributionnelles à l'œuvre dans les premières productions des jeunes enfants, qu'elles soient phono-prosodiques, sémantiques ou lexico-syntaxiques. D'autres études se sont quant à elles

focalisées davantage sur l'influence des contextes pragmatico-discursifs au sein desquels sont produits fillers et déterminants, afin notamment d'expliquer la variation observée dans les corpus en position pré-nominale. Nous tenterons dans la prochaine section (3.3) de faire un état des lieux de ces recherches.

### **3.3 Production des formes pré-nominales et contextes pragmatico-discursifs**

#### **3.3.1 Influence du contexte discursif sur la production des proto-morphèmes**

Les différentes études qui se sont intéressées aux fillers et à l'évolution de leur rôle dans les énoncés de l'enfant se sont appuyées pour leurs analyses sur des propriétés formelles ou grammaticales du langage de l'enfant. On trouve cependant dans la littérature quelques travaux orientés vers la dimension pragmatico-discursive de l'usage des fillers (notamment, Salazar Orvig et Morgenstern, 2002). Inscrites dans une perspective interactionniste, ces études considèrent, comme nous l'avons vu à la section 2.3 du premier chapitre, que la grammaire ne doit pas être envisagée comme un système autonome, déconnecté du processus de communication et qu'au contraire, l'enfant utilise d'abord le langage à des fins communicationnelles et mobilise la grammaire en conséquence. Dès lors, l'acquisition de la grammaire apparaît comme la conséquence de l'usage du discours, et non comme son but premier. De la même manière que les proto-noms sont utilisés au sein d'activités référentielles et les proto-verbes à des fins prédicatives, n'impliquant pas pour autant la maîtrise du statut grammatical de nom ou de verbe (*cf.* par exemple Veneziano, 2003), on peut en effet s'interroger sur le rôle et le fonctionnement des éléments proto-grammaticaux qui leur sont associés. Comme l'ont montré certaines études sur lesquelles nous reviendrons à la section suivante (*i.a.* Gundel *et al.*, 2007 ; Salazar Orvig *et al.*, 2005, 2013), l'enfant fait très précocement preuve de compétences pragmatiques et de sensibilité au contexte discursif et attentionnel lors de la production de ses énoncés, notamment en ce qui concerne le maniement des expressions référentielles, et particulièrement des déterminants.

Dans la continuité de ces constats, Salazar Orvig et Morgenstern ont été à l'initiative d'une série de travaux (Salazar Orvig et Morgenstern, 2002 ; Salazar Orvig *et al.*, 2008 ; Morgenstern, 2009b) ayant cherché à expliquer la fluctuation en position pré-nominale par des facteurs pragmatico-discursifs. Leurs travaux se sont concentrés sur l'analyse de deux corpus longitudinaux, ceux de Daniel et de Léonard, âgés entre 1;8 et 2;4. Pour chacun des corpus, les auteures ont relevé tous

les proto-noms contenus dans les énoncés de l'enfant, en étant attentives d'une part à la place du référent dans le dialogue et d'autre part à la présence ou à l'absence de fillers ou de déterminants. Leurs résultats ont montré que, selon le corpus observé, les premiers fillers étaient dans un cas associés à des référents déjà mentionnés, la première mention n'étant accompagnée d'aucune forme pré-nominale (dans le corpus de Léonard) et dans l'autre cas, que la première mention était marquée par la production d'un filler et les mentions suivantes par celle d'une forme nulle (dans le corpus de Daniel). Le propre des fillers étant de ne pas posséder de valeur conventionnelle dans le proto-système de l'enfant, la variation dans les usages qu'en font les deux enfants dans les études de Salazar Orvig et de Morgenstern pourrait s'expliquer par le fait qu'ils ne leur attribuent pas la même valeur et créent chacun, pour une courte période, leur propre système de correspondance forme-fonction.

Dans les deux cas, les fillers ne sont pas produits de façon aléatoire mais en alternance avec les omissions et plus tard avec les formes adultes, dans le cadre de contrastes référentiellement significatifs, qui pourraient, d'après les auteures, témoigner d'une sensibilité précoce de l'enfant à la dimension discursive :

« Si les observations faites étaient confirmées par d'autres travaux et d'autres données, il serait envisageable de penser que si les fillers possèdent un statut proto-morphologique, ce n'est pas uniquement parce qu'ils occupent une position syntaxique ou présentent des propriétés structurelles similaires à celles des déterminants adultes mais bien plutôt parce que l'enfant y a recours pour marquer les fonctions discursives et pragmatiques des expressions nominales. Dans ce cas, nous pourrions supposer que la valeur pragmatique et discursive précède les valeurs grammaticales ou syntaxiques. Les fillers seraient par conséquent une voie d'accès aux déterminants non seulement en tant qu'élément syntaxique obligatoire mais surtout comme un indicateur de la valeur pragmatique ou discursive du nom » (Salazar Orvig *et al.*, 2008 : 74, *notre traduction*).

Ces observations ont été ultérieurement confirmées dans le cadre du travail de Salazar Orvig *et al.* (2013) sur l'acquisition des définis et des indéfinis chez 24 enfants francophones de 1;10 à 3;0 ans. Les auteurs ont en effet constaté qu'au stade où les indéfinis ne sont pas encore produits de façon systématique, les enfants ont recours à des fillers ou à des omissions - et non à des définis - dans les contextes où des indéfinis seraient attendus (dans des contextes non référentiels notamment). Là encore, on relève donc que les premiers contrastes dans les productions des formes pré-nominales sont pertinents du point de vue référentiel.

Des constats similaires ont été faits par Nashawati (2010) dans le cadre de sa thèse. Elle a pu observer que les fillers étaient associés à des objets de discours au statut référentiel spécifique. Dans le corpus de Clémence (de 2;2 à 2;7), les fillers sont davantage produits dans des cas de référenciations secondes, et dans le corpus de Mathilde (de 2;1 à 2;7), dans des cas de référenciations premières. Les résultats de ce travail reflètent donc de nouveau l'importance de la variation individuelle dans l'usage des fillers.

Dans un article de 2009, Heurdier s'est aussi intéressé à cette question du statut pragmatico-discursif des fillers. A travers l'observation d'exemples de productions de deux enfants âgées de 2;3, il montre que les fillers peuvent aussi bien apparaître avec des référents nouveaux que donnés dans le discours, et que les contextes discursifs dans lesquels ils sont produits peuvent varier d'un enfant à l'autre. Dans un cas, suite à des questions partielles de l'adulte, l'enfant a tendance à produire des omissions ou des fillers et dans l'autre, les fillers semblent être produits davantage dans le cadre de réintroductions à distance des référents.

Le recours à des outils d'analyse pragmatico-discursifs nous laisse donc à penser que la fonction à attribuer aux fillers est davantage soumise à des tendances individuelles qu'à des usages discursifs spécifiques et partagés par tous les enfants. Ce constat vient donc étayer l'idée que les fillers sont des morphèmes en construction, qui ne sont pas encore conventionnalisés et dépendent donc du proto-système que chaque enfant aura construit à un temps donné de son développement. Par ailleurs, les premiers contrastes constatés par ces travaux dans l'utilisation des formes pré-nominales indiquent que, comme pour les morphèmes adultes, le contexte discursif semble influencer le choix des formes produites par l'enfant.

D'autres études, qui se sont cette fois focalisées davantage sur les formes produites dans le discours que sur les caractéristiques pragmatico-discursives des référents auxquelles elles sont associées, ont montré que l'enfant est également sensible aux formes pré-nominales produites par son interlocuteur, et que pour sa part, l'interlocuteur adapte ses productions aux capacités de l'enfant. Bertin (2011) a par exemple travaillé dans le cadre de sa thèse sur le rôle de l'interaction entre adulte et enfant dans la mise en place de la morphologie et traité notamment la question de l'influence des reprises sur la production des déterminants, chez cinq enfants francophones entre 1 et 3 ans. Elle a pu observer que le processus n'était pas en tous points comparable chez tous les enfants, mais qu'il était tout de même possible de dessiner un parcours général similaire chez les

cinq enfants du corpus. Elle a pu noter d'une part que ces enfants s'appuient davantage sur les productions et les reprises ou reformulations de l'adulte au cours des premières étapes de leur développement morphologique, et que progressivement, lorsque les enfants commencent à s'approprier les formes adultes, ils se détachent des offres de l'adulte. Ce détachement n'est cependant pas linéaire chez tous les enfants et peut au contraire évoluer sous la forme d'allers retours. Elle a pu relever d'autre part que l'adulte s'ajuste aux tâtonnements de l'enfant en lui offrant des séquences adaptées à son développement, étayantes ou potentiellement acquisitionnelles, et ce sont ces séquences qui favoriseraient la mise en place des formes adultes de déterminants chez l'enfant.

### **3.3.2 Une sensibilité précoce aux fonctions pragmatico-discursives des déterminants ?**

Si l'influence du contexte discursif sur la production des proto-morphèmes a été relativement peu étudiée, on trouve dans la littérature un grand nombre de travaux sur l'acquisition des fonctions pragmatico-discursives des déterminants. Deux questions ont été soulevées à de nombreuses reprises dans ces travaux : i) à quel âge les enfants maîtrisent-ils les différentes fonctions pragmatiques qui peuvent être portées par les déterminants, et ii) comment acquièrent-ils les associations forme-fonction de la langue ? Les réponses apportées à ces questions divergent : pour certains auteurs (Gundel *et al.*, 2007 ; Rozendaal et Baker, 2008 ; Salazar Orvig *et al.*, 2013), morphologie et pragmatique ne se développent pas de façon isolée mais conjointe - le développement pragmatique pouvant même précéder le développement des formes - et la plupart des valeurs référentielles des déterminants sont acquises très tôt, à 3 ans environ. Pour d'autres auteurs (*i.a.* Karmiloff-Smith, 1979), la plurifonctionnalité des déterminants n'est pleinement acquise que tardivement, entre 8 et 12 ans, cette perspective impliquant donc que les formes soient acquises avant certaines fonctions. Pour Coene (2005) par exemple, les premiers déterminants définis que l'enfant produit là où seraient parfois attendus des indéfinis sont ce qu'elle appelle des 'imposteurs', produits sans valeur référentielle. Les premiers déterminants portant une valeur pragmatique sont les indéfinis produits au moment où le taux de définis 'imposteurs' diminue.

Les résultats des études sur l'acquisition de la référence chez l'enfant - et en particulier de l'opposition entre définis et indéfinis - se déclinent entre travaux expérimentaux et travaux sur

des données naturelles. Les études basées sur des données recueillies en situation naturelle s'accordent à dire - pour la plupart - que les enfants font précocement preuve de maîtrise des combinaisons, ou d'une partie des combinaisons forme-fonction que nous avons évoquées au point 1.3.3 de ce chapitre. Cet aspect de la littérature sur la maîtrise des usages des définis et indéfinis en situation naturelle sera développé au point 3.3.2.2. Avant cela, nous nous intéresserons (au point 3.3.2.1) aux résultats des études menées en situation expérimentale, qui nous le verrons, obtiennent des résultats plus contrastés sur cette question.

De la même façon que dans la première section de ce chapitre, nous faisons le choix ici de limiter cette revue de travaux à l'acquisition de l'opposition entre définis et indéfinis, d'une part parce qu'il s'agit de l'opposition la plus documentée dans la littérature, de par son caractère fondamental pour l'établissement et le maintien de la référence en discours, et d'autre part, parce que c'est ce contraste qui nous intéressera plus particulièrement lorsque nous analyserons les usages des déterminants dans la troisième partie de ce travail.

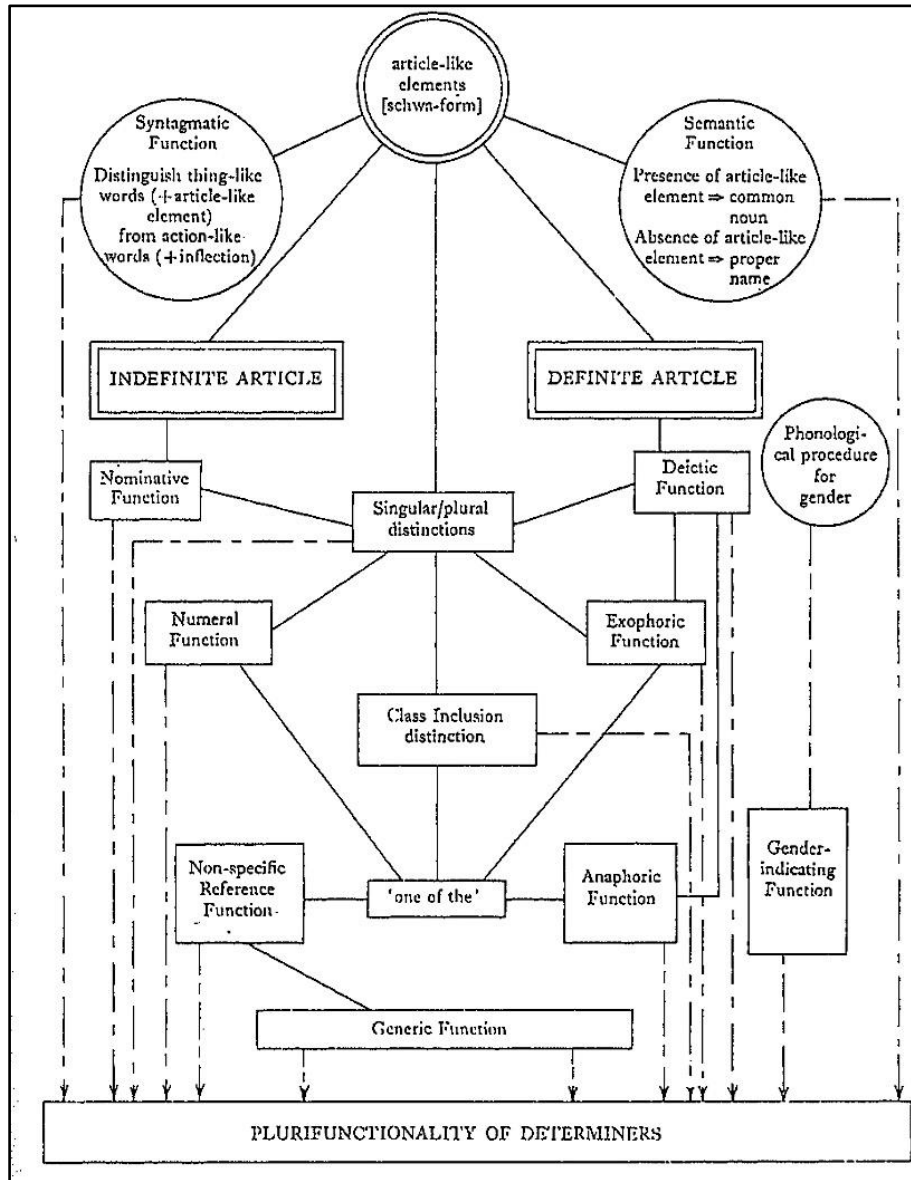
### *3.3.2.1 Analyse des usages des définis et indéfinis en situation expérimentale*

En 1976, Maratsos constate dans le cadre de tâches de compréhension et de production chez des enfants de 3 et 4 ans que même les enfants de 3 ans sont capables de différencier définis et indéfinis en fonction du caractère spécifique ou non spécifique des référents ; mais il note également que certains usages du défini au lieu de l'indéfini révèlent que la prise en compte de l'interlocuteur est plus tardive.

Quelques années plus tard, en 1979, Karmiloff-Smith publie une étude sur l'acquisition des déterminants chez des enfants francophones de 3 à 12 ans, en focalisant son analyse sur l'acquisition de leur polyvalence fonctionnelle. Elle prend en considération dans son analyse les facteurs cognitifs nécessaires au développement linguistique d'une part, et les facteurs spécifiquement linguistiques impliqués dans l'acquisition d'autre part (notamment le fait que les entités linguistiques puissent avoir plusieurs fonctions, et qu'une fonction puisse être réalisée par plusieurs formes). Elle met en évidence dans son travail l'émergence de fonctions telles que le marquage du genre et du nombre, de la référence générique, spécifique, ou encore anaphorique. L'ordre d'acquisition des différentes fonctions de l'indéfini et du défini est synthétisé par le schéma que nous avons reproduit ci-dessous (schéma 4).



Schéma 4 - Ordre d'acquisition de la plurifonctionnalité des déterminants (Karmiloff-Smith, 1979 : 215)



Karmiloff-Smith constate, dans son étude, qu'après une période transitoire de production de proto-formes, les définis et les indéfinis semblent se développer de façon simultanée, et qu'ils sont rarement intervertis par les enfants : « [...] the articles are not initially confused. Each one has a clearly defined, separate function for the small child » (Karmiloff-Smith, 1979 : 219). Dans une première étape, l'article indéfini serait utilisé pour la dénomination et le défini - souvent accompagné d'un pointage - pour sa fonction déictique, pour attirer l'attention de l'interlocuteur. On retrouve cette idée dans les travaux de Bronckart (1977) pour qui les définis « servent

essentiellement à attirer l'attention sur l'élément désigné par le nom » (Bronckart, 1977 : 293) et les indéfinis « explicitent en quelque sorte l'opération de désignation » (*ibid.*). Dans la même lignée, Hickmann (2003) propose également que le recours aux indéfinis pour l'introduction des référents chez les enfants de 4 ans soit lié au processus de dénomination, et non à la prise en compte du savoir de l'interlocuteur, facteur qui n'interviendrait que plus tardivement dans la production des déterminants.

En parallèle du développement respectif des fonctions de dénomination et déictique de l'indéfini et du défini, la distinction entre singulier et pluriel se met en place. Vient ensuite l'acquisition du genre des déterminants au singulier, qui, d'après Karmiloff-Smith, ne dérive pas des articles mais se baserait essentiellement sur la terminaison des mots (d'où son statut différencié sur le schéma). Un peu plus tard, en même temps que se développe la fonction numérale de l'indéfini, l'enfant acquiert la fonction exophorique du défini. Une fois cette fonction acquise, le défini est produit lorsqu'un objet est seul et l'indéfini pour faire référence à un objet parmi plusieurs objets identiques. Lors de l'étape suivante se met en place, d'après l'auteur, l'usage de l'indéfini pour la référence non spécifique et du défini pour la référence anaphorique. En dernier lieu, l'enfant acquiert la fonction de référence générique, qui serait la plus difficile à maîtriser : « the acquisition data collected in this study allow us to hypothesize nonetheless that the strength ontogenetically of the deictic function of the definite article, as well as the deictic component of the anaphoric function, must make it very difficult for the child to use the definite article generically » (Karmiloff-Smith, 1979 : 222).

Suite à l'analyse des données des différents groupes d'âge de son étude, Karmiloff-Smith établit trois étapes de développement dans l'acquisition de la plurifonctionnalité des déterminants. La première se situe autour de 3 à 5 ans. A cet âge, les enfants produisent un grand nombre de déterminants et à première vue, les constructions enfantines sont très élaborées du point de vue de l'usage qui est fait des déterminants. Or, d'après l'auteure, une analyse plus fine permet d'établir que les enfants n'auraient pas à ce stade conscience de la plurifonctionnalité des déterminants. Ce n'est que dans un second temps, de 5 à 8 ans environ, que ces compétences se développent. Durant cette période, l'enfant utilise un morphème pour chaque fonction qu'il souhaite exprimer. Un mot équivaut alors à une seule signification. Durant cette même période, l'auteure relève la création de formes légèrement agrammaticales dans le but de distinguer les

deux fonctions d'un même mot. La distinction entre les fonctions descriptive et déterminative est faite par l'enfant, mais elle reste difficile à marquer linguistiquement. Il en est de même pour la référence anaphorique : le défini n'est pas toujours produit mais d'autres marqueurs sont utilisés en compensation. On peut en conclure que l'enfant de 5 à 8 ans prend conscience du statut plurifonctionnel des différentes formes de déterminants, contrairement à celui de 3 à 5 ans qui utilise un grand nombre de morphèmes mais seulement de manière unidimensionnelle. Ces conclusions peuvent être appuyées par l'observation des corrections spontanées des enfants à la fin de cette deuxième période. La dernière phase couvrirait la période de 8 à 12 ans. Le statut plurifonctionnel des déterminants est alors acquis et le système des déterminants est organisé.

Enfin, Karmiloff-Smith observe que - malgré l'antériorité du développement en compréhension dans certains contextes - la tendance générale est globalement la même qu'il s'agisse de compréhension ou de production, le développement de la maîtrise des différentes fonctions des déterminants s'avérant, dans ses données, similaire dans les deux cas.

Les conclusions faites par Karmiloff-Smith sur l'acquisition tardive des différentes fonctions des déterminants sont partagées avec celles d'autres travaux, comme ceux de Kail et Hickmann (1992). Ces auteures ont cherché à observer chez des enfants francophones de 6, 9 et 11 ans à quel âge le marquage des référents nouveaux était conditionné par le critère de savoir partagé entre l'enfant et son interlocuteur. Impliqués dans une tâche de narration, les enfants sont confrontés à deux types de situations : la narration à un interlocuteur qui a accès à l'imagier proposé aux enfants (situation de connaissance partagée), et la narration à un interlocuteur aux yeux bandés (situation de connaissance non partagée). Les résultats de l'étude montrent d'une part que dans tous les groupes, les indéfinis sont plus nombreux dans les cas où l'interlocuteur ne partage pas le savoir de l'enfant, et d'autre part que les usages sont différents d'un groupe d'âge à l'autre. A 6 ans, les enfants produisent en situation de connaissance non partagée aussi bien des indéfinis que des définis ; à 9 ans, les indéfinis sont produits dans le contexte où l'interlocuteur ne partage pas la connaissance de l'enfant, et les définis dans la situation où la connaissance est au contraire partagée ; et enfin, à 11 ans, les indéfinis sont produits non seulement lorsque l'interlocuteur n'a pas de visibilité sur l'album, mais aussi dans le cas où il partage la lecture avec l'enfant.

D'après les auteures, ces résultats montrent que l'enfant fait une acquisition des règles d'introduction des référents en trois étapes : dans un premier temps, aucune règle systématique n'est appliquée en l'absence de connaissance partagée, dans un deuxième temps, l'enfant produit des déterminants appropriés au contexte, et dans une troisième étape, il fait l'acquisition d'une règle supplémentaire et généralise la production des indéfinis à toutes les situations de narration qui le requièrent, qu'importe le degré de connaissance partagée.

La question soulevée dans cet article de Kail et Hickmann (1992) est reposée quelques années plus tard par Hickmann *et al.* (1996). Les auteurs ont comparé le marquage des référents nouveaux en anglais, en français, en allemand et en chinois chez des enfants de maternelle, chez des enfants de 7 ans et 10 ans ainsi que chez des adultes. Les résultats montrent que - bien que les outils pour référer à ce qui est nouveau dans le discours diffèrent d'une langue à l'autre - les constats quant au degré de maîtrise de la référence sont les mêmes : le marquage systématique du caractère nouveau d'un référent émerge tardivement, autour de 7 ans.

Pour Schaeffer et Matthewson (2005), les erreurs relevées dans la production des déterminants, qui pourraient indiquer un manque de prise en compte des connaissances de l'interlocuteur par l'enfant, seraient liées à un système pragmatique lacunaire. Les enfants n'auraient pas encore la compétence leur permettant de distinguer leurs propres croyances de celles de leur interlocuteur. Inscrites dans le cadre théorique de la Grammaire Universelle, les auteures considèrent que tous les principes de la grammaire sont acquis par l'enfant et que les paramètres sont mis en place dès que l'input pertinent pour les acquisitions de l'enfant est accessible. Dès lors, les éléments du langage de l'enfant s'écartant de la grammaire de la langue cible seraient le résultat de l'immaturité de leur système pragmatique.

De Cat (2011) s'est elle aussi penchée sur les productions de déterminants non attendues et leur maîtrise, que certains auteurs considèrent comme tardive, des contrastes entre définis et indéfinis. Elle a notamment proposé que dans les cas où l'enfant produit un défini lorsque le référent a déjà été mentionné et dans ceux où il associe au nom un défini pour renvoyer à un référent nouveau dans le discours, l'enfant ne serait pas confronté à des restrictions de compétences d'ordre linguistique mais bien plutôt d'ordre cognitif. D'après l'auteure, l'enfant aurait tendance à considérer que dans des contextes de narration, l'interlocuteur partage avec lui un savoir commun plus vaste que ce que pourrait envisager un adulte. Les difficultés rencontrées

par l'enfant dans les associations forme-fonction seraient ainsi davantage liées à un manque de compétences globales dans l'activité de narration.

Dans un travail daté de 2013, De Cat s'est à nouveau intéressée aux erreurs dites 'égocentriques' (autrement dit, aux contextes dans lesquels l'enfant produit un défini lorsque le référent est nouveau) chez 45 enfants francophones âgés de 2;3 à 5;6 environ. Deux tâches de narration ont été proposées aux enfants et les erreurs repérées au sein de ces narrations ont ensuite été confrontées aux résultats d'un test de théorie de l'esprit. Aucune corrélation significative n'a été observée entre les usages non attendus de définis et les scores au test, ceci suggérant que ces erreurs ne peuvent pas seulement être dues à des facteurs cognitifs et que d'autres facteurs seraient à prendre en considération pour pleinement les comprendre et les expliquer. De Cat relève par exemple que dans plusieurs cas, les erreurs des enfants seraient la « manifestation [d'une] tendance (probablement inconsciente) à compter sur leur interlocuteur pour trouver l'information dans le contexte visuel » (2013 : 67, *notre traduction*).

Les études menées en contexte expérimental que nous venons de citer considèrent globalement que l'enfant ne prend en compte que tardivement les connaissances de son interlocuteur, et que de ce fait, les déterminants ne peuvent pas être pleinement acquis avant 7 à 10 ans. On trouve cependant dans la littérature un certain nombre d'études, conduites elles aussi en situation expérimentale, qui montrent que les enfants de 3 ans choisissent et interprètent les expressions référentielles en se fondant sur le degré de connaissance de l'interlocuteur. Si ces études ne portent pas spécifiquement sur la maîtrise des déterminants, mais plus largement sur le choix des expressions référentielles en fonction du contexte, elles peuvent néanmoins nous éclairer sur la question des compétences pragmatiques requises pour intégrer les différentes fonctions des déterminants.

Dans une étude de 2005, Wittek et Tomasello ont analysé le choix des expressions référentielles fait par deux groupes d'enfants germanophones de 2;5 et 3;5 à qui un expérimentateur posait trois types de questions : certaines impliquant que celui-ci connaissait l'objet sur lequel portait la question, d'autres supposant que l'adulte ne connaissait pas l'objet auquel il était fait référence, et enfin, une série de questions indiquant que l'expérimentateur se faisait une idée erronée de l'objet dont il était question. Les auteurs ont fait l'hypothèse que dans le cas d'une prise en compte de la connaissance de l'interlocuteur, l'enfant ferait un usage contrasté des formes référant à l'objet de

discours : pronom ou forme nulle dans le premier cas, nom ou pronom dans le deuxième et nom dans le dernier cas de figure. En accord avec cette hypothèse, les résultats montrent en effet qu'au travers de la question posée par l'adulte, l'enfant peut appréhender le degré de familiarité de l'expérimentateur avec l'objet et y répond en conséquence.

Dans la lignée du travail de Wittek et Tomasello, Matthews *et al.* (2006) ont eux aussi testé l'impact de l'accessibilité - visuelle ou discursive - du référent pour l'interlocuteur sur le choix des expressions référentielles fait par l'enfant. Les auteurs ont travaillé sur un corpus de 100 enfants monolingues anglophones répartis en trois groupes de 2, 3 et 4 ans, et mené deux expérimentations. Les résultats de la première étude montrent que les enfants de 3 et 4 ans font un usage contrasté des expressions référentielles (noms *vs.* pronoms) selon que l'interlocuteur a visuellement accès ou non au référent. A l'inverse, aucun contraste n'a été relevé dans les productions des enfants de 2 ans. Les résultats de la seconde étude indiquent clairement que les groupes d'enfants de 3 et 4 ans ont recours à des expressions référentielles différenciées en fonction du caractère nouveau ou donné du référent dans le discours de l'interlocuteur. Les données montrent par ailleurs que les réponses des enfants de 2 ans impliquent davantage de noms lorsque le référent n'a pas été mentionné au préalable. Ces résultats suggèrent, d'après les auteurs, que bien que le contraste entre ce qui est donné et nouveau pour l'interlocuteur est complexe et relativement long à maîtriser, les enfants semblent faire cette distinction plus tôt lorsqu'elle est basée sur des indices discursifs que lorsqu'elle requiert la compréhension d'indices perceptuels.

Une étude de O'Neill (1996) est également régulièrement citée dans la littérature qui s'intéresse aux compétences pragmatiques des jeunes enfants. Dans ce travail, l'auteure s'est focalisée non sur les productions verbales mais sur les gestes réalisés par les enfants en fonction de la connaissance que l'interlocuteur a du référent. Les résultats mettent en évidence qu'à 2 ans, les enfants semblent peu sensibles au savoir de l'interlocuteur mais que quelques mois plus tard, à 2;5, ils se montrent très influencés par le fait que l'objet de discours soit nouveau ou donné pour l'adulte. Les enfants ont en effet fait référence à l'objet de l'expérimentation avec plus de gestes - et souvent, une désignation verbale - lorsque le parent avait été absent au moment où l'objet avait été introduit à l'enfant et qu'il n'en connaissait donc pas la localisation.

Si ces trois études n'ont pas spécifiquement expérimenté l'influence de la familiarité de l'interlocuteur avec le référent sur le choix des déterminants par l'enfant, elles impliquent néanmoins que l'on s'interroge sur les résultats de ces travaux, considérant qu'à défaut de prise en compte de l'interlocuteur, la plurifonctionnalité des définis et indéfinis ne peut être acquise à 3 ans.

Les études sur la référence menées sur des données expérimentales sont très nombreuses et les quelques résultats des études proposés ici ne constituent qu'un état des lieux non exhaustif des travaux existants dans la littérature. Les conclusions listées ici nous permettent cependant de constater que pour une part de ces travaux, la référence est acquise très tôt (à 3 ans environ) et que pour d'autres, certaines fonctions des déterminants, telle que le marquage de la référence spécifique et non spécifique, semblent acquises à 3 ans et d'autres fonctions (marquage de ce qui est nouveau ou donné dans le discours, prise en compte de l'interlocuteur) sont acquises plus tardivement. Dans cette perspective, la plurifonctionnalité des déterminants n'est acquise que progressivement.

### *3.3.2.2 Analyse des usages des définis et indéfinis en situation naturelle*

Les travaux pionniers de Brown (1973) sur l'acquisition de la référence en anglais ont notamment montré qu'à 3 ans, Adam, Eve et Sarah - les trois enfants de son corpus longitudinal - faisaient un usage attendu des indéfinis dans des contextes de dénomination et de référence non spécifique. Il signale néanmoins que les contrastes entre indéfinis et définis sont moins nets dans les contextes d'introduction des référents et de maintien de la référence.

Dans plusieurs de ses travaux, Gundel s'est elle aussi interrogée sur la maîtrise de la référence portée par les déterminants définis et indéfinis (ainsi que par les démonstratifs et les pronoms) chez des enfants en cours d'acquisition de l'anglais et de l'espagnol de 1;5 à 3;5 (Gundel et Page, 1998, Gundel *et al.*, 1999, Gundel *et al.*, 2001). Les conclusions de ces différentes études montrent que chez les enfants anglophones comme hispanophones, tout le champ des expressions référentielles est utilisé de façon appropriée à 3 ans (ou plus tôt). En ce qui concerne les déterminants, les enfants utilisent de façon attendue les indéfinis pour référer à ce qui est nouveau et les définis pour référer à ce qui est donné dans le discours.

Des constats similaires sont faits par Salazar Orvig *et al.* (2005) qui ont pu observer, sur des données à la fois longitudinales et transversales recueillies en situation naturelle, un contraste très clair dans l'usage des définis et des indéfinis par les enfants. Les définis sont utilisés pour des référents déjà mentionnés par l'enfant et/ou son interlocuteur et les indéfinis pour des référents mentionnés pour la première fois dans le dialogue. Certains définis sont néanmoins produits en première mention, mais dans la grande majorité des cas, pour faire référence à des éléments connus de l'interlocuteur. À l'inverse, les enfants ont recours à l'indéfini lorsque la connaissance du référent n'est pas partagée par l'interlocuteur. Pour les auteurs, les déterminants seraient bien plurifonctionnels dès les premières productions, et l'enfant ferait l'acquisition des aspects pragmatiques et grammaticaux de la langue de façon simultanée. L'enfant ne ferait pas l'acquisition des formes dans un premier temps et l'acquisition des fonctions de ces formes dans un second temps, mais bien plutôt l'acquisition des formes pour répondre à des besoins de marquage d'oppositions référentielles ou discursives.

Afin de contrôler l'influence possible des associations entre nom et déterminant (voir Pine et Lieven, 1997 ; *cf.* section 3.3), Salazar Orvig *et al.* proposent quelques années plus tard (Salazar Orvig *et al.*, 2013) une nouvelle analyse des fonctions associées aux déterminants, en prenant en compte le taux d'overlap et les usages des noms au sein de la dynamique du discours et du dialogue. Les résultats montrent que les indéfinis sont produits dans des contextes d'usages non référentiels des noms et de premières mentions des référents, et les définis dans des contextes beaucoup plus diversifiés : usages non référentiels, premières mentions, réactivations du référent, réponses. Par ailleurs, lorsque les noms sont produits pour développer un thème, ils sont associés à un défini ; les indéfinis ne remplissant pas cette fonction. Ces résultats suggèrent donc à nouveau que l'enfant associerait rapidement les déterminants à des fonctions spécifiques et que ces formes seraient utilisées d'emblée de façon plurifonctionnelle.

Dans la lignée des travaux menés par l'équipe de Salazar Orvig *et al.*, Nashawati (2010) explore dans le cadre de sa thèse l'usage des noms et des pronoms dans des dialogues mère-enfant et constate elle aussi que les syntagmes nominaux indéfinis servent à référer à ce qui est nouveau et introduit pour la première fois dans le discours et qu'une fois connus et partagés dans le discours, les référents sont repris par un pronom ou un syntagme nominal défini. Ce contraste



dans l'usage des syntagmes nominaux définis et indéfinis marque pour Nashawati la capacité de l'enfant à distinguer informations nouvelles et données.

Des conclusions similaires ont été faites dans le cadre des travaux de Rozendaal et Baker (2008), ainsi que par Rozendaal (2008), déjà cités dans la première section de ce chapitre. Dans l'objectif de comparer l'usage des déterminants chez l'adulte et chez l'enfant, les auteures ont répertorié cinq fonctions pragmatiques différentes observées dans l'input : la dénomination, la référence non spécifique, la référence à un élément déjà mentionné dans le discours, la référence à un élément absent du discours mais connu des locuteurs, et enfin, la référence à un élément absent à la fois du discours et des connaissances partagées des locuteurs. L'analyse des combinaisons forme-fonction chez l'enfant présente des tendances similaires dans les trois langues observées (français, anglais, néerlandais). A partir de 2;3, les auteures remarquent des associations entre déterminants et fonctions pragmatiques nettement comparables à celles de l'input. Les indéfinis sont très largement utilisés pour la dénomination, ainsi que pour la référence non spécifique. Ces fonctions peuvent aussi parfois être remplies par le défini et le démonstratif, mais surtout chez les enfants francophones, ce qui pourrait s'expliquer, d'après les auteures, par des usages plus fréquents de définis et de démonstratifs dans des contextes de référence non spécifique au sein de l'input français. Les déterminants définis et démonstratifs sont le plus souvent associés aux référents qui ont déjà été mentionnés dans le discours, mais on les trouve également dans des contextes où le référent n'a pas été mentionné, et cela, y compris lorsque la connaissance du référent n'est pas partagée entre les interlocuteurs. Les indéfinis sont utilisés de façon significativement moins fréquente pour les référents donnés dans le discours que pour les référents nouveaux, quel que soit l'âge observé. A 2;3, la distinction dans l'usage des formes en fonction des connaissances partagées ou non entre les locuteurs n'est pas perceptible, mais les auteures suggèrent qu'à 3 ans le contraste entre les deux est en cours d'acquisition, malgré quelques emplois non attendus du défini pour des référents inconnus. D'après les auteures, ces résultats suggèrent que très tôt, les enfants font preuve d'une capacité à distinguer les fonctions pragmatiques des déterminants - quant à ce qui est spécifique/non spécifique ainsi que donné/nouveau dans le discours - et à associer formes et fonctions des déterminants.

Au-delà des associations appropriées que font les enfants entre formes et fonctions et que nous venons de mentionner, Rozendaal (2008) relève que l'acquisition de la référence n'est pas liée à

l'âge mais au moment d'apparition des formes. Les valeurs référentielles des formes sont, semble-t-il, maîtrisées par l'enfant dès le moment où ces formes sont disponibles dans sa langue, qu'il s'agisse des déterminants ou des pronoms. Les enfants francophones sont ceux qui acquièrent les déterminants le plus tôt et témoignent d'une sensibilité plus importante au point de vue de l'interlocuteur. Enfin, de même que dans l'article que nous citons plus haut, Rozendaal signale l'importance considérable des combinaisons forme-fonction de l'input dans l'acquisition de la référence.

Les résultats de Rozendaal et Baker (2008) et ceux de Rozendaal (2008) concordent avec un certain nombre d'études sur des enfants bilingues, parmi lesquelles celle de Serratrice (2000), qui a travaillé dans le cadre de sa thèse sur l'acquisition des catégories fonctionnelles chez des enfants en cours d'acquisition de l'anglais et de l'italien, et a pu constater une différenciation précoce entre usages spécifiques - marqués par le défini - et non spécifiques - marqués par l'indéfini - des déterminants. Des conclusions semblables ont été faites par Kupisch (2006) chez des bilingues allemand-italien et allemand-français. D'après l'auteure, le contraste entre spécifique et non spécifique est acquis très tôt et les usages qui sont faits des déterminants par les enfants sont appropriés au contexte. Le défini est cependant produit dans des contextes où le référent n'est pas connu de l'interlocuteur et où un indéfini serait attendu. Ces erreurs dites 'égocentriques' témoignent, d'après l'auteure, du caractère plus tardif de l'acquisition du trait de 'familiarité' ; la découverte de ce trait pourrait coïncider, d'après Kupisch, avec celle de la production systématique des déterminants devant les noms.

Hickmann *et al.* (2015) rappellent en partie les conclusions qui viennent d'être exposées puisque les auteurs affirment qu'une part des résultats porte à croire qu'à 3 ans, l'enfant maîtrise pleinement la référence. Peu d'erreurs sont relevées dans les corpus, et les enfants semblent faire preuve de sensibilité au savoir de l'interlocuteur, puisque des marques des connaissances présumées de ce dernier sont relevées dans leur discours. *A contrario*, les auteurs signalent que d'autres résultats montrent que dans des contextes de production spontanée et de récits, les enfants de moins de 3 ans ne font pas nécessairement la distinction entre définis et indéfinis dans les contextes de première mention de référents nouveaux et que l'utilisation des déterminants dans des contextes de maintien de la référence n'est véritablement stabilisée que chez les enfants plus âgés. Contrairement aux travaux de Brown (1973) ou de Maratsos (1976) qui suggèrent que

les enfants font des usages appropriés de l'opposition entre définis et indéfinis pour marquer le caractère spécifique/non spécifique et nouveau/connu du référent, les travaux menés par Hickmann (notamment, 2003) dans des contextes de narrations laissent à penser que si ces déterminants sont produits de façon précoce, la maîtrise de l'opposition entre 'nouveau' et 'donné' n'est acquise que tardivement, autour de 7, voire 10 ans. Chez les enfants plus jeunes, si certains usages peuvent être appropriés, c'est, d'après Hickmann *et al.* (2015), en raison de stratégies reflétant des principes généraux du discours selon lesquels certains types d'entités (entités animées ou principales), certains rôles (agent, sujet) et certains contextes discursifs (coréférence thématique) contribuent à la production d'unités linguistiques marquant le maintien de la référence. Seuls les enfants plus âgés seraient capables de repérer certains contextes plus complexes de continuité référentielle, tels que la réintroduction du référent dans des contextes non coréférentiels ou le changement thématique dans des contextes coréférentiels.

Les conclusions concernant la maîtrise des fonctions des déterminants par l'enfant sont donc contradictoires. Au sein des travaux menés en situation naturelle comme en situation expérimentale, les auteurs ne s'accordent pas sur le degré de maîtrise des différentes fonctions pragmatiques des définis et des indéfinis chez le jeune enfant (à 3 ans environ) et les fonctions considérées comme maîtrisées à cet âge ne sont pas les mêmes chez tous les auteurs.

Plusieurs différences de résultats obtenus dans les études que nous venons de mentionner aux points 3.3.2.1 et 3.3.2.2 pourraient, d'après Gundel *et al.* (2007), s'expliquer par des différences méthodologiques. Il pourrait en effet être envisagé que la situation - naturelle *vs.* expérimentale - et que l'activité en cours et les genres de discours déployés dans l'activité puissent constituer un biais non négligeable dans ce que l'enfant donne à voir de ses compétences (Salazar Orvig *et al.*, à paraître).

Par ailleurs, on peut aussi penser que certaines différences pourraient provenir de la comparaison faite par les chercheurs entre ce qui est attendu chez l'enfant et ce qui l'est chez l'adulte. Dans un travail comparant les productions des syntagmes nominaux définis et indéfinis chez 25 enfants de 3;1 à 5;8 ans et 6 adultes, Van Hout *et al.* (2009) ont ainsi mis en évidence que l'usage non attendu du défini chez les enfants pour renvoyer à des référents non mentionnés pouvait s'expliquer par l'accessibilité visuelle des référents concernés. A l'inverse de ce qui est observé chez les adultes, le contexte visuel pourrait, d'après les auteurs, suffire à créer chez l'enfant un

antécédent à un syntagme nominal défini, et rendre ainsi le référent déjà donné pour l'enfant. Ces conclusions - qui font écho au travail de Matthews *et al.* (2006) que nous avons précédemment mentionné (section 3.3.2.1) - permettent donc à nouveau de penser que les résultats proposés par certaines études - et particulièrement celles concernant l'utilisation de supports visuels - pourraient être reconsidérées à la lumière de ces questionnements à la fois théoriques et méthodologiques.

#### **4. Enjeux d'une étude sur l'acquisition du paradigme des déterminants**

La revue de littérature effectuée au sein de ce chapitre nous permet de constater que le processus d'acquisition des déterminants fait l'objet d'une littérature très riche, sur laquelle nous pourrions nous appuyer pour préparer l'analyse des données et l'interprétation des résultats de cette thèse. Il nous semble néanmoins que certains aspects de ce processus restent à confirmer, voire à explorer.

Nous avons notamment pu observer que certains travaux, qui tentent d'expliquer l'émergence des déterminants en ne prenant en compte qu'un seul facteur, ne semblent rendre compte que partiellement des productions enfantines, soit parce qu'à un temps du développement de l'enfant, toutes les productions pré-nominales ne concordent pas avec l'hypothèse proposée par les auteurs, soit parce que cette hypothèse ne couvre pas toutes les périodes du développement. Par ailleurs, nous avons pu relever qu'au sein de chacune des approches développées dans la littérature, l'observation des données ne convergeait pas systématiquement vers les mêmes résultats et conclusions. Nous avons également pu remarquer que tous les auteurs qui se sont intéressés à cet aspect du développement langagier n'ont pas pris en considération les omissions et les fillers, négligeant ainsi certains indices fondamentaux et certaines étapes dans l'émergence des formes adultes. Enfin, très peu d'études ayant porté sur les éventuelles fonctions pragmatico-discursives portées par les fillers, cette interface entre formel et fonctionnel reste à explorer pour pouvoir compléter la description de ces proto-formes.

#### 4.1 Insuffisance des approches monodimensionnelles

Nous avons, en effet, pu constater que peu d'études ont analysé de façon croisée les facteurs susceptibles d'influencer la production des premières formes produites par l'enfant en position pré-nominale. Or, les résultats de certains travaux nous permettent de penser que l'observation séparée de l'influence de chaque facteur ne semble pas expliquer de façon suffisamment précise et systématique le processus d'émergence des déterminants.

A titre d'exemple, si l'approche phono-prosodique développée par Demuth - qui fait l'hypothèse, rappelons le, que l'enfant tend à produire des formes pré-nominales devant les noms monosyllabiques et à l'inverse, à omettre ces formes devant des noms bisyllabiques - semble expliquer, d'après la littérature, une grande partie des données, on peut cependant s'interroger sur la présence dans les productions enfantines de syntagmes nominaux combinant un filler ou un déterminant et un nom de deux syllabes, ou encore, l'omission d'une forme pré-nominale devant des noms d'une seule syllabe (Le Mené et Yamaguchi, 2016). Ces observations suggèrent qu'à elle seule l'hypothèse phono-prosodique ne permet pas d'expliquer toutes les productions des enfants.

Cette remarque est d'ailleurs valable pour toutes les hypothèses développées dans la littérature. Dans la mesure où seules quelques études ont confronté sur les mêmes corpus plusieurs de ces hypothèses, il semble difficile de savoir si un facteur influence effectivement plus que les autres l'acquisition des déterminants ou si *a contrario*, toutes les hypothèses qui ont pu être faites expliquent une part du phénomène de façon complémentaire et simultanée ou à des périodes distinctes du développement.

Signalons par ailleurs que toutes les approches ne semblent pas rendre compte du processus complet d'acquisition des déterminants. Il est par exemple avancé dans la littérature que l'analyse du facteur phono-prosodique est particulièrement pertinente lorsque l'enfant en est au stade des premières combinaisons et que les fillers sont pré-morphologiques (*cf.* par exemple Bassano *et al.*, 2008 ou Veneziano et Sinclair, 2000) ; néanmoins, toutes les études argumentant en faveur de cette influence n'expliquent pas quels facteurs peuvent ensuite expliquer que certains noms continuent au stade proto-morphologique à être précédés d'omissions et d'autres, de fillers et de déterminants.

En s'intéressant à l'influence du sémantisme des noms, Bassano *et al.* (2008) argumentent quant à eux en faveur d'une influence phono-prosodique dans une première étape du développement, et d'une influence sémantique dans un second temps, dans la mesure où la production des formes pré-nominales ne semble pas influencée par le caractère animé ou inanimé des noms dans le premier groupe d'âge (1;8), mais seulement dans le deuxième (2;6) et le troisième (3;3). Ce constat nous paraît particulièrement intéressant pour illustrer la nécessité d'une analyse multidimensionnelle qui permettrait d'expliquer l'émergence des déterminants sur toute sa durée.

On peut également mentionner le travail de Veneziano (2001, 2003, 2014b), Veneziano et Sinclair (1997, 2000) ou Kilani-Schoch (1998) qui ont elles aussi largement contribué à l'enrichissement de la littérature sur le développement grammatical du nom en français, dans une perspective multidimensionnelle. Ces auteurs ont croisé développement morphologique et contraintes prosodiques, et établi que les enfants produisent des fillers dont le statut évolue, allant du statut d'entités prosodiques à celui de précurseurs des morphèmes adultes. D'après Veneziano (notamment, 2014b), il existerait en effet plusieurs types de fillers ou plusieurs périodes dans leur développement, dont une première période, dite pré-morphologique, pendant laquelle les fillers seraient majoritairement produits pour remplir des fonctions phono-prosodiques. Le recours aux outils de la phonologie permet donc à l'auteure d'expliquer une partie de l'acquisition progressive du déterminant sans pour autant qu'il soit envisagé que le facteur phono-prosodique puisse expliquer toutes les étapes de sa mise en place.

#### **4.2. Nécessité de la prise en compte des formes de transition**

Comme nous l'avons mentionné à plusieurs reprises au sein de ce chapitre, les chercheurs ayant eu pour objectif d'expliquer la mise en place du paradigme des déterminants chez l'enfant n'ont pas systématiquement pris en considération dans leur analyse les noms produits seuls, donc les omissions et/ou les fillers. On observe par exemple dans les travaux ancrés exclusivement dans la dimension phono-prosodique que les formes pré-nominales sont envisagées comme un ensemble, sans que la distinction soit faite entre fillers et déterminants. Les analyses se focalisent en effet sur la présence ou l'absence de formes devant les noms, mais pas précisément sur le type de forme produit par l'enfant. Cette approche ne permet donc pas à elle seule d'expliquer pourquoi l'enfant produit, en alternance avec les formes adultes, des formes de transition, et

quels mécanismes à l'œuvre dans le développement langagier pourraient expliquer la diminution et la disparition de ces formes.

Le même constat peut être fait à la lecture de la plupart des études qui ont travaillé sur la dépendance entre déterminants et noms ou cadres syntaxiques. Qu'il s'agisse de l'étude de Pine et Martindale (1996), de celle de Pine et Lieven (1997) ou encore des travaux d'Ambridge et Lieven (2015), on observe en effet que seuls les déterminants sont conservés pour l'analyse, les formes de transition - omissions et fillers - n'étant pas prises en compte. Or, pour rendre compte de façon suffisamment précise de l'influence des cadres lexico-syntaxiques sur les premiers syntagmes nominaux produits par l'enfant, il semblerait pertinent de ne pas analyser exclusivement les syntagmes comprenant des formes adultes mais aussi ceux comprenant des omissions ou des fillers. L'acquisition des déterminants n'étant pas amorcée seulement lorsque les enfants deviennent capables d'en produire, cette analyse permettrait de renseigner l'influence des cadres sur la production des formes au véritable commencement du processus.

On sait désormais des nombreux travaux sur l'émergence des déterminants que cette acquisition est un processus graduel qui peut difficilement se passer du traitement des proto-formes produites avant que les déterminants adultes deviennent majoritaires. Les fillers et les omissions constituent pour beaucoup une fenêtre sur les processus d'acquisition des mots grammaticaux, et notamment des déterminants, et sur la façon dont l'enfant traite et s'approprie l'input linguistique. A la lumière des travaux publiés depuis l'étude fondatrice de Peters et Menn (1993), il paraît non seulement pertinent mais nécessaire de s'appuyer sur ces proto-formes pour comprendre le parcours que l'enfant emprunte avant d'atteindre la maîtrise des formes adultes.

### **4.3 Des résultats parfois contradictoires**

Nous l'avons déjà signalé plusieurs fois, Bassano a par exemple établi que dans ses données, les déterminants et les fillers étaient davantage produits à 2;6 et 3;3 devant des noms inanimés que devant des noms animés (*i.a.* Bassano *et al.*, 2008). Nashawati (2010) a également soumis son corpus à l'analyse de ces contraintes sémantiques, en considérant trois catégories de noms : les noms de la catégorie animée, les noms de la catégorie inanimée, et les noms d'actions et d'événements. L'auteure observe, comme Bassano, que les déterminants sont produits majoritairement devant des noms inanimés, mais elle n'observe cependant pas un emploi

privilegié des fillers pré-nominaux devant un type spécifique de nom. Par ailleurs, dans leur travail sur les déterminants, Pine et Lieven (1997) constatent que dans leur corpus, les enfants semblent utiliser des déterminants devant des noms sémantiquement très hétérogènes, ce qui leur permet de remettre en question la pertinence d'une approche sémantique des premières productions de l'enfant. Cette divergence dans les résultats obtenus rend difficile la formulation d'une conclusion indiscutable sur l'influence potentielle du sémantisme des noms sur la production des déterminants.

Il en est de même si l'on s'intéresse au débat sur l'émergence de la catégorie 'déterminant' opposant Valian et Pine et Lieven (*cf.* section 3.2.2 de ce chapitre). Dans une perspective générativiste, Valian considère que la catégorie 'déterminant' est acquise d'emblée chez l'enfant, dès les premières productions de déterminants. Dans une approche usage-based de l'acquisition du langage, Pine et Lieven considèrent pour leur part que la catégorie n'est acquise que progressivement et que les premiers déterminants sont d'abord produits dans des cadres lexico-syntaxiques spécifiques.

Dans le cadre de la section 3.3.2 de ce chapitre, nous avons également eu l'occasion de constater que les résultats des études menées sur l'acquisition des fonctions pragmatiques des déterminants sont contradictoires, tous les auteurs ne situant pas la maîtrise de ces fonctions au même âge. Pour certains, une grande partie des fonctions portées par les déterminants sont acquises autour de la troisième année alors que pour d'autres, certains contrastes entre définis et indéfinis ne sont acquis que très tardivement, autour de 11 ou 12 ans.

Tant sur le plan des aspects formels que fonctionnels, les débats ne sont pas clos et ces quelques exemples de conclusions divergentes témoignent assez nettement de la nécessité de poursuivre les recherches sur l'acquisition des déterminants. Nous reviendrons sur ces différents questionnements lors de la discussion des résultats des analyses menées dans le cadre de cette thèse.

#### **4.4 Fillers et fonctions pragmatico-discursives : une interface à explorer**

Certains auteurs, considérant que l'enfant n'apprend pas à produire les formes indépendamment de leurs fonctions, se sont penchés sur les éventuelles fonctions pragmatico-discursives portées par les fillers en français. On trouve parmi ces études celles de Salazar Orvig et Morgenstern



(2002), Salazar Orvig *et al.* (2008) et Morgenstern (2009b) qui ont cherché à observer si les contextes pragmatico-discursifs dans lesquels les noms sont produits peuvent conditionner l'usage des fillers. D'après les auteurs, bien que les résultats divergent d'un enfant à l'autre - les fillers étant selon le corpus soit associés à des référents donnés ou nouveaux, et les omissions à des référents nouveaux ou donnés - les enfants semblent sensibles aux contraintes pragmatico-discursives de la langue et cela, avant même de maîtriser les morphèmes adultes.

Nashawati (2010) a elle aussi abordé cette question dans le cadre de sa thèse et a fait le même constat de variation individuelle puisque selon les enfants, les fillers étaient associés à des référents déjà mentionnés ou nouveaux dans le discours, mais elle a également pu observer que les enfants tendaient à produire ces formes de transition pour remplir des fonctions pragmatico-discursives spécifiques.

Cette interface entre fillers et fonctions pragmatiques est également abordée dans le travail de Salazar Orvig *et al.* (2013) auquel nous avons fait référence à de nombreuses reprises et qui au-delà de dresser le panorama des formes produites devant les noms et de leur évolution a confirmé l'influence de contraintes liées au discours dans la production des fillers. Les auteurs constatent en effet que dans les contextes où la production d'indéfinis aurait été appropriée - dans les contextes non référentiels par exemple - mais que ceux-ci ne sont pas produits, les noms sont précédés non pas par des définis mais par des omissions ou des fillers. Ce contraste entre définis et proto-formes suggère, d'après les auteurs, que les fillers sont, en contraste avec les omissions d'une part et avec les déterminants d'autre part, au moins en partie produits pour marquer une fonction différente du syntagme nominal dans le discours.

Ces différents travaux combinant une analyse des fillers et des contraintes fonctionnelles permettent d'élargir le panorama des facteurs susceptibles de conditionner la production des fillers et de comprendre la fluctuation observée entre omissions, fillers et déterminants en position pré-nominale. Il nous semble cependant que cette interface gagnerait à être explorée de façon plus approfondie.

En effet, si les travaux sur l'acquisition des fonctions pragmatiques des déterminants sont nombreux, rares sont ceux qui ont pris en compte les fillers. La plupart des études que nous avons eu l'occasion de citer au point 3.3.2 de ce chapitre ont pour objet l'âge auquel l'enfant maîtrise les différentes combinaisons entre déterminants et fonctions pragmatiques, ainsi que le

parcours suivi par l'enfant avant de maîtriser cette compétence, mais il n'a pas été observé si certaines des fonctions portées par les formes adultes pouvaient déjà être portées par les formes de transition. Il nous semble que les auteurs de ces études se sont intéressés aux fonctions pragmatiques des déterminants davantage pour ce qu'elles apportent comme éclairage sur la maîtrise de la référence par l'enfant que comme un facteur pouvant influencer la production des formes pré-nominales, et expliquer la présence concomitante d'omissions, de fillers et de déterminants.

Il existe donc peu de travaux sur cette question et nous avons pu voir que les conclusions apportées par les seules études existantes témoignent d'une variation individuelle importante. Les fillers ne possédant pas de valeur conventionnelle, on peut imaginer que tous les enfants ne leur attribuent pas la même fonction dans leur proto-système (Salazar Orvig et Morgenstern, 2002). On pourrait également envisager que chacun des enfants dont les productions ont été étudiées dans ces travaux se trouve à une période différente de son développement morphologique, et qu'à chaque étape puissent correspondre des contrastes distincts dans les usages qui sont faits des formes de transition et des formes adultes. Il reste donc difficile d'établir avec certitude si, avant même la maîtrise des formes adultes, les enfants se montrent déjà sensibles - au travers des usages qui sont faits des omissions et des fillers - aux relations entre formes et fonctions pragmatico-discursives.



Cette première partie a été consacrée d'une part à la présentation des principales approches qui composent le domaine de l'acquisition du langage, ainsi qu'aux choix théoriques pour lesquels nous avons optés pour cette thèse, et d'autre part, à l'état de l'art sur l'émergence des déterminants et des fillers chez l'enfant. Nous aborderons dans la partie suivante de cette thèse les questions qui ont guidé notre travail de recherche, ainsi que la méthodologie pour laquelle nous avons optée pour collecter nos données, les transcrire et les analyser.

# PARTIE II

## *Questions de recherche et méthodologie*

---

**CHAPITRE III** Questions de recherche

**CHAPITRE IV** Corpus et méthodologie



# **CHAPITRE III**

## *Questions de recherche*

---



Nous avons pu remarquer dans le cadre du chapitre précédent que les études sur l'acquisition des déterminants sont nombreuses et que celles-ci ont adopté des approches multiples pour traiter les questions relatives à l'émergence des formes adultes et de transition. De cette littérature riche et variée découlent un certain nombre de questions de recherche, sur lesquelles nous nous appuierons pour conduire les analyses qui seront ultérieurement présentées dans ce travail.

Nos premiers questionnements porteront sur la distribution et la description des formes pré-nominales dans notre corpus et les suivants sur les conditions qui pourraient expliquer leur présence/absence devant le nom. Nous avons pu constater au chapitre précédent que les travaux qui se sont penchés sur les contextes de production des premiers fillers et déterminants dans les productions enfantines ont abordé cette question au travers de prismes théoriques et méthodologiques différents. Certains auteurs (*i.a.* Bassano *et al.*, 2008 ; Demuth, 2001 ; Pine et Lieven, 1997) se sont focalisés sur des contraintes que l'on a appelées « distributionnelles », en analysant les formes produites par les enfants au vu de facteurs phono-prosodiques, sémantiques, lexicaux, ou syntaxiques. D'autres (*i.a.* Morgenstern, 2009b ; Salazar Orvig *et al.*, 2008 ; Salazar Orvig *et al.*, 2013) se sont plus particulièrement attachés à mettre en perspective les productions pré-nominales et les valeurs pragmatico-discursives des référents associés aux noms produits par l'enfant. Dans la perspective d'une approche multidimensionnelle sur le développement du paradigme des déterminants, notre réflexion sera marquée de façon complémentaire par ces deux approches, et donc guidée à la fois par des contraintes distributionnelles et fonctionnelles.

Nos questions seront donc construites autour de trois axes majeurs :

- la description des formes produites devant les noms ;

- les contraintes distributionnelles en jeu au cours du développement du paradigme des déterminants ;
- l'impact des facteurs fonctionnels sur la production des formes pré-nominales.

## 1. Description des formes

En nous appuyant sur les conclusions de travaux préalablement cités et conduits par des chercheurs comme Bassano (2010) ou Veneziano (2003), nous nous interrogerons donc dans un premier temps sur la façon dont les déterminants émergent dans le langage de l'enfant, afin d'une part de confronter nos données aux conclusions existantes, et d'autre part, de tenter d'apporter certaines réponses à des questions restant encore en suspens.

Nous chercherons à savoir à quelle période du développement langagier apparaissent les premières formes de fillers et de déterminants, et comment ces formes - ainsi que les omissions - se distribuent dans notre corpus au fil des séances. Ces premières observations nous amèneront à traiter deux questions sous-jacentes : ces formes sont-elles produites successivement par l'enfant ou conformément aux constats faits dans la littérature, de façon concomitante ? Le développement du paradigme est-il un phénomène linéaire qui laisse à voir une augmentation progressive de la part des formes adultes ou bien au contraire non linéaire et marqué par une augmentation discontinue des déterminants ?

Nous nous demanderons également comment évoluent formes de transition et formes adultes au cours du développement linguistique : est-il préférable de traiter déterminants et fillers comme des catégories uniformes quels que soient l'âge ou le degré de maîtrise langagière de l'enfant, ou d'envisager plutôt une variété de déterminants et de fillers produits entre le stade des premières productions pré-nominales et le stade auquel le déterminant adulte devient l'usage préférentiel de l'enfant ? En d'autres termes, a-t-on affaire exclusivement à une évolution dans la distribution des catégories produites devant le nom ou bien aussi à une évolution au sein de chacune de ces catégories ?

Nos observations nous permettront par ailleurs de répondre à une autre question déjà soulevée dans la littérature : la mise en place du paradigme des déterminants est-elle semblable, aussi bien



en termes de rythme que de style, chez tous les enfants de notre corpus ou bien au contraire très dépendante du fonctionnement langagier de chaque enfant ?

## **2. Contraintes distributionnelles**

Nous l'annonçons préalablement, notre deuxième axe d'analyse sera articulé autour de questions portant sur les éventuelles contraintes distributionnelles à l'œuvre dans les premières productions pré-nominales des jeunes enfants. Nous aurons recours pour cela à des outils méthodologiques empruntés aux approches phono-prosodiques, sémantiques, lexicales et syntaxiques du développement des fillers et/ou des déterminants.

### **2.1 Structure syllabique du nom**

Dans la lignée des travaux de Veneziano et Sinclair (1997), Demuth (2001), ou Tremblay (2005) décrits au sein du chapitre II, nous chercherons à voir si la structure syllabique du nom peut influencer la production des formes pré-nominales. Nous avons vu plus tôt qu'en français, le schéma prosodique préférentiel pour le syntagme nominal serait composé d'un pied binaire (et iambique) ; par conséquent, un nom monosyllabique serait précédé d'un filler ou d'un déterminant, à l'inverse d'un nom composé de deux syllabes qui serait quant à lui précédé d'une omission. Afin de savoir si l'absence ou la présence de formes devant le nom peuvent être conditionnées par des facteurs phono-prosodiques, nous nous demanderons donc quelle proportion des omissions, des fillers et des déterminants produits par l'enfant le sont devant des noms mono- et bisyllabiques.

### **2.2 Sémantisme du nom**

Nous avons par ailleurs - dans la continuité des travaux menés par Bassano (2010) ou Nashawati (2010) - choisi de nous interroger sur le rôle joué par le statut sémantique du nom sur l'emploi des formes pré-nominales. Nous chercherons en effet à définir si les productions de l'enfant peuvent être conditionnées par des critères sémantiques tels que le caractère animé/inanimé ou concret/abstrait du nom. En d'autres termes, nous nous interrogerons sur la possibilité d'une corrélation dans nos données entre un type sémantique de nom spécifique et l'emploi chez l'enfant d'une forme pré-nominale préférentielle.

## **2.3 Contextes lexicaux et syntaxiques**

Enfin, nous avons vu préalablement que les travaux menés dans le cadre des approches usage-based considèrent que l'enfant n'acquiert pas d'emblée les catégories syntaxiques adultes mais bien au contraire, les construit progressivement en prenant appui sur les usages les plus fréquents de sa langue et de celle de ses interlocuteurs. L'acquisition du déterminant est ainsi considérée comme dépendante des noms et des constructions syntaxiques produits le plus souvent par l'enfant et dans l'input qu'il reçoit (*i.a.* Pine et Lieven, 1997). Nous examinerons cette possible influence des facteurs lexicaux et syntaxiques dans nos données en nous demandant d'une part si les productions pré-nominales sont associées de façon privilégiée à des noms spécifiques, et d'autre part si elles le sont dans le cadre de constructions syntaxiques particulières.

Si des associations particulières se dessinent dans les premières productions des enfants de notre corpus, nous chercherons à voir si elles peuvent être liées à l'usage qui est fait par l'enfant et par ses interlocuteurs des lexèmes et des constructions syntaxiques à chaque étape de son développement langagier : à quelle fréquence sont-ils employés et précédés de quelles formes ?

Nous nous interrogerons ensuite sur la façon dont se développent ces usages, afin de définir notamment si la dépendance à des lexèmes ou des constructions spécifiques s'estompe pour laisser place à une généralisation des usages des formes. Signalons que si cette perspective développementale nous semble particulièrement fondamentale pour le traitement des facteurs lexicaux et syntaxiques, nous l'adopterons également lors de l'analyse des facteurs phono-prosodiques et sémantiques.

## **3. Influences des facteurs fonctionnels**

L'analyse des contraintes distributionnelles d'ordre lexical et syntaxique laisse entrevoir, au travers de la question de la récurrence des formes produites par l'enfant et par son environnement, la place du discours de l'autre et de l'interaction dans le développement langagier de l'enfant. Bien que ces aspects soient partie intégrante du paysage théorique des approches usage-based, ils ne sont pas analysés en tant que tels, mais d'autres études ont quant à elles réinvesti cette question du développement des déterminants en se penchant spécifiquement sur l'impact des facteurs pragmatico-discursifs sur la production des formes pré-nominales.

### 3.1 Référence et discours

Dans la lignée des travaux conduits par Salazar Orvig (Salazar Orvig *et al.*, 2013 ; Salazar Orvig *et al.*, 2008) nous ferons l'hypothèse que des facteurs distributionnels - qu'ils soient d'ordre phonoprosodique, sémantique, lexical ou syntaxique - ne peuvent pas à eux seuls expliquer la production des fillers et rendre compte de la mise en place et de la maîtrise des déterminants chez l'enfant. Nous chercherons par conséquent à observer si les productions de l'enfant peuvent être, avant même la maîtrise des formes adultes, associées à certains contextes discursifs spécifiques. Nous nous interrogerons pour cela sur le rôle joué d'une part par le caractère référentiel ou non référentiel de l'expression nominale, et d'autre part, dans le cas des usages référentiels, sur le type de référence, le statut attentionnel et discursif du référent, ainsi que sa familiarité pour les interlocuteurs.

### 3.2 Rôle des reprises

Par ailleurs, en nous appuyant sur les nombreux travaux sur le phénomène de reprise (*i.a.* Bertin, 2011 ; Salazar Orvig, 2000 ; Bernicot, Salazar Orvig et Veneziano, 2006) qui témoignent d'une part de l'importance capitale et de la complexité du processus de reprise dans le développement de la capacité dialogique de l'enfant et d'autre part de l'influence des reprises et plus largement de la conversation sur l'acquisition de nouvelles formes (et notamment celle des morphèmes grammaticaux associés aux noms et aux verbes), nous chercherons à voir si les reprises que fait l'enfant de son propre discours ou de celui de l'adulte peuvent influencer le développement des formes pré-nominales.

## 4. Mise en perspective des facteurs distributionnels et fonctionnels

En dernier lieu, nous mettrons en perspective les premières réponses obtenues en traitant les questions liées aux contraintes distributionnelles et fonctionnelles pour nous interroger sur le poids respectif de ces facteurs dans le développement du paradigme des déterminants. Peut-on avant tout observer dans nos données la prédominance d'un facteur spécifique ou bien au contraire une influence combinée des facteurs ? Si la mise en place des formes pré-nominales semble marquée par la complémentarité de plusieurs facteurs, peut-on parler d'une influence successive - et liée aux périodes du développement - ou simultanée ? *A contrario*, si le constat est

fait d'une influence majoritaire d'un des facteurs tout au long de la période étudiée, a-t-on affaire davantage à une prépondérance de la dimension distributionnelle ou bien fonctionnelle ?



Pour traiter les questions que nous venons de lister, nous avons choisi de travailler sur un corpus longitudinal de quatre enfants, âgés de 1;6 pour le plus jeune au début des séances sélectionnées à 3;5 pour le plus âgé à la fin. Le choix de cette période nous a permis de couvrir la mise en place presque complète du paradigme chez tous les enfants. Ce corpus, ainsi que la présentation de la méthodologie de recueil, de transcription et d'analyse adoptée dans le cadre de cette thèse seront introduits au lecteur dans le prochain chapitre.

# **CHAPITRE IV**

## *Corpus et méthodologie*

---



Au sein de ce chapitre IV, consacré au corpus de la thèse et à la méthodologie de transcription et d'analyse, nous aborderons en premier lieu les questionnements qui nous ont servi à établir notre protocole de recueil (section 1) ainsi que la question du choix que nous avons fait de recueillir notre propre corpus à l'heure du partage de données (section 2). Nous présenterons ensuite les enfants qui constituent le corpus de cette thèse (section 3) et les situations d'enregistrement (section 4), après quoi nous discuterons au sein de la section 5 de notre place d'observatrice pendant le recueil. La section 6 sera consacrée aux questions de transcription et en dernier lieu (section 7), nous exposerons notre méthodologie d'analyse.

## **1. Questionnements préalables à la constitution du corpus**

Le corpus recueilli pendant cette thèse est au cœur de notre travail, des premiers questionnements à la discussion de nos résultats. Cette place centrale du corpus dans la réflexion implique de nombreuses interrogations préalables, concernant notamment le type de données à recueillir, la durée de la collecte ainsi que l'âge et le nombre d'enfants à observer.

### **1.1 Le type de corpus à recueillir**

Depuis les années soixante, les études en acquisition du langage se sont penchées sur les enregistrements des productions des enfants et de leurs interlocuteurs, non plus seulement sur des notes issues de comptes rendus ou de journaux parentaux, consignées le plus souvent par les chercheurs eux-mêmes, grâce notamment à l'arrivée et l'utilisation d'outils d'enregistrements sonores. Les travaux de Brown (1973) sur les corpus longitudinaux d'Adam, Eve et Sarah ont considérablement élargi les perspectives sur le recueil des données. Si d'autres questions -

soulevées par exemple par les méthodes de recueil ou la transcription des données<sup>1</sup> - se posent au sujet du corpus recueilli, cette nouvelle méthode de collecte écarte les problèmes liés à l'impossibilité de revenir aux données, à l'appréciation des productions des enfants dans les rapports parentaux (sur ou sous-estimation des productions), ou encore à la sélection des informations qu'impliquait le travail par prise de notes. Avec l'apparition des outils d'enregistrements vidéo, les chercheurs ont pu se concentrer sur d'autres aspects du développement de l'enfant tels que les gestes ou les regards<sup>2</sup> et ouvrir la réflexion sur une dimension multimodale de l'acquisition.

Dans la lignée de très nombreux travaux depuis l'étude pionnière de Brown, nous avons opté pour un recueil vidéo longitudinal. Notre objet d'étude impliquait nécessairement l'enregistrement de données, puisque nous avons besoin de nous saisir de la façon la plus fine possible des productions d'enfants en cours d'acquisition des déterminants. Si l'enregistrement audio eut été peut-être plus discret, notre choix s'est malgré tout porté sur l'enregistrement vidéo. La connaissance de la situation pouvait nous aider à désambigüiser certains éléments du corpus, lorsque les productions des enfants ou des adultes étaient difficilement compréhensibles ou lorsque les référents n'étaient pas rendus explicites dans le dialogue. Par ailleurs, comme nous savons que l'enfant peut communiquer à propos d'objets ou de personnes grâce aux mots mais aussi grâce aux gestes, et particulièrement aux pointages (Bruner, 1975, 1983), le choix de l'enregistrement vidéo nous permet une approche multimodale indispensable pour notre objet d'étude. Dans la mesure où notre travail porte en grande partie sur le rôle de l'interaction et de l'interlocuteur dans la mise en place des déterminants, nous avons souhaité également pouvoir répondre à des questions telles que la place, le positionnement des participants dans l'échange ou les regards portés sur l'autre et les objets. A titre d'exemple, il n'aurait pas été envisageable de traiter la question du statut attentionnel et discursif du référent (qui sera discuté au chapitre VII) sans avoir au moins partiellement accès à la direction des regards des participants. Quand bien même nous aurions étayé nos données audio de notes extrêmement exhaustives, nous n'aurions pu rendre compte de toutes les dimensions dont nous fait bénéficier la vidéo. De plus, notre rôle

---

<sup>1</sup> Ces questions feront l'objet d'une discussion au sein de ce chapitre (sections 5 et 6).

<sup>2</sup> Voir Kail et Bassano (2000) pour un état de l'art sur les méthodes de recueil en acquisition du langage.



d'observatrice, participante ou active (sur lequel nous reviendrons au point 5 de ce chapitre), ne nous permettait pas une prise de notes pendant l'enregistrement.

Notre objet d'étude impliquait par ailleurs un recueil permettant d'appréhender le développement d'un paradigme, des premières formes produites devant le nom à la maîtrise des morphèmes adultes ; c'est pourquoi nous avons procédé à un recueil de type longitudinal. Chaque enfant a été filmé pendant environ un an et demi, et le recueil s'est déroulé d'octobre 2011 à avril 2014. Nous avons opté pour des enregistrements mensuels, parfois bimensuels lorsque cela a été possible. Cet espacement, justifié par le souhait de respecter les disponibilités des familles, nous a permis de recueillir le corpus nécessaire à notre recherche. Notre étude s'intéressant à un phénomène très fréquent dans le langage de l'enfant, ces séances mensuelles nous laissent à voir quand et comment se développent les formes pré-nominales. Si le développement langagier de l'enfant au quotidien nous échappe, et que nous aurions pu essayer de le saisir grâce à un corpus plus dense, nous ne perdons pas pour autant les grandes étapes de la mise en place du langage, et plus spécifiquement de la mise en place des déterminants. Nous restons néanmoins consciente que la part des formes produites et entendues par l'enfant dans nos enregistrements est restreinte et ne nous aurait probablement pas suffi pour l'analyse de phénomènes apparaissant plus rarement dans le langage de l'enfant (Tomasello et Stahl, 2004).

## **1.2 La population à étudier**

Le corpus initialement constitué est composé des données de trois enfants âgés de 1;5 à 1;8 au début des enregistrements et de 2;8 à 3;2 à la fin des enregistrements. Pour plusieurs raisons, nous avons préféré limiter notre collecte à un corpus de trois enfants. Des contraintes de temps et de faisabilité du projet ne nous auraient pas permis de consacrer davantage de temps au terrain, ainsi qu'au travail de transcription et d'analyse des données. Par ailleurs, si une étude sur un grand nombre d'enfants nous aurait aidée à partir d'une démarche statistique et quantitative, à établir de grandes tendances, nous n'aurions pas pu apprécier qualitativement l'évolution langagière des enfants suivis. De plus, dans l'hypothèse d'une grande variabilité entre les enfants dans le traitement des fillers et des déterminants, nous n'avons pas la certitude qu'une plus grande cohorte nous aurait apporté un éclairage plus net sur le phénomène étudié.

### **1.3 Le choix d'un recueil 'naturel'**

Avant la constitution du corpus, nous nous sommes également penchée sur la question du type de recueil, et nous avons préféré collecter nos données dans le cadre de 'scènes quotidiennes', avec lesquelles l'enfant était déjà familiarisé. La place que nous avons eue en tant qu'observatrice soulève inévitablement des questions sur l'aspect naturel, écologique de nos données (notamment, Labov, 1972a ; Gadet, 2000) - et nous y consacrerons une partie de notre développement méthodologique - mais nous avons essayé d'être au plus près des productions habituelles des enfants que nous avons suivis, ainsi que de celles de leurs parents et frères ou sœurs. Il ne s'agit pas là de dire que nos données sont tout à fait représentatives des productions des participants hors du champ de la caméra, et sans la présence d'un observateur mais notre présence fréquente et familière dans l'environnement des familles nous permet d'envisager les conduites langagières de leurs membres comme les plus courantes possible au vu du cadre dans lequel elles ont été recueillies.

## **2. Le recueil de corpus à l'heure du partage de données**

### **2.1 Le partage de données**

Les avancées technologiques ont permis le renouvellement du type de données en rendant accessibles aux chercheurs le recueil de données audio et vidéo mais elles ont également favorisé de nouvelles pratiques de recherche en permettant le partage des données. Depuis la création par B. MacWhinney et C. E. Snow en 1984 de la base de données CHILDES (Child Language Data Exchange System), les chercheurs en acquisition disposent désormais de plus de 130 corpus dans 32 langues différentes accessibles en ligne. A cette heure, plus de 3200 publications sont issues d'analyses de données disponibles sur CHILDES (MacWhinney, 2010).

Au vu du temps consacré à la création d'un corpus (collecte des données, numérisation et transcription), l'intérêt des corpus partagés est considérable. Par ailleurs, plusieurs dimensions du développement de l'enfant (*e.g.* dimensions phonologique, prosodique, syntaxique, lexicale, gestuelle etc.) peuvent être abordées sur les mêmes corpus et ainsi permettre des perspectives plus diversifiées et plus riches sur le développement des enfants dont les corpus sont partagés.

Nous y reviendrons lors de la présentation détaillée de notre corpus mais c'est dans cette perspective d'analyse multidimensionnelle que nous avons commencé à travailler sur le corpus d'Adrien (recueilli par N. Yamaguchi dans le cadre de sa thèse<sup>3</sup>) ajouté aux données traitées dans ce travail. C'est également dans ce même objectif que nous mettrons à disposition à la fin de notre thèse les deux corpus pour lesquels nous avons des autorisations de partage signées par les parents des enfants.

De nombreux corpus transcrits et pour lesquels les vidéos des enregistrements ont été mises à disposition par des équipes de chercheurs auraient pu nous être utiles pour rendre compte du développement des déterminants chez le jeune enfant. Nous pensons tout particulièrement au corpus *Paris* (Morgenstern et Parisse, 2012) recueilli dans le cadre des projets *Léonard*<sup>4</sup> (projet ANR jeunes chercheurs) et *CoLaJE*<sup>5</sup> (projet ANR), coordonnés par A. Morgenstern (et impliquant également sans tous les citer F. Bourdoux, S. Caët, C. Parisse, ou M. Sekali), ou encore au corpus *Lyon* (Demuth et Tremblay, 2008), recueilli par l'équipe d'H. Jisa, grâce notamment à un financement attribué à K. Demuth<sup>6</sup>. Nous avons néanmoins souhaité recueillir nous-même la majeure partie du corpus de cette thèse (trois des quatre corpus qui constituent la base de nos analyses), et cela malgré son impact sur la durée du travail. Plusieurs éléments ont motivé ce choix : nous souhaitions être au plus près de notre objet d'étude et nous initier au travail de terrain, mais aussi enregistrer des données comparables. Il nous fallait pour cela collecter un corpus spécifique à notre travail de thèse. L'initiation du jeune chercheur passant aussi par la transcription du corpus recueilli, nous avons souhaité y être confrontée sans l'influence d'une transcription préalable. Enfin, dans l'objectif de rendre compte de la diversité des enfants, il nous a semblé intéressant de ne pas travailler que sur des enfants qui avaient déjà fait l'objet d'études antérieures et d'ajouter d'autres données à celles existantes.

## 2.2 Le recueil d'un corpus spécifique à la thèse

---

<sup>3</sup> Yamaguchi, N. (2012), *Parcours d'acquisition des sons du langage chez deux enfants francophones*, Thèse de doctorat en phonétique-phonologie soutenue à l'Université Sorbonne Nouvelle-Paris 3.

<sup>4</sup> <http://anr-leonard.ens-lyon.fr/>

<sup>5</sup> <http://colaje.scicog.fr/>

<sup>6</sup> Demuth, K. et Tremblay, A. (2008), « Prosodically-conditioned variability in children's production of French determiners », *Journal of Child Language*, 35, pp. 99-127

Nous avons souhaité nous approprier notre objet d'étude, en prenant le temps de l'observation directe, afin d'appréhender le plus fidèlement possible tous les niveaux impliqués dans le développement de l'enfant, au-delà de la seule observation des acquisitions langagières. Si notre travail porte sur un aspect spécifique du langage de l'enfant, il nous a paru essentiel de comprendre par ailleurs la multitude des acquisitions que fait simultanément l'enfant dans la période que nous avons capturée. Le recueil de données plus que l'utilisation de corpus existants nous apportait cette vision globale du développement de l'enfant. Nous avons pu constater que le recueil rendait possible le contact avec l'enfant hors du champ de la caméra, dans des situations où adulte et enfant ne sont plus contraints par la présence parfois dérangeante du matériel d'enregistrement et des activités impliquées par nos visites. Nous avons eu la possibilité d'apprécier à de très nombreuses reprises l'évolution générale des enfants que nous avons suivis. Ceci nous a aidée non seulement à concevoir ce qu'était la matière première de notre travail mais nous a aussi permis de considérer l'ampleur et la diversité des interactions offertes à l'enfant en dehors de notre heure mensuelle ou bimensuelle d'enregistrement. Pour des raisons matérielles, il n'aurait pas été envisageable de filmer davantage mais le temps passé auprès des enfants sans caméra, largement supérieur à la durée des enregistrements, nous apporte des éclairages importants dont nous n'aurions pas pu bénéficier sans s'accorder ce temps de collecte. Nous avons pu par exemple prendre la mesure ponctuellement et qualitativement de l'impact du développement moteur de l'enfant ou de son entrée à la crèche ou à l'école sur ses productions langagières.

Nous avons également souhaité recueillir des données comparables, et pour cela nous avons besoin d'un corpus le plus homogène possible. Nous avons donc initialement cherché trois garçons, que nous avons commencé à filmer autour d'un an et demi. En ayant connaissance des travaux menés sur l'influence de la place dans la fratrie sur le développement langagier (Bernicot et Roux, 1998), nous avons tenu à ce que les trois enfants du corpus aient le même rang dans la fratrie : ils ont tous un aîné mais pas de cadet. Nous avons par ailleurs désiré dans le protocole d'origine enregistrer des situations semi-contrôlées en filmant les mêmes activités pour toutes les dyades parent-enfant, mais pour des raisons qui seront exposées ultérieurement (point 4.2), nous avons *in fine* préféré ne pas contraindre nos participants à utiliser des supports imposés.

Enfin, l'étude des fillers soulevant des questions importantes de transcription (dont nous discuterons dans une partie dédiée, *cf.* point 6.5.3 de ce chapitre), nous avons souhaité nous confronter aux problèmes liés à la transcription de l'oral, sans l'influence d'une transcription initiale. Il est de toute évidence très enrichissant de pouvoir contribuer à plusieurs à la transcription d'un corpus, qui gagne indubitablement en exactitude et en finesse. Néanmoins, dans la mesure où nous estimons que la transcription du corpus implique déjà une grande part d'analyse (notamment, Ochs, 1979 ; Parisse et Morgenstern, 2010), nous avons préféré la mener sans base existante.

### **3. Présentation des enfants**

Nous l'avons mentionné précédemment, le corpus de cette thèse est composé de quatre suivis longitudinaux, trois que nous avons recueillis nous-même et un quatrième, celui d'Adrien, recueilli par N. Yamaguchi. Nous commencerons cette présentation détaillée des corpus par les trois premiers suivis, et nous consacrerons la partie suivante de ce chapitre à la description du corpus d'Adrien.

#### **3.1 Le corpus de Lubin**

Le corpus de Lubin est le premier que nous avons recueilli. Les enregistrements ont débuté le 6 octobre 2011 et se sont arrêtés le 24 janvier 2013. Lubin est né le 20 avril 2010 et avait donc 1;5.17 an lors du premier enregistrement. Nous avons en tout effectué dix-huit enregistrements de Lubin, que nous avons en moyenne espacés de deux semaines à un mois, et parfois davantage lorsque la disponibilité des parents ne le permettait pas ou lors des périodes de vacances. Toutes les données du corpus n'ont pas été utilisées dans le cadre de cette thèse<sup>7</sup> mais le corpus de Lubin représente en tout 13,5 heures d'enregistrement. Sur les dix-huit séances d'enregistrement, huit d'entre elles ont été conservées pour la présente recherche<sup>8</sup>.

---

<sup>7</sup> Nous reviendrons dans la suite du chapitre sur les raisons qui ont motivé la sélection des séances retenues au sein de chaque corpus.

<sup>8</sup> Dans les trois tableaux à suivre (Tableaux 1, 2 et 3), les séances retenues pour les analyses de cette thèse sont en gras.

Lubin vit avec ses parents et sa sœur Léonie, qui est de trois ans son aînée. Ils vivent en région parisienne. Lubin a été gardé par ses parents (père et mère en alternance selon leurs impératifs professionnels) à la maison jusqu'à ses deux ans. Il a ensuite été gardé en crèche de un à trois jours par semaine environ.

**Tableau 1 - Corpus de Lubin**

Date d'enregistrement	Âge de Lubin	Durée de l'enregistrement
06-10-11	1;5.17	33'
20-10-11	1;6.00	32'
03-11-11	1;6.14	30'
21-11-11	1;7.02	33'
12-12-11	1;7.23	32'
<b>11-01-12</b>	<b>1;8.23</b>	<b>45'</b>
26-01-12	1;9.07	58'
<b>28-03-12</b>	<b>1;11.08</b>	<b>27'</b>
05-04-12	1;11.16	30'
<b>03-05-12</b>	<b>2;0.14</b>	<b>46'</b>
<b>07-06-12</b>	<b>2;1.19</b>	<b>56'</b>
20-06-12	2;2.01	54'
31-08-12	2;4.12	58'
<b>20-09-12</b>	<b>2;5.02</b>	<b>60'</b>
<b>23-10-12</b>	<b>2;6.04</b>	<b>62'</b>
14-11-12	2;6.26	57'
<b>11-12-12</b>	<b>2;7.23</b>	<b>36'</b>
<b>24-01-13</b>	<b>2;9.06</b>	<b>60'</b>
	<b>Total</b>	13h30

### 3.2 Le corpus de Naël

La première séance d'enregistrement de Naël a eu lieu le 15 janvier 2012. Naël est né le 24 juillet 2010, il était donc âgé de 1;5.23 an. Lors du dernier enregistrement, Naël avait 2;8.14 ans. Les séances ont eu lieu à un mois d'intervalle environ ; elles sont plus espacées parfois, pour les mêmes raisons de disponibilité et de vacances que celles précédemment évoquées pour Lubin. Sur les treize séances que nous avons filmées, huit d'entre elles ont été utilisées pour nos analyses.

Naël vit en région parisienne avec ses parents et son aîné Tristan, né 2 ans avant lui. Naël a été gardé par sa mère lorsque l'emploi du temps de celle-ci le lui permettait mais la plupart du temps il a été gardé par une assistante maternelle.

Tableau 2 - Corpus de Naël

Date d'enregistrement	Âge de Naël	Durée de l'enregistrement
<b>15-01-12</b>	<b>1;5.23</b>	<b>34'</b>
<b>11-03-12</b>	<b>1;7.18</b>	<b>58'</b>
31-03-12	1;8.08	40'
<b>11-04-12</b>	<b>1;8.19</b>	<b>60'</b>
<b>13-05-12</b>	<b>1;9.20</b>	<b>61'</b>
<b>17-06-12</b>	<b>1;10.25</b>	<b>56'</b>
04-07-12	1;11.11	51'
<b>19-09-12</b>	<b>2;1.28</b>	<b>49'</b>
05-10-12	2;2.13	35'
<b>01-12-12</b>	<b>2;4.09</b>	<b>58'</b>
06-01-13	2;5.15	56'
<b>02-02-13</b>	<b>2;6.11</b>	<b>53'</b>
06-04-13	2;8.14	46'
	<b>Total</b>	11h

### 3.3 Le corpus d'Antonin

Antonin a été filmé du 13 octobre 2012 au 9 avril 2014. Il avait donc 1;8.01 an lors de notre première séance, et 3;1.26 la dernière fois que nous l'avons filmé. A l'exception des mois d'été, les séances ont toutes lieu à un mois d'intervalle. Nous avons effectué seize enregistrements, et nous en avons conservé huit pour ce travail de thèse.

Antonin, ses parents et son frère vivent à Paris. Solal, le frère d'Antonin est de trois ans son aîné. Pendant la période des enregistrements, Antonin était gardé à son domicile par une nourrice. Toutefois, dans la dernière période des enregistrements, il a aussi été gardé en crèche.

Tableau 3 - Corpus d'Antonin

Date d'enregistrement	Âge d'Antonin	Durée de l'enregistrement
13-10-12	1;8.01	44'
<b>10-11-12</b>	<b>1;8.29</b>	<b>43'</b>
15-12-12	1;10.03	50'
<b>26-01-13</b>	<b>1;11.14</b>	<b>47'</b>
28-02-13	2;0.17	36'
<b>30-03-13</b>	<b>2;1.17</b>	<b>40'</b>
27-04-13	2;2.14	34'
<b>08-06-13</b>	<b>2;3.26</b>	<b>32'</b>
06-07-13	2;4.23	29'
<b>07-09-13</b>	<b>2;6.25</b>	<b>38'</b>
05-10-13	2;7.23	34'
<b>09-11-13</b>	<b>2;8.28</b>	<b>28'</b>
07-12-13	2;9.25	39'
<b>18-01-14</b>	<b>2;11.06</b>	<b>28'</b>
03-03-14	3;0.19	38'
<b>09-04-14</b>	<b>3;1.26</b>	<b>28'</b>
	<b>Total</b>	9h48

Les trois enfants que nous venons de présenter vivent dans des familles monolingues francophones et se sont avérés ne pas avoir de retard de langage. Lors des derniers enregistrements, Lubin, Naël et Antonin étaient tous les trois scolarisés en petite section de maternelle.

Enfin, les autorisations parentales<sup>9</sup> nous permettent de diffuser et de partager les corpus de Lubin et d'Antonin, mais pas le corpus de Naël<sup>10</sup>.

### 3.4 Le corpus d'Adrien

Le corpus d'Adrien a été recueilli par N. Yamaguchi, au sein du projet CoLaJe précédemment mentionné, et dans le cadre de sa thèse de doctorat en phonologie. L'auteur y a analysé le développement phonologique - et plus spécifiquement la mise en place des consonnes dans les mots lexicaux - de deux enfants francophones, Adrien et Madeleine. Ces corpus ont ensuite fait l'objet d'un travail de comparaison de l'acquisition des consonnes dans les mots lexicaux et dans les mots grammaticaux (Yamaguchi, 2012 ; 2013).

<sup>9</sup> Cf. modèle en annexe.

<sup>10</sup> Pour cette raison, nous n'avons pas fait figurer dans les annexes les corpus vidéo de Naël, mais seulement les transcriptions.



Nous avons commencé à analyser ce corpus après avoir effectué notre propre recueil, à l'occasion d'un travail collectif<sup>11</sup> comparant les facteurs (phono-prosodiques, syntaxiques et pragmatiques) en jeu dans l'acquisition des déterminants, et nous l'avons ensuite ajouté à notre corpus de thèse. Ce corpus initialement transcrit sous PHON est désormais disponible sur la base de données CHILDES, au format CHAT<sup>12</sup>.

Le corpus a été recueilli d'avril 2006 à septembre 2009, à raison d'un enregistrement par mois. Il existe trente-sept enregistrements d'Adrien, et nous en avons utilisé huit pour notre étude.

A la différence des trois autres enfants de notre corpus, Adrien est enfant unique. Avant d'entrer à l'école, il était gardé avec d'autres enfants par une assistante maternelle, au domicile de cette dernière.

**Tableau 4 - Enregistrements d'Adrien retenus pour l'analyse**

Date d'enregistrement	Âge d'Adrien	Durée de l'enregistrement
20-02-07	2;1.23	55'
26-04-07	2;3.29	57'
20-06-07	2;5.23	60'
07-08-07	2;7.10	53'
10-09-07	2;8.13	57'
09-12-07	2;11.11	51'
10-04-08	3;3.13	52'
12-06-08	3;5.15	55'
<b>Total</b>		<b>7h20</b>

Le matériel ayant servi aux enregistrements est le même pour les trois premiers corpus (caméra Sony ACC-FH70). Le corpus d'Adrien a été filmé avec une caméra Sony HDV 1080i et la piste audio a également été enregistrée à l'aide d'un microphone sans fil (Sennheiser, modèle EW 112-p G2), qu'Adrien portait la plupart du temps dans une banane attachée en bandoulière. Lors des scènes de bain, le microphone est porté par le parent.

A l'exception d'Adrien pour qui nous disposons déjà d'informations sur son développement phonologique, les corpus utilisés ayant été spécifiquement recueillis pour cette thèse, ils n'ont pas fait l'objet d'études préalables à celle présentée ici. Nous faisons donc le choix de commenter

<sup>11</sup> Le Mené, M., Yamaguchi, N., Fox, G., Jullien, S. et Salazar Orvig, A. (en préparation), « Acquiring a paradigm : the case of determiners ». In A. Salazar Orvig, G. de Weck, R. Hassan, et A. Riialand (eds), *Constructing reference in dialogue*, Amsterdam / Philadelphia : John Benjamins Publishing Company.

<sup>12</sup> Corpus Yamaguchi (<http://chilides.psy.cmu.edu/data/Romance/French/Yamaguchi.zip>)

certains éléments relatifs aux compétences langagières des enfants - lexicales, phonologiques ou morfo-syntaxiques - dans les chapitres suivants, dans la partie de la thèse consacrée aux résultats, et non lors de la présentation des enfants.

## **4. Les situations d'enregistrement**

### **4.1 Les participants lors des enregistrements**

Tous les enregistrements se sont déroulés au domicile des enfants. Le nombre de participants varie selon les séances et les corpus.

Lubin a été filmé le plus souvent en interaction individuelle avec son père ou sa mère. La sœur aînée, Léonie, étant le plus souvent à l'école lors des enregistrements, elle n'apparaît que rarement sur nos vidéos.

Nous avons le plus souvent filmé Naël en interaction avec sa mère, le père n'est que rarement présent sur nos enregistrements. Naël ayant été systématiquement filmé le week-end, il est la plupart du temps filmé avec son frère. Sa tante qui vit dans une maison mitoyenne est également présente lors de certains enregistrements.

Le corpus d'Antonin a été recueilli presque exclusivement le samedi matin, lors des activités extrascolaires de Solal, son frère aîné. Toutes les séances sont filmées avec la mère, le frère apparaît seulement à quelques reprises dans notre corpus, à la fin des enregistrements. Si le père était également présent au domicile de l'enfant à la fin des séances, il n'apparaît pas du tout sur nos vidéos.

Dans la plupart des enregistrements, Adrien est en interaction avec son père mais quelques séances ont été filmées avec la mère, ou avec les deux parents.

Dans ces quatre corpus, les deux observatrices peuvent ponctuellement devenir l'interlocutrice principale de l'enfant, lorsque les parents quittent la pièce par exemple. Mais lorsque les enfants sont en interaction avec les membres de leurs familles, elles restent la majeure partie du temps silencieuses.

## 4.2 Les types d'activités

Nous avons dans le protocole de départ envisagé un découpage des séances d'enregistrement en deux parties : une première partie lors de laquelle les participants auraient été filmés avec un support apporté par l'observatrice (différent pour chacune des séances) et une seconde partie pendant laquelle les participants auraient été libres de continuer l'activité avec ce support imposé ou de poursuivre avec une autre activité de leur choix. Nous aurions donc recueilli des séances comparables pour ainsi pouvoir par exemple mesurer chez tous les enfants l'impact des activités et des types de discours sur la production des formes pré-nominales. Les premiers enregistrements de Lubin se sont déroulés avec des supports que nous avons amenés mais malgré l'énergie déployée par les parents, l'enfant ne s'y est pas intéressé et a préféré jouer avec ses propres jouets. Afin de ne pas contraindre et gêner parents et enfants, nous avons donc finalement décidé de faire le recueil le plus 'écologique' qui soit et de laisser nos participants libres de choisir les activités lors de nos visites.

Les activités que nous avons enregistrées sont des activités du quotidien. Nous avons donc filmé des goûters et des déjeuners dans les corpus de Lubin et Naël ainsi qu'un certain nombre de séances en extérieur (toboggan, bac à sable) dans le corpus de Naël. Nous avons par ailleurs pu constater que tous les enfants - y compris Adrien - sont impliqués au cours de certaines séances dans des activités de lecture ou de jeu (de société, de construction, de voitures, etc.). Les livres et les jeux ne sont pas identiques d'une famille à l'autre mais les types d'activités sont semblables, si bien qu'une étude sur l'impact des genres de discours associés aux activités sur la production des formes pré-nominales a pu être envisagée dans nos perspectives de travail ultérieures.

## 5. La place de l'observateur

Nous l'avons mentionné préalablement, la place que nous avons eue en tant que chercheuse, observatrice ou participante dans les familles que nous avons filmées nous a beaucoup interrogée pendant toute la durée du recueil. Nous connaissions bien la mère d'Antonin avant de procéder au recueil de corpus, mais nous n'avions rencontré qu'une fois les parents de Lubin et de Naël, le mois qui a précédé les premiers enregistrements. Nous avons toutefois rencontré ces deux familles par l'intermédiaire d'une amie commune, ce qui a probablement rendu plus facile la prise de contact entre observatrice et observés, ou pour reprendre les mots de Bourdieu, favorisé la

mise en place d'une « communication 'non violente' » (1993 : 1395). La place que nous avons eue dans le quotidien de ces familles lors de nos enregistrements, et l'impact que notre présence et celle de la caméra ont pu avoir sur leurs pratiques ont néanmoins suscité de nombreuses interrogations de notre part tout au long de la thèse.

Si la pratique des enregistrements vidéo effectués par un observateur-chercheur est devenue courante dans le champ des sciences du langage, elle n'en soulève pas moins de questions. En 1972, Labov écrivait dans son ouvrage intitulé *Sociolinguistic patterns* « the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed ; yet we can only obtain these data by systematic observation » (1972a : 209), et c'est là la définition de ce qu'il a appelé le 'paradoxe de l'observateur'. Comment envisager que la présence de l'observateur et de la caméra ne puisse pas être un frein aux pratiques 'habituelles' ? Nous pouvons difficilement considérer que de très jeunes enfants s'hypercorrigent en notre présence, mais qu'en est-il des parents ? Est-ce qu'en apprenant que nos recherches portaient sur le développement du langage de l'enfant, les parents n'ont pas modifié leur comportement et insisté plus qu'à l'habitude sur les reprises, les reformulations des productions de leur enfant ? Notre présence dans un premier temps peu familière n'a-t-elle pas eu pour conséquence une inhibition particulière de l'enfant ? Ces questions entraînent une remise en cause légitime du caractère représentatif de nos données, des compétences des enfants et des pratiques des parents. Cette interrogation faisait déjà suite à un premier constat sur le peu de données que peuvent finalement représenter les suivis longitudinaux, mais elle n'en est que plus fondée si l'on s'intéresse au rôle joué par l'observateur. Nous ferons état dans les pages ci-après des différents constats que nous avons pu faire sur notre statut d'observatrice.

### **5.1 Le statut de l'observateur dans les familles**

Accueillir un chercheur à domicile n'est pas chose aisée ou naturelle pour les participants observés, qui doivent accepter la présence d'un autre extérieur à la famille, mais elle ne l'est guère plus pour le chercheur débutant. Comme l'a signalé Morgenstern, « cette position d'observateur, il faut s'y sentir bien, savoir l'occuper, en connaître les limites » (2009a : 89). Plusieurs facteurs peuvent contribuer à faciliter les rôles de chacun dans cette démarche d'observation. Il nous a semblé dans notre recueil que la symétrie dans le statut des parents des familles et celui de

L'observatrice a pu permettre une intégration plus rapide du chercheur dans les familles, et de la même façon, l'intégration des familles dans la sphère du chercheur. Par 'symétrie' nous entendons ici des correspondances générationnelles, sociales, ou encore culturelles. Si l'asymétrie de départ impliquée par le format de l'observation - qui suppose que l'observateur soit celui qui donne les règles du jeu (Bourdieu, 1993) - est inévitable, elle peut être partiellement dissipée si l'observateur fait partie intégrante du groupe social observé, et que par exemple, les statuts socio-éducatifs ne sont pas à l'origine d'une autre asymétrie. Le chercheur est situé socialement (Mondada, 1998), il est le reflet d'un certain groupe qui peut être en accord ou non avec la population observée.

L'acceptation du chercheur par les participants repose aussi sur la description du rôle que nous avons en tant qu'observateur, et en ce qui nous concerne, du rôle que nous avons en tant que linguiste, celui de décrire les faits et non de les évaluer. Il nous a semblé important d'accentuer l'idée que notre but n'était pas de travailler sur de 'bons' enfants ou de 'bons' parents mais sur des 'façons de faire ordinaires'. Relevons également que les parents de notre corpus ont accepté la présence d'un observateur à leur domicile et ont accepté d'être filmés. Si nous nous devons de respecter le 'contrat' passé au début des enregistrements (durée du recueil, respect de la disponibilité des parents et des intervalles entre chaque enregistrement) pour que le recueil puisse être poursuivi, cet accord initial impliquait d'emblée un terrain favorable à notre travail. Parmi les raisons qui ont motivé les parents à participer au projet, outre celle de rendre service et de contribuer à une recherche universitaire, figure un certain état d'esprit des parents. Nous avons en effet eu le sentiment que les parents filmés étaient à l'aise avec leur identité de parents, et en apparence peu sujets à des inquiétudes, des remises en question sur les méthodes éducatives ou les façons de s'adresser aux enfants. Sans doute cela s'explique-t-il en partie parce que nous les avons principalement filmés en interaction avec leur deuxième et non leur premier enfant, ce qui a peut-être contribué à l'effacement d'une crainte d'être jugés. Par ailleurs, l'observatrice n'ayant pas d'enfants, les adultes de notre corpus restent, parmi les participants, les plus experts de la parentalité. En cela, la relation n'est plus symétrique mais cette asymétrie a, d'après nous, conditionné positivement notre recueil.

Nous l'avons mentionné précédemment, la durée du recueil est aussi un facteur à prendre en compte, d'après nous, pour définir la place que nous avons eue dans les familles de notre corpus.

Pour Lallier, qui a entre autres travaillé sur les situations de tournage de documentaires, l'acceptation de ceux qui sont filmés ne repose pas nécessairement sur un rapport de confiance instauré entre ces derniers et celui qui filme (Lallier, 2011/2). En effet, l'immersion de l'observateur-filmant et la relation privilégiée entre les deux parties n'est pas une condition nécessaire aux enregistrements mais nous avons cependant la certitude que le recueil longitudinal, outre son intérêt pour la thématique de notre recherche, nous a donné à voir davantage et a laissé à tous les participants impliqués le temps nécessaire pour que familiarité et confiance s'installent.

Enfin, nous tenons à préciser également que deux des mères de notre recherche sont chercheuses, connaissent donc les problématiques de terrain et de recueil, et ont aussi peut-être plus que d'autres parents une perspective éclairée sur leur rôle de participantes observées, ce qui implique une acceptation des rôles à attribuer au chercheur et aux participants peut-être plus rapide. La troisième mère de notre corpus est comédienne, ce facteur a, d'après nous, facilité notre tâche lors des enregistrements puisqu'elle était déjà habituée à être observée dans ses pratiques professionnelles quotidiennes. Par ailleurs, une des mères chercheuses est spécialisée dans le domaine de l'acquisition du langage et a également été enseignante de classe primaire. Nous ne ferons pas l'hypothèse *a priori* que son comportement langagier est différent de celui des autres mères mais sa formation ainsi que son statut de participante 'informée' pourront néanmoins être à prendre en compte lors du traitement des résultats.

## 5.2 Une observation directe mais périphérique

Nous avons fait le choix lors des enregistrements d'une observation directe. Nous avons donc assisté aux séances filmées pour d'une part en avoir une compréhension intégrale, et d'autre part laisser libres les participants de mouvements éventuels, ce qui n'aurait pas été possible si la caméra avait été installée sur un trépied à distance de l'observateur. Mais nous n'avons pas, malgré notre présence lors des enregistrements, décidé de nous placer au centre de l'activité. Nous avons en effet préféré nous placer « hors-jeu », « à l'écart des interactions observées » (Lallier, 2011/2 : 123), et cela pour deux raisons. Les biais à considérer en travaillant sur des données 'naturelles' sont déjà nombreux, nous ne souhaitons pas en ajouter de nouveaux en étant à la fois la chercheuse et l'interlocutrice principale de l'enfant. Notre participation occasionnelle aux échanges avec l'enfant est donc liée le plus souvent à des questions de

disponibilité des parents mais n'avait pas été prévue initialement. Par ailleurs, nous n'avons pas souhaité intervenir lors des enregistrements pour ne pas entraîner le désengagement des participants - parents comme enfants - de la situation filmée.

A la suite de Brougère, nous considérons que notre participation aura été périphérique et non centrale puisqu'elle associait « implication (marginale) et distance (limitée) » de notre part (2006 : 6) mais légitime dans la mesure où elle était acceptée par les participants<sup>13</sup>. Cette relation entre observateur périphérique et participants centraux reste tout à fait paradoxale puisque comme le souligne Lallier, « chacune des parties joue à s'ignorer mutuellement, imitant en quelque sorte l'indifférence réciproque, sur le modèle de la *non-interaction* » (2011/2 : 117). Le cadre de ces interactions est donc défini au préalable et accepté par les participants et l'observateur, bien qu'il ne repose pas sur une pratique interactionnelle 'naturelle', mais bien sur une pratique 'représentée'.

### 5.3 La vidéo, reflet infidèle de la réalité

Le choix de l'enregistrement vidéo pour notre recueil repose sur le besoin de fixer les productions des participants pour les analyser, en saisissant à la fois ce qui s'est dit dans la situation mais aussi ce qui s'est 'fait' (gestes et regards). Nous avons néanmoins conscience que ce qui s'est joué devant nous lors des enregistrements ne se retrouve pas à l'identique dans nos données vidéo. D'après Mondada, « l'enregistrement des données vise explicitement à préserver, pour les étudier, les traits caractéristiques de l'interaction ; il ne constitue pas pour autant un "reflet" fidèle d'un événement "réel" - conception qui naturaliserait les données en objectivant les technologies qui ont permis de les saisir » (1998 : 61). En effet, la vidéo ne peut être que le reflet de ce qui se déroule dans le champ de la caméra, le hors-champ n'étant pas systématiquement pris en compte dans les données que nous analysons (Joly-Rissoan, 2004).

Par ailleurs, la caméra est un objet qui favorise peut-être plus encore que la seule présence de l'observateur la mise en scène de l'événement vécu par les participants. Gibson qui a travaillé sur des vidéos tournées sans la présence de l'observateur (ou plutôt en présence d'un observateur absent), considère que les vidéos comme d'autres types de données « are instances of meaning-

---

<sup>13</sup> Notons que cette notion de *participation périphérique légitime* (« Legitimate Peripheral Participation ») est à l'origine issue des travaux de Jean Lave et Etienne Wenger (1991).

making events in which participants do identity work to present a particular kind of self » (2005 : 41). La vidéo peut essayer de capturer les usages des participants mais ces usages ne restent que des représentations qui offrent à voir les 'performances' des acteurs des interactions filmées.

Nous garderons donc à l'esprit que « l'image filmée occupe de fait le même statut qu'une question dans un questionnaire ou un entretien : elle est un regard sur le social, un mode d'entrée particulier qui éclaire certaines dimensions, et en néglige d'autres » (Joly-Rissoan, 2004 : 3). Si pour les raisons que nous venons d'évoquer, notre recours à la caméra et à la vidéo s'est fait par défaut, il convient tout de même de signaler qu'aucun autre outil d'observation ne pouvait nous apporter toutes les perspectives indispensables à notre étude et offertes par l'image filmée.

#### **5.4 Le dépassement du paradoxe de l'observateur**

Depuis les travaux pionniers de Labov, de nombreux chercheurs se sont penchés sur la question du rôle de l'observateur. Ces chercheurs ne se sont pas contentés de constater et déplorer le paradoxe inhérent au recueil de données mais ils ont essayé de le dépasser.

Plusieurs solutions ont été proposées afin de limiter l'impact de la présence du chercheur, dans le cadre des entretiens sociologiques notamment ou des travaux de sociolinguistique visant à collecter la langue vernaculaire des échanges quotidiens. Il a notamment été établi que la construction d'un rapport familier avec les personnes observées pouvait améliorer le degré de spontanéité des échanges entre les participants (par exemple, Judd *et al.*, 1991, Mondada, 1998). Si le linguiste est membre - ou est accepté comme membre - de la communauté observée (dans le cadre des observations participantes), les biais liés à l'intimidation des observés ou à l'acclimatation au chercheur seraient moins nettement marqués. D'autres, comme Gumperz (1982) ou Labov (1972b), ont mis en avant l'intérêt des enregistrements de sessions de groupes, qui limiteraient la conscience des participants d'être observés. Wolfson note néanmoins que si les interactions en groupe peuvent en effet éloigner l'attention des participants de l'enregistreur ou de l'observateur, la question de l'effet de l'observation reste encore posée (1976 : 199). Par ailleurs, si les 'recettes' proposées par les uns et les autres fonctionnent dans certains cas de figure, elles ne s'appliqueront pas nécessairement à tous les contextes de recueil. Nous avons néanmoins le sentiment que l'effet quelque peu 'dépolarisant' des sessions de groupe a pu être à l'œuvre dans nos données également, quand bien même nos 'groupes' étaient composés de deux



ou trois participants seulement. Le rapport à l'enregistrement reste différent lorsque l'on interagit avec un autre, qui plus est lorsque l'on cumule les activités (communiquer mais aussi jouer ou manger).

Enfin, quels que soient les domaines dans lesquels la réflexion sur la place de l'observateur et son paradoxe ont pu opérer, l'idée de la prise de conscience des biais inhérents à la recherche sur les données a finalement semblé plus importante que la possibilité de contourner ces biais, comme l'écrivait Bourdieu dans *La misère du monde* :

« Le rêve positiviste d'une parfaite innocence épistémologique masque en effet que la différence n'est pas entre la science qui opère une construction et celle qui ne le fait pas, mais entre celle qui le fait sans le savoir et celle qui, le sachant, s'efforce de connaître et de maîtriser aussi complètement que possible ses actes, inévitables, de construction et les effets qu'ils produisent tout aussi inévitablement » (1993 : 905).

Certains chercheurs en analyse conversationnelle, tels que Gubrium et Holstein (1997) ont même établi que les contextes de recueil de données étant profondément sociaux et interactionnels, le rôle du linguiste n'était pas tant de décrire la réalité que de la co-construire avec les participants au recueil. A ce titre, « attempts to control bias may not only be futile, but may stifle the very features of interaction that are theoretically interesting » (Speer, 2002 : 512). Si les biais sont inévitables, ils n'en sont pas moins intéressants pour le chercheur, à condition qu'ils soient connus et maîtrisés.

Ce rapide état des lieux des moyens de dépasser le paradoxe de l'observateur vient confirmer la conviction que nous avons eue pendant et après la collecte des données, celle d'avoir - partiellement du moins - réussi à limiter le biais que nous représentions. L'avis devrait revenir aux familles que nous avons suivies mais là encore, nous serions confrontée à la subjectivité de l'enquête.

Nous avons pendant tout le temps du recueil, et particulièrement dans les deux familles de Lubin et de Naël que nous ne connaissions pas, passé beaucoup de temps avec les parents et les enfants. Il nous semble que notre image de chercheur susceptible d'évaluer les bonnes ou les mauvaises pratiques s'en est trouvée atténuée. Nous avons participé aux activités dans ces foyers, en dehors des jours d'enregistrement, pour que les temps passés en présence de la caméra finissent par devenir des temps marginaux. En essayant de nous intégrer à la vie des familles enregistrées et en participant aux échanges, parfois pendant mais surtout en dehors du champ de

la caméra, nous avons, nous semble-t-il, réussi à dépasser le handicap que nous représentions pour la collecte des données (Gadet, 2000).

Nous n'avons évidemment pas la certitude que ces constats soient partagés par les familles, nous en faisons donc seulement l'hypothèse. S'il est simple de trouver des traces dans les enregistrements de la négociation en jeu lors de la présence d'un observateur (regards, mentions de l'observateur ou de la caméra, discours adressé à l'observateur), il est beaucoup plus complexe voire impossible de savoir si les participants ont 'oublié' la présence de la caméra (Caronia, 2014), d'autant que l'observateur peut être mentionné dans deux cas de figure différents : lorsque la présence de l'observateur n'est pas encore négociée, que celui-ci reste un membre extérieur à l'activité en cours, mais aussi lorsqu'il est impliqué dans la situation, et potentiel destinataire des productions des participants. Dans ce dernier cas, tous les signes verbaux et gestuels traitant de l'observateur ou adressés à lui ne sont plus au même titre que dans le premier cas des marques de la conscience que les participants ont de l'observation.

Au-delà de la difficulté de renoncer à trancher sur le degré de spontanéité de nos participants, c'est la connaissance que nous avons des biais inhérents au recueil de corpus qui nous semble de loin l'élément le plus important pour cette présente recherche et les travaux qui seront menés ultérieurement sur ces données. Nous avons conscience de ce que représentent la participation d'un observateur et la présence d'une caméra, « intrusion toujours un peu arbitraire qui est au principe de l'échange » entre observateur et observés (Bourdieu, 1993 : 1392). En effet, si les conditions d'enregistrement nous semblent avoir été favorables à un recueil de données les plus fidèles possibles aux habitudes des familles, nous ne pourrions cependant pas écarter l'impact de notre présence sur les données. En avoir conscience et conscience nous préservera de conclusions hâtives sur des 'pratiques quotidiennes' et nous fera préférer la notion de 'pratiques lors des enregistrements'.

La suite de ce chapitre sera consacrée à deux autres points méthodologiques de la thèse : la transcription et le traitement des données. Nous y présenterons dans une première partie notre outil de transcription, les conventions adoptées et les questions générales soulevées par la transcription du langage enfantin, ainsi que les questions spécifiques posées par la transcription des fillers. Dans une seconde partie, nous exposerons les axes d'analyse retenus pour notre étude,

qui nous ont permis d'une part de confronter cette dernière aux résultats d'études précédentes et d'autre part, de soumettre nos hypothèses à l'épreuve du corpus.

## **6. La transcription du corpus**

Une fois le corpus recueilli, nous l'avons numérisé grâce au logiciel iMovie et nous l'avons ensuite transcrit. Nous reviendrons ici sur les raisons qui ont motivé la sélection des données que nous avons transcrites et conservées pour l'analyse ainsi que sur les conventions de transcription adoptées et les questions - nombreuses - posées par la transcription du langage enfantin.

### **6.1 La sélection des données à analyser**

Nous l'avons mentionné lors de la présentation du corpus, toutes les séances enregistrées n'ont pas été retenues pour l'analyse. Pour les corpus que nous avons recueillis, nous avons fait le choix de conserver pour chaque enfant l'une des deux dernières séances enregistrées (la dernière pour Lubin et Antonin et l'avant-dernière dans le cas de Naël), lorsque la majorité des formes pré-nominales sont conformes à la cible adulte, et d'analyser ensuite parmi les séances antérieures, sept autres séances, sélectionnées à environ deux mois d'intervalle. Les séances analysées couvrent ainsi le temps du recueil et nous donnent à voir toute la trajectoire des enfants, jusqu'à la production quasi-systématique des formes adultes devant les noms.

Signalons toutefois que d'autres facteurs, liés à la situation d'enregistrement, ont également pu conditionner notre sélection des séances. Dans le corpus de Lubin, la sœur est régulièrement présente mais pas systématiquement. Le temps de parole de Lubin étant considérablement moins long lorsque la sœur participe aux mêmes activités que lui, nous avons dans certains cas privilégié les séances pendant lesquelles la sœur est absente, au détriment de l'intervalle de deux mois que nous avons initialement fixé. Des contraintes du même ordre ont guidé notre sélection des séances pour Naël. Par exemple, dans la dernière séance enregistrée, l'enfant est non seulement filmé avec son frère (contexte d'enregistrement le plus fréquent) mais aussi avec les enfants des voisins. Sans négliger l'intérêt de cette séance, notamment pour l'analyse des interactions polyadiques entre pairs, nous avons préféré ne pas l'inclure dans notre corpus de travail, du fait du peu d'interventions de l'enfant cible. Par ailleurs, plusieurs séances de ce même corpus se déroulent en extérieur (dans la cour devant la maison). La qualité de l'enregistrement étant moins

bonne que lorsque l'enfant est filmé en intérieur, nous avons sélectionné autant que possible des séances ayant lieu dans la maison, et cela même si notre première sélection s'en trouvait modifiée.

Les critères de sélection des séances du corpus d'Adrien ont été quelque peu différents et ont reposé davantage sur des marques de développement phonologique. Dans la mesure où ce corpus fait l'objet d'un travail collectif étudiant notamment l'influence de certains facteurs phonologiques dans la mise en place des déterminants, nous avons d'abord sélectionné les séances à analyser en fonction du type de fillers produits par Adrien : fillers vocaliques ou consonantiques. Nous n'avons donc pas retenu les séances à analyser en fonction de l'âge ou du niveau de développement linguistique de l'enfant mais avons opté plutôt pour le choix de séances représentatives des moments charnières du développement phonologique et morphologique d'Adrien. Nous avons procédé dans un premier temps à la sélection de six séances choisies en fonction des formes produites en position pré-nominale : omissions, fillers (vocaliques et consonantiques) ou déterminants de forme adulte. Sur ces six séances, nous avons sélectionné trois séances contenant une majorité d'omissions et des fillers surtout vocaliques, deux séances contenant devant les noms une majorité de fillers consonantiques mais aussi des omissions et des déterminants de forme adulte, et enfin une séance avec une majorité de formes adultes. L'analyse de ces six séances a révélé la nécessité de prendre en compte pour notre travail deux autres séances ; une entre 2;5 et 2;8, moment où les tendances s'inversent, avant que les déterminants ne deviennent préférentiellement produits, et une autre après 3;3 pour voir totalement disparaître les proto-formes.

Nous avons utilisé pour nos analyses trente minutes de chacune des séances retenues pour les corpus d'Antonin, Lubin et Naël et les séances complètes d'Adrien, soit environ une heure de corpus par séance.

Bien que le choix que nous avons fait implique que les séances ne soient pas comparables au regard de l'âge des enfants et de leur niveau global de développement langagier, les séances sont tout de même équivalentes d'un enfant à l'autre du point de vue des formes produites en position pré-nominale. Nous retrouvons en effet pour chacun des enfants le développement complet de cette position, de la première à la dernière séance analysée, de la production des

proto-formes à celle systématique des formes adultes. Cet aspect sera bien sûr développé lors de la présentation des résultats.

## **6.2 L'outil utilisé pour la transcription**

Les données numérisées ont été transcrites à l'aide du logiciel CLAN (Computerized Language Analysis), logiciel de transcription et d'analyse conçu par B. MacWhinney (MacWhinney, 2000) dans le cadre du projet CHILDES précédemment évoqué<sup>14</sup>. Notre choix s'est porté sur cet outil pour plusieurs raisons : le logiciel permet la segmentation de la vidéo et son alignement avec le fichier de transcription, ce qui a non seulement facilité le travail de transcription mais aussi le travail d'analyse lorsque nous avons eu besoin de revenir à certains extraits de corpus et de visualiser à la fois la vidéo et la transcription. Le travail d'analyse a ainsi pu être mené sans s'éloigner des données initiales, qui ne sont pas les données écrites mais bien les enregistrements vidéo. Outre l'alignement entre la vidéo et la transcription et la recontextualisation en images des analyses, ce logiciel nous a également intéressée parce qu'il permet de conduire un certain nombre d'analyses quantitatives sur les corpus, notamment sur le niveau de développement langagier des enfants, la diversité lexicale des productions des locuteurs ou encore sur la fréquence de combinaisons de certains items dans les corpus. Par ailleurs, notre objectif étant à terme de partager une partie de notre corpus, nous avons opté pour le format de transcription le plus fédérateur au sein de la communauté de chercheurs en acquisition du langage, et qui nous permettait de plus d'homogénéiser le format de transcription de nos quatre corpus, puisqu'après avoir été transcrit sous PHON<sup>15</sup>, le corpus d'Adrien a été converti au format CHAT.

## **6.3 Les conventions de transcription**

Le choix des conventions de transcription a découlé de celui de notre logiciel de transcription et à nouveau, de la perspective future d'un partage de données. Les conventions adoptées sont celles du projet CHILDES, soit les conventions de transcription associées au format CHAT (Codes for the Human Analysis of Transcripts). Nous avons eu recours pour nous aiguiller dans

---

<sup>14</sup> <http://childes.psy.cmu.edu/clan/>

<sup>15</sup> Logiciel de transcription et d'analyse phonologique conçu par Yvan Rose (<http://childes.psy.cmu.edu/phon/>)

la transcription au manuel disponible sur le site de CHILDES<sup>16</sup> et surtout au guide - plus simple d'utilisation - élaboré au sein des projets Léonard et CoLaJe<sup>17</sup>. La partie de ce guide traitant des conventions de transcription CHAT est reproduite en annexe.

La transcription de chaque session débute par les entêtes de présentation du corpus, de l'enfant et de la séance. Un exemple est reproduit ci-dessous.

**Exemple 1 - En-têtes de début de transcription (corpus de Lubin, 2;5.02)**

```
@Begin
@Languages: fra
@Participants: CHI lubin Target_Child, FAT Father, MOT Mother, OBS Observer
@ID: fra|Le Mené_Corpus_Lubin|CHI|2;05.02|male|||Target_Child|||
@ID: fra|Le Mené_Corpus_Lubin|FAT||male|||Father|||
@ID: fra|Le Mené_Corpus_Lubin|MOT||female|||Mother|||
@ID: fra|Le Mené_Corpus_Lubin|OBS||female|||Observer|||
@Media: LUBIN-2_05_02, video
@Location: Lubin's home
@Situation: CHI et FAT sont dans le salon. CHI est allongé dans le canapé avec son doudou et FAT demande à CHI à quoi il aimerait jouer.
```

Le découpage du corpus a été fait énoncé par énoncé, en suivant les critères de découpage des énoncés établis dans le cadre de l'ANR DIAREF<sup>18</sup>. Conformément à ces conventions, nous avons délimité les énoncés des participants en fonction de critères à la fois prosodiques et micro-syntaxiques<sup>19</sup>. Les productions des tout jeunes enfants étant relativement courtes, nous avons la plupart du temps pu considérer qu'un énoncé était équivalent à un tour de parole, et le découpage de leurs énoncés n'a pas soulevé beaucoup de questions. A l'inverse, le découpage des productions des parents s'est avéré parfois plus problématique, mais nous avons tenu malgré cela à conserver un découpage en énoncés, plus pertinent pour, entre autres, notre analyse ultérieure des cadres lexico-syntaxiques.

Afin de respecter le format de transcription établi par les conventions CHAT, nous avons pour chaque énoncé transcrit d'une part une ligne principale orthographique normée débutant par les trois premières lettres du statut du locuteur (CHI pour *child* ou FAT pour *father*, par exemple) et

---

<sup>16</sup> <http://chilides.psy.cmu.edu/manuals/CHAT.pdf>

<sup>17</sup> <http://chilides.psy.cmu.edu/intro/leonard.doc>

<sup>18</sup> Le projet DIAREF - *L'acquisition en dialogue des expressions référentielles : approches multidimensionnelles* - mené de 2009 à 2013 a été coordonné par Anne Salazar Orvig et financé grâce au programme ANR « Enfants et enfance » (ANR 09 - ENFT - 055).

<sup>19</sup> Ces critères de délimitation des énoncés sont reproduits en annexe.

se terminant par un marqueur de fin d'énoncé (la liste complète des marqueurs est présentée au sein des conventions de transcription jointes en annexe). A la ligne principale sont associées des lignes secondaires de transcription phonétique large (pour l'enfant), d'informations non verbales (les actions et les gestes des participants) ou de commentaires et d'interprétations du transcripteur.

**Exemple 2 - Transcription ligne principale et lignes secondaires (corpus d'Antonin, 2;3.26)**

\*MOT: c'est le corps du cheval avec ses pattes .

\*CHI: yyy i@fs saute yyy nə@fs y@c .

%pho: e i sət ã nə y

%int: de nouveau <y> pour cheval

\*MOT: i(l) saute ?

Si la transcription orthographique du corpus était nécessaire pour un travail sous CLAN, l'analyse a été réalisée en observant la ligne de phonétique. La transcription phonétique, plus fidèle aux productions des enfants en bas âge, nous offre une vision ni trop 'optimiste' ni trop 'pessimiste' des productions des enfants (Blanche-Benveniste, 2010) et nous laisse la possibilité de ne pas trancher sur la nature grammaticale de tous les éléments produits, à un stade où les éléments du discours sont en cours d'acquisition. Nous pouvons néanmoins regretter l'impact que peut avoir l'interprétation parfois très subjective de la forme adulte (transcrite sur la ligne orthographique) sur l'observation des données. Ces constats seront développés dans la section 6.5.

La ligne consacrée aux actions et aux gestes, spécifique à chaque participant, nous a permis lors de l'analyse ultérieure du corpus de contextualiser les énoncés produits, notamment lorsque les référents ne sont pas mentionnés explicitement mais font seulement l'objet de manipulations (*cf.* exemple 3), de regards des locuteurs ou de pointages (*cf.* exemple 4). Nous n'avons repéré dans le corpus que les gestes pertinents du point de vue de la communication ou utiles au lecteur pour la compréhension de la situation. Ce repérage serait donc à affiner pour un travail spécifique sur les gestes dans nos corpus, ou encore sur l'association entre les formes nominales et les gestes.

**Exemple 3 - Utilisation de la ligne %act pour signaler les manipulations d'objets (corpus de Lubin, 2;5.02)**

\*CHI: ah veux Boubou .

%pho: a vø bubu

%act: attrape doudou

**Exemple 4 - Utilisation de la ligne %act pour signaler les pointages des participants (corpus de Lubin, 2;5.02)**

\*CHI: ça !

%pho: sa

%act: pointe réveil hors-champ

Les lignes de commentaires ont également permis de désambiguïser certains éléments de la situation, lorsque la référence à un objet ou une personne est faite par un pronom par exemple et n'est pas associée de façon claire à un geste ou un regard, comme dans l'exemple ci-dessous.

**Exemple 5 - Utilisation de la ligne %com pour désambiguïser la référence (corpus de Naël, 1;10.25)**

\*MOT: oh qu'est ce qui est sur ton pouce ?

%com: parle du pion rouge au bout du pouce de CHI

Ces lignes nous ont également été utiles pour coder certaines informations paralinguistiques, pour signaler par exemple l'intensité avec laquelle sont produits les énoncés (*cf.* exemple 6) ou encore pour donner des informations sur certains éléments du discours qui ne seraient pas clairement identifiables par des lecteurs de notre corpus, observateurs extérieurs à la situation (*cf.* exemple 7).

**Exemple 6 - Utilisation de la ligne %com pour le marquage d'informations paralinguistiques (corpus de Naël, 1;9.20)**

\*CHI: xxx café maman ?

%pho: X kafe mamã

%com: chuchote

**Exemple 7 - Utilisation de la ligne %com pour le marquage d'informations extralinguistiques (corpus d'Antonin, 1;11.14)**

\*MOT: oh@i regarde ça bouge !

\*MOT: toi t'as déjà pris le bus avec Bibi ?

%com: Bibi est la nourrice d'Antonin

\*MOT: tu te souviens ?

Malgré de nombreuses écoutes, nous n'avons pas toujours pu transcrire phonétiquement les productions de l'enfant, ou les associer à une cible adulte sur la ligne orthographique. Ces éléments sont repérés dans la transcription par des symboles spécifiques référencés dans les conventions de transcription. Mais nous avons également souhaité signaler à plusieurs reprises les doutes que nous avons pu rencontrer quant aux choix de transcription faits sur la ligne principale (orthographique), et nous avons ainsi indiqué notre degré d'incertitude sur une ligne



secondaire dédiée (commençant par %int). Ces repérages sur la ligne d'interprétation nous ont aidée à ne jamais perdre de vue les limites de l'interprétation du transcripteur, à être plus prudente lors des codages, et permettront également aux futurs utilisateurs du corpus de considérer les interprétations de la ligne orthographique avec plus de prudence.

Enfin, une ligne plus générale d'informations sur la situation (%sit) nous a permis de pouvoir comprendre la situation en cours sans un recours systématique à la vidéo. Nous y avons noté des indications sur l'activité, les actions conjointes des participants (telles que les déplacements d'une pièce à l'autre), que nous ne pouvions attribuer davantage à un participant qu'à un autre, ou encore, les événements extérieurs ayant un impact sur les échanges entre les participants (cf. exemple ci-dessous).

**Exemple 8 - Utilisation de la ligne %sit (corpus de Lubin, 1;11.08)**

\*FAT: i(l) r(es)semble un peu à un pingouin c(e) toucan .  
 \*CHI: pingouin .  
 %pho: wêpwê  
 %sit: un jeu musical se met en marche dans la caisse de jouets  
 \*FAT: tiens un qui s(e) réveille .  
 \*CHI: 0 .  
 %act: cherche autour de lui d'où provient le bruit  
 \*FAT: nan i(l) vient pas d(e) là le bruit .

**6.4 La transcription du corpus d'Adrien**

La transcription du corpus d'Adrien a quant à elle d'abord été effectuée sous PHON, logiciel plus adapté à la transcription et aux analyses phonétiques et phonologiques, et le corpus a ensuite été converti au format CHAT pour être ajouté à la base de données CHILDES. Cette conversion n'ayant été effectuée qu'en 2014, les transcriptions que nous avons utilisées pour nos codages sont les transcriptions au format PHON. Tous les énoncés de l'enfant sont transcrits en phonétique entre crochets, et la cible adulte est signalée entre barres obliques. Les énoncés des adultes sont transcrits en orthographe mais aussi parfois en phonétique. Nous avons néanmoins eu recours également aux transcriptions CHAT pour faire des recherches dans le corpus et mener certaines analyses automatiques. Notons que seul le format de la transcription diffère d'un logiciel à l'autre puisque le corpus n'a pas été modifié ou révisé mais seulement converti<sup>20</sup>.

---

<sup>20</sup> Les fichiers joints en annexe de cette thèse sont les transcriptions disponibles sur la base de données CHILDES, et adaptées donc au logiciel CLAN et non PHON.

## 6.5 Les questions soulevées par la transcription

Tout comme le recueil des données a pu susciter chez nous de nombreuses interrogations, l'étape de la mise à l'écrit des séances enregistrées a également soulevé un certain nombre de questions. Nous reviendrons ici sur le caractère interprétatif de l'activité de transcription, qui implique déjà une grande part d'analyse, ainsi que sur les questions propres à l'utilisation du format choisi pour la transcription, et enfin, sur les problèmes spécifiques posés par la transcription des fillers.

### 6.5.1 La transcription : un premier pas dans l'analyse

Si la transcription peut être perçue comme une simple procédure technique visant à reproduire à l'écrit le matériau audio ou vidéo recueilli par le chercheur, celui qui transcrit s'aperçoit rapidement qu'il ne s'agit pas là d'une procédure dénuée de prises de décisions. Quel niveau de détail choisir pour transcrire les différents niveaux de l'interaction ? Comment traiter les problèmes de compréhension des productions verbales des participants ? Sous quel format représenter ces productions ? Ces questions soulevées dans de nombreux travaux (notamment, Bailey, 2008 ; Dannequin, 1989 ; Morgenstern et Parisse, 2007) témoignent de l'importance de la prise en compte de la subjectivité dans le travail de transcription.

De nombreux biais sont en effet à prendre en compte lorsque nous transcrivons des enregistrements audio ou vidéo, à commencer par ceux liés à la discipline, à l'approche théorique et aux objectifs de la recherche (Green *et al.*, 1997 ; Ochs, 1979). Selon les disciplines ou les sous-disciplines, les conventions de transcription utilisées peuvent varier, et l'accent peut être mis sur des aspects différents de l'interaction. A titre d'exemple, une recherche étudiant le développement du langage de l'enfant dans sa dimension multimodale entraînera l'élaboration de transcriptions reflétant l'usage que font les enfants des ressources à la fois verbales et non verbales, alors qu'à l'inverse, une étude traitant également de l'acquisition du langage de l'enfant mais cette fois dans le champ de la phonétique acoustique supposera une transcription bien moins détaillée de la dimension non verbale mais beaucoup plus fine du point de vue des productions verbales. Toutes les données secondaires sur lesquelles sont basées les recherches sont construites par le chercheur dans un cadre disciplinaire et un but particuliers. Nous mentionnions préalablement le choix du format de transcription que nous avons fait. Ce choix

découlant d'une volonté de partage de notre corpus avec les chercheurs de notre communauté, nous avons opté pour le format le plus fédérateur et permettant un traitement des données plus simple. Notre choix est donc guidé par les pratiques de notre communauté scientifique. Les décisions que nous avons prises lors des transcriptions ne sont pas seulement déterminées par les conventions de transcription propres à ce format mais aussi par les objectifs de notre recherche et il est probable que pour une autre recherche nous n'aurions pas mis en avant ou en retrait les mêmes éléments de l'interaction. Une sélection a été opérée en fonction des besoins de notre étude.

Ochs (1979) souligne également que ce processus de sélection des données est lié aux enjeux théoriques du chercheur, et donne notamment pour étayer son argumentation l'exemple des transcriptions en colonnes d'interactions adulte-enfant ou enfant-adulte. D'après elle, le fait d'opter pour une transcription des énoncés ou tours de parole de l'adulte à gauche et de l'enfant à droite « peut non seulement refléter mais véritablement renforcer l'idée de l'adulte dominant » (1979 : 50, *notre traduction*). Ceci est dû pour Ochs à la notion de priorité temporelle attribuée à la partie gauche dans les sociétés occidentales, et par conséquent à l'idée que les points de départ des séquences de l'interaction seraient exclusivement initiés par l'adulte. Ce format de transcription serait donc lié à des *a priori* culturels et conceptuels et biaiserait ensuite la lecture et les déductions du lecteur. Ochs invite donc à repenser la transcription en colonnes en plaçant le verbal et le non verbal de l'adulte à droite et non plus à gauche, afin d'équilibrer la représentation de l'interaction.

L'activité de transcription implique par ailleurs un processus de compréhension préalable, sans lequel il est parfois impossible de cerner ce qui est dit par les participants. Ce processus de compréhension repose sur une interprétation subjective des énoncés produits, et probablement plus subjective encore lorsqu'il s'agit de l'interprétation de données enfantines. Le transcripateur face aux données peut être conditionné par ce que comprennent ou reprennent les interlocuteurs de l'enfant ou par sa vision 'adulto-centrée' du langage, et ainsi fabriquer ce qui n'a pas été produit (Morgenstern et Parisse, 2007). Cette distance entre langage représenté et langage produit donne lieu à une surestimation ou une sous-estimation des compétences des locuteurs (Roberts, 1997), qui peuvent de fait biaiser les analyses ultérieures. Au-delà des biais pour l'analyse, la sous-évaluation des productions des locuteurs peut également entraîner la dévalorisation de la langue

des locuteurs ayant accepté de participer aux recherches (Blanche-Benveniste et Jeanjean, 1987), or il semble important de ne jamais perdre de vue que « nous transcrivons [aussi] des individus lorsque nous transcrivons de la parole » (Roberts, 1997 : 170, *notre traduction*).

Par ailleurs, comme le signalent Roberts (1997) ou Bailey (2008), le travail de transcription suppose de prendre en considération plusieurs aspects parmi lesquels le degré d'interprétation et la façon de représenter les données, que nous avons déjà mentionnés, mais ce travail requiert également un effort de la part du transcripteur pour rendre le texte lisible et lui donner du sens. Cette contrainte de lisibilité n'est pas sans conséquences sur la façon de présenter les données. Le transcripteur peut par exemple limiter la quantité d'informations à faire figurer sur la transcription, ne pas décrire toutes les actions ou tous les gestes des participants, ou bien encore tendre à une interprétation plus importante des productions verbales ou à une représentation conforme à la norme pour faciliter la lecture. Ce besoin de lisibilité peut nuire à la fidélité de la transcription si les choix du transcripteur ne sont pas réfléchis en amont mais peut également répondre de façon tout à fait satisfaisante aux besoins du chercheur. S'il peut exister autant de transcriptions différentes d'un même enregistrement qu'il existe de chercheurs ou de recherches, l'essentiel n'est pas toujours l'harmonisation ou l'uniformisation du style de la transcription mais plutôt la lisibilité en fonction non seulement des besoins de l'étude mais aussi du public visé.

Pour reprendre les propos de Dannequin, nous pouvons dire ici que « le chercheur ne peut pas toujours éviter le filtrage qu'il impose aux données qu'il transcrit mais il doit être conscient des distorsions qu'il fait subir à la réalité » (1989 : 28). En effet, nous ne pouvons pas nous départir du caractère interprétatif et subjectif des transcriptions et nous devons au contraire accepter que cette activité fasse déjà partie de l'analyse des données. Nous pouvons néanmoins nous donner autant d'outils que possible pour étayer nos décisions (Morgenstern et Parisse, 2007) et permettre à ceux qui utiliseront notre corpus d'avoir également conscience de la part d'interprétation du chercheur en jeu dans l'activité de transcription. S'il nous est impossible d'être absolument certain de ce que l'enfant aura voulu produire, nous pouvons étoffer suffisamment nos transcriptions pour justifier les décisions que nous avons prises, en transcrivant par exemple les productions de tous les participants dans l'environnement de l'enfant, susceptibles d'influencer les productions de l'enfant (et pas seulement celles de ses interlocuteurs), en donnant

suffisamment d'informations sur la situation, sur le non verbal (actions, gestes, regards), ou encore sur la prosodie si la compréhension du corpus en est facilitée. Ces mesures ne nous épargnent cependant pas d'allers retours constants entre les données filmées et le corpus écrit et analysé, seule véritable clef pour limiter les représentations que le chercheur peut se faire du corpus.

### **6.5.2 Questions posées par le choix du format de transcription**

Le choix du logiciel nous a contraint à adopter le format de transcription qui en dépend, et ceci nous a semblé pour les raisons évoquées plus haut être un choix satisfaisant. Nous souhaitons malgré cela revenir ici sur quelques aspects qui ont pu nous interpeler lors du travail sur le corpus.

Le format de transcription sous CLAN implique l'enchaînement linéaire des énoncés des participants ; ceux-ci apparaissent donc à la suite les uns des autres. Il nous a semblé que ce format n'était pas le plus à même de rendre compte de la dynamique de l'échange, de la participation des locuteurs et particulièrement des chevauchements entre ces derniers, comme pourrait le faire une transcription présentée en colonnes présentant face à face le verbal et le non verbal des participants. Des conventions existent pour signaler ces aspects du dialogue mais la lisibilité de la transcription s'en trouve affectée et reste moins représentative de ce qui s'est fait et dit lors de l'interaction filmée. C'est pour cette raison d'ailleurs que nous avons fait le choix de présenter les exemples de la thèse (à l'exception de ceux choisis pour illustrer les conventions de transcription) dans un format différent de celui des transcriptions figurant en annexe.

Nous l'avons déjà mentionné plus tôt, il nous a semblé également que la transcription d'une ligne orthographique bien qu'obligatoire pour le traitement automatique des données par le logiciel n'était pas souhaitable pour de très jeunes enfants comme ceux de notre corpus. Si la nécessité d'interpréter les productions de l'enfant peut être bénéfique pour une meilleure compréhension de l'interaction, cette ligne peut conditionner la transcription des productions réelles des participants, et faire tendre la transcription phonétique vers une transcription de parole plus normée que celle réellement produite. Relevons de plus que même si cette ligne reste finalement factice dans notre cas puisque les codages ultérieurs se font d'après la ligne phonétique, nous pouvons nous demander néanmoins dans quelle mesure l'interprétation n'influence pas l'analyse

en surévaluant parfois les compétences langagières de l'enfant. Par ailleurs, cette ligne implique une prise de décision sur le statut grammatical des formes produites par l'enfant. Or, jusqu'à un certain âge, variable d'un enfant à l'autre de notre corpus, les mots de la langue ne sont pas figés, il est donc difficile parfois de les organiser en classes syntaxiques et par conséquent de les transcrire en orthographe.

Si la ligne d'interprétation orthographique peut poser problème au transcripateur, la ligne phonétique suscite également quelques interrogations. De nouveau, pour le bon fonctionnement de la transcription doivent apparaître sur la ligne phonétique autant d'éléments que sur la ligne principale. Cette contrainte implique donc une segmentation de la parole en mots. Là encore il peut être difficile de délimiter la parole de l'enfant en éléments distincts. C'est le cas par exemple dans les contextes de productions figées ou lors de la production d'éléments proto-grammaticaux comme ceux que nous analysons qui peuvent apparaître toujours associés aux mêmes noms, donc rattachés à un noyau lexical et pas encore traités de façon indépendante dans le système de l'enfant. De nouveau, cette segmentation entraîne nécessairement une approche adulte du langage de l'enfant que nous aurions préféré éviter au vu de l'âge des enfants de notre corpus et de leur niveau de développement langagier.

Notons finalement qu'aucun format de transcription ne peut être pleinement satisfaisant et que celui que nous avons adopté nous a paru être le meilleur compromis pour les besoins de partage, d'alignement et de traitement automatique précédemment mentionnés.

### **6.5.3 La transcription des fillers**

Toute transcription peut amener celui qui transcrit à s'interroger sur le degré d'interprétation déjà à l'œuvre dans cette activité ou sur le format à choisir pour rapporter au mieux l'interaction enregistrée. Ceci est d'autant plus vrai lorsque l'on s'intéresse au langage de l'enfant, en construction et parfois encore difficile à appréhender. Notre objet d'étude est particulièrement représentatif de ce langage en mouvement puisque les fillers sont des éléments présents de façon relativement éphémère dans le langage de l'enfant, et sans équivalent dans la langue adulte.

Bien que les fillers aient très tôt suscité l'intérêt des chercheurs (Grégoire, 1937 ; Braine 1963), leur étude n'a véritablement pris de l'importance que tardivement, dans les années 1990, et ceci est probablement lié, comme le signale Morgenstern (2009b), à l'imprécision des transcriptions

jusqu'à alors. L'orthographe étant le plus couramment utilisée pour transcrire le langage enfantin, les fillers peu accentués n'étaient pas transcrits ou bien alors, transcrits comme des formes adultes. La transcription des fillers est en effet rendue difficile par l'instabilité de leur forme et leur faible marquage prosodique. Le travail du transcripteur implique donc d'une part, de ne pas omettre le repérage de ces formes et d'autre part, de ne pas être tenté de transcrire une cible adulte adéquate là où est produite une forme de transition.

Tout le travail de notre thèse reposant sur ces formes, nous y avons accordé une attention particulière lors de la transcription. Néanmoins, nous devons admettre là encore qu'une part d'interprétation est en jeu dans les transcriptions que nous avons faites, et qu'il nous est difficilement possible d'affirmer avec certitude que certains choix que nous avons faits ne seraient pas discutés par un autre transcripteur. Ces cas sujets à discussion sont listés dans les paragraphes suivants.

Nous avons rencontré des difficultés particulières de distinction entre articles indéfinis et fillers. La frontière phonétique entre l'article indéfini masculin singulier /*ɛ̃*/ et certains fillers vocaliques tels que /*ɛ*/, /*a*/ ou /*ə*/ est tellement ténue (dans le premier cas, seul le trait de nasalité distingue le filler de la forme adulte) qu'il nous a parfois été difficile de trancher sur le statut de la forme produite par l'enfant. Cette difficulté s'estompe en partie lorsque l'enfant commence à produire des fillers contenant des consonnes. Il devient alors plus évident de discriminer morphèmes adultes et proto-formes.

La transcription des mots tronqués a également soulevé un certain nombre d'interrogations concernant la présence ou l'absence de formes pré-nominales. Dans certains cas, il nous a été difficile de savoir s'il était préférable de considérer que l'enfant avait seulement omis une ou plusieurs syllabes (en simplifiant phonologiquement ou non l'initiale du mot effectivement produit) ou s'il avait, en plus de l'omission, produit un filler ou un déterminant (semblable ou non à une syllabe du mot cible) en initiale du mot tronqué (*cf.* énoncé ci-dessous, produit par Lubin à 1;11.08). Nous reviendrons sur ces cas dans la troisième partie de cette thèse, lors de la présentation de notre méthodologie de catégorisation des formes pré-nominales (*cf.* section 1.2 du chapitre 5).

### Exemple 9 - Transcription des mots tronqués

\*FAT: là .  
%act: pointe endroit du cube  
\*CHI: (é)léphant .  
%pho: lefã  
\*FAT: oui un éléphant .

La distinction entre proto-présentatifs et proto-déterminants s'est aussi parfois avérée compliquée. Dans les contextes dans lesquels une forme comme 'c'est' pouvait faire sens dans le discours de l'enfant et qu'une seule forme précédait le nom, nous n'avons pas toujours pu donner une interprétation fiable de ce qui avait été produit par l'enfant (*cf.* exemple 10). Ces cas ont été signalés sur la ligne d'interprétation. Si aucun indice ne nous permettait de trancher entre proto-présentatif et filler pré-nominal, les noms ont ensuite été exclus de l'analyse.

### Exemple 10 - Distinction entre proto-présentatif et filler

\*CHI: yyy manteau .  
%pho: te pato  
%act: pointe manteau en regardant OBS  
%int: ici incertitude entre proto-c'est et filler pré-nominal

Dans les quatre corpus de la thèse, les quelques formes s'apparentant phonologiquement à la cible adulte mais produites trop précocement pour être considérées comme de véritables morphèmes adultes ont été transcrites comme des déterminants, sans le symbole réservé aux fillers mais ces cas ont été repérés lors du codage et sont discutés ultérieurement dans la thèse.

Dans les rares cas où la forme produite par l'enfant correspond à une forme de la grammaire adulte mais que le genre de la forme n'est pas celui attendu, nous avons considéré que nous avons affaire à des fillers et signalé leur présence sur la ligne principale. Nous reviendrons sur ces quelques occurrences dans le premier chapitre de présentation des résultats. Cette décision a par ailleurs été suivie également pour les productions inadéquates trouvées devant les verbes.

Quelle que soit la nature du problème rencontré, les cas nous ayant posé de trop grandes difficultés ont été transcrits comme ininterprétables, sur la ligne orthographique et phonétique si nécessaire. Notons néanmoins que notre intérêt principal dans le cadre de cette thèse étant de pouvoir distinguer les cas d'omissions, les cas où sont produits des fillers et ceux dans lesquels l'enfant produit un déterminant de forme adulte, les éventuelles inexactitudes dans la transcription des fillers ont une incidence plus restreinte sur nos analyses que si notre objectif



avait été de pouvoir faire une description très fine des différentes formes phonologiques des fillers dans notre corpus.

De par la plus grande variété des éléments pouvant être produits devant le verbe en français, le repérage et la transcription des formes produites en position pré-verbale ont posé davantage de difficultés. Dans les contextes où sont produits des participes passés ou des candidats à participes passés, la position immédiatement précédente peut en effet être occupée par un proto-pronom, un proto-auxiliaire ou encore un proto-présentatif. Notre étude ne portant pas sur les fillers pré-verbaux, nous avons seulement mentionné la présence d'un filler lorsque cela nous a paru pertinent. Toutefois, pour mener un travail sur les premières formes produites par l'enfant en position pré-verbale, une réflexion méthodologique importante serait à prévoir pour soit associer ces proto-formes à des cibles adultes spécifiques lorsque cela est possible, soit au contraire, les traiter comme un ensemble unique d'amalgames précédant les participes passés.

Notons enfin que les pronoms clitiques de troisième personne pour lesquels le /l/ est éliidé n'ont pas été considérés comme des fillers mais comme des formes adultes. Leur réalisation n'est certes pas conforme à la production standard de ces pronoms mais ce phénomène est suffisamment fréquent en français oral et chez les interlocuteurs des enfants du corpus pour que nous envisagions que la cible adulte à acquérir par l'enfant soit une cible avec omission du /l/.

A la suite de cet état des lieux sur les questions soulevées par la transcription du corpus, il nous semble important de répéter que nous avons conscience que les transcriptions ne sont pas le reflet exact des enregistrements, et encore moins des scènes réelles qui se sont produites devant la caméra. La transcription bien que nécessaire pour notre travail ne reste qu'une entrée possible et non exclusive dans le corpus. Elle représente déjà le début de l'analyse et peut sans cesse être remise en question.

Cet exercice témoigne également pour nous de façon très nette de l'intérêt des recherches collectives. Plusieurs regards et écoutes différentes enrichissent le déchiffrage et la lecture du corpus, et permettent incontestablement de s'approcher au plus près des productions des participants enregistrés. Dans le cadre d'un projet collectif, il aurait été bénéfique et recommandé de faire réviser toutes les transcriptions par une personne dédiée ou par les chercheurs impliqués dans le projet et travaillant sur les mêmes données, mais une telle vérification du corpus n'a pas été possible dans le cadre de cette thèse, sauf pour le corpus d'Adrien. Dans la mesure où celui-ci

avait déjà été intégralement transcrit par un autre chercheur, nous avons pu être le second regard sur la transcription, réécouter les enregistrements, et signaler si nécessaire tout désaccord. Une partie du corpus (4,5%) avait par ailleurs déjà été vérifiée - lors de la thèse de N. Yamaguchi - par un second transcripateur et des analyses acoustiques avaient été effectuées pour identifier les cas incertains ou les cas de désaccord entre les deux transcripateurs.

Si la transcription suscite des interrogations, retenons cependant que toutes ces questions n'ont fait qu'alimenter la réflexion sur la mise en place des morphèmes que nous étudions, et nous ont permis de connaître et de nous approprier notre corpus.

## 7. L'analyse du corpus

La suite de ce chapitre est consacrée à la présentation globale de la grille d'analyse construite pour répondre aux questions que nous nous sommes posées dans ce travail. Le détail des catégories retenues pour chacun de nos axes d'analyse sera exposé au lecteur ultérieurement, avant la discussion des résultats présentés au sein des chapitres V à VII.

### 7.1 Identification et caractérisation des formes analysées

L'analyse d'un énoncé débute par l'identification et la caractérisation de la forme produite en position pré-nominale (lorsqu'un nom est produit par l'enfant). Par défaut, nous parlons de position pré-nominale mais la possibilité de conversion syntaxique en français nous donne également à voir dans nos corpus des déterminants précédant des adjectifs, des adverbes ou des pronoms produits en fonction nominale (*cf.* chapitre V, section 1.1).

Cette partie du codage nous a notamment servi à repérer et identifier la proto-forme ou le morphème grammatical précédant le nom ou l'équivalent du nom (omission, filler ou déterminant) et à préciser la forme et la nature du morphème lexical qui lui est associé. Le repérage des morphèmes lexicaux a facilité un codage *ad hoc* de l'alternance des formes devant les noms lorsque les mêmes noms étaient produits plusieurs fois par le même enfant ("overlaps"). Un des objectifs de cette thèse étant de pouvoir décrire les premières oppositions entre déterminants dans le langage de l'enfant, nous avons également spécifié la nature des morphèmes adultes produits. C'est également à ce stade de l'analyse que nous avons décidé de ne pas analyser certaines formes lorsque l'incertitude quant à leur transcription ou leur statut était trop grande.

## 7.2 Analyse phono-prosodique des formes pré-nominales

Nous avons souhaité pouvoir confronter nos données à celles des travaux sur la dimension phono-prosodique du développement des fillers et des premiers déterminants (entre autres, Veneziano et Sinclair, 1997 ; Veneziano, 2014b ; Demuth, 2001 ; Demuth et Tremblay, 2008 ; Demuth *et al.*, 2012). Rappelons que dans le cadre de ces travaux la production des fillers par l'enfant pourrait être un moyen pour celui-ci de préserver la structure syllabique et rythmique de la langue. Pour Demuth en particulier, le développement des déterminants en français serait conditionné par une unité prosodique, le *pied*, correspondant à l'ensemble des syllabes qui constituent le nom en tant qu'unité rythmique. Les premiers mots des enfants auraient tendance à se conformer au pied le plus commun dans la langue à laquelle ils sont exposés. Les premiers morphèmes grammaticaux seraient majoritairement inclus dans cette hiérarchie prosodique et beaucoup plus rarement en dehors des schémas préférentiels de la langue. En français, les fillers seraient donc produits davantage dans le but de préserver la structure bisyllabique (et iambique) de la langue et seraient donc majoritairement produits, de même que les déterminants devant des noms monosyllabiques (*i.a.*, Demuth et Tremblay, 2008).

Nous avons donc repris pour notre étude la méthodologie mise en place dans le cadre de ces recherches et décompté pour chaque occurrence de nom le nombre de syllabes du nom et du syntagme phonologique. Nous avons donc ainsi pu observer si ces formes sont effectivement produites dans notre corpus en majorité devant les mots d'une seule syllabe.

Nous tenons à préciser que la méthodologie adoptée pour ce codage a été réfléchie en concertation avec N. Yamaguchi dans le cadre de travaux collectifs que nous menons actuellement sur l'acquisition du paradigme des déterminants.

## 7.3 Analyse du statut sémantique des noms

Tous les noms relevés dans notre corpus ont ensuite fait l'objet d'un codage sémantique. A la suite des travaux de Bassano (notamment, Bassano, 2008 ; Bassano *et al.*, 2008, Bassano, 2010) sur la relation de dépendance entre développement grammatical (ou grammaticalisation) des noms et développement lexical, nous avons voulu tester - comme l'a fait Nashawati dans sa thèse (2010) - l'effet de la propriété sémantique sur l'acquisition des déterminants. Pour cela nous

avons procédé à un codage concernant le caractère animé/inanimé et concret/abstrait des noms produits dans notre corpus.

Rappelons que pour Bassano, si la prosodie influence le plus la production des fillers dans une première période du développement langagier de l'enfant, ce sont les critères lexicaux qui deviennent ensuite prépondérants puisque les enfants produiraient un nombre plus important de fillers et de déterminants devant des noms inanimés concrets. Nashawati (*ibid.*) a quant à elle observé que les trois enfants de ses corpus (âgés de 1;11 à 3;0 ans) ne faisaient pas d'emploi privilégié des fillers devant un type de nom spécifique mais que les premiers déterminants de forme adulte apparaissaient majoritairement devant des noms inanimés.

Cette analyse a été menée sur tous les noms et pour les quatre enfants du corpus. Les résultats ont ensuite été mis en regard avec la fréquence de production des différents types de noms dans nos données.

#### **7.4 Analyse de la dépendance au lexème et aux structures syntaxiques**

Nous avons par ailleurs souhaité mesurer l'impact du contexte lexical et de la structure syntaxique sur la production des formes précédant le nom. Nous avons vu précédemment que les approches constructivistes dites 'usage-based' du développement de la grammaire de l'enfant (entre autres, Pine et Lieven, 1997 ; Tomasello, 2000a) ont montré que ce dernier n'a pas accès d'emblée aux catégories syntaxiques adultes (contrairement à ce qui avait été avancé par Valian, 1986) et qu'au contraire, en ce qui concerne le déterminant notamment (Pine et Lieven, 1997), ou le pronom (Lieven *et al.*, 1997), la catégorie se construit seulement progressivement, à partir des usages les plus fréquents de la langue de l'enfant et de ceux de ses interlocuteurs.

Contrairement aux études de Bassano précédemment citées, les noms devant lesquels sont produits les déterminants seraient pour Pine et Lieven (1997) sémantiquement très hétérogènes mais les déterminants produits par l'enfant seraient d'abord liés à des noms spécifiques ou des constructions syntaxiques particulières et leur production serait ensuite seulement généralisée à tous les noms.

C'est donc ce que nous avons mesuré en analysant d'une part l'alternance des formes devant les noms produits à plusieurs reprises par l'enfant et d'autre part les structures syntaxiques des énoncés contenant des noms.

### **7.5 Caractérisation pragmatico-discursive des référents**

A la suite des travaux menés par Salazar Orvig *et al.* (notamment, Salazar Orvig et Morgenstern, 2002 ; Salazar Orvig *et al.*, 2008, et Salazar Orvig *et al.*, 2013) et de leurs constats sur le rôle des aspects fonctionnels dans les premières productions de mots grammaticaux, nous avons également décidé de prendre en compte dans nos analyses certains facteurs liés à l'interaction et au discours, afin de savoir si les formes produites par l'enfant pouvaient, avant même la maîtrise des morphèmes adultes, être associées à des usages discursifs spécifiques.

Nous avons ainsi caractérisé chacun des référents associés aux noms selon notamment leur type de référence, leur statut attentionnel et discursif ou leur familiarité pour l'interlocuteur de l'enfant. Nous avons également systématiquement repéré la mention immédiatement précédente du référent associé au nom afin de savoir quel locuteur l'avait produite et sous quelle forme.

### **7.6 Analyse des phénomènes de reprise**

Enfin, notre recherche ne pouvait se dispenser d'une analyse des phénomènes de reprise et de leur rôle dans le développement de la position pré-nominale. Dans la lignée des nombreux travaux sur la reprise (parmi lesquels on peut citer Bertin, 2011 ; Salazar Orvig, 2000 et Bernicot, Salazar Orvig et Veneziano, 2006), qui témoignent notamment de l'impact des reprises et plus largement du dialogue sur l'acquisition des morphèmes grammaticaux ou lexicaux, nous avons souhaité pouvoir observer si l'enfant s'appuie sur les productions nominales de l'adulte et les reprend pour construire ses énoncés.

Pour chacun des noms produits par les enfants, nous avons donc repéré l'éventuelle mention précédente du même nom dans la séance en cours et précisé dans chaque cas la source, la distance de la reprise et la forme pré-nominale de la mention précédente.

## 7.7 L'outil utilisé pour l'analyse

Une fois la grille d'analyse élaborée, les codages ont été effectués manuellement dans une grille Excel (*cf.* exemple 11). L'utilisation d'un tableur pour le traitement de nos données nous permettait de croiser facilement les énoncés analysés (apparaissant verticalement dans notre grille) avec chacun des aspects que nous souhaitions renseigner (apparaissant cette fois horizontalement), en rendant plus faciles la lecture et la manipulation de nos codages. Nous avons fait figurer sur la même feuille Excel toutes les séances d'un même enfant, afin de mesurer rapidement certains phénomènes sur l'intégralité du corpus d'un enfant. Les codages figurant sur chaque feuille suivent le déroulement de l'interaction au sein de chaque séance ainsi que l'ordre chronologique des séances. Nous devons le choix de cette méthodologie de codage au projet DIAREF, que nous avons déjà mentionné au sein des chapitres précédents et auquel nous avons participé au cours de cette thèse.

Chaque grille est constituée d'une série de colonnes de description du corpus. Ces métadonnées renseignent sur le corpus, l'âge de l'enfant, son MLU ainsi que sur la situation en cours. La colonne suivante est réservée aux énoncés produits par les participants, en phonétique et en orthographe pour l'enfant et en orthographe seulement pour l'adulte, sauf dans le cas du corpus d'Adrien pour lequel les énoncés des interlocuteurs de l'enfant ont également fait l'objet d'une transcription phonétique.

Les colonnes suivantes sont le reflet de l'approche multidimensionnelle que nous avons souhaité adopter sur le développement de la position pré-nominale chez l'enfant, et s'articulent donc autour de plusieurs aspects. Chaque nom ou syntagme nominal produit par l'enfant a été repéré et analysé selon les critères à la fois prosodiques, sémantiques, syntaxiques et pragmatiques que nous venons d'évoquer et sur lesquels nous reviendrons de façon plus détaillée dans les chapitres à venir. Tous les énoncés contenant plusieurs noms ont été dédoublés pour que chaque nom puisse être analysé sur une ligne consacrée.

## Exemple 11 - Grille de codage sous Excel (corpus d'Adrien)

Record #	ENF	Speaker	Orthography	IPA Actual	raison	Analyzabilit	Item analysé	Nature du mot analysé	PREPPOS	Catégorie	SPEC DET	Morph. lex. associé	Type	Harmonies	Lib.sélf.com
9	8	2:05 Adrien	[Bateau]	[bato]	a	ANA	NEXP	NOMM	SEUL	NEXP	NPER	bateau	ZERO		DEUX
12	11	2:05 Adrien	[Peur ,jai peur , bateau ]	[xpcal la pø: bato]	a	ANA	NEXP	NOMM	PREP	NEXP	NPER	peur	ZERO		UN
13	11	2:05 Adrien	[Peur ,jai peur , bateau ]	[xpcal la pø: bato]	b	ANA	NEXP	NOMM	POST	NEXP	NPER	peur	ZERO		UN
14	11	2:05 Adrien	[Peur ,jai peur , bateau ]	[xpcal la pø: bato]	c	ANA	ZERO	NOMM	POST	ZERO	NPER	bateau	ZERO		DEUX
18	15	2:05 Adrien	[L' eau]	[l o]	a	ANA		NOMM	SEUL	DETE	DDEF	eau	invoyle		UN
27	24	2:05 Adrien	[Non papa ]	[nø papa]	a	ANA	NEXP	SMNP	SEUL	NEXP	NPER	papa	V		NPER
28	25	2:05 Adrien	[xxx lit papa ]	[bo li lu a papa]	a	ANA	NEXP	SMNP	SEUL	NEXP	NPER	papa	V		NPER
30	27	2:05 Adrien	[Lis papa ]	[li papa]	a	ANA	NEXP	SMNP	POST	NEXP	NPER	papa	ZERO		NPER
58	54	2:05 Adrien	[Piscine papa ]	[pisiil papa]	a	ANA	NEXP	SMNP	SEUL	NEXP	NPER	piscine			DEUX
59	54	2:05 Adrien	[Piscine papa ]	[pisiil papa]	b	ANA	NEXP	SMNP	SEUL	NEXP	NPER	papa			NPER
62	57	2:05 Adrien	[Piscine ! Piscine ]	[pisiic pisiic]	a	ANA	ZERO	NOMM	SEUL	NEXP	NPER	piscine	ZERO		DEUX
63	57	2:05 Adrien	[Piscine ! Piscine ]	[pisiic pisiic]	b	ANA	ZERO	NOMM	SEUL	NEXP	NPER	piscine	ZERO		DEUX
65	59	2:05 Adrien	[Piscine ]	[pisi:]	a	ANA	ZERO	NOMM	SEUL	NEXP	NPER	piscine	ZERO		UN
67	61	2:05 Adrien	[x@pm peur bateaux   x@pm peur x@pm bateaux ]	[o pur bato o pu de bato]	a	ANA	NEXP	NOMM	POST	NEXP	NPER	peur			UN
68	61	2:05 Adrien	[x@pm peur bateaux   x@pm peur x@pm bateaux ]	[o pur bato o pu de bato]	b	ANA	ZERO	NOMM	POST	ZERO	NPER	bateau	ZERO		DEUX
69	61	2:05 Adrien	[x@pm peur bateaux   x@pm peur x@pm bateaux ]	[o pur bato o pu de bato]	c	ANA	NEXP	NOMM	POST	NEXP	NPER	peur			UN
70	61	2:05 Adrien	[x@pm peur bateaux   x@pm peur x@pm bateaux ]	[o pur bato o pu de bato]	d	ANA	de	NOMM	POST	FILLER	NPER	bateau	CøVc		DEUX
81	72	2:05 Adrien	[Parc ]	[paʁ]	a	ANA	NEXP	NOMM	SEUL	NEXP	NPER	parc			UN
91	82	2:05 Adrien	[Naomi]	[nami]	a	ANA	NEXP	NOPR	SEUL	NEXP	NPER	naomi			UN
94	85	2:05 Adrien	[Oui papa ]	[wi papa]	a	ANA	NEXP	SMNP	SEUL	NEXP	NPER	papa			NPER
96	87	2:05 Adrien	[Oui , Naomi ]	[wi mi]	a	ANA	NEXP	SMNP	SEUL	NEXP	NPER	naomi			UN
99	90	2:05 Adrien	[xxx film Naomi film Naomi ]	[akwi fi nami fi na:mi]	a	ANA	NEXP	NOPR	SEUL	NEXP	NPER	naomi			DEUX
100	90	2:05 Adrien	[xxx film Naomi film Naomi ]	[akwi fi nami fi na:mi]	b	ANA	NEXP	NOPR	SEUL	NEXP	NPER	naomi			DEUX



Suite à l'état des lieux théorique et méthodologique effectué dans les quatre chapitres précédents, nous commenterons dans la partie à venir (partie III) les résultats des analyses que nous venons de décrire. Comme préalablement annoncé en introduction, le prochain chapitre (chapitre V) sera consacré à la description des formes produites devant les noms du corpus. Les chapitres VI et VII traiteront respectivement de l'influence des facteurs distributionnels et fonctionnels sur la mise en place des déterminants. Le dernier chapitre avant la conclusion nous permettra de mettre en perspective les différents facteurs considérés au sein des chapitres précédents, afin de définir si un ou plusieurs facteurs prédomine(nt) tout au long du processus de construction des déterminants, ou si *a contrario* tous les facteurs influencent l'acquisition de façon complémentaire.





# PARTIE III

## *Analyses et résultats*

---

- CHAPITRE V**      Distribution et description des formes pré-nominales
- CHAPITRE VI**     Influence des facteurs distributionnels sur la mise en place des déterminants
- CHAPITRE VII**    Influence des facteurs fonctionnels sur la mise en place des déterminants
- CHAPITRE VIII**   Observation croisée des influences distributionnelles et fonctionnelles



# CHAPITRE V

*Distribution et description des formes pré-nominales*

---



Ce premier chapitre de présentation et de discussion de nos résultats porte sur la distribution et l'évolution des formes pré-nominales chez les quatre enfants. Dans la première section, nous décrirons la façon dont nous avons d'une part identifié les positions dites 'pré-nominales', et d'autre part, caractérisé les formes produites dans ces positions. Nous ferons ensuite un état des lieux de leur développement chez les quatre enfants du corpus, et en dernier lieu, nous discuterons de l'évolution de la position pré-nominale dans nos données, en abordant trois points que nous avons jugés particulièrement notables : la variation, de rythme comme de style, dans la mise en place des formes, la non-linéarité du développement de ces formes, et enfin, le caractère évolutif des fillers et des déterminants au cours de la mise en place du paradigme des déterminants.

## **1. Identification des positions dites 'pré-nominales' et catégorisation des formes**

Afin d'analyser le développement du paradigme des déterminants, nous avons au préalable identifié les positions susceptibles d'être caractérisées par la production d'un déterminant dans la langue adulte, et ensuite, déterminé la nature des éléments produits par l'enfant. La méthodologie adoptée pour notre codage, ainsi que les questions qu'elle a pu engendrer, sont exposées dans les deux points qui suivent.

### **1.1 Identification des positions pré-nominales**

Notre étude reposant sur l'analyse des facteurs pouvant influencer la mise en place des déterminants, nous avons cherché dans un premier temps à dessiner un panorama des formes produites devant les noms par chacun des enfants. Une fois le corpus transcrit, nous avons

analysé chaque énoncé produit et relevé toutes les formes pouvant être étiquetées comme des noms. Nous avons ensuite procédé au repérage des positions dans lesquelles un déterminant pouvait être attendu.

Si la démarche peut sembler paradoxale pour une période où le processus de grammaticalisation du nom est tout juste amorcé, et pendant laquelle nous n'envisageons pas que l'enfant puisse avoir acquis les catégories grammaticales (et en particulier les noms, dans la mesure où le déterminant n'est pas produit de façon systématique), elle était néanmoins nécessaire pour pouvoir observer les contextes dans lesquels l'enfant n'avait produit aucune forme devant le nom. Dans le cas où nous aurions seulement cherché à observer les occurrences de déterminants et de fillers, nous aurions pu utiliser ces formes comme voie d'entrée dans le corpus, mais la recherche des absences de formes justifiait le besoin d'une approche méthodologique différente. Afin d'alléger la rédaction, nous parlerons des 'noms' produits dans le corpus et de position 'pré-nominale', même s'il aurait été préférable d'avoir recours à une terminologie plus fidèle au caractère labile de la construction des catégories chez l'enfant et parler en conséquence de 'proto-noms' et de position 'pré-proto-nominale', chaque fois que le stade de développement langagier de l'enfant le justifiait.

Dans la mesure où CLAN impliquait de faire figurer les productions des participants en orthographe sur la ligne principale (*cf.* point 6.2 du chapitre IV), une grande partie du travail de 'catégorisation' des éléments du discours de l'enfant avait été fait en amont, et le repérage des candidats à noms n'a pas soulevé de difficultés notables au moment de l'analyse. La plupart des énoncés et des segments considérés comme non interprétables au moment de la transcription ont également été codés comme non interprétables lors de l'analyse. Il est toutefois arrivé à quelques reprises qu'un élément pour lequel nous n'avions pas pu faire d'interprétation préalable, et donc d'hypothèse orthographique, puisse - lors de l'analyse et d'un second examen de l'interprétation (éventuelle) de l'adulte et du contexte - être finalement examiné. Tous les énoncés interprétables mais ne contenant pas de nom ont été conservés dans la grille mais n'ont de toute évidence pas été traités, et pour tous les énoncés contenant des noms, la cible adulte du lexème a été reportée (au singulier) dans une colonne spécifique, afin de repérer ultérieurement la présence d'associations préférentielles entre lexèmes et formes pré-nominales (*cf.* point 3 du chapitre suivant sur l'influence du nom sur la production des formes pré-nominales).

Notons que si le déterminant est attendu en français devant les noms communs, nous avons cependant considéré les noms propres et les semi-noms propres (tels que ‘maman’, ‘papa’ ou ‘doudou’) comme des contextes pouvant occasionnellement admettre un déterminant et nous ne les avons, par conséquent, pas exclus de l’analyse. De même, certains éléments préférentiellement utilisés comme des adjectifs (exemple : ‘petit’) ou comme des pronoms (exemple : ‘autre’) mais produits dans notre corpus en fonction nominale ont été catégorisés comme des lexèmes pouvant être associés à des déterminants. En parcourant nos données, aussi bien lors de la transcription que du codage, nous avons par ailleurs pu observer quelques occurrences de fillers ou de déterminants devant des pronoms dans des contextes où ceux-ci n’étaient donc pas attendus (exemple : ‘le ça’ produit par Adrien à 3;3). Notre intérêt se portant sur l’usage des formes pré-nominales plus que sur les noms en eux-mêmes, ces cas ont été repérés et intégrés au répertoire des syntagmes analysés, malgré le caractère non conventionnel de l’association entre forme pré-nominale et pronom. Enfin, précisons également que dans les quatre premières séances du corpus d’Antonin, l’enfant fait référence aux voitures et à certains animaux (pingouin, vache, mouton, ...) avec des termes idiosyncrasiques ([buba] pour ‘voiture’ par exemple ou [kala] pour ‘pingouin’). Dans la mesure où ces formes sont produites de façon systématique, nous les avons considérées comme des noms communs bien qu’elles ne constituent pas des lexèmes conventionnels dans la langue adulte. Ces mots inventés par l’enfant sont signalés par le codage ‘@c’ sur la ligne orthographique de la transcription. Si nous avons tenu au début de cette section à signaler le caractère discutable de l’emploi de ‘nom’ à un âge où les catégories grammaticales ne sont pas acquises, nous pouvons ici aussi remarquer que nous y aurons recours par extension, pour couvrir aussi bien des noms que des formes non conventionnelles ou des formes utilisées comme des substantifs mais dont les usages les plus fréquents dans la langue sont autres.

Outre le fait d’avoir noté la cible adulte de chacun des lexèmes, nous avons également identifié la catégorie, ou plus précisément, la proto-catégorie (nom commun, nom propre, semi-nom propre, adjectif, adverbe ou pronom) de l’élément précédé ou pouvant être précédé d’un déterminant.

## 1.2 Catégorisation des formes pré-nominales

Une fois les positions pré-nominales sélectionnées, nous avons cherché à caractériser les formes produites devant les noms. Les cinq cas de figure distingués sont présentés dans le tableau ci-dessous (tableau 1). Tous les exemples sont issus du corpus d'Adrien.

**Tableau 1 - Caractérisation des formes pré-nominales**

<b>Absence de forme non attendue (ZERO)</b>	Aucune forme pré-nominale n'est produite, dans un contexte où elle serait attendue dans la langue adulte.	ENF (2;7) : [Ø <u>gi</u> la] 'Ø chaise là' (2;7)
<b>Absence de forme attendue (NEXP)</b>	Aucune forme pré-nominale n'est produite, dans un contexte où elle ne serait pas non plus attendue dans la langue adulte : <ol style="list-style-type: none"> <li>1 - noms propres</li> <li>2 - semi-noms propres</li> <li>3 - devant des composants de prédicats complexes</li> <li>4 - dans des contextes dialogiques ne nécessitant pas la production d'un déterminant (suite à une amorce de l'adulte ou à la production d'un SN sans déterminant).</li> </ol>	<ol style="list-style-type: none"> <li>1 - ENF (3;3) : [t a vy Ø akemi] 't'as vu, Ø Akemi ?'</li> <li>2 - ENF (3;3) : [de me a jinə Ø mamã] 'F mets F jean Ø maman'</li> <li>3 - MER : Adrien, on va faire dodo ? ENF (2;8) : [nɔ̃ pa Ø dodo] 'non pas Ø dodo'</li> <li>4 - MER : vas-y dis ! piscine ! ENF (2;5) : [Ø pisil papa] 'Ø piscine papa'</li> </ol>
<b>Filler pré-nominal (FILLER)</b>	L'enfant produit un filler pré-nominal.	ENF (2;7) : [da duç] 'F douche'
<b>Déterminant (DETE)</b>	Un déterminant est produit devant le nom.	ENF (2;7) : [lə lu] 'le loup'
<b>Cas incertain (NINT)</b>	La forme produite par l'enfant ne peut pas être caractérisée.	ENF (2;7) : [X bato] 'xxx bateau'

Nous avons donc pour chaque nom (ou autre élément utilisé comme un nom) relevé la forme le précédant - forme phonétique pour les fillers et déterminants, ZERO pour les absences de formes non attendues, NEXP pour les absences attendues et NINT pour les cas incertains - et dans la colonne suivante, nous avons fait figurer la catégorie de la forme produite : FILLER pour



signaler la présence d'un filler pré-nominal, DETE pour la présence d'un déterminant, et pour les autres catégories, les codes utilisés dans la colonne précédente ont été reproduits.

Nous l'avons déjà mentionné en abordant les difficultés que nous avons rencontrées lors de la transcription (*cf.* point 6.5.3 du chapitre précédent), certaines formes pré-nominales ont soulevé des questionnements quant à leur catégorisation. Trois cas de figure ont particulièrement retenu notre attention : les cas des premiers déterminants associés à des noms dans des usages figés, les déterminants dont le genre ou la forme ne sont pas appropriés au nom associé, et les phonèmes produits devant des noms tronqués.

Nous avons en effet relevé dans nos données certaines formes correspondant phonologiquement à des déterminants de forme adulte mais associées à des lexèmes spécifiques et produites à des âges où les déterminants ne sont pas encore présents dans d'autres contextes. C'est le cas par exemple du défini élidé 'l' associé au nom 'eau' dans le syntagme 'l'eau' fréquemment produit dans notre corpus. Il nous semble assez net que l'enfant produit ce syntagme de façon figée, sans que l'on puisse véritablement envisager ici qu'il s'agisse d'un déterminant. Dans le corpus d'Adrien par exemple, c'est la seule forme pré-nominale adéquate relevée à 2;1. Toutefois, afin d'éviter d'une part une surinterprétation du développement langagier de l'enfant à chaque stade, et d'autre part, de devoir ultérieurement trancher sur le moment auquel ces formes pourraient être considérées comme des déterminants à part entière, nous avons opté pour un critère exclusivement formel et avons donc catégorisé ces premières formes phonologiquement adéquates comme des déterminants. Cette catégorisation ne nous empêchera toutefois pas de revenir sur ces cas particuliers, notamment lors de la discussion de l'évolution de la catégorie 'déterminant' au point 3.3 de ce chapitre.

Dans quelques cas, lorsque la forme pré-nominale correspondait à une cible adulte mais que celle-ci n'était pas conforme au genre du nom auquel elle était associée, nous avons considéré que nous avions affaire à un filler plutôt qu'à un déterminant. Ce choix, comme le précédent, est très certainement contestable, mais il nous a semblé préférable d'envisager que ces formes puissent être le reflet du processus en cours, de construction de la catégorie 'déterminant', et donc le propre des proto-formes, plutôt que la marque d'une 'erreur' de genre. Nous avons par ailleurs pu observer, en particulier dans les dernières séances d'Antonin, quelques cas dans lesquels l'enfant produit une préposition suivie d'un défini (ex : 'à le bras' produit par Antonin à

3;2) dans un contexte où un défini contracté serait attendu, ou une forme pleine au lieu d'un déterminant élidé ('la autre voiture', produit également par Antonin à 2;9) Ces formes marquent l'existence de schémas de type 'préposition + défini + nom' et 'déterminant féminin + nom féminin' et attestent de leur productivité. Bien que ces généralisations ne puissent pas être considérées comme des proto-formes au même titre que celles que l'enfant produit à un stade moins avancé de son développement, nous les avons ici caractérisées comme des fillers pour les distinguer des cas où la forme produite est conforme à celle qui serait produite dans la langue adulte. Nous y reviendrons lors de la présentation du développement de la position pré-nominale chez Antonin (*cf.* section 2.3 de ce chapitre).

La caractérisation des formes produites devant les noms dont la ou les première(s) syllabe(s) est/sont tronquée(s) a également fait l'objet d'un certain nombre de questionnements. Trois cas différents ont été distingués :

- les cas comme [apɛ̃] ou [apo] souvent cités dans la littérature (*cf.* notamment Veneziano et Sinclair, 2000) pour lesquels la voyelle initiale est conforme à la voyelle de la première syllabe du mot cible ([a] pour [la] et [a] pour [ʃa]). Dans ces contextes, nous avons considéré que la consonne initiale n'était pas produite et que le nom était précédé d'une absence de forme.
- les cas dans lesquels la seconde syllabe du mot cible est précédée d'une voyelle qui n'est pas conforme à la voyelle du mot cible (exemple : [əpɛ̃] pour [lapɛ̃]). Si le nom est repris plusieurs fois dans la même séance et que la voyelle produite à l'initiale est identique, nous avons considéré qu'il s'agissait d'une simplification phonologique de la cible et donc, d'un nom précédé d'une omission du déterminant. A défaut de pouvoir comparer l'usage qui est fait par l'enfant du lexème dans d'autres contextes, ce codage a également été adopté lorsque le nom n'est produit qu'une fois.
- A l'inverse, lorsque la voyelle initiale n'est pas conforme à celle du mot cible et que le nom est produit plusieurs fois avec une alternance vocalique, autrement dit, précédé de voyelles différentes, nous avons considéré, comme l'ont fait Kilani-Schoch (1998) ou Wauquier et Yamaguchi (2013), que le nom était tronqué et précédé d'un filler.

Pour chaque occurrence de déterminant, nous avons par ailleurs spécifié le type de déterminant produit, en suivant la classification présentée au premier point du deuxième chapitre et en choisissant donc parmi neuf types de déterminants susceptibles d'être produits en français (*cf.*

Riegel *et al.*, 1994) : le défini, le démonstratif, le possessif, l'indéfini, le partitif, le numéral, le quantifieur, l'interrogatif, et enfin, l'exclamatif. Nous verrons ultérieurement que toutes ces catégories ne sont pas aussi pertinentes les unes que les autres pour l'analyse des déterminants dans le langage de l'enfant.

Afin d'analyser l'évolution éventuelle des fillers, nous avons, de la même façon que pour les déterminants, tenu à préciser le type de filler produit par l'enfant. Pour cela, nous avons distingué les fillers de type 'voyelle' (V), 'consonne' (C) et 'consonne-voyelle' (CV). Une catégorie 'autres' (AUTRES) a également été utilisée pour classer les quelques cas de fillers dont le schéma phonologique diffère des trois schémas que nous venons de mentionner.

Les différents codages nécessaires pour d'une part, identifier les positions pré-nominales et d'autre part, catégoriser les formes produites dans ces positions, sont récapitulés dans le tableau ci-dessous (tableau 2).

**Tableau 2 - Récapitulatif des codages réalisés pour l'identification des positions pré-nominales et la caractérisation des formes**

Énoncé	Item analysé	Catégorie item	Spéc. Dét. ou Fill.	Morphème lexical associé	Nature morphème
Adrien (2;7) : [Ø ɛi la] 'Ø chaise là'	ZERO	ZERO	Non pertinent	chaise	NOM
Mère : vas-y dis ! piscine ! Adrien (2;5) : [Ø pisil papa] 'Ø piscine papa'	NEXP	NEXP	Non pertinent	piscine	NOM
Adrien (2;7) : [da duç] 'F douche'	da	FILLER	CV	douche	NOM
Adrien (2;7) : [lə lu] 'le loup'	lə	DETE	DEF	loup	NOM

Ces premières analyses nous ont permis d'observer la façon dont se distribuent les formes devant les noms dans nos données, et ainsi, de pouvoir décrire leur développement au fil des séances analysées, pour les quatre enfants du corpus. Ces résultats sont présentés et commentés au point 2 ci-dessous.

## **2. Développement des formes pré-nominales chez les quatre enfants du corpus**

Des sections 2.1 à 2.4, nous décrivons la distribution des formes produites en position pré-nominale chez chacun des enfants, et au point 3, nous reviendrons sur les caractéristiques communes aux quatre corpus mais aussi, sur les caractéristiques individuelles de la mise en place du paradigme des déterminants. Signalons que sur toutes les catégories codées de formes pré-nominales, nous n'avons retenu dans cette section que les cas de figure les plus pertinents pour étudier le développement des formes pré-nominales : les absences ou omissions non attendues, les fillers et les déterminants. Ont donc été exclus de l'analyse les cas d'absences attendues de forme et les cas incertains.

### **2.1 Dans le corpus de Lubin**

Avant de faire l'état des lieux du développement des formes pré-nominales dans nos corpus, nous avons souhaité donner au lecteur une indication plus globale du niveau langagier des enfants à chacune des séances retenues pour l'analyse. Nous avons ainsi fait figurer, dans les tableaux 3, 4, 5 et 7, la longueur moyenne des énoncés produits par chaque enfant à chaque âge. Bien que cette mesure ne donne qu'un aperçu quantitatif et largement insuffisant de l'acquisition du langage de l'enfant, elle nous permet toutefois de faire quelques constats préliminaires et surtout, de comparer l'évolution des formes pré-nominales chez les quatre enfants au regard de leur développement langagier général.

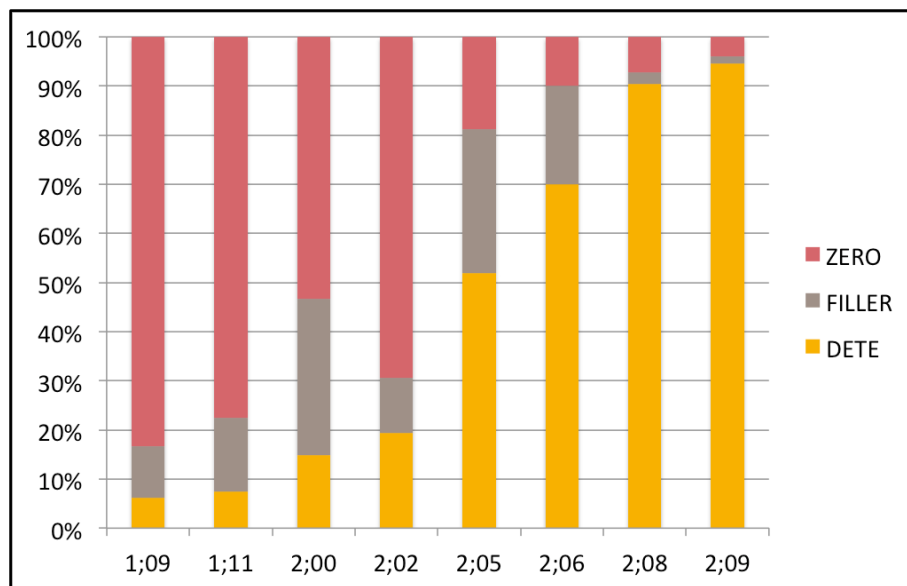
Dans le corpus de Lubin, l'observation de la longueur moyenne des énoncés (LME) nous permet de constater entre autres que les énoncés produits par l'enfant à 1;8.23, lors de la première séance analysée, sont constitués d'un terme en moyenne, et de trois termes environ lors de la dernière séance, à 2;9.06. Les énoncés s'allongent de façon systématique, mais l'augmentation de la LME n'est toutefois pas régulière. On peut par exemple observer qu'elle est minimale entre 1;11.08 et 2;0.14 et qu'au contraire, elle est plus marquée à partir de 2;05.02, à une période où les séances retenues ne sont espacées que d'un mois environ.

**Tableau 3 - Longueur moyenne des énoncés dans le corpus de Lubin**

Âge	1;8.23	1;11.08	2;0.14	2;1.19	2;5.02	2;6.04	2;7.23	2;9.06
LME	1.181	1.435	1.455	1.654	2.009	2.305	2.603	3.009

De même que la LME de Lubin augmente nettement de la première séance à la dernière, le graphique 1 ci-dessous nous permet de constater que la part des formes adultes devant les noms augmente elle aussi considérablement de 1;9 à 2;9<sup>1</sup>. De façon attendue, les déterminants - représentant 6,25% des formes produites lors la première séance - deviennent la forme quasiment exclusive à 2;9 (94,59% des formes en position pré-nominale). *A contrario*, le pourcentage d’omissions diminue et passe de 83,33% à 1;9 à 4,05% dans la dernière séance. D’un point de vue purement formel, nous pouvons donc en conclure que le processus de construction de la catégorie ‘déterminant’ est quasiment achevé à 2;9.

**Graphique 1 - Distribution des formes pré-nominales dans le corpus de Lubin**



Par ailleurs, nous pouvons signaler que les fillers ne sont à aucun stade du développement de l’enfant utilisés comme la forme privilégiée devant les noms. Ceux-ci sont surtout observables à

<sup>1</sup> Précisons ici que l’âge de l’enfant sera l’indicateur retenu pour référer à chaque séance dans la suite de cette troisième partie de notre travail mais que celui-ci sera limité - afin d’alléger la lecture - à l’année et au mois, et donc arrondi.

2;0 et 2;5, séances lors desquelles ils sont produits dans respectivement 31,82% et 29,13% des cas. A la différence des omissions encore visibles à 2;8 et 2;9, les fillers sont quasiment absents des données dans ces deux dernières séances (2 occurrences à 2;8 et 1 occurrence à 2;9).

Outre l'augmentation progressive des déterminants et la diminution globale des omissions, il nous semble intéressant de constater que le processus n'est pas en tous points régulier. La part des déterminants augmente d'une séance à l'autre mais il apparaît clairement que cette augmentation est plus rapide de 2;5 à 2;8 que de 1;9 à 2;2. Il aurait de toute évidence été souhaitable de pouvoir observer comment l'enfant manie les formes pré-nominales entre 2;2 et 2;5. Dans ce but, nous avons initialement transcrit la séance à 2;4, qui se déroule au retour des vacances d'été de la famille de l'enfant. Néanmoins, nous l'avons signalé plus tôt dans notre chapitre de méthodologie, cette séance enregistrée en présence de la sœur de Lubin ne présente qu'une part minimale d'interventions de l'enfant et n'a *in fine* pas pu être conservée pour l'étape d'analyse de nos données.

La différence entre ces deux périodes de quatre séances chacune se manifeste parallèlement dans la proportion que représentent les formes de transition, nettement plus nombreuses dans la période 1 que dans la période 2. Les omissions diminuent progressivement de 1;9 à 2;0 puis de 2;5 à 2;9 mais on note également qu'elles représentent une part plus importante des formes à 2;2 qu'à 2;0. Elles sont en effet présentes devant 53,41% des noms à 2;0 et redeviennent la forme largement privilégiée (69,44% des productions de l'enfant) deux mois plus tard. Ce type de 'régression' déjà évoquée dans la littérature (*cf.* point 2.2.3 du chapitre II), s'observe aussi simultanément pour les fillers. Leur proportion augmente à mesure que la part des omissions diminue de 1;9 à 2;0, puis décroît à 2;2 avant d'augmenter de nouveau à 2;5.

## 2.2 Dans le corpus de Naël

Dans le corpus de Naël, comme dans celui de Lubin, la longueur moyenne des énoncés (*cf.* tableau 4) est globalement croissante - à 1;5.23, les énoncés de l'enfant sont composés d'un terme et demi en moyenne et de plus de trois termes à 2;6.11 - mais n'est cependant pas régulière. La LME est quasiment identique dans les deux premières séances et elle augmente ensuite de 0.3 point environ entre 1;7.18 et 1;8.19. Dans les trois séances de 1;8.19 à 1;10.25 (couvrant une période de deux mois), l'augmentation n'est que de 0.1 point entre chaque séance,

puis les énoncés s'allongent nettement plus rapidement entre 1;10.25 et 2;1.28 (correspondant à la période d'été), puis de nouveau entre 2;1.28 et 2;4.09. Pendant cette période de cinq mois, les énoncés de Naël passent de deux termes en moyenne à plus de trois et demi. Enfin, de 2;4.9 à 2;6.11, la LME diminue de 0.4 point.

**Tableau 4 - Longueur moyenne des énoncés dans le corpus de Naël**

Âge	1;5.23	1;7.18	1;8.19	1;9.20	1;10.25	2;1.28	2;4.09	2;6.11
LME	1.571	1.561	1.828	1.905	2.091	2.698	3.623	3.254

L'observation des données tendant à montrer que le langage de l'enfant est d'une complexité très comparable dans ces deux séances, nous faisons ici l'hypothèse que cette diminution puisse s'expliquer par les activités en cours, par ce que mettent en œuvre les participants dans ces activités, ainsi que par le type de dialogue dans lequel ils sont impliqués. Dans la première partie de la dernière séance, l'enfant et son père lisent un livre. Naël est beaucoup sollicité, au travers de questions qui ne nécessitent pas systématiquement de réponses syntaxiquement complètes ou autonomes (*cf.* exemple 1*a* ci-dessous). Dans la séance précédente, Naël joue à la poupée avec son frère, et sa mère étant occupée, c'est l'enfant qui la sollicite et non l'inverse. Ceci implique peut-être que l'enfant produise des énoncés - davantage en initiative qu'en réponse - qui soient plus complets (se rapprochant d'une structure canonique Sujet-Verbe-Objet), et par conséquent plus longs (*cf.* exemple 1*b* ci-dessous). Cette hypothèse n'ayant pas été vérifiée, nous ne pouvons cependant pas exclure que d'autres facteurs puissent expliquer cette diminution de la LME entre 2;4.09 et 2;6.11.

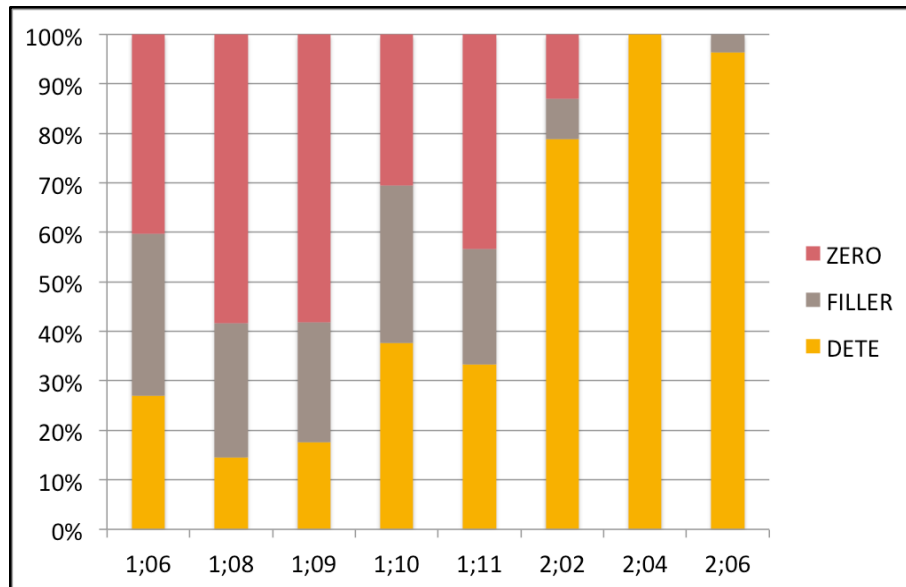
**Exemples 1 - Énoncés produits par Naël dans le cadre de deux activités différentes à 2;4 (b) et 2;6 (a)**

- (a) Père : comment s'appelle cet arbre ?  
 Naël : [lə] 'le +/.'  
 Père : palétuvier.  
 Naël : [paletyvje] 'palétuvier.'
- (b) Naël : [mamã ty pø lə mɛt sa ʋɔb] 'maman tu peux F met(tre) sa robe ?  
 Mère : nan pas si tu me parles comme ça.  
 Naël : [ty pø lə mɛt sa ʋɔb sil tə plɛ] 'tu peux F met(tre) sa robe s'il te plaît ?'  
 Mère : ah ça c'est mieux.

Si l'on considère que la LME nous fournit un indice sur le développement global du langage de l'enfant, nous pouvons constater que chez Naël, celle-ci a lieu sur une période assez courte. En comparant ces données à celles des autres enfants de notre corpus, nous avons notamment pu remarquer que les premiers mots apparaissent plus précocement (d'où notre choix de conserver pour l'analyse ultérieure la séance enregistrée à 1;6) et les énoncés atteignent trois termes plus rapidement (six mois plus tôt que dans le corpus de Lubin et environ un an plus tôt que chez Antonin et Adrien). Nous pouvons donc faire l'hypothèse que parallèlement, le développement des formes pré-nominales sera non seulement rapide mais également plus précoce que chez les trois autres enfants.

Et en effet, le graphique 2 ci-dessous nous confirme qu'en moins d'un an seulement - de 1;6 à 2;2 - les déterminants deviennent la forme la plus fréquemment utilisée devant les noms. On observe par ailleurs que comme chez Lubin, l'acquisition des formes adultes n'est pas régulière : de 1;6 à 1;11, les omissions et fillers sont majoritaires, puis de 2;2 à 2;6, les déterminants représentent au minimum 78,82% des formes pré-nominales.

**Graphique 2 - Distribution des formes pré-nominales dans le corpus de Naël**



Bien que ces deux périodes se dessinent clairement dans les données de Naël, les séances qui les composent ne sont pas absolument comparables. Dans la première période, la distribution des trois formes est presque identique à 1;8 et 1;9 mais la proportion de déterminants augmente de presque 20% entre 1;9 et 1;10. Un mois plus tard, à 1;11, la part des déterminants est similaire



mais la part des fillers diminue au profit des omissions. Une diminution de la part des fillers peut également être relevée entre la première et la deuxième séance, de même qu'une diminution de la part des déterminants, représentant 26,92% des formes en position pré-nominale à 1;6 contre 14,58% des formes deux mois plus tard. Il nous paraît néanmoins important de signaler que la séance à 1;6 ne compte que peu de noms devant lesquels un déterminant est attendu (52 au total) et que 11 déterminants sur les 14 produits à cet âge sont associés au nom 'eau'. L'observation des autres corpus de cette thèse et de données analysées dans le cadre d'autres recherches nous laisse penser qu'il s'agit là d'un usage figé récurrent chez l'enfant ('de l'eau') que nous avons d'un point de vue formel considéré comme un emploi conforme à la cible adulte mais qui ne peut pas être étiqueté comme un déterminant au même titre que les formes produites avec d'autres lexèmes.

Dans la période 2, que nous situons entre 2;2 et 2;6, les omissions et les fillers représentent à eux deux 21,18% des formes produites devant les noms à 2;2, puis disparaissent complètement à 2;4. Trois formes que nous avons considérées comme des fillers sont produites dans la dernière séance. Ces trois cas sont présentés et discutés ci-dessous (exemples 2).

### Exemples 2 - Formes considérées comme des fillers dans la séance 8 du corpus de Naël

- (a) Père : nan mais là qu'est ce qu'il y a là ?  
Naël : [ɔ vjø pɛʃœʁ] 'F vieux pêcheur'
- (b) Naël : [i mɛt de kadju dɑ̃ l dɑ̃ l panje] 'i(ls) mettent des cailloux dans l(e) [/] dans l(e) panier +...  
Père : pour quoi faire ?  
Naël : [a lɔ vjø pɛʃœʁ] '+, F vieux pêcheur'
- (c) Père : ne peux tu pas +..?  
Naël : [fœʁ dɑ̃ bwi] '++ , faire F bruit'  
Père : faire moins de bruit ?

Dans les deux premiers exemples, Naël et son père lisent ensemble un livre que l'enfant connaît bien. Dans l'exemple *a*, le personnage du pêcheur est introduit pour la première fois. Dans la mesure où à cet âge, Naël produit des définis, des indéfinis, des possessifs mais plus de fillers vocaliques, la présence de ce [ɔ] nous a semblé surprenante. Toutefois, si la forme attendue devant 'vieux pêcheur' pourrait être 'un' ou 'le', le défini contracté 'au' est à exclure dans ce contexte, d'où le choix que nous avons fait de caractériser cette voyelle comme un filler. L'exemple *b* illustre un phénomène très différent. Dans le deuxième énoncé de Naël (complétant son énoncé précédent), l'enfant produit une combinaison associant la préposition 'à' et le défini

'le', dans un contexte où le défini 'du' serait attendu. Nous l'avons déjà mentionné dans la première section de ce chapitre, catégoriser ce type de forme comme un filler peut être largement discutable dans la mesure où cette combinaison n'est pas le reflet de ce que nous pouvons communément classer dans la catégorie des protoformes. Toutefois, sans pour autant en faire une catégorie spécifique, nous avons préféré distinguer ces cas de figure des autres formes correspondant du point de vue de la forme à des productions adultes. Ce type de production méritant d'être discuté plus amplement, nous y reviendrons en abordant le développement des formes dans le corpus d'Antonin, au sein duquel plusieurs de ces approximations ont pu être relevées (*cf.* point 2.3 ci-dessous). Dans l'exemple *c*, les deux participants sont en train de lire un second livre. Le père de Naël prend la voix d'un des personnages s'adressant au personnage principal. Suite à l'amorce de l'adulte, Naël produit devant le lexème 'bruit' la forme [də] là où le partitif 'du' serait attendu. Cependant, la réponse conforme au récit (et verbalisée par le père dans l'énoncé qui suit) étant 'faire moins de bruit', on peut également envisager ici que la forme produite par l'enfant puisse être conforme au déterminant adulte mais que l'adverbe 'moins' ait été omis.

Enfin, nous pouvons signaler que de la même façon que chez Lubin - mais à l'inverse de ce que nous pourrions ultérieurement observer chez Antonin et Adrien - les déterminants de forme adulte alternent plus fréquemment dans le corpus de Naël avec des omissions qu'avec des fillers. A l'exception de la dernière séance (*cf.* exemples ci-dessus) et de la séance à 1;10 dans laquelle l'enfant a recours aux fillers dans les mêmes proportions que les omissions, les fillers apparaissent moins souvent que les omissions devant les noms de 1;6 à 1;9 puis à 1;11 et 2;2. Toutes séances confondues, la part des omissions représente 30,63% des formes produites et les fillers, 18,41% des formes observées en position pré-nominale. Nous y reviendrons au point 3.1 lorsque nous aborderons la question de la variation de style dans l'acquisition des déterminants.

### **2.3 Dans le corpus d'Antonin**

L'observation de la longueur moyenne des énoncés dans le corpus d'Antonin nous permet de constater que conformément à ce qui peut être attendu, les énoncés de l'enfant s'allongent de deux mots en moyenne entre 1;8.29 et 3;1.26. L'évolution entre ces deux âges est globalement progressive. On repère cependant que l'augmentation entre 1;11.14 et 2;1.17 n'est que de 0.57

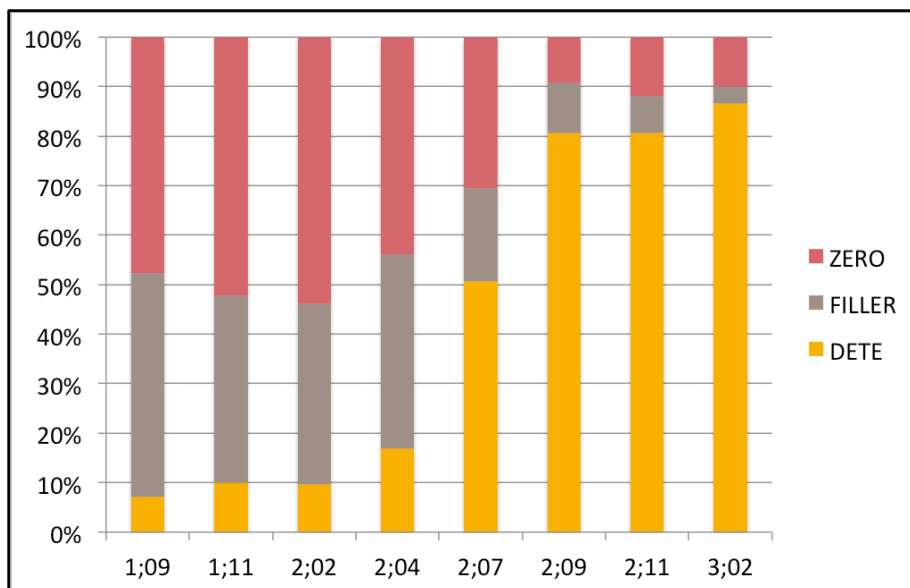
point, et que la LME à 2;11.06 est même plus basse (2.512) qu'à 2;8.28 (2.560). Si les énoncés d'Antonin sont plus longs et plus complexes lors de la dernière séance que lors de la première, la période d'un an et cinq mois couverte entre ces deux âges laisse donc apparaître une évolution qui n'est pas constamment régulière et linéaire.

**Tableau 5 - Longueur moyenne des énoncés dans le corpus d'Antonin**

Âge	1;8.29	1;11.14	2;1.17	2;3.26	2;6.25	2;8.28	2;11.06	3;1.26
LME	1.340	1.696	1.753	2.197	2.380	2.560	2.512	2.961

Contrairement à la LME qui augmente de 1;9 à 2;4, le graphique 3 nous montre que la position pré-nominale est stable pendant ces quatre premières séances et que l'enfant a recours aux omissions, fillers et déterminants dans des proportions comparables. Pendant cette période, les noms sont précédés d'omissions dans 49,40% des cas en moyenne. Les fillers représentent quant à eux 39,68% des formes produites devant les noms et les déterminants, 10,91% des formes en position pré-nominale. Leur proportion augmente toutefois sensiblement de 2;2 (9,62% des formes) à 2;4 (16,89%).

**Graphique 3 - Distribution des formes pré-nominales dans le corpus d'Antonin**



Le panorama des formes produites entre 2;7 et 3;2 est différent puisque les formes de transition sont nettement moins fréquentes et les déterminants deviennent majoritaires. Omissions et fillers représentent à eux deux 83,11% des productions à 2;4 et plus que 49,39% des formes à 2;7.

Parallèlement, la proportion de déterminants entre ces deux séances augmente de 33,72%, pour atteindre 50,61% des formes en position pré-nominale à 2;7. Si l'augmentation des déterminants entre 2;4 et 2;7 est notable, elle l'est tout autant entre 2;7 et 2;9. Puis on constate qu'elle se stabilise lors des deux séances suivantes. A 3;2, les omissions représentent encore un dixième des formes, les fillers ont quant à eux presque disparu et les déterminants atteignent 86,67% des productions.

Il nous a semblé intéressant de constater par ailleurs qu'à la fin de la période étudiée, les déterminants contrastent plus souvent avec des omissions qu'avec des fillers. Ceci est d'autant plus vrai si l'on considère que certains des fillers produits entre 2;9 et 3;2 ne sont pas des fillers au même titre que ceux produits lors des séances précédentes. Rappelons que pendant cette période, certaines des formes codées comme des fillers constituent ce qui pourrait représenter le dernier stade d'approximation des déterminants attendus, tant du point de vue de la forme que de la fonction. La différence entre ces formes et les formes adultes repose le plus souvent sur le genre inapproprié de la forme produite ou sur l'usage d'une forme pleine plutôt qu'élidée ou contractée. Quelques exemples de ces cas de figure sont présentés ci-dessous.

**Tableau 6 - Exemples d'approximations de déterminants produits dans le corpus d'Antonin entre 2;9 et 3;2**

Âge	Exemples
2;9	[dɛ zɔt watyʁ] 'F aut(re)s voitures' [la ɔt vwatyʁ] 'F aut(re) voiture' [ɥn kɔk] 'F coq'
2;11	[ẽ pti pø də lə zɔn] 'un p(e)tit peu de F jaune'
3;2	[d̃ ʒu ẽ ti pø le vwatyʁ] 'on joue un petit peu F voitures' [ãkɔʁ a lə bʁa] 'encore à F bras'

Dans les deux premiers exemples tirés de la séance à 2;9, Antonin produit deux déterminants de forme adulte ('des' et 'la') mais dans des contextes où leur forme élidée serait attendue. Il semble que l'enfant n'ait pas encore acquis de façon systématique l'usage de l'indéfini ou du défini élidé devant 'autre' lorsqu'il est utilisé comme adjectif devant un nom. En revanche, lorsqu'il n'est pas suivi d'un nom et donc utilisé comme un pronom, le lexème 'autre' est utilisé très tôt associé à une forme adulte. Sur les trente occurrences du pronom observées dans les données, 25 d'entre elles le sont précédées d'un défini élidé, 4 d'un indéfini et seulement une d'un filler. Nous

pouvons donc faire l'hypothèse que dans les cas où l'enfant a recours à cette forme en tant qu'adjectif, le déterminant produit est adapté non pas à 'autre' mais au nom qui suit.

Dans l'exemple qui suit ([yn kɔk]), le déterminant attendu est masculin ('un coq') et la forme produite n'est donc pas appropriée au genre du lexème. Dans le contexte où ce syntagme est observé, l'enfant manipule une petite poule en plastique (qu'il a préalablement caractérisée comme une poule) et demande à sa mère si il s'agit bien d'une poule ou plutôt d'un coq. Nous avons ici l'impression que l'enfant combine le genre féminin du référent (la poule) et le lexème masculin 'coq'.

A 2;11, l'enfant suggère à sa mère d'ajouter un peu de jaune à la caserne de pompiers qu'ils sont tous les deux en train de construire et combine l'usage du quantifieur 'un peu de' et du défini 'le'. Cette forme 'le' n'est donc pas non plus ce que nous pourrions caractériser véritablement comme un filler mais davantage comme un usage inadéquat de déterminant.

Enfin, à 3;2, Antonin produit d'une part une forme phonologiquement identique à un défini pluriel 'les' et d'autre part, une préposition ('à') suivie de la forme 'le', dans deux contextes où un défini contracté serait attendu ('aux voitures' et 'au poignet'). Dans la mesure où seuls quatre déterminants de ce type sont produits dans le corpus (dont deux en reprise immédiate de l'adulte et deux associés au même lexème) il nous est difficile de pouvoir dire si par ailleurs, le défini contracté semble acquis par l'enfant.

## **2.4 Dans le corpus d'Adrien**

Dans le corpus d'Adrien, le développement du langage s'étend sur une période relativement longue. La longueur moyenne des énoncés nous indique notamment que l'enfant produit à 2;1 des énoncés constitués d'un terme en moyenne, et de trois termes environ lors de la dernière séance, à 3;5. De façon comparable à ce que nous venons d'observer dans le corpus d'Antonin, la LME n'augmente pas de façon constante. La progression est plus importante de la première à la deuxième séance, séparées de deux mois, qu'entre les autres séances pour lesquelles l'intervalle de temps peut être plus important. Par ailleurs, la LME diminue à 2;8 - 2.091 contre 2.124 un mois plus tôt - avant d'augmenter de nouveau à 2;11. Si ce phénomène 'd'aller-retour' n'exclut pas que le langage de l'enfant puisse toutefois se complexifier entre 2;7 et 2;8, il peut également signifier

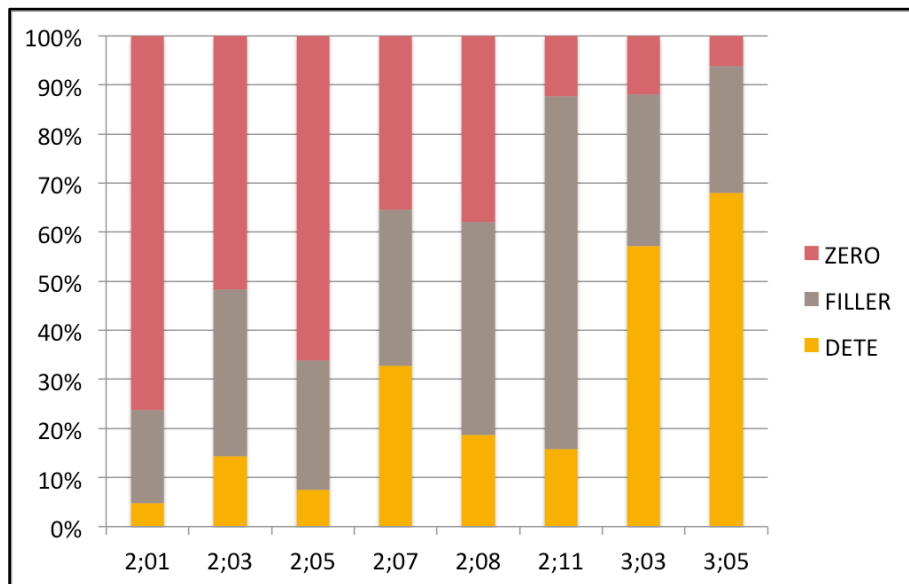
que cette période de fluctuation se manifestera dans la nature du langage produit par l'enfant, et pour ce qui nous concerne, dans la nature des formes produites en position pré-nominale.

**Tableau 7 - Longueur moyenne des énoncés dans le corpus d'Adrien**

Âge	2;1	2;3	2;5	2;7	2;8	2;11	3;3	3;5
LME	1.128	1.661	1.832	2.124	2.091	2.345	2.639	2.997

L'analyse de la distribution des formes devant le nom montre que l'apparition des déterminants chez Adrien est plutôt tardive (*cf.* graphique 4 ci-dessous). Si à 3;5 les formes adultes sont bien majoritaires devant les noms (68,04% des productions pré-nominales), les omissions et les fillers restent assez largement présents puisqu'ils représentent tout de même à eux deux 31,96% des formes produites.

**Graphique 4 - Distribution des formes pré-nominales dans le corpus d'Adrien**



On observe par ailleurs que globalement, de 2;1 à 3;5, la proportion des déterminants augmente : à 2;1, seul un nom est précédé d'un déterminant (dans l'usage figé et en répétition 'l'eau' auquel nous faisons référence précédemment), contre 68,04% à 3;5. Ces formes adultes deviennent la forme prédominante à 3;3 en atteignant 57,14% des formes produites devant les noms. Parallèlement à l'augmentation des formes adultes, la part des omissions diminue de façon importante : à 2;1, celles-ci représentent 76,19% des formes, soit une très large majorité, et à 3;5, elles ne sont plus observées que dans 6,19% des contextes pré-nominaux. Les fillers suivent un

développement différent puisque leur proportion tend à rester assez stable jusqu'à 2;7, puis augmente ensuite entre 2;8 et 2;11 avant de diminuer nettement entre 2;11 et 3;3. Dans la dernière séance, ils représentent encore un quart (25,77%) des formes pré-nominales. L'examen de ce graphique montre donc qu'à la fin de la période observée, Adrien n'a pas recours systématiquement aux déterminants, néanmoins le processus de construction de la catégorie est bien avancé, et sur le plan exclusivement formel, bientôt achevé.

Enfin, nous pouvons repérer que si la distribution des formes produites devant le nom est en évolution constante et tend, d'une séance à l'autre, à révéler un accroissement des formes adultes et une réduction de la part des formes de transition, le développement n'est toutefois pas complètement linéaire. Certaines séances sont en effet marquées, comme chez Lubin, par des 'régressions' dans la mise en place des formes adultes. Par exemple, à 2;7, Adrien produit des déterminants plus fréquemment qu'à 2;8 ou 2;11, et à l'inverse, à 2;5, le pourcentage d'omissions est plus important qu'à 2;3.

### **3. Bilan comparatif sur l'évolution de la position pré-nominale dans le corpus**

#### **3.1 Des variations de rythme et de style**

Nous avons pu constater dans la section précédente que dans tous les corpus, les déterminants apparaissent dans les dernières séances retenues pour l'analyse comme la forme largement majoritaire mais nous avons pu remarquer également que le processus d'acquisition des formes adultes peut s'étendre sur une période dont la durée varie d'un enfant à l'autre. Dans le corpus de Lubin, les fillers et omissions ne représentent plus que 5,41% des formes devant les noms à 2;9. Chez Naël, ces formes de transition ne sont plus observées dès 2;4 et seules trois occurrences de production non conformes sont relevées à 2;6. Dans les données d'Antonin, l'acquisition formelle des déterminants peut être considérée comme très largement avancée dans la dernière séance à 3;2 mais encore environ une forme sur dix en position pré-nominale est un filler ou une omission (13,33% de formes de transition). Enfin, dans le corpus d'Adrien, le processus est plus long puisqu'à 3;5, les fillers et omissions restent produits devant 31,96% des noms.

Conformément à ce que nous pouvions envisager, ces constats peuvent être mis en perspective avec les indications fournies par l'observation de la longueur moyenne des énoncés. Chez les

quatre enfants, les énoncés sont en moyenne constitués de trois termes lors de la dernière séance. Celle-ci correspondant à des âges très différents d'un enfant à l'autre, nous pouvons d'une part signaler que ces quatre enfants acquièrent le langage à des rythmes très variables mais aussi et surtout, que le développement des formes pré-nominales peut être considéré comme corrélé à l'allongement des énoncés. Naël, qui d'après nos analyses est l'enfant qui acquiert les déterminants le plus rapidement, est aussi celui dont la LME est la plus haute à 2;6. Par ailleurs, le stade de développement des formes pré-nominales est semblable chez Lubin, Naël et Antonin lorsque leurs énoncés atteignent une longueur de trois termes en moyenne. Le processus est moins avancé chez Adrien mais on relève toutefois que ce dernier produit une large majorité de formes adultes lorsque la longueur moyenne de ses énoncés avoisine 3.

Ces quelques remarques nous permettent de confirmer que la maîtrise des déterminants va - de façon évidente - de pair avec l'allongement des énoncés<sup>2</sup> et le développement général du langage de l'enfant. La grammaticalisation du nom est à envisager comme incluse dans une complexification du langage à plusieurs niveaux, et nous pourrions en particulier remarquer dans la suite de ce travail que le développement des pronoms a une incidence sur la production des noms et des formes qui le précèdent.

Outre la gestion de la variabilité de rythme dans l'acquisition des déterminants, nos données confirment un phénomène déjà largement décrit dans la littérature et que nous retrouvons chez les quatre enfants de notre corpus, celui de la concomitance des formes de transition et des formes adultes. La position pré-nominale n'est pas périodiquement ou successivement marquée par la production d'omissions, puis de fillers, et en dernier lieu, de déterminants, mais bien occupée de façon simultanée par ces trois types de formes. Signalons néanmoins que dans la mesure où le processus de grammaticalisation est déjà entamé chez tous les enfants dès la première séance analysée, il nous est impossible de dire avec certitude si dans une période antécédente, les noms ne sont pas précédés exclusivement d'omissions ou si dès lors que les noms apparaissent, ils le sont accompagnés de fillers ou de déterminants (inclus pour la plupart dans des constructions figées).

---

<sup>2</sup> Comme attendu, une mesure de la corrélation entre la LME et la proportion des déterminants produits à chaque séance (effectuée grâce au test de Spearman) montre une corrélation forte et positive entre les deux variables (corpus de Lubin :  $r_s=1$  et  $p=4,9^{-5}$  ; corpus de Naël :  $r_s=0.92$  et  $p=1.1^{-3}$  ; corpus d'Antonin :  $r_s=0.98$  et  $p=4.0^{-4}$  ; corpus d'Adrien :  $r_s=0.90$  et  $p=4.5^{-3}$ ).



Nous avons pu observer par ailleurs que pour tous les enfants, le filler est une voie d'entrée dans la grammaticalisation du nom et dans le maniement des déterminants mais tous n'y ont pas recours dans les mêmes proportions d'une part et pendant les mêmes périodes d'autre part. Les fillers sont largement observés chez Antonin et Adrien : durant les quatre premières séances, Antonin produit autant de fillers que d'omissions, et dans le corpus d'Adrien, ceux-ci sont même la forme préférentielle dans deux séances sur huit (ils représentent en particulier 70,18% des formes observées devant les noms à 2;11). *A contrario*, les fillers sont bien moins fréquents dans les corpus de Lubin et Naël, et cela au profit des omissions durant une première période et des déterminants dans une seconde. A titre d'exemple, les omissions représentent en moyenne 70,92% des formes pré-nominales produites par Lubin dans les quatre premières séances, et sont donc considérablement plus fréquentes que les fillers qui ne représentent que 17,09% des productions pendant cette même période. Nous pouvons donc conclure de ces premières observations que la place à attribuer à la catégorie transitionnelle des fillers peut varier d'un enfant à l'autre. Nous pouvons faire l'hypothèse que ces différences de statut et de proportions pourront donner lieu à des usages variés des fillers (en contraste avec les omissions et les déterminants) sur le plan fonctionnel. Nous reviendrons sur cette hypothèse de façon plus détaillée lors de la présentation des différents types de fillers produits dans les quatre corpus au point 3.3.

### **3.2 Une mise en place progressive mais non linéaire des proto-formes et des formes adultes**

Conformément à ce que nous pouvions attendre, la proportion des formes adultes augmente au fil du temps jusqu'à représenter une très large majorité des productions en position pré-nominale à la fin de la période observée. Parallèlement, les dernières séances sont marquées par une diminution nette des omissions et des fillers. Bien que cette évolution puisse, comme nous l'avons vu, suivre un rythme différent, tous les enfants de notre corpus produisent dans la séance 8 davantage de formes adultes que de formes de transition. Le développement des déterminants peut donc être considéré chez chacun d'entre eux comme progressif et comparable à l'évolution canonique décrite dans la littérature.

Au-delà de ce constat somme toute très attendu, c'est la façon dont évoluent les trois formes qui a particulièrement retenu notre attention. Le développement des déterminants ne suit pas une

évolution régulière mais semble plutôt se dérouler par phases ou périodes. Dans les corpus de Lubin et d'Antonin, nous avons en effet pu repérer deux périodes distinctes, regroupant chacune quatre séances : de 1;9 à 2;2 puis de 2;5 à 2;9 chez Lubin, et de 1;9 à 2;4 puis de 2;7 à 3;2 chez Antonin. La première période est marquée par une distribution assez stable des omissions, des fillers et des déterminants (largement minoritaires), et la seconde, par une augmentation rapide des formes adultes et une stabilisation des productions. Le processus est en voie d'achèvement (sur le plan formel du moins) et la position pré-nominale est presque exclusivement occupée par des formes adultes. Seuls quelques rares fillers et omissions sont encore observables. Des constats comparables peuvent être faits au sujet du corpus de Naël mais les deux périodes sont réparties différemment. Nous pouvons considérer une première période de 1;6 à 1;11, comprenant cinq séances, et une seconde, de 2;2 à 2;6, constituée de trois séances. De même, deux périodes se dessinent également dans les données d'Adrien mais la seconde apparaît plus tardivement. Dans les six premières séances (période 1), les déterminants sont nettement minoritaires puis leur proportion augmente largement à 3;3 et ceux-ci deviennent la forme privilégiée devant les noms dans les deux dernières séances (période 2).

Nous avons par ailleurs pu remarquer que l'évolution des formes produites en position pré-nominale n'est pas strictement linéaire, autrement dit, que certains 'allers retours' ou 'régressions' peuvent être observés dans la distribution des omissions, fillers et déterminants. Chez Lubin, les fillers sont moins fréquents à 2;2 qu'à 2;0 et 2;5, et en parallèle, les omissions sont plus fréquentes à 2;2 qu'à 2;0. Dans le corpus d'Adrien, nous avons pu relever une plus grande proportion d'omissions (ainsi qu'une moindre proportion de déterminants) à 2;5 qu'à 2;3, et plus de déterminants à 2;7 qu'à 2;8 et 2;11 (séances auxquelles le filler est la forme préférentiellement utilisée par l'enfant).

Au travers du corpus de Naël et de la diminution des déterminants constatée de la première à la deuxième séance, nous avons pu voir que les régressions apparentes dans les premières séances pourraient, en ce qui concerne les déterminants, s'expliquer par la présence de figements associant une forme phonologiquement identique à un déterminant adulte et un lexème adulte récurrent. Nous y reviendrons à propos du corpus d'Adrien lorsque nous ferons, dans les pages qui suivent, l'analyse des types de déterminants produits par chaque enfant. Plus globalement, il peut également être envisageable que le développement des formes pré-nominales interagisse

avec d'autres niveaux du développement langagier et que certaines des progressions rapides ou au contraire des régressions observées dans nos données puisse s'expliquer par le fait que l'enfant fait en parallèle une multiplicité d'autres acquisitions.

### **3.3 Fillers et déterminants : des catégories (transitionnelle et adulte) en évolution ?**

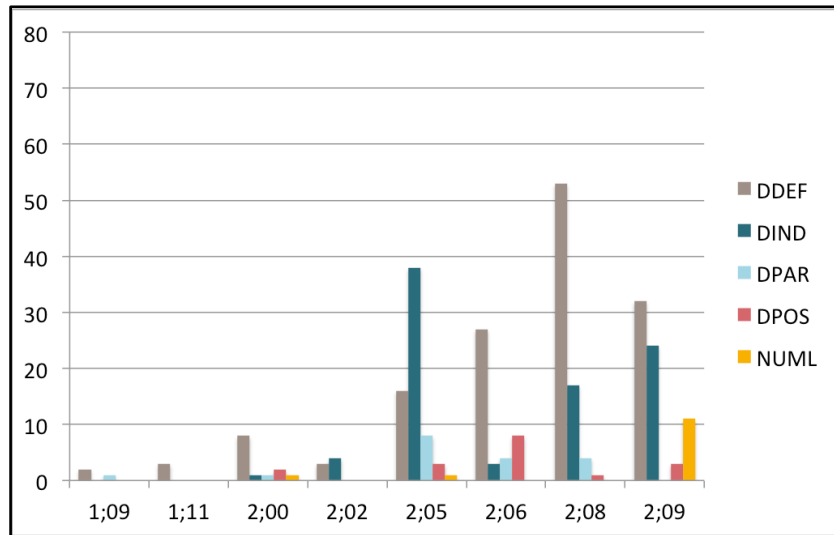
Afin d'observer et de comparer l'évolution éventuelle de la catégorie 'déterminant' chez chacun des enfants, nous avons - comme nous l'annoncions au point 1 de ce chapitre - catégorisé chaque forme adulte produite en fonction de sa nature. Au sein de ce point 3.3, nous nous interrogerons entre autres sur l'ordre d'apparition des déterminants dans nos données, ainsi que sur la fréquence de production des différents types de formes, et en particulier des définis et des indéfinis.

Les quatre graphiques ci-dessous (5 à 8) représentent la distribution des différents types de déterminants produits par chaque enfant et à chaque séance. La lecture des résultats aurait probablement été facilitée par un affichage en pourcentages. Toutefois, au vu du faible nombre d'occurrences de formes adultes dans les premières séances, nous avons ici préféré présenter les résultats en valeurs absolues.

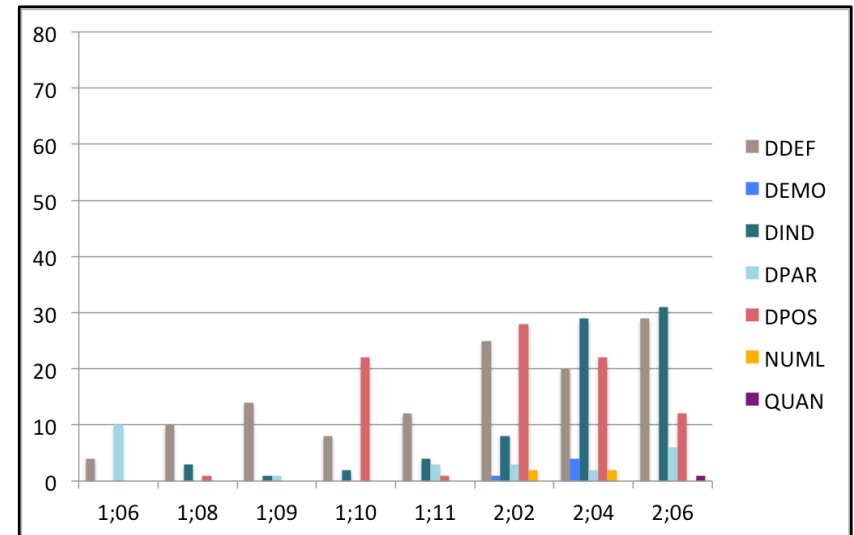
En premier lieu, nous pouvons constater que toutes séances confondues, le défini est dominant dans les quatre corpus, conformément à ce qui a pu être décrit dans la littérature. Le défini est largement dominant chez Antonin à toutes les séances (et, nous le verrons plus tard, en partie parce que les dénominations se font majoritairement avec le défini, et non l'indéfini). C'est également la forme la plus visible chez Lubin dans six séances sur huit mais l'indéfini est aussi largement utilisé à 2;5, 2;8 et 2;9. Dans les données de Naël, le défini est la forme la plus fréquente (38,01% du total des déterminants), mais les possessifs et les indéfinis sont également très représentés (respectivement 26,79% et 24,30% des déterminants), et surtout dans les deux dernières séances en ce qui concerne l'indéfini. Dans le corpus d'Adrien, lorsque les déterminants deviennent majoritaires - donc à 3;3 et 3;5 - le défini est aussi la forme privilégiée.

## Graphiques 5 à 8 - Distribution des types de déterminants dans les quatre corpus

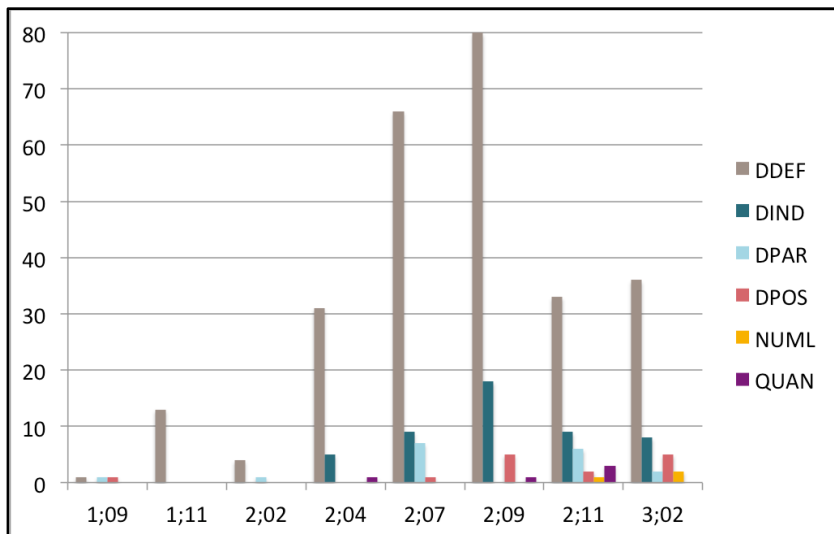
Graphique 5 - Corpus de Lubin



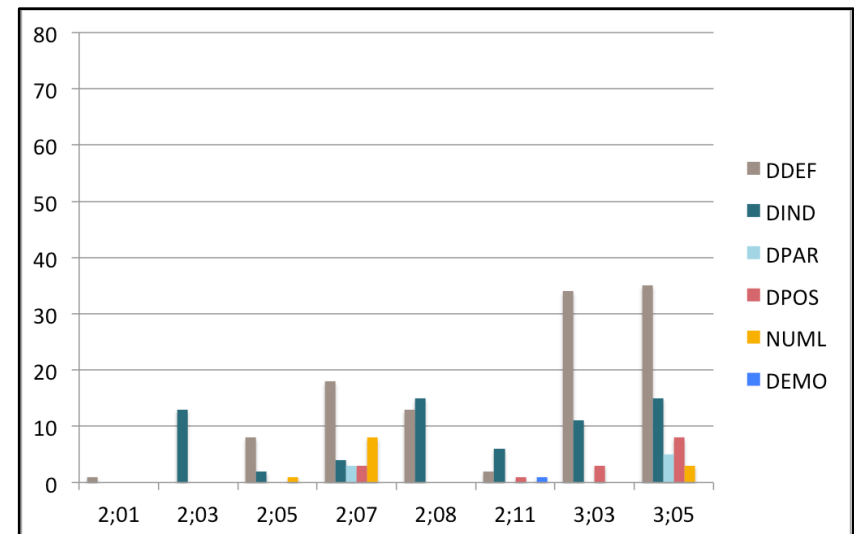
Graphique 6 - Corpus de Naël



Graphique 7 - Corpus d'Antonin



Graphique 8 - Corpus d'Adrien



Il peut sembler étonnant que l'indéfini soit la seule forme produite par Adrien à 2;3. Il nous paraît important de relever qu'il s'agit là de 13 occurrences de l'indéfini masculin singulier 'un', toutes produites associées au lexème 'autre'. Dès lors qu'aucun autre indéfini n'est produit dans cette séance (associé à un autre lexème), et que les fillers et omissions sont de loin les plus fréquents à cet âge (85,71% si on les traite comme un ensemble), il apparaît légitime de penser que l'enfant fait un usage figé du segment 'un autre', en concurrence avec seulement deux occurrences de type 'filler + autre' ([a not]).

Outre la présence forte de définis, on observe que les formes produites sont plus diversifiées et que la part des déterminants autres que le défini est plus importante chez Lubin, Naël et Antonin à partir de la période 2. Nous avons notamment pu voir que l'indéfini se met en place à partir de la séance 5 chez Lubin. Bien que le défini soit nettement plus produit dans toutes les séances d'Antonin, l'indéfini apparaît à partir de 2;4 comme la deuxième forme la plus observée. Dans le corpus de Naël, l'indéfini est plus utilisé que le défini à 2;4 et 2;6 et l'enfant a également largement recours au possessif à 1;10 ainsi que dans les trois dernières séances. Dans ce corpus, les séances ont été enregistrées en présence du frère, qui participe aux activités au même titre que Naël. Nous pouvons faire l'hypothèse que ce recours massif au possessif puisse avoir pour fonction de distinguer pour l'enfant ce qui lui appartient de ce qui appartient à son frère et ceci pourrait être confirmé par l'absence de possessifs à 1;11, séance à laquelle Tristan, le frère de Naël, ne prend pas part puisqu'il regarde un dessin animé. Quelques exemples de ces emplois de possessifs à 2;2 sont présentés et discutés ci-dessous.

### Exemples 3 - Cas de possessifs produits dans le corpus de Naël à 2;2

- (a) Observatrice : qu'est ce que t'as d'autre alors ?  
 Naël : [ʒɛ ʁuve mɔ̃ mɔ̃ ʃəvalje] 'j'ai (t)rouvé mon [//] mon chevalier.' (sort chevalier de la caisse et le montre à OBS)
- (b) Observatrice : alors euh ça va sur l(e) château fort ça ? (OBS fait référence à l'hélicoptère que Naël a dans les mains)  
 Naël : [nɑ̃] 'nan.'  
 Observatrice : non ?  
 Naël : [dɑ̃ mɔ̃ ʃato fɔʁ sa va pa] 'dans mon château fort ça va pas.'
- (c) Naël : [ma lɑ̃s l e pa ɔm a mɔ̃ ʃəvalje] 'ma lance F est pas yyy [//] à mon chevalier.'  
 Mère : nan ta lance on sait pas du tout où elle est passée.  
 Mère : on la retrouvera un jour peut être.
- (d) Naël : [X ʃəval də tɔ̃ ʃəvalje] 'xxx cheval de ton chevalier.' (sort cheval de la caisse et essaie de récupérer chevalier des mains de son frère)

Frère : ah nan [ʃ] nan.

Naël : [ʃa tɔ̃ ʃəval] ‘*tiens ton cheval.*’ (pose cheval entre les jambes de son frère)

Dans cette séance à 2;2, Naël montre à l’observatrice le château fort qu’il a reçu en cadeau à son anniversaire. Pour jouer avec ce château, l’enfant cherche dans une caisse de jouets des personnages (chevalier, roi, princesse) et des objets (lance, épée, ...). Naël et son frère partagent la même chambre, les mêmes caisses de jouets et possèdent chacun un château fort, des chevaliers et des accessoires comparables. On peut donc imaginer que dans l’exemple *a*, l’enfant relève au travers de l’emploi du possessif que parmi les chevaliers présents dans la caisse de jouets, il a réussi à trouver un chevalier qui lui appartient, et non un de ceux appartenant à son frère. Ce recours au possessif nous semble d’autant plus intéressant que la question de l’observatrice pourrait donner lieu à la production d’un indéfini, dans la mesure où en ce début de séance, l’objet est introduit pour la première fois dans l’activité et dans le discours, et que par ailleurs, l’observatrice n’a encore jamais vu ce personnage. Dans l’exemple *b*, l’enfant cherche à placer un hélicoptère sur la tour du château et de nouveau, contrairement à l’adulte qui emploie un défini dans sa question, l’enfant fait référence au château avec un possessif de première personne. Dans l’exemple *c*, on retrouve le syntagme ‘mon chevalier’. Cette fois l’enfant cherche la lance habituellement accrochée à ce chevalier. Il s’agit donc de la lance du personnage mais ici encore, c’est avec le possessif ‘mon’ qu’elle est introduite dans le discours. Si le possessif de première personne est fréquemment utilisé, on retrouve aussi celui de deuxième personne lorsque l’enfant s’adresse à son frère. Dans l’exemple *d*, Naël cherche un cheval sur lequel installer le chevalier de son frère et fait référence aux deux entités avec le possessif ‘ton’. Il aurait été intéressant de voir comment l’enfant introduit le référent ‘cheval’ dans la première mention (*cf.* premier énoncé de l’exemple *d*), mais cette partie de son énoncé étant inaudible, nous n’avons pas pu transcrire cette forme.

Ces observations concernant l’activité de jeu avec les chevaliers à 2;2 pourraient être transposées à d’autres séances du corpus, et en particulier à la séance suivante à 2;4 lorsque l’enfant et son frère jouent à la poupée. Si ces quelques exemples n’expliquent probablement pas entièrement pourquoi les possessifs sont si nombreux dans les données de Naël si on les compare aux données des trois autres enfants, elles nous permettent cependant de souligner l’importance d’une observation ‘contextualisée’ des formes pré-nominales. Les productions de l’enfant ne

peuvent semble-t-il pas être analysées dans l'absolu mais seulement en considérant aussi leurs contextes de production (participants impliqués, types d'activités, types de discours privilégiés dans ces activités, etc.), et cela même lorsque l'on s'intéresse exclusivement à leur distribution ou à leur dimension formelle.

Par ailleurs, bien que les formes se diversifient dans une seconde période, nos données, et en particulier celles de Lubin, Antonin et Adrien, montrent que le panel des types de déterminants produits dans les dernières séances n'est pas si étendu et que le défini continue à être plus utilisé que les autres formes. On ne trouve par exemple pas de démonstratif dans les corpus de Lubin et d'Antonin, et seulement une occurrence de ce type de forme dans celui d'Adrien, dans le cadre particulier d'une comptine répétée, et donc pas dans une production spontanée de l'enfant ([*la se mesə*] pour '*à ces messieurs*', produit par Adrien à 2;11). La présence majoritaire de formes adultes ne veut donc pas nécessairement signifier que la diversité du paradigme est acquise et que le processus d'acquisition des déterminants est achevé, et cela même si l'on se focalise exclusivement sur la forme, sans considérer l'acquisition de la dimension fonctionnelle des déterminants.

L'affichage des données brutes permet également de constater que les déterminants sont nettement plus nombreux chez Antonin que chez les autres enfants. Antonin n'étant pas plus avancé que Lubin et Naël dans le développement des formes, ces chiffres sont surtout liés à un usage plus important de noms dans ce corpus. Il nous semble d'ailleurs intéressant de relever que les noms sont plus nombreux à 2;4 (219 occurrences), 2;7 (164 occurrences) et 2;9 (129 occurrences) qu'à 2;11 (67 occurrences) et 3;2 (60 occurrences). N'ayant pas fait d'analyse du développement des pronoms, il nous est impossible de fournir de données chiffrées mais nos observations du corpus nous laissent penser que cette diminution du nombre de noms coïncide avec une augmentation de la part des pronoms, qui viendraient se substituer aux syntagmes nominaux (*cf.* exemples 4)<sup>1</sup>.

---

<sup>1</sup> Une analyse de la fréquence avec laquelle les lexèmes sont produits par Antonin (effectuée grâce au logiciel CLAN) montre par exemple que dans les séances à 2;11 et 3;2, le pronom démonstratif '*ça*' est produit respectivement 34 et 47 fois, contre seulement 13 et 16 fois dans les deux séances qui précèdent, à 2;7 et 2;9.

**Exemples 4 - Cas de reprises par un nom à 2;7 (a) et par un pronom à 3;2 (b) dans le corpus d'Antonin**

- (a) Antonin : [atũsɔ̃ a boku batyɕ] ‘attention a beaucoup voitures.’  
 [...]
   
Antonin : [avã̃s e batyɕ] ‘avance F voitures.’
   
Antonin : [avã̃s lɛ batyɕ] ‘avance les voitures.’
   
[...]
   
Antonin : [lɛ batyɕ osi ɔ̃ pø pa demae] ‘les voitures aussi on peut pas déma(rr)er.’
- (b) Antonin : [ɛ̃ ɛ̃ ɛ mal] ‘elle elle est mal [=! sourit].’
   
Mère : eh oui elle elle est mal oui [=! rires].
   
Mère : le quat(re) quatre.
   
Antonin : [i pø ale vit] ‘i(l) peut aller vite.’
   
Antonin : [i pø sɔ̃ktɪɕ də la ɕut] ‘i(l) peut sortir de la route.’

Cette diminution de la part des noms nous amène à considérer de nouveau le rôle joué par le contexte, et en particulier, par l'activité. Dans le corpus d'Antonin, les activités impliquant préférentiellement de la dénomination sont absentes des deux dernières séances - et au contraire, très présentes dans les séances précédentes -, ce qui pourrait partiellement expliquer que le recours aux noms soit moins fréquent pour verbaliser les référents. L'influence de l'activité pourrait donc se manifester dans la nature des expressions référentielles produites par l'enfant, mais aussi dans le type de déterminant produit. Par exemple, tous les numéraux produits dans le corpus de Lubin à 2;9 le sont dans le cadre d'un jeu en cours dans cette séance, impliquant de compter des éléments (dés, points ou boutons) et donnant lieu à des productions telles que celles présentées dans les exemples 3 ci-dessous.

**Exemples 5 - Déterminants numéraux produits dans le corpus de Lubin à 2;9**

- (a) [tɕwa de] ‘trois dés’
   
(b) [si pɔ̃ɛ̃] ‘six points’
   
(c) [tɕwa butɔ̃ puɕ mwa] ‘trois boutons pour moi’
   
(d) [mɔ̃ɛ̃ tɕwa butɔ̃] ‘moins trois boutons’
   
(e) [dø butɔ̃ puɕ lybɛ̃] ‘deux boutons pour Lubin’
   
(f) [dø butɔ̃ puɕ mamã̃] ‘deux boutons pour maman’
   
(g) [tus ẽ̃ butɔ̃] ‘tous un bouton’

Pour faire la synthèse des observations que nous avons pu faire sur les différents déterminants produits dans notre corpus, nous pouvons dire d'une part que le défini est globalement majoritaire chez Lubin, Naël et Antonin, et dominant également dans le corpus d'Adrien dans les deux séances pour lesquelles le déterminant est la forme préférentielle. Si les quatre premières



séance de chaque enfant sont compilées, on observe que le défini est la forme la plus fréquente, ce qui tend à montrer que l'entrée dans le déterminant se fait d'abord au travers du défini dans une première période, avant que le paradigme se diversifie et/ou que la proportion des autres types de déterminants augmente dans une seconde période. D'autre part, nous avons pu observer, dans les données de Naël et Adrien en particulier, que les premières formes étiquetées comme des déterminants correspondent plus probablement à des constituants de syntagmes figés qu'à de véritables éléments pré-nominaux autonomes. Enfin, dès lors que certains contextes d'enregistrement ou certaines activités semblent dans nos données entraîner une production plus importante de noms ou de types spécifiques de déterminants, il nous paraît important de soulever que l'influence de ces facteurs ne devra pas être négligé pour nous approcher d'une compréhension plus précise des formes produites par l'enfant.

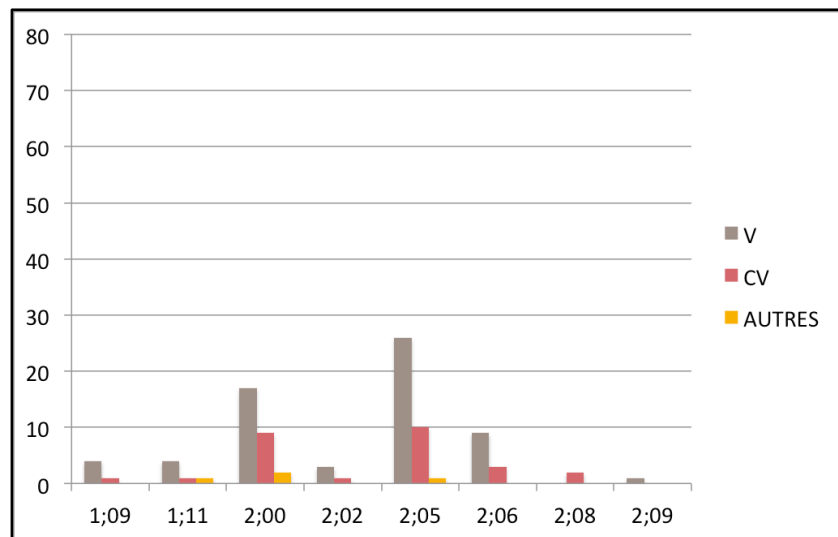
Nous avons par ailleurs catégorisé le type phonologique des fillers produits par chacun des enfants (consonne seule, voyelle seule, ou combinaison consonne-voyelle), afin de définir, si certains changements apparaissent dans les formes observées de la première à la dernière séance analysée. Nous avons en particulier cherché à examiner si la forme phonologique des fillers se diversifie au fil du développement de l'enfant, et par ailleurs, si un type de forme prédomine ou si au contraire, tous les types de fillers sont utilisés de façon comparable, pendant une étape ou sur toute la durée de la période étudiée.

Les graphiques ci-dessous (9 à 12) illustrent la distribution des différents types de fillers chez les quatre enfants et pour chaque séance. Comme pour les déterminants, les résultats sont affichés ici en valeurs absolues.

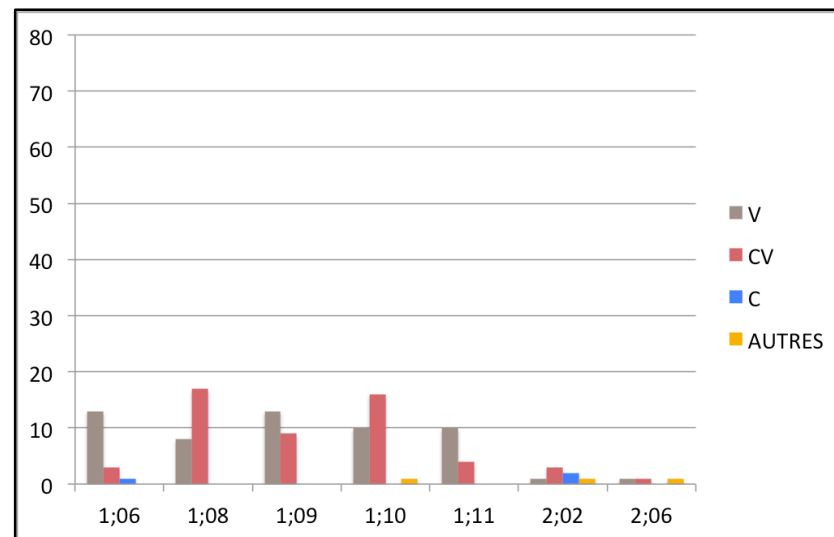
Les fillers de type 'voyelle seule' sont - toutes séances confondues - largement prédominants dans les corpus de Lubin et d'Antonin et apparaissent ensuite les fillers de type 'consonne-voyelle'. On observe également quelques occurrences de consonnes seules dans le corpus d'Antonin, et chez les deux enfants, quelques occurrences de fillers 'autres', dont la forme ne correspond donc ni à une voyelle ou consonne seules, ni à un schéma de type 'consonne-voyelle'.

## Graphiques 9 à 12 - Distribution des types de fillers dans les quatre corpus

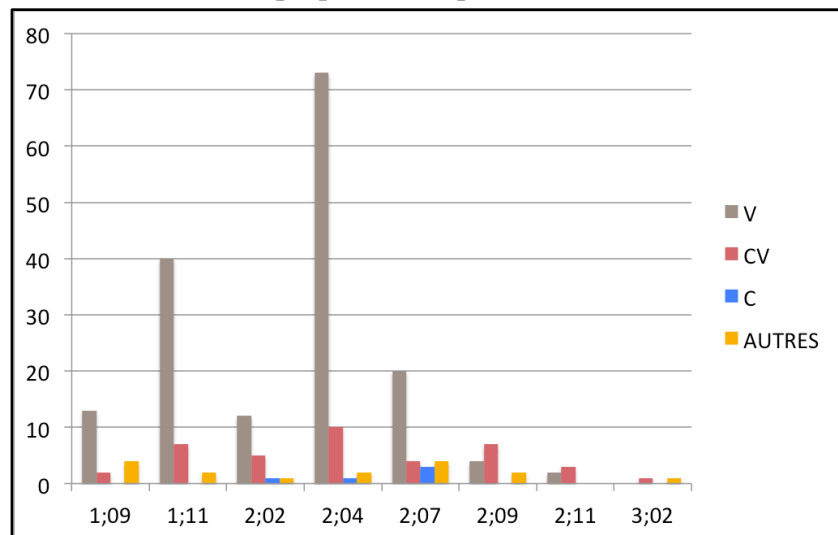
Graphique 9 - Corpus de Lubin



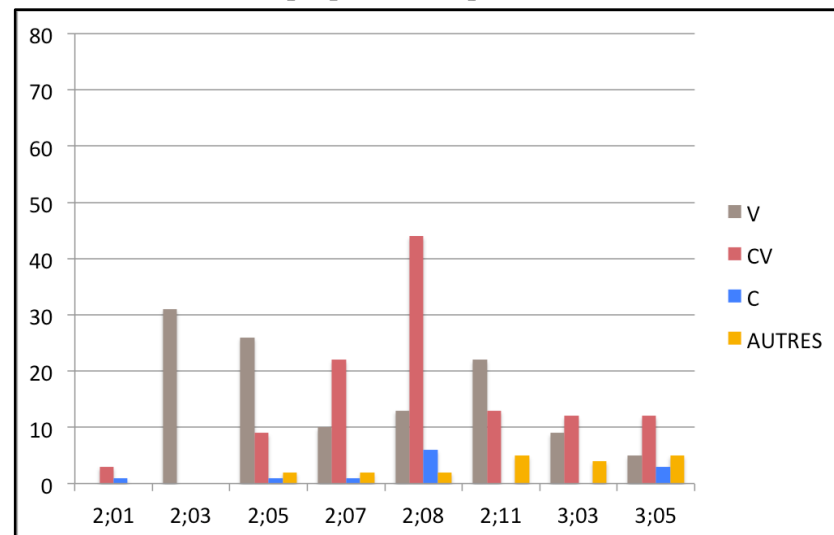
Graphique 10 - Corpus de Naël



Graphique 11 - Corpus d'Antonin



Graphique 12 - Corpus d'Adrien



Le tableau est quelque peu différent dans le corpus d'Adrien : les fillers vocaliques sont plus nombreux à 2;3 et 2;5, puis ce sont les fillers de type 'consonne-voyelle' qui dominent à 2;7, 2;8, 3;3 et 3;5. A 2;11, les fillers vocaliques sont majoritaires, ce qui peut paraître inattendu durant cette période où le filler préférentiel semble être le filler CV. L'observation détaillée des données nous montre que sur les 22 fillers vocaliques produits, 9 le sont associés au lexème 'jean' (toujours avec le filler [a]), et 7 avec le lexème 'autre' (devant lequel alternent les fillers [a] et [e]). Ce constat nous amène à faire deux hypothèses : d'une part, nous pouvons envisager que l'ensemble 'filler+jean' puisse être produit de façon figée, ou que ce lexème soit préférentiellement associé à une forme pré-nominale spécifique, ce qui tend à être confirmé par le fait que ce lexème n'apparaisse par ailleurs qu'une seule fois dans la séance avec une autre forme (que nous avons catégorisée comme un indéfini). D'autre part, il semble qu'à ce stade, certaines des productions de l'enfant se 'défigent' puisque le pronom 'autre' qui était antérieurement (à 2;3) exclusivement utilisé avec une forme phonologiquement identique à un indéfini l'est à 2;11 avec deux indéfinis 'un' et deux types de fillers. L'hypothèse du figement que nous avons formulée plus tôt à propos des indéfinis observés à 2;3 semble ainsi être confirmée.

Les résultats des analyses concernant le corpus de Naël illustrent un emploi des différents types de fillers plus proche de ce que nous observons chez Adrien que dans les deux autres corpus. Les premiers fillers produits sont vocaliques, et à partir de la deuxième séance, les fillers CV sont utilisés dans des proportions comparables aux fillers V, et cela dans les quatre séances qui suivent, jusqu'à ce que les fillers ne soient presque plus produits. A titre indicatif, les fillers vocaliques représentent en moyenne 49,89% des fillers produits de 1;8 à 1;11, et les fillers de type 'consonne-voyelle' représentent quant à eux 49,18% des fillers dans ces mêmes séances.

Notons de plus que la distribution des fillers reflète nécessairement le développement général des formes et peut être représentative de l'utilisation que les enfants font par ailleurs des déterminants. On observe notamment que les fillers sont très rares dans les dernières séances de Lubin, Naël et Antonin (et même absents à 2;6 dans les données de Naël) dans la mesure où les formes adultes sont en place à ce stade et au contraire, les fillers restent plus visibles dans le corpus d'Adrien, pour qui la mise en place des formes adultes est plus tardive. Ces graphiques traduisent également les stratégies que nous avons pu repérer plus tôt : les fillers sont plus

fréquents dans les corpus d'Antonin et d'Adrien que chez Lubin et Naël, et de fait, ces formes de transition sont nettement moins nombreuses chez ces derniers.

Ces graphiques et l'observation de la fréquence avec laquelle sont produits les différents types de fillers nous amènent à nous interroger également sur l'éventuel statut fonctionnel de ces formes chez chacun des enfants. Nous essaierons de voir entre autres au sein du chapitre VII si les fillers vocaliques, majoritaires dans les données de Lubin et d'Antonin, peuvent être produits dans des contextes pragmatiques distincts de ceux dans lesquels sont produits les premiers définis, et si par conséquent, ils pourraient être assimilés à des précurseurs d'indéfinis. Toutefois, dans la mesure où les fillers vocaliques sont plus nombreux que les définis, en particulier dans les séances de 1;9 à 2;4 du corpus d'Antonin, on peut imaginer qu'en parallèle, le filler vocalique puisse constituer une 'archi-*proto-forme*' de déterminant, utilisée dans des contextes où seraient attendus dans la langue adulte aussi bien des définis que des indéfinis. Adrien ayant quant à lui recours aux fillers sur une période plus longue, on peut se demander si du point de vue fonctionnel, il les utilise de la même façon que les autres enfants. Il serait par exemple envisageable que du point de vue fonctionnel, les fillers qu'il produit s'approchent davantage de formes 'dégradées' ou 'approximatives' de déterminants et qu'ainsi, les fillers vocaliques puissent être produits dans des contextes différents de ceux impliquant des fillers CV, ou bien encore en opposition avec les premiers définis adultes. De même, dans le corpus de Naël, les fillers CV étant aussi nombreux que les fillers vocaliques, on cherchera dans la suite de ce travail à savoir si l'enfant y a recours dans des contextes pragmatiques spécifiques.

La catégorie transitionnelle des fillers n'est donc pas uniforme. Nous avons vu que les premiers fillers sont d'abord vocaliques, et cela chez tous les enfants. Si ceux-ci prédominent dans les données d'Antonin et de Lubin, des fillers de type 'consonne-voyelle' sont également observables. Ils sont même utilisés à part égale des fillers vocaliques chez Naël et Adrien. Ces constats nous permettent d'avancer également que le recours aux fillers est différent d'un enfant à l'autre. Outre leur fréquence et la période pendant laquelle l'enfant les produit, il semble que les mêmes types de formes n'apparaissent pas dans les mêmes proportions chez tous les enfants. Enfin, ces résultats seront à mettre en perspective avec le développement des déterminants et de leur dimension fonctionnelle d'une part, et des fillers d'autre part. Peut-on imaginer que les fillers vocaliques constituent une 'archi-*proto-forme*' de déterminant dans les données de Lubin et

Antonin, ou bien que ceux-ci contrastent avec les premiers définis ? Peut-on faire l'hypothèse que les fillers CV soient utilisés dans des syntagmes dont les fonctions pragmatiques diffèrent de celles qui seraient associées aux syntagmes impliquant des fillers vocaliques dans les productions de Naël et Adrien ? Nous reviendrons sur ces premières interrogations au sein du chapitre VII.



L'analyse de la distribution des formes devant les noms dans nos données nous a permis de confirmer entre autres que l'acquisition des déterminants est un phénomène que l'on peut considérer comme globalement progressif - dans la mesure où la part des déterminants augmente et que celle des formes de transition diminue - mais dont l'évolution peut ne pas être strictement régulière. Nous avons également pu observer, conformément aux constats de la littérature, que tous les enfants n'entrent pas dans le processus de grammaticalisation au même rythme, et le font parfois avec des stratégies quelque peu différentes. En examinant de plus près les types de déterminants et de fillers présents dans notre corpus, nous avons aussi pu mettre en évidence la primauté du défini et la diversification graduelle des déterminants lorsque ceux-ci deviennent la forme privilégiée devant les noms, ainsi que le caractère hétérogène et évolutif des fillers. Ce constat donnera lieu dans la suite de ce travail à l'exploration de potentiels contrastes fonctionnels entre les différents types de fillers, de même qu'entre fillers et déterminants. Enfin, nous n'avons pu que corroborer ce que de nombreuses études citées au chapitre II ont signalé avant nous : l'enfant produit en alternance, au sein des mêmes séances, aussi bien des formes adultes que des fillers et des omissions. Il existe donc une variation importante en position pré-nominale et c'est ce phénomène que nous chercherons en particulier à cerner au chapitre suivant.



## CHAPITRE VI

*Influence des facteurs distributionnels sur la  
mise en place des déterminants*

---





Nous avons fait le constat au chapitre précédent d'une production concomitante de déterminants, fillers et omissions au sein de chacune des sessions analysées. Plusieurs hypothèses distributionnelles ont été proposées et testées pour expliquer cette variation pré-nominale, et de façon plus générale, pour expliquer la mise en place progressive des formes adultes (*cf.* chapitre 2). On trouve notamment dans la littérature des travaux sur la dimension phono-prosodique du développement des fillers et des déterminants (Veneziano et Sinclair, 1997 ; Kilani-Schoch et Dressler, 2000 ; Demuth, 2001), ainsi que des études sur l'influence du statut sémantique du nom (Bassano *et al.*, 2008) ou du contexte lexico-syntaxique (Pine et Lieven, 1997) sur la production des formes pré-nominales.

Dans une approche phono-prosodique, Demuth considère entre autres que le déterminant ne peut pas se développer hors du pied binaire, ce qui expliquerait d'après elle la présence simultanée d'omissions et de déterminants. En se focalisant spécifiquement sur les fillers, Veneziano et Sinclair (1997, 2000) ou Kilani-Schoch et Dressler (2000) ont également mis en évidence la sensibilité de l'enfant au contexte phono-prosodique, en particulier dans la première période dite 'pré-morphologique' et dans une moindre mesure au sein de la seconde période, dite 'proto-morphologique'. Ces conclusions sont partagées par Bassano qui envisage quant à elle une mise en place des déterminants en deux temps, d'abord soumise à des contraintes prosodiques, puis à des facteurs sémantiques. La distribution des formes pré-nominales serait, dans cette perspective, prédéterminée par le nombre de syllabes du nom, et plus tard, par son statut sémantique (caractère animé/inanimé, concret/abstrait du nom). Enfin, dans une approche lexico-syntaxique, le travail de Pine et Lieven a montré qu'en anglais, les déterminants sont produits d'abord dans le cadre de structures figées et associés de façon privilégiée à des noms ou des structures syntaxiques spécifiques.

En nous appuyant sur les méthodes adoptées dans le cadre de ces travaux, nous chercherons à voir dans nos données si i) la structure syllabique du nom et du syntagme nominal peuvent expliquer la présence vs. absence de forme devant le nom ii) le statut sémantique du nom conditionne le type de forme pré-nominale produite iii) le contexte lexical et syntaxique influence le type de déterminant produit, et plus largement, le type de forme pré-nominale.

Ce chapitre est organisé en trois parties. Dans une première section, nous nous intéresserons à l'influence de la structure syllabique, dans une seconde, à l'influence du statut sémantique du nom associé à la forme pré-nominale et dans une dernière partie (section 3), nous nous pencherons sur l'impact du contexte lexical et syntaxique dans lequel le syntagme nominal est produit.

## 1. Influence de la structure syllabique

Nous l'annonçons en introduction de ce chapitre, les travaux qui se sont penchés sur la dimension prosodique du développement des fillers et des déterminants (*cf.* par exemple Veneziano et Sinclair, 2000 ; Veneziano, 2014b ; Demuth, 2001, Demuth et Tremblay, 2008) ont proposé que la production des formes (qu'il s'agisse de déterminants ou de fillers) puisse être conditionnée par une préférence pour un pied binaire. Demuth fait notamment l'hypothèse que la hiérarchie prosodique (telle qu'elle a été développée par Nespor et Vogel (1986) ou Selkirk (1984)) ne serait acquise par l'enfant que progressivement et que les constituants prosodiques ne seraient pas tous accessibles dans un premier temps du développement. Les déterminants ou les fillers seraient ainsi prosodifiés directement au niveau du pied (binaire) puis seulement ultérieurement prosodifiés au niveau du syntagme phonologique.

Dans la lignée de ces travaux que nous venons de citer, nous chercherons dans ce premier point du chapitre à observer si la variation en position pré-nominale peut s'expliquer par une préférence de l'enfant pour une structure bisyllabique. Si l'on s'en tient par exemple aux résultats de Demuth et Tremblay (*ibid.*) sur le français, nous pouvons faire deux hypothèses : i) les fillers et déterminants seront avant tout produits devant des noms d'une ou trois syllabe(s) ii) les noms de deux et quatre syllabes généreront quant à eux une proportion plus importante d'omissions.

Cette première section sera organisée comme suit : en premier lieu, nous ferons un point sur la méthode à laquelle nous avons eu recours pour décompter le nombre de syllabes des noms et des

syntagmes phonologiques. Au point 1.2 sera traitée la question du nombre de syllabes contenues dans les syntagmes phonologiques et aux points 1.3 et 1.4, nous discuterons respectivement de la distribution des formes pré-nominales devant les noms mono- et bisyllabiques et de la distribution des formes devant les noms de plus de deux syllabes. Un bilan de nos observations - axé sur la question de la préférence pour un pied binaire - sera proposé au point 1.5.

## **1.1 Décompte du nombre de syllabes des noms et des syntagmes phonologiques**

Afin de vérifier les hypothèses que nous venons de présenter et ainsi comparer nos données à celles des travaux existants, nous avons d'une part décompté le nombre de syllabes du nom et d'autre part, le nombre de syllabes du syntagme phonologique. Nos lectures nous avaient dans un premier temps amenée à évaluer également l'impact de la structure rythmique (et en particulier, iambique) sur la production des formes pré-nominales mais cette question ne sera pas développée dans cette thèse. Pour une plus grande fiabilité de nos analyses, ce travail sera mené dans le cadre d'une collaboration avec une phonologue, et les analyses développées ici ne porteront donc que la structure syllabique.

A l'inverse de Demuth, nous n'avons exclu ni les mots à initiale vocalique, ni les mots tronqués. Nous nous sommes appuyée pour ce codage sur les productions effectives de l'enfant. Dans le cas de mots tronqués (exemple : [la b<sup>w</sup>u:t] pour 'la bouteille', produit par Adrien à 2;8), nous avons donc décompté le nombre de syllabes produites et non le nombre de syllabes qui seraient produites dans la langue adulte.

Par ailleurs, étant donné que le syntagme phonologique peut inclure d'autres éléments que les fillers ou déterminants et les noms (exemples : adjectifs, prépositions), nous avons donc - et contrairement à Demuth - décompté le nombre de syllabes des syntagmes phonologiques en prenant en compte ces éléments additionnels. Bien que la hiérarchie prosodique soit construite sur d'autres critères, on notera que la plupart du temps, celle-ci coïncide avec les frontières des groupes syntaxiques.

Enfin, signalons que les noms devant lesquels un déterminant était attendu mais pour lesquels la forme précédente n'a pas pu être interprétée n'ont pas été retenus pour l'analyse (*cf.* exemple *a* ci-dessous). Ont également été exclus les cas de syntagmes devant lesquels ont été produits des

formes au statut indécidable (cf. [e ʁo] dans le dernier énoncé de l'exemple b), et les cas de syntagmes répétés suite à un tâtonnement de l'enfant (ex. c).

### Exemples 1 - Cas indécidables non retenus pour l'analyse phono-prosodique

- (a) Lubin (2;5) : [X məsʝə ati fɛ] 'xxx monsieur yyy fait ?'  
(b) Naël (1;8) : [sa] 'ça ?'  
Mère : ça c'est un drôle d'animal qui s'appelle une haie.  
Naël : [e ʁo ɛ] 'yyy yyy haie ?'  
(c) Antonin (2;9) : [kəmə la wat la ɔt vwatyʁ] 'comme la voit(ure) [/] la aut(re) voiture.'

## 1.2 Longueur syllabique des syntagmes phonologiques

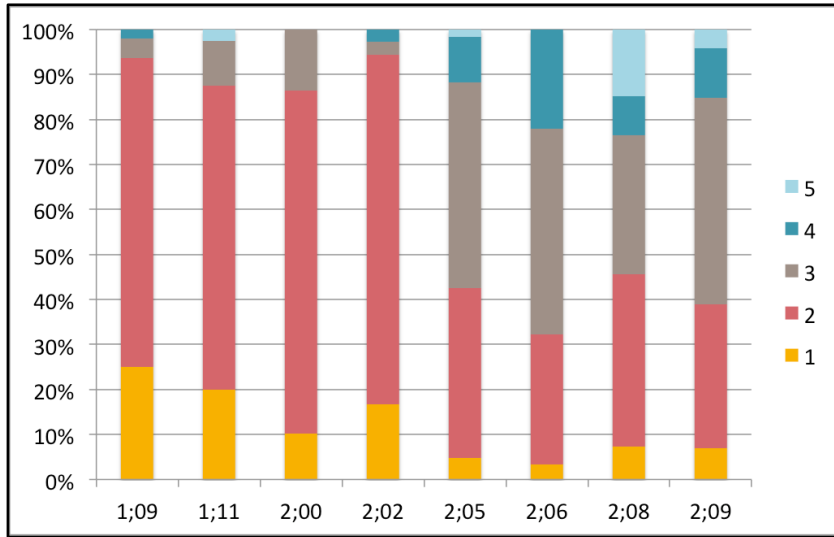
Avant de nous concentrer spécifiquement sur l'emploi des formes pré-nominales devant les noms, nous passerons brièvement en revue la question de la longueur - autrement dit, du nombre de syllabes - des syntagmes phonologiques produits par les quatre enfants. Seuls les syntagmes phonologiques au sein desquels des déterminants peuvent être attendus ont été retenus pour cette analyse.

Les quatre graphiques ci-dessous (1 à 4) représentent pour chacun des enfants le nombre de syllabes contenues dans les syntagmes phonologiques, à chaque séance analysée. Rappelons que dans la perspective phono-prosodique, on devrait s'attendre à observer une majorité de syntagmes phonologiques de deux syllabes, ou multiples de deux syllabes.

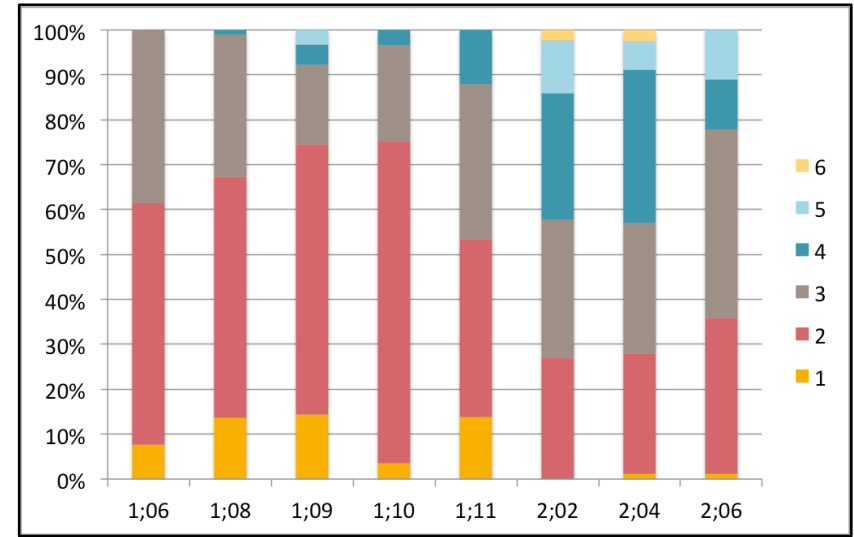
Ces graphiques nous montrent en premier lieu que le nombre de syllabes contenues dans les syntagmes phonologiques évolue de façon comparable chez Lubin et Adrien d'une part, et chez Naël et Antonin d'autre part. Dans le cas de Lubin et Adrien, les syntagmes phonologiques sont d'abord presque exclusivement constitués d'une et deux syllabe(s), puis la part des syntagmes d'une syllabe diminue et celle des syntagmes phonologiques de plus de deux syllabes augmente.

**Graphiques 1 à 4 - Nombre de syllabes contenues dans les syntagmes phonologiques (chez les quatre enfants)**

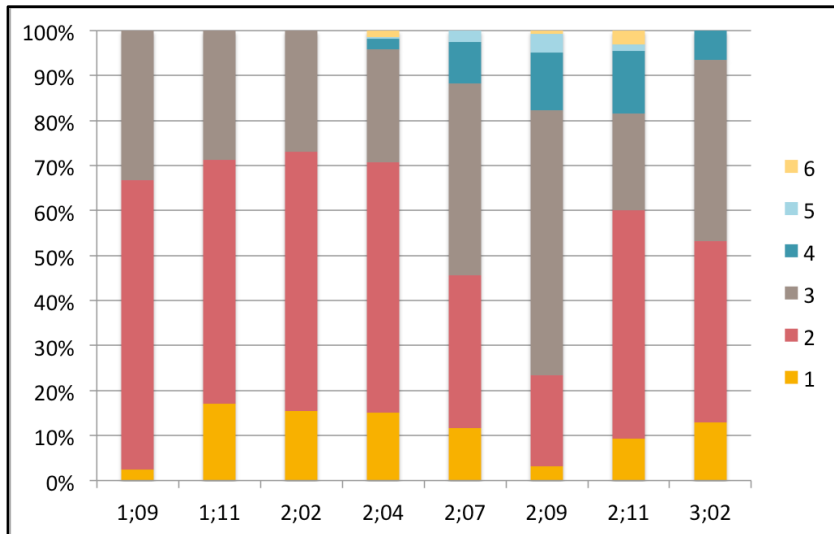
*Graphique 1 - Corpus de Lubin*



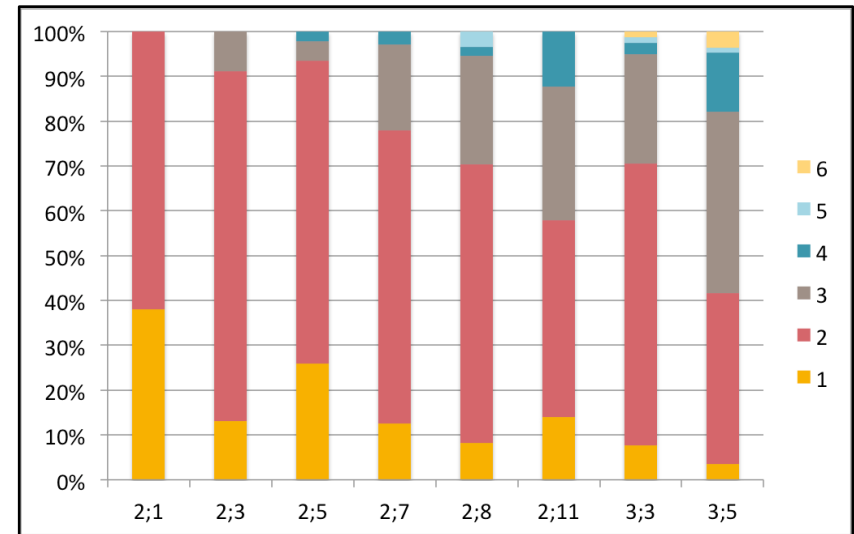
*Graphique 2 - Corpus de Naël*



*Graphique 3 - Corpus d'Antonin*



*Graphique 4 - Corpus d'Adrien*



De façon plus détaillée, les syntagmes de deux syllabes représentent en moyenne 75,54% des productions dans les quatre premières séances du corpus de Lubin, et 34,21% des syntagmes dans les quatre dernières séances. Cette diminution nette des syntagmes bisyllabiques coïncide donc avec les deux périodes préalablement relevées au chapitre suivant lors de l'analyse de la distribution des formes pré-nominales. Parallèlement, la part des syntagmes phonologiques de trois syllabes augmente à 2;5 et passent de 7,65% des cas analysés dans la période 1 à 42,03% durant la seconde période. Les syntagmes phonologiques de quatre et cinq syllabes restent peu fréquents, et cela même dans la période 2 puisqu'ils ne représentent en moyenne que 18,14% des productions de 2;5 à 2;9.

Comme chez Lubin, la proportion des syntagmes phonologiques d'une syllabe diminue dans le corpus d'Adrien mais la diminution des syntagmes de deux syllabes et l'augmentation de ceux de trois syllabes sont moins nettes. Dans la dernière séance, à 3;5, les syntagmes bisyllabiques représentent 38,10% des productions et les trisyllabiques, 40,48% (contre 31,94% et 45,83% dans le corpus de Lubin).

Dans les données de Naël et Antonin, les syntagmes de deux syllabes sont aussi plus fréquents dans la première période (couvrant, rappelons le, les quatre premières séances dans le corpus d'Antonin et les cinq premières dans celui de Naël) que dans la seconde. Ceux-ci équivalent en moyenne à 55,79% du total des syntagmes en P1 chez Naël contre 29,40% en P2, et à 57,94% dans la première période chez Antonin contre 36,30% dans la seconde. Les syntagmes trisyllabiques sont plus fréquents chez ces deux enfants durant la première période que ce que nous avons pu observer dans les données de Lubin et Adrien, chez qui ils sont presque inexistantes dans les premières séances. Notons également que les syntagmes de quatre syllabes sont plus présents dans le corpus de Naël que dans tous les autres corpus (31,21% en moyenne dans les séances 6 et 7). On relève par ailleurs que la part des syntagmes d'une syllabe fluctue chez Antonin plus qu'elle ne diminue. Ce constat est également valable lorsque l'on observe le graphique de Naël mais seulement en ce qui concerne les cinq premières séances puisque ces syntagmes disparaissent ensuite dans la période 2.

Au vu de ces graphiques et de la préférence globale (mais non exclusive) pour les syntagmes de deux syllabes, nous pouvons d'une part faire l'hypothèse que cette structure bisyllabique influencera les productions pré-nominales des enfants. On peut en effet s'attendre à ce que les

productions des quatre enfants soient le plus souvent (en ce qui concerne les noms d'une et deux syllabe(s)) des combinaisons de type 'omission + nom bisyllabique' ou 'filler/déterminant + nom monosyllabique'. D'autre part, la proportion des bisyllabiques diminuant chez tous les enfants entre la période 1 et 2, nous pouvons imaginer que cette influence sera moins marquée dans la seconde période que dans la première. Toutefois, la diminution de la part de ces syntagmes de deux syllabes étant moins prononcée chez Adrien (à l'exception de la dernière séance), on peut penser que les combinaisons entre forme pré-nominale et nom que nous observerons chez ce dernier respecteront davantage la structure bisyllabique que celles des autres enfants à la fin de nos observations. Enfin, signalons également que les syntagmes de quatre syllabes atteignant environ 30% des syntagmes produits à 2;2 et 2;4 dans le corpus de Naël, une diminution de la part des bisyllabiques ne signifiera pas pour autant une diminution globale du schéma syllabique binaire.

### 1.3 Distribution des formes devant les noms mono- et bisyllabiques

Dans une étape ultérieure, nous avons cherché à observer de façon croisée la longueur syllabique des noms (et non plus des syntagmes phonologiques) devant être précédés d'un déterminant et le type de forme pré-nominale produite. Dans la mesure où notre intérêt dans cette partie du travail est porté sur la structure phono-prosodique, nous n'avons pas distingué les fillers des déterminants et n'avons considéré que deux catégories distinctes : la présence *vs.* l'absence de forme.

Afin de rendre plus claire la lecture des graphiques, nous avons séparé la présentation des résultats concernant les noms d'une et deux syllabe(s) de celle des noms de trois et quatre syllabes. Nous aborderons pour commencer la distribution des formes devant les noms mono- et bisyllabiques.

Sur les quatre graphiques ci-dessous (5 à 8) sont présentées pour chaque enfant et chaque séance la part des productions de type 'déterminant/filler + nom bisyllabique', 'omission + nom monosyllabique', 'déterminant/filler + nom monosyllabique', et enfin, la part des productions de type 'omission + nom bisyllabique'. Les deux dernières combinaisons, présentées dans la partie inférieure des graphiques répondent à l'hypothèse phono-prosodique, et les deux premières, présentées quant à elles dans la partie supérieure des graphiques, ne sont au contraire pas

attendues puisqu'elles génèrent la production d'un pied non binaire. Notons que tous les pourcentages sont calculés sur le nombre total de noms d'une et deux syllabe(s) produits à chaque séance.

L'observation de ces graphiques nous amène à faire trois constats majeurs. Dans un premier temps, on remarque que comme pour la distribution des syntagmes phonologiques, la distribution des formes devant les noms mono- et bisyllabiques est semblable chez Lubin et Adrien d'une part et chez Naël et Antonin d'autre part. Les graphiques 5 (Lubin) et 8 (Adrien) donnent à voir une diminution de la part des noms d'une et deux syllabe(s) précédés d'une omission et une augmentation de la proportion des noms d'une et deux syllabe(s) précédés d'un filler ou d'un déterminant. Le patron représenté dans les graphiques 6 (Naël) et 7 (Antonin) est comparable à celui que nous venons de décrire mais on observe toutefois davantage de noms précédés de fillers ou de déterminants dans la première période.

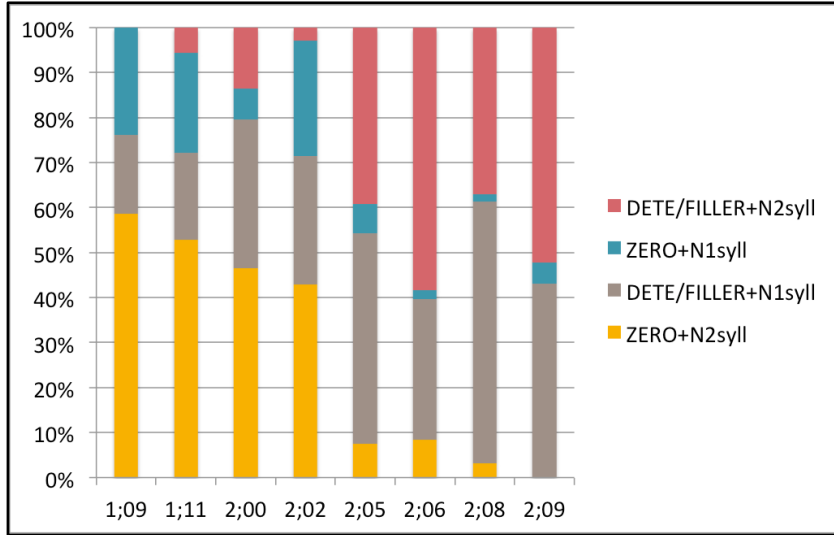
Dans un second temps, nous pouvons relever que le développement global des formes pré-nominales se reflète dans la distribution des omissions et fillers ou déterminants devant les noms mono- et bisyllabiques, puisque la part des omissions diminue et celle des fillers et des déterminants augmente au fil des séances.

Enfin, nous pouvons conclure de la lecture de ces graphiques qu'une grande partie des productions des enfants concordent avec l'hypothèse phono-prosodique telle qu'elle a été développée entre autres par Demuth. En considérant à la fois les cas d'omissions devant les noms de deux syllabes et les fillers/déterminants devant les noms d'une syllabe, on atteint 74,82% et 73,38% des syntagmes impliquant des noms mono- et bisyllabiques dans les corpus de Lubin et Adrien dans la première période, puis 49,84% et 62,54% dans la seconde. Ces productions attendues - si l'on s'en tient à l'hypothèse d'une préférence pour le pied binaire - représentent 61,56% des productions de Naël et 59,17% dans celles d'Antonin dans la première période, puis 50,18% et 52,73% dans la période 2.

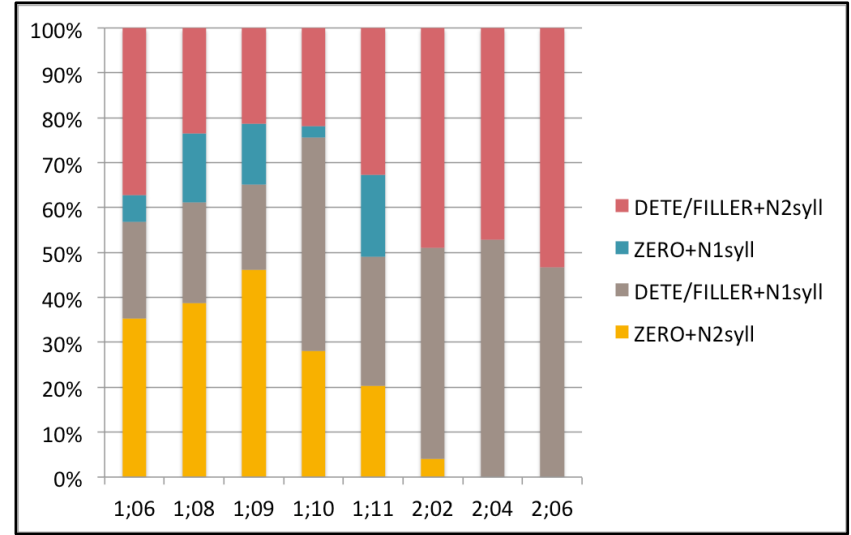


**Graphiques 5 à 8 - Distribution des formes pré-nominales devant les noms mono- et bisyllabiques (chez les quatre enfants)**

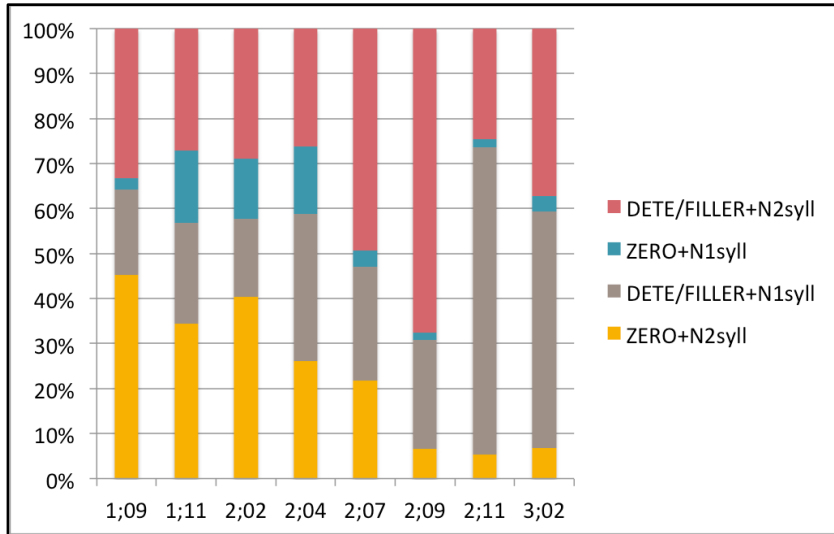
*Graphique 5 - Corpus de Lubin*



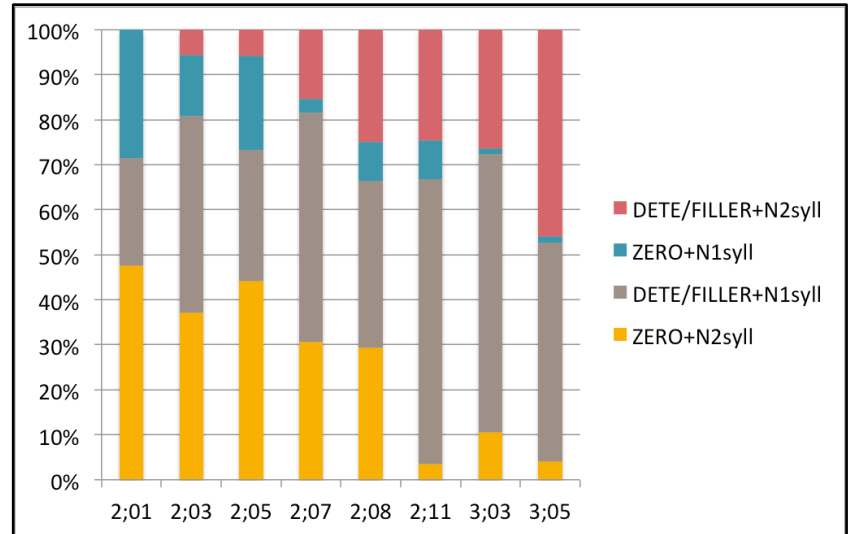
*Graphique 6 - Corpus de Naël*



*Graphique 7 - Corpus d'Antonin*



*Graphique 8 - Corpus d'Adrien*



Si ces cas sont majoritaires, la part représentée par les cas non attendus est cependant loin d'être minime et augmente avec l'âge de l'enfant. La présence de ces cas qui ne s'expliquent pas du point de vue de l'hypothèse phono-prosodique - dans la mesure où le pied n'est pas binaire - pourrait s'expliquer par l'influence d'autres facteurs et leur augmentation pourrait signifier, comme cela a déjà pu être signalé entre autres dans les travaux de Veneziano sur les fillers, que la sensibilité à la structure syllabique serait plus marquée dans une première période que lorsque le processus d'acquisition des déterminants est déjà quasiment achevé (comme c'est le cas en période 2 dans le corpus de Naël). Nous reviendrons sur ces questions au chapitre VIII, lorsque nous confronterons l'influence des différents facteurs étudiés dans cette thèse, et notamment, le facteur phono-prosodique.

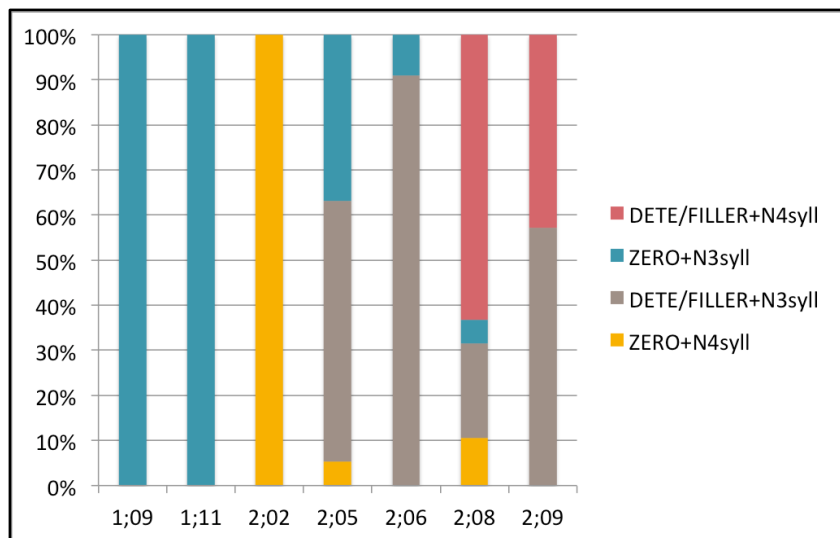
#### **1.4 Distribution des formes devant les noms de plus de deux syllabes**

Les noms n'étant que rarement, voire jamais, composés de plus de quatre syllabes, nous avons limité notre analyse des noms multisyllabiques aux noms de trois et quatre syllabes. De la même façon que pour les noms mono- et bisyllabiques, nous avons représenté ici la présence *vs.* l'absence de forme dans quatre contextes (lorsque ceux-ci étaient présents dans les données) : 'déterminant/filler + nom de quatre syllabes', 'omission + nom de trois syllabes', 'déterminant/filler + nom de trois syllabes' et 'omission + nom de quatre syllabes'. Dans les deux derniers cas de figure, les productions de l'enfant sont en accord avec l'hypothèse syllabique de Demuth et dans les deux premiers, elles ne le sont pas. Comme au point 1.3, les pourcentages sur la base desquels les graphiques sont élaborés correspondent à la part de chaque combinaison sur le total des noms de trois et quatre syllabes dans chaque corpus.

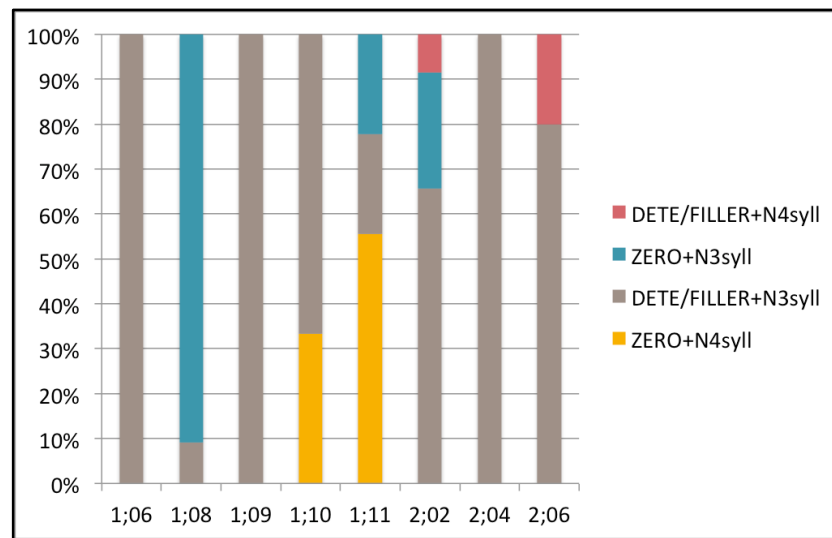
Avant de commenter les quatre graphiques ci-dessous, signalons que toutes les séances n'y sont pas représentées. Nous n'avons en effet pas observé de noms de plus de deux syllabes dans le corpus de Lubin à 2;0, dans le corpus d'Antonin à 1;9 et 2;2, ainsi qu'à 2;1 et 2;11 dans celui d'Adrien. Dans ce dernier corpus, nous n'avons par ailleurs pas non plus observé de noms de plus de trois syllabes devant lesquels un déterminant était attendu. Plus globalement, le nombre de noms de trois et quatre syllabes étant nettement inférieur au nombre de noms d'une et deux syllabe(s), nous resterons prudente dans les constats et conclusions que nous dresserons.

**Graphiques 9 à 12 - Distribution des formes pré-nominales devant les noms de trois et quatre syllabes (chez les quatre enfants)**

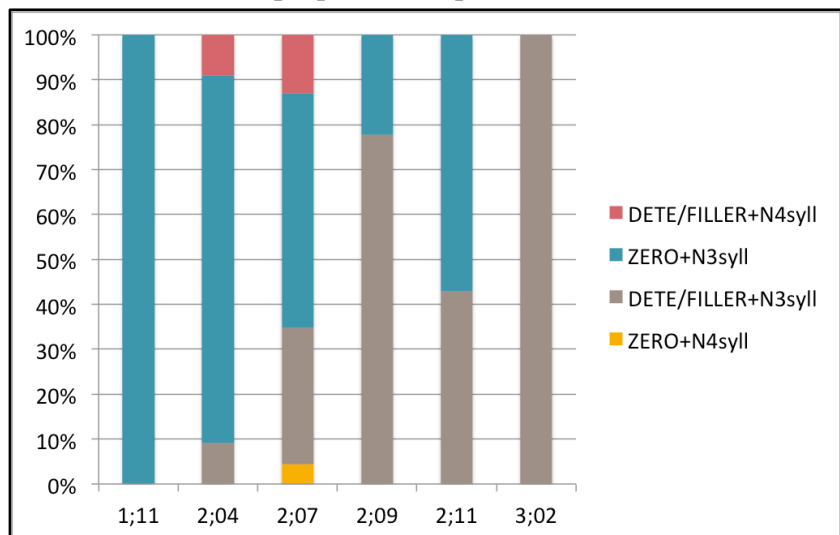
*Graphique 9 - Corpus de Lubin*



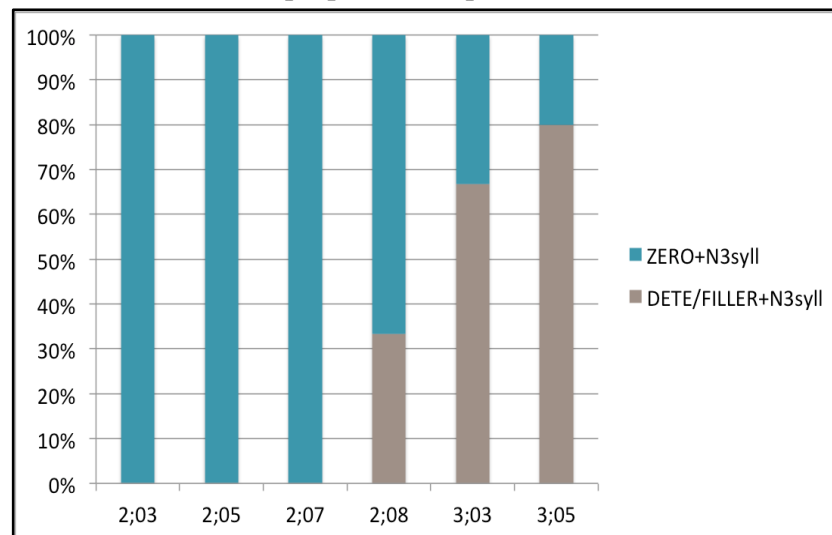
*Graphique 10 - Corpus de Naël*



*Graphique 11 - Corpus d'Antonin*



*Graphique 12 - Corpus d'Adrien*



Dans le corpus d'Adrien, le seul constat qu'il nous est possible de faire est celui de la diminution des noms de trois syllabes précédés d'une omission à 2;8, 3;3 et 3;5, ainsi que de l'apparition et de l'augmentation des fillers/déterminants devant les noms de trois syllabes. Nous pouvons par conséquent noter que pour les quatre premières séances observées dans cette section, l'intégralité - de 2;3 à 2;7 - ou la majorité - à 2;8 - des productions de l'enfant ne vont pas dans le sens de l'hypothèse phono-prosodique, et que cette tendance s'inverse dans les deux dernières séances.

Dans les données de Lubin, les fillers ou les déterminants apparaissent devant les noms de trois syllabes à 2;5, puis à 2;8 devant ceux de quatre syllabes. Ceci implique que les productions de l'enfant s'alignent avec l'hypothèse d'une préférence pour la structure syllabique binaire avec les noms de trois syllabes et *a contrario*, les productions impliquant des noms de quatre syllabes ne respectent plus cette structure dans les deux dernières séances. On peut donc s'interroger ici, en ce qui concerne les noms de trois et quatre syllabes, sur le rôle joué par le facteur phono-prosodique, mais aussi sur la part à attribuer au développement global des formes et/ou à d'autres facteurs.

L'apparition des noms de trois ou quatre syllabes précédés de fillers ou de déterminants est plus précoce dans le corpus d'Antonin que dans celui de Lubin mais les constats que nous venons de faire au sujet de l'hypothèse syllabique sont comparables. Les noms de trois syllabes précédés d'une omission sont majoritaires jusqu'à 2;7, puis de nouveau à 2;11, donc si l'on compile toutes les séances, les productions qui ne sont pas conformes à un schéma syllabique binaire sont majoritaires mais leur proportion diminue dans la période 2, au profit des combinaisons 'déterminant/filler + nom de trois syllabes'.

Enfin, dans le corpus de Naël, dans lequel on observe le plus de noms de plus de deux syllabes, il nous paraît surtout intéressant de relever que la combinaison impliquant un filler ou un déterminant et un nom de trois syllabes est - toutes séances confondues - largement majoritaire. Par ailleurs, si ces résultats sont mis en perspective avec ceux obtenus plus tôt pour les noms mono- et bisyllabiques, on remarque que la diminution des cas attendus observée dans la période 2 n'est pas valable pour les noms de plus de deux syllabes, ce qui pourrait partiellement remettre en question notre constat d'un impact moindre de la structure syllabique à ce stade de développement.

## 1.5 Une préférence pour le pied binaire ?

A des fins de synthèse, nous avons rassemblé dans le tableau ci-dessous toutes les productions attendues et non attendues - dans l'hypothèse d'une préférence pour la structure bisyllabique - chez chacun des enfants, et pour chacune des deux périodes préalablement observées. Les pourcentages que nous avons faits figurer dans ce tableau regroupent donc pour tous les noms analysés (de une à quatre syllabes), la part que représentent à chaque session les combinaisons dites 'préférentielles' (ou attendues) et la part représentée par les combinaisons dites 'non préférentielles' (ou non attendues)<sup>1</sup>.

**Tableau 1 - Part des productions attendues et non attendues avec les noms de une à quatre syllabes en P1 et P2 (chez les quatre enfants)**

Contexte	Période	Lubin	Naël	Antonin	Adrien
Préférentiel	1	74,40%	62,46%	57,37%	71,86%
	2	51,78%	54,29%	47,83%	64,42%
Non préférentiel	1	25,60%	37,54%	42,63%	28,14%
	2	48,22%	45,71%	52,17%	35,58%

Conformément à ce que nous avons pu remarquer avec les noms d'une et deux syllabe(s), on constate que tous les enfants privilégient des combinaisons binaires en période 1. Les productions ne sont cependant pas toutes conformes à ce qui pourrait être attendu dans l'hypothèse syllabique. Le minimum de cas non attendus est observé chez Lubin en P1 et représente tout de même 25,60% des combinaisons impliquant des noms d'une et quatre syllabes. En période 2, la proportion de noms produits dans un contexte préférentiel est inférieure à celle des noms produits dans un contexte non préférentiel chez Antonin, et légèrement supérieure seulement dans les données de Lubin et Naël. Seul Adrien privilégie encore nettement la structure bisyllabique.

<sup>1</sup> Pour rappel, les combinaisons dites 'préférentielles' correspondent aux quatre combinaisons suivantes : 'omission + nom bisyllabique', 'filler/déterminant + nom monosyllabique', 'omission + nom de quatre syllabes' et 'filler/déterminant + nom de trois syllabes'. Les combinaisons dites 'non préférentielles' regroupent quant à elles les quatre cas de figure qui suivent : 'omission + nom d'une syllabe', 'filler/déterminant + nom de deux syllabes', 'omission + nom de trois syllabes' et 'filler/déterminant + nom de quatre syllabes'.

Nos résultats concernant la première période tendent donc à confirmer les conclusions de travaux antérieurs sur la sensibilité de l'enfant à la structure syllabique (*i.a.* Demuth, 2001 ; Demuth et Tremblay, 2008). En période 2, l'augmentation des combinaisons non préférentielles est probablement liée en grande partie à la diminution des occurrences de noms de deux syllabes précédés d'une omission, mais il est possible d'envisager également, comme nous le signalions plus tôt, que pour certains enfants au moins, le rapport au facteur phono-prosodique soit plus marqué dans les premières étapes du développement des déterminants. Ces résultats concorderaient alors avec les conclusions de Veneziano et Sinclair (1997, 2000) ou Kilani-Schoch et Dressler (2000) sur l'évolution du statut des fillers, d'abord entités prosodiques avant de devenir des précurseurs des morphèmes adultes. Toutefois, pour que nos résultats puissent être pleinement comparés à des travaux antérieurs, nos analyses devront être complétées par une observation de la structure rythmique des syntagmes nominaux. Nous pourrions ainsi évaluer si les fillers et les déterminants sont préférentiellement produits dans des contextes spécifiques, et en particulier, dans des contextes de type 'Syll. 1 V non accentuée + Syll. 2 CV accentuée', comme cela a pu être montré dans la littérature.

Malgré la nette préférence observée pour la structure bisyllabique en période 1, nos données montrent néanmoins que même durant cette première période, le facteur phono-prosodique n'explique pas à lui seul toutes les productions de l'enfant et par ailleurs, que tous les enfants n'y sont pas sensibles de la même façon ou durant la même période. Si la variation individuelle est plus visible en période 2, on voit également que durant la période 1, Lubin et Adrien tendent à respecter davantage le schéma syllabique binaire que Naël et Antonin.

Enfin, d'un point de vue méthodologique, il nous semble que d'autres éléments - qui ne sont pas inclus dans le syntagme phonologique - pourraient être pris en compte dans le décompte syllabique, et en particulier, les éléments pivots dans les cas d'énoncés à deux termes. Dans les exemples 2 ci-dessous, les trois enfants - Lubin, Naël et Adrien - produisent des syntagmes phonologiques que nous avons considérés comme des syntagmes d'une syllabe et qui ne répondent donc *a priori* pas à l'hypothèse phono-prosodique. Or, on constate que ces syntagmes sont précédés de lexèmes que nous pouvons à ces âges considérer comme des pivots ('pas', 'c'est', 'veux') et qui ne sont constitués que d'une seule syllabe. L'énoncé complet de l'enfant est par conséquent bisyllabique et nous pouvons donc nous interroger ici sur la pertinence des

frontières du syntagme phonologique pour mesurer la préférence pour une structure syllabique binaire lorsque la LME ne dépasse pas 2 et que les énoncés sont proto-syntaxiques.

### Exemples 2 - Syntagmes phonologiques d'une syllabe précédés d'éléments pivots

- (a) Lubin (1;9) : [pa pɔm] 'pas pomme.'
- (b) Naël (1;9) : [se vas] 'est vache ?'
- (c) Adrien (2;3) : [vo pah] 'veux pâ(tes).'

Ce constat est également valable pour les cas de figure qui ont été caractérisés comme des syntagmes phonologiques de deux syllabes et donc conformes à l'hypothèse syllabique (*cf.* exemples 3).

### Exemples 3 - Syntagmes phonologiques de deux syllabes précédés d'éléments pivots

- (a) Lubin (1;11) : [sa zazɛ] 'ça gaze(lle).'
- (b) Naël (1;8) : [te kefa] 'est cochon ?'

Dans les deux exemples ci-dessus, Lubin et Naël produisent des noms bisyllabiques, ici aussi précédés de pivots ('ça' et 'est'). Nous pourrions envisager dans ces exemples que le rapport de l'enfant à la structure syllabique puisse se situer au niveau de l'énoncé entier plutôt qu'au niveau du syntagme phonologique, rendant ainsi ces cas non adéquats pour une interprétation prosodique.

## 2. Influence du statut sémantique du nom associé à la forme pré-nominale

Outre l'influence de la structure syllabique, nous avons cherché à évaluer l'impact éventuel du statut sémantique du nom sur la production des formes pré-nominales. Nous avons vu plus tôt (*cf.* point 3.2.1 du chapitre 2) que d'après Bassano (*cf.* par exemple Bassano, 2005 ou Bassano *et al.*, 2008), les propriétés lexico-sémantiques du nom peuvent expliquer la distribution des formes devant les noms. L'auteure a notamment pu observer que l'omission est plus représentée devant les noms renvoyant à des entités animées et les déterminants devant ceux renvoyant à des entités inanimées concrètes. La catégorie des inanimés abstraits serait quant à elle la catégorie privilégiée pour la production des fillers. Les constats concernant les déterminants sont partagés par Nashawati (2010) mais à l'inverse, les résultats relatifs aux fillers et aux omissions ne confirment pas les conclusions de Bassano. Ces travaux induisent trois questions principales : i) peut-on

observer des différences d'emploi des formes pré-nominales devant les noms dits 'animés' et 'inanimés' ? ii) qu'en est-il de la distribution des omissions, fillers et déterminants devant les noms renvoyant à des entités concrètes et abstraites ? iii) de façon globale, est-il possible de considérer qu'il existe des combinaisons préférentielles entre formes pré-nominales et statuts sémantiques des noms ?

Cette section 2 est construite en quatre parties. Le premier point (2.1) nous permettra de rendre compte des différentes classes sémantiques retenues pour caractériser chacun des noms. Avant de discuter de l'influence des propriétés sémantiques du nom sur les formes produites, nous ferons au point 2.2 un état des lieux de l'emploi des différents types de noms dans notre corpus. Aux points 2.3 et 2.4, nous nous intéresserons d'une part à la distribution des formes devant les noms d'entités animées et inanimées, ainsi que devant les noms marquant un événement ou une action, et d'autre part, à la distribution des formes devant les noms concrets et abstraits. De l'observation de ces premières oppositions découlera au point 2.5 une discussion sur l'existence de combinaisons préférentielles entre formes pré-nominales et statuts sémantiques des noms.

## **2.1 Caractérisation sémantique des noms du corpus**

Sur la base des travaux que nous venons de citer, nous avons dans un premier temps identifié les noms de la catégorie des animés, par opposition aux noms de la catégorie des inanimés. Les premiers renvoient à des entités (humaines ou non) animées, présentes ou non dans la situation de communication et pouvant être réelles ou fictives. Les seconds renvoient quant à eux à des entités inanimées (objet, lieu, aliment, ...). Cette seule opposition n'étant pas pleinement satisfaisante dans la mesure où elle ne permet pas de couvrir tous les types de productions de l'enfant, nous avons, comme l'avait fait Nashawati (2010) au préalable, ajouté une troisième catégorie : celle des noms renvoyant à une action ou à un événement. Notons par ailleurs que dès lors que l'enfant confère aux personnages de livre, playmobil ou peluches des caractéristiques propres aux êtres humains, tous les noms les caractérisant ont été intégrés à la catégorie des entités animées.

Un second codage a été effectué, opposant cette fois les noms dont les référents sont concrets et ceux dont les référents sont abstraits. Les noms dits concrets désignent, d'après les grammairiens (*cf.* Riegel *et al.*, 1994), des éléments qui peuvent être perçus par nos sens (exemples : 'table',



‘chanson’, ‘parfum’) et *a contrario*, les noms dits abstraits désignent des éléments qui ne peuvent pas être perçus par nos sens mais seulement par l’esprit, tels que les idées, les sentiments ou les émotions (exemples : ‘peur’, ‘paix’, ‘tristesse’). Ces noms renvoyant à des entités abstraites ne peuvent pas être quantifiés.

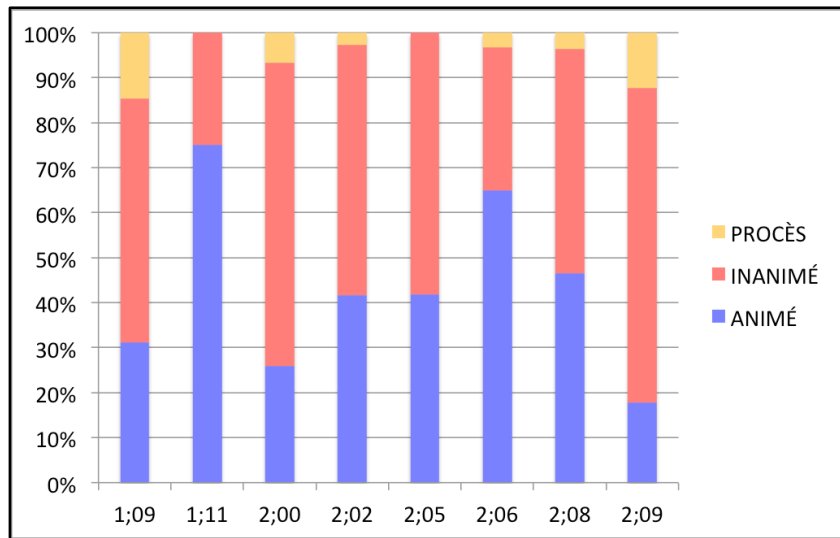
## 2.2 Emploi des classes sémantiques de noms dans le corpus

Avant de commenter les résultats des analyses portant sur la distribution des formes devant les différents types sémantiques de noms, nous nous pencherons dans cette section 2.2 sur l’emploi de ces classes sémantiques par les quatre enfants. Afin de nous figurer la représentation de chacune des catégories présentées au point précédent, nous passerons d’abord en revue la distribution des noms animés, inanimés, et des noms de procès dans les quatre corpus (graphiques 13 à 16), ainsi que la distribution des noms concrets et abstraits (tableau 2). Les premiers résultats sont calculés en pourcentages sur la part totale des noms pouvant être précédés d’un déterminant, pour chaque enfant et séance par séance. En revanche, nos données ne comptant que quelques occurrences de noms abstraits, les résultats relatifs à l’opposition entre noms concrets et abstraits sont quant à eux présentés en valeurs absolues.

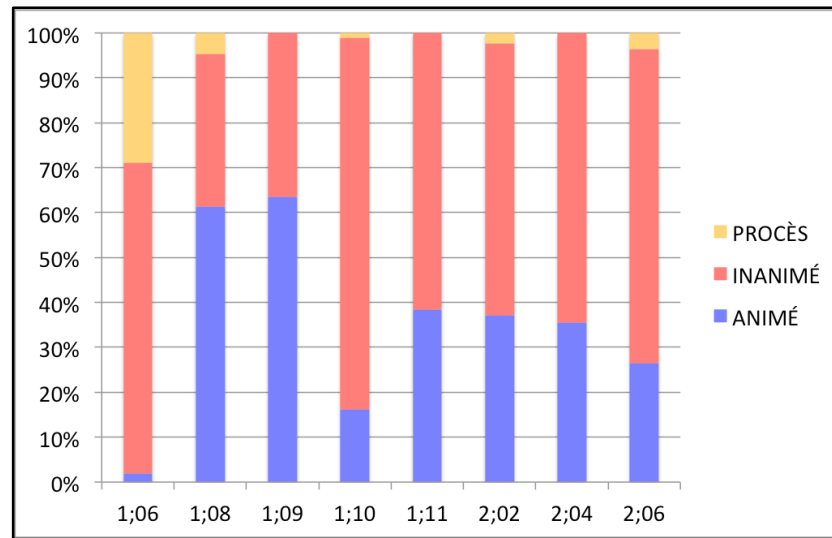
Les résultats montrent que dans le corpus de Lubin, les noms référant à des entités inanimées sont les plus fréquents si l’on observe toutes les séances à la fois (54,20%) mais les noms renvoyant à des entités animées sont aussi largement produits dans ce corpus (40,79%, toutes séances confondues). Les animés sont même majoritaires à 1;11, et cela en raison de l’activité en cours dans cette séance : le père et l’enfant jouent à un puzzle représentant des animaux. Les inanimés sont plus nettement majoritaires dans les données de Naël, à l’exception des séances à 1;8 et 1;9. Cette proportion plus importante d’animés s’explique comme chez Lubin par l’activité de l’enfant. A 1;8, Naël joue avec des magnets sur lesquels figurent des animaux et à 1;9, avec un puzzle sur lequel figurent les animaux de la ferme. Dans les données d’Antonin, la part des noms renvoyant à des entités animées est faible dans les séances 1 et 5 à 8, et beaucoup plus importante de 1;11 à 2;4. Ici également ces pourcentages sont à mettre en lien avec les activités et les supports utilisés lors de ces séances (puzzle mettant en scène des animaux, imagiers d’animaux, jeux d’associations entre animaux jeunes et adultes, cubes à assembler pour recomposer des animaux, ...).

**Graphiques 13 à 16 - Distribution des noms animés, inanimés, et des noms de procès dans les quatre corpus**

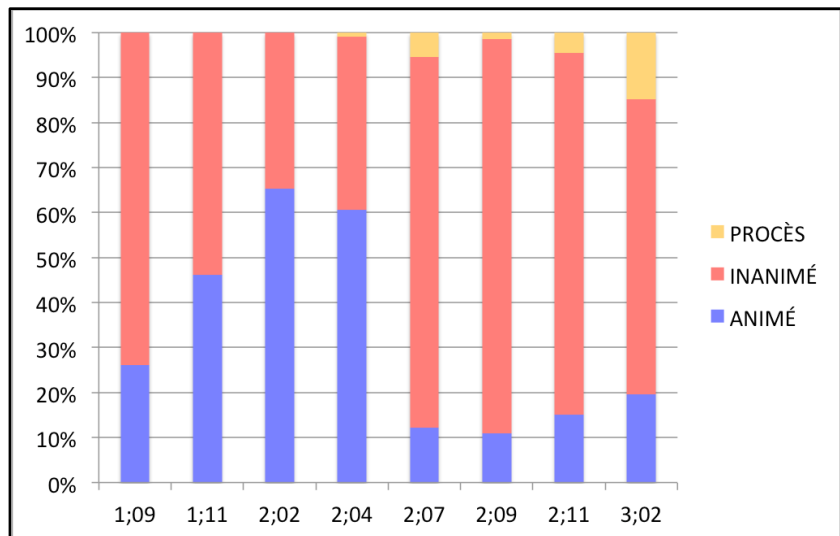
*Graphique 13 - Corpus de Lubin*



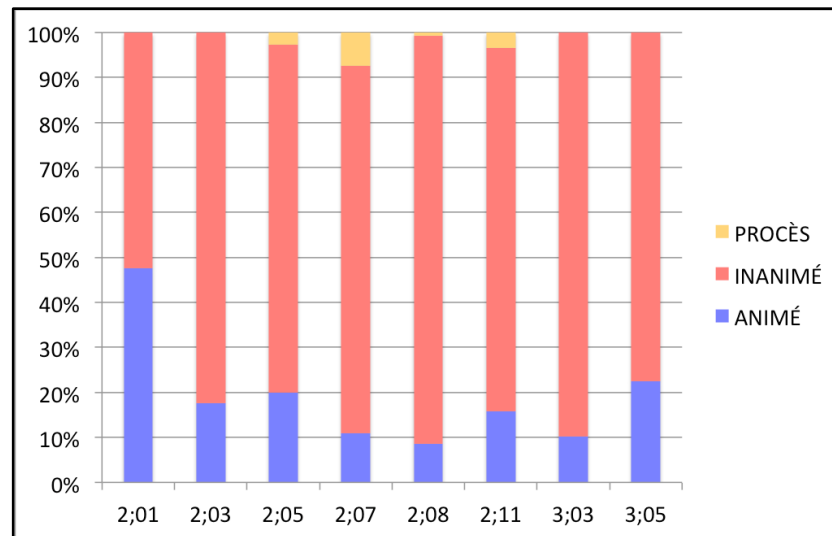
*Graphique 14 - Corpus de Naël*



*Graphique 15 - Corpus d'Antonin*



*Graphique 16 - Corpus d'Adrien*



Le corpus d'Adrien compte quant à lui nettement moins d'animés que d'inanimés, sauf en séance 1, et cela en raison d'une référence répétée à la tortue en plastique que manipule Adrien dans son bain.

Nous n'observons donc pas d'évolution particulière de la distribution des noms animés et inanimés dans nos données, mais plutôt une corrélation entre le type d'entité et l'activité en cours ainsi que le support de cette activité.

Conformément à ce que nous pouvions attendre, nous n'observons que quelques occurrences de noms renvoyant à une action ou un événement ('procès') et cela, chez tous les enfants. Ces noms sont le plus souvent comparables d'un enfant à l'autre. On retrouve par exemple à plusieurs reprises dans nos données des productions de type 'faire des bêtises', 'faire un gâté', 'prendre le goûter', 'chanter une chanson'.

En ce qui concerne la distinction entre noms concrets et abstraits, on observe de façon évidente que les noms concrets représentent la quasi-intégralité des noms de notre corpus et qu'au contraire, les noms abstraits ne sont visibles que dans quelques séances chez chaque enfant, et en très petit nombre (*cf.* tableau 2 ci-dessous).

**Tableau 2 - Distribution des noms concrets et abstraits dans les quatre corpus**

		1	2	3	4	5	6	7	8
<b>CONCRETS</b>	<b>Lubin</b>	48	40	88	36	129	60	83	68
	<b>Naël</b>	51	106	96	87	60	86	76	80
	<b>Antonin</b>	42	130	52	218	164	128	66	59
	<b>Adrien</b>	21	91	141	103	150	56	67	89
<b>ABSTRAITS</b>	<b>Lubin</b>	---	---	---	---	---	---	1	3
	<b>Naël</b>	1	---	---	---	---	---	3	3
	<b>Antonin</b>	---	---	---	---	---	---	---	2
	<b>Adrien</b>	---	---	---	6	---	---	---	---

Ces résultats s'expliquent d'après nous par la nature du premier lexique de l'enfant, qui se construit d'abord au travers de ce qui est concret (et par ailleurs, présent dans la situation). Il est donc probable que les enfants que nous avons observés soient trop jeunes pour produire des noms abstraits. Ce constat implique que nous privilégions plutôt au point 2.4 une observation

qualitative des noms abstraits produits dans notre corpus plutôt qu'une recherche quantitative des éventuels contrastes dans les usages des formes associées aux noms concrets et abstraits, comme cela avait été initialement envisagé.

### **2.3 Distribution des formes devant les noms animés, inanimés et de procès**

Au sein de ce point 2.3 seront présentés les résultats obtenus pour les noms animés et inanimés, ainsi que pour les noms de procès. Afin de ne pas lisser un phénomène qui évolue et peut donc donner lieu à des résultats divergents d'une séance à l'autre, nous avons préféré ne pas compiler les séances et présenter la distribution des formes séance par séance. Les quatre séries de trois graphiques qui suivent présentent la distribution des formes pré-nominales chez chaque enfant, lorsque les noms sont animés, inanimés et enfin, lorsqu'ils renvoient à des actions ou des événements. Pour les noms animés et inanimés, les graphiques présentent la répartition - en pourcentages - des formes sur le nombre total de noms de chaque catégorie sémantique produits à chaque séance. Dans la mesure où la dernière catégorie sémantique (noms de procès) n'est que très peu représentée dans nos données, ce sont les valeurs absolues que nous avons fait figurer dans le troisième graphique de chaque série. Contrairement aux analyses phono-prosodiques, nous avons cette fois distingué les fillers des déterminants.

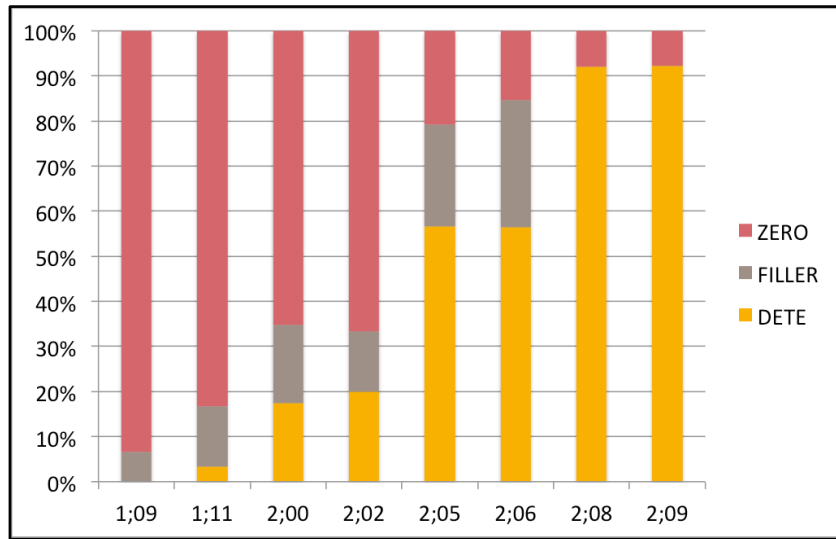
La première remarque que nous pouvons faire suite à l'observation des résultats concernant le corpus de Lubin (graphiques 17 à 19) - et qui sera également valable pour les autres corpus - porte sur le lien évident entre la distribution des formes dans les différents contextes sémantiques et le développement général de la position pré-nominale. L'intérêt de ces résultats repose donc, outre ce premier constat, sur la recherche d'aspects spécifiques à chaque statut sémantique.

Dans le corpus de Lubin, si l'on compare les deux statuts sémantiques les plus fréquents, les animés et les inanimés, on n'observe pas de différence notable entre les deux, en particulier dans la proportion des formes produites à 2;2, 2;5, 2;8 et 2;9.

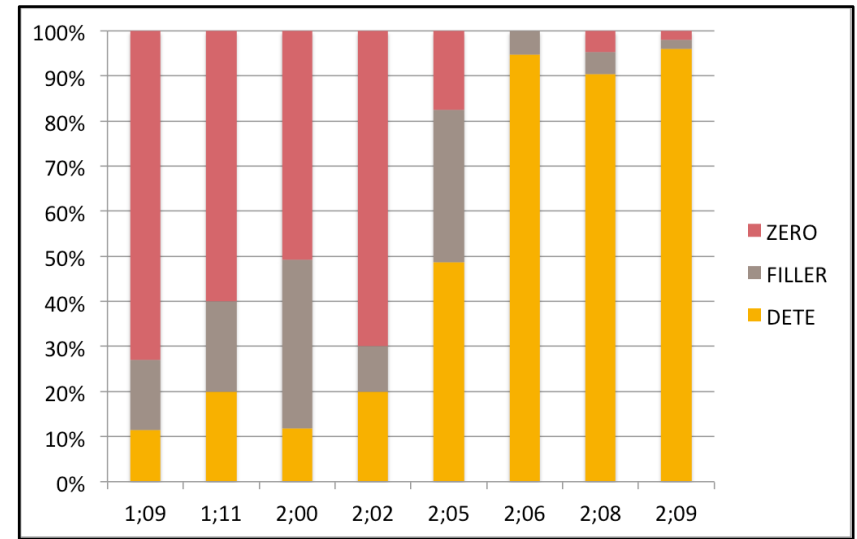
Certains éléments semblent cependant différencier ces deux contextes. De 1;9 à 2;0, les fillers et les déterminants sont plus fréquents avec les noms renvoyant à des entités animées en séance 1 (26,92% contre 6,67% pour les inanimés), en séance 2 (40,00% contre 16,67%), et en séance 3 (49,15% contre 34,78%).

**Graphiques 17 à 19 - Distribution des formes pré-nominales devant les noms animés, inanimés et les noms de procès chez Lubin**

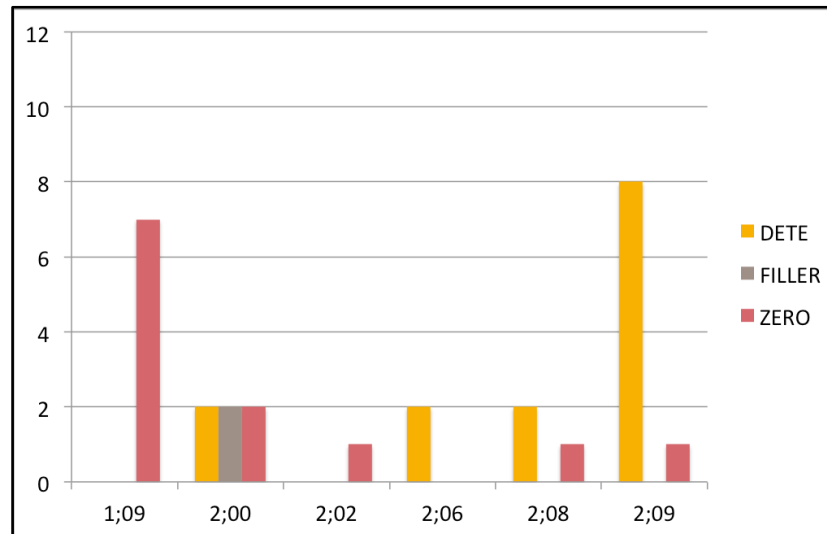
*Graphique 17 - Noms animés*



*Graphique 18 - Noms inanimés*



*Graphique 19 - Noms de procès*



En parallèle, les absences de formes sont plus fréquentes devant les noms renvoyant à des entités animées. A 2;6, les noms inanimés sont presque exclusivement associés à des formes adultes (94,74%) et noms animés présentent une diversité plus grande dans les formes qui les précèdent : 56,41% de déterminants, 28,21% de fillers et 25,38% d'omissions.

Ces constats pourraient signifier que les noms inanimés sont privilégiés par l'enfant pour le développement des formes mais il est plus pertinent semble-t-il de considérer dans ce corpus un ensemble 'présence de forme' *vs.* 'absence de forme' plutôt que de considérer les fillers et les déterminants de façon distincte.

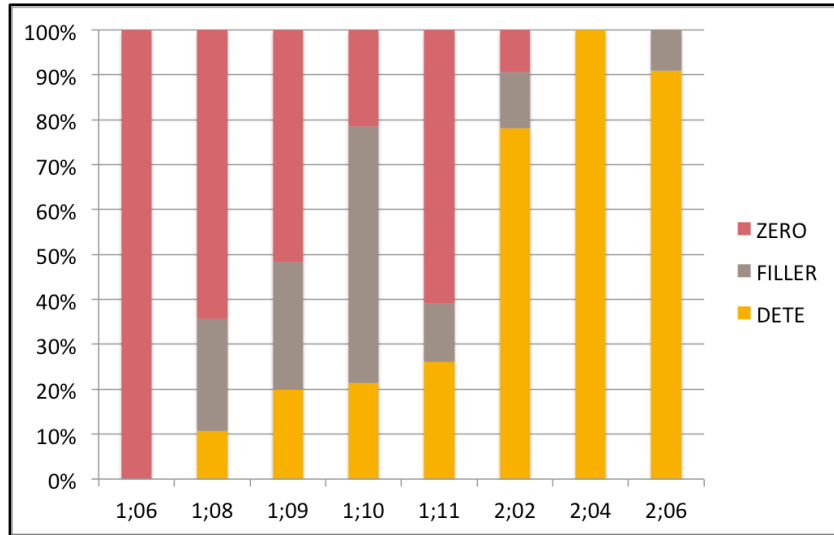
Enfin, les constats concernant les noms renvoyant à des actions ou des événements sont limités par le peu d'occurrences de noms de ce type. Outre le fait évident que les omissions sont plus visibles de 1;9 à 2;2 que de 2;6 à 2;9, on peut néanmoins constater que les noms de procès en semblent pas favoriser la production de fillers.

Les résultats concernant les données de Naël (graphiques 20 à 22) ne donnent pas non plus à voir de différences remarquables entre les noms renvoyant à des entités animées et ceux renvoyant à des entités inanimées. On peut cependant relever quelques différences dans la distribution des formes à 1;6, 1;10 et 1;11. Dans la mesure où un seul nom animé est produit à 1;6, la comparaison n'est en fait pertinente qu'à 1;10 et 1;11. A 1;10, les animés sont surtout précédés de fillers (57,14%) et à part égale, d'omissions et de déterminants. Devant les noms inanimés, les fillers ne représentent que 26,76% des formes produites, les fillers, 32,39%, et contrairement à ce que nous pouvons observer avec les animés, les déterminants sont majoritaires (40,85%). A 1;11, les déterminants représentent 26,09% des formes pré-nominales associées aux animés, contre 36,11% avec les inanimés. Les omissions sont plus présentes avec les animés (60,87%) qu'avec les inanimés (33,33%) En revanche, les fillers sont plus fréquents dans cette séance avec les inanimés (30,56%, contre 13,04% avec les inanimés).

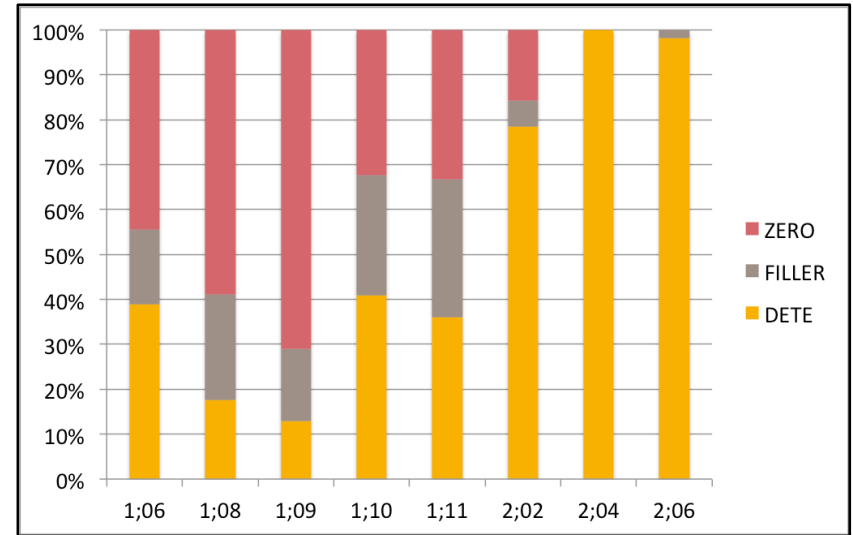
Ces résultats tendent à montrer que si le statut sémantique du nom a un rôle à jouer dans le type de forme pré-nominale que l'enfant produit, il se manifeste seulement chez Naël par une production plus précoce et plus fréquente de déterminants avec les inanimés.

**Graphiques 20 à 22 - Distribution des formes pré-nominales devant les noms animés, inanimés et les noms de procès chez Naël**

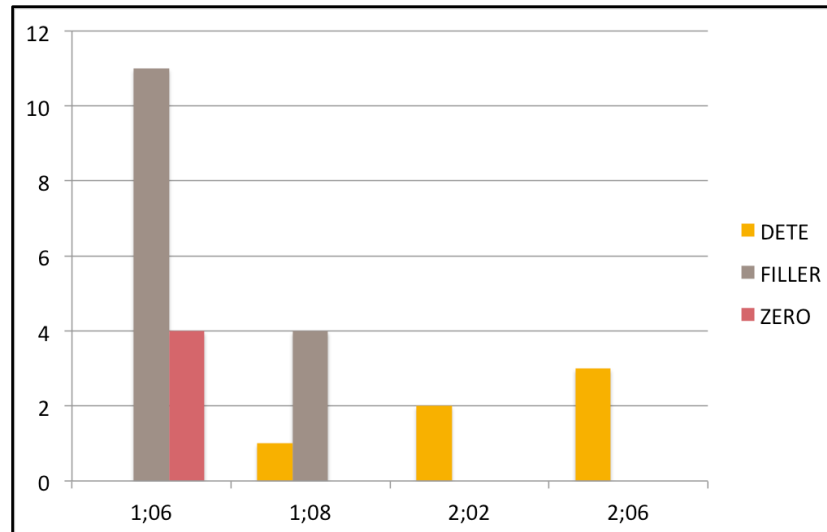
*Graphique 20 - Noms animés*



*Graphique 21 - Noms inanimés*



*Graphique 22 - Noms de procès*



Ce que nous observons par ailleurs avec les noms d'actions ou d'événements (procès) nous permet seulement de dire que les formes produites par l'enfant sont conformes à celles qu'il produit quel que soit le contexte sémantique. Si les noms de procès ne semblent pas favoriser les omissions au-delà de la première séance (pourtant largement présentes à 1;8), la part de ces noms est beaucoup trop minime pour que nous puissions en tirer de quelconques conclusions.

Dans le corpus d'Antonin (graphiques 23 à 25), la distribution des formes est comparable devant les noms animés et inanimés dans une première période, de 1;9 à 2;4. La seule différence que l'on observe concerne la première séance. En effet, les seules occurrences de déterminants produits par l'enfant durant cette session le sont devant des noms renvoyant à des entités inanimées.

A 2;7, la proportion de déterminants est nettement plus importante devant les noms inanimés (56,30%) que devant les noms animés (25,00%) et à l'inverse, les fillers sont privilégiés devant les noms animés.

Le panorama des formes s'inverse ensuite de 2;9 à 3;2 puisque les déterminants deviennent majoritaires devant les noms animés et les fillers et les omissions devant les noms inanimés.

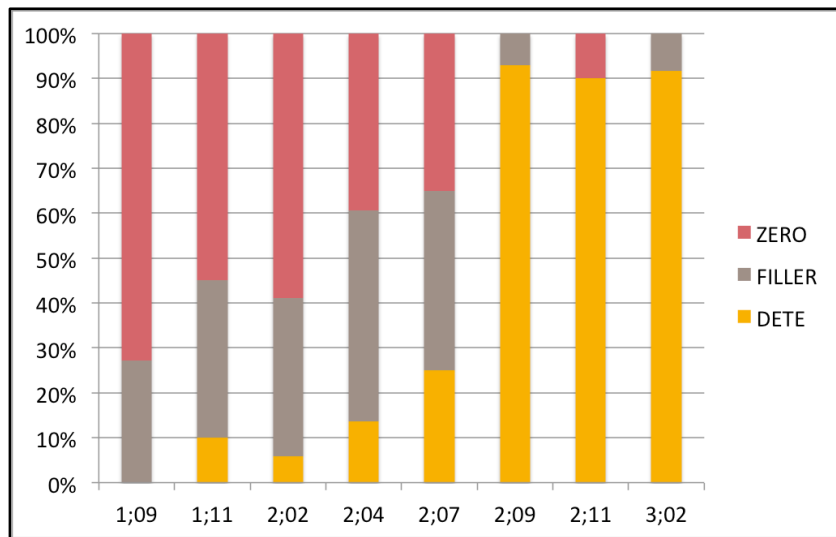
Les différences que nous observons ici dans l'usage que l'enfant fait des formes devant les noms animés et inanimés ne sont pas suffisamment nombreuses (ou pas assez constantes) pour pouvoir affirmer ici que le statut sémantique du nom joue un rôle dans la distribution des omissions, fillers et déterminants, et cela quelque soit la période étudiée.

Enfin, les occurrences de noms de procès étant ici encore peu nombreuses, il nous est difficile de faire des constats étayés mais on peut toutefois relever que comme chez Lubin, les fillers - par ailleurs très présents dans les données d'Antonin - ne semblent pas être privilégiés devant des noms renvoyant à des actions ou à des événements.

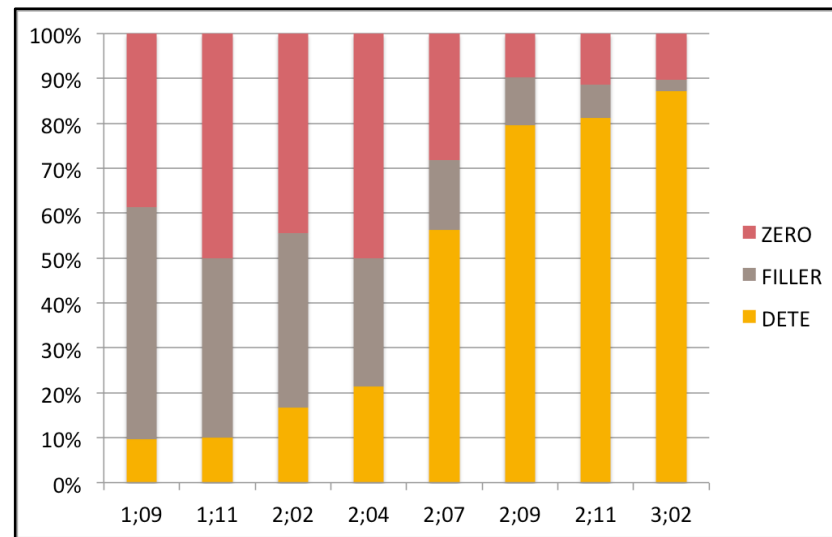


**Graphiques 23 à 25 - Distribution des formes pré-nominales devant les noms animés, inanimés et les noms de procès chez Antonin**

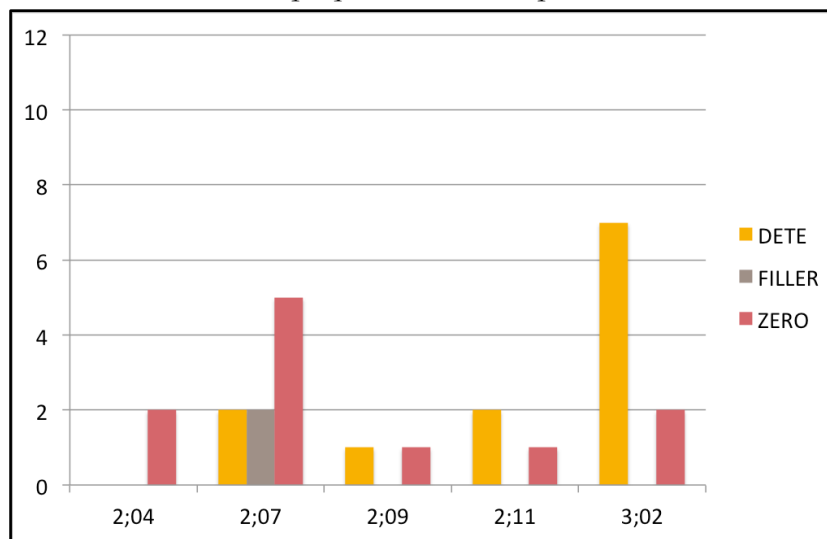
*Graphique 23 - Noms animés*



*Graphique 24 - Noms inanimés*



*Graphique 25 - Noms de procès*



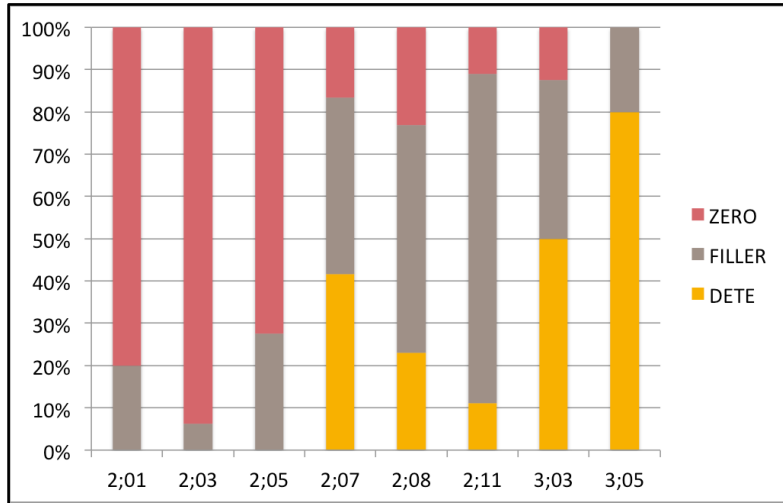
Les graphiques 26 et 27 ci-dessous montrent que globalement, chez Adrien comme chez les trois autres enfants, la distribution des formes devant les noms animés et inanimés suit l'évolution de la position pré-nominale. On observe également que les trois formes - omission, filler et déterminant - sont employées dans les deux contextes de façon comparable. Un récapitulatif toutes séances confondues le confirme : les déterminants sont présents devant 24,79% des noms animés et 28,57% des noms inanimés ; les fillers sont produits dans 31,62% des contextes animés et 35,80% des contextes inanimés, et parallèlement, les omissions sont présentes devant 43,59% des noms animés et 35,63% des noms inanimés.

On peut toutefois relever quelques différences : les seuls déterminants utilisés dans les premières séances (à 2;1, 2;3 et 2;5) sont produits devant des noms inanimés et à l'inverse, les omissions sont plus largement présentes lors de ces séances devant les noms animés qu'inanimés. Dans la continuité de ce que nous observons pour les déterminants, nous pouvons aussi signaler que les noms inanimés produits à 2;3 sont plus souvent précédés de fillers que les noms animés au même âge.

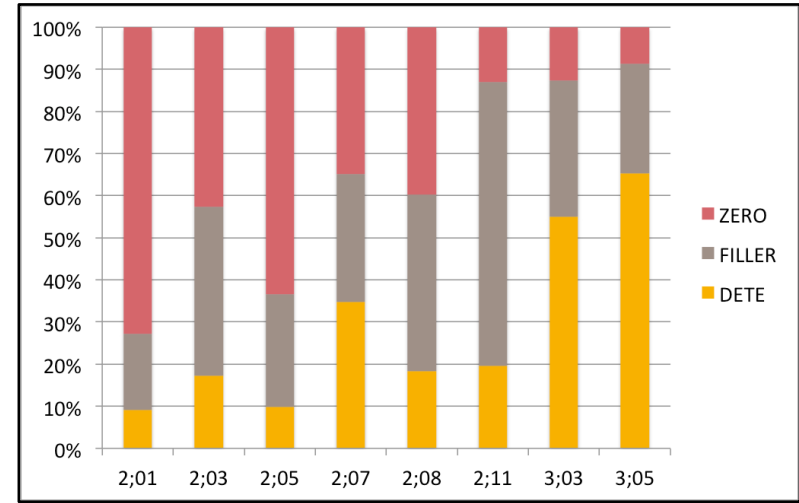
En ce qui concerne les noms renvoyant à des procès, aucun constat solide n'est possible dans le corpus d'Adrien, comme dans les autres corpus, du fait du peu d'occurrences de noms de ce type (présents dans seulement quatre des huit séances). On peut cependant remarquer qu'aucun déterminant n'est produit devant ce type de noms, et que comme pour les noms animés et inanimés, les formes produites concordent avec le développement global des formes puisque les noms sont d'abord majoritairement précédés d'omissions avant d'être précédés de fillers.

**Graphiques 26 à 28 - Distribution des formes pré-nominales devant les noms animés, inanimés et les noms de procès chez Adrien**

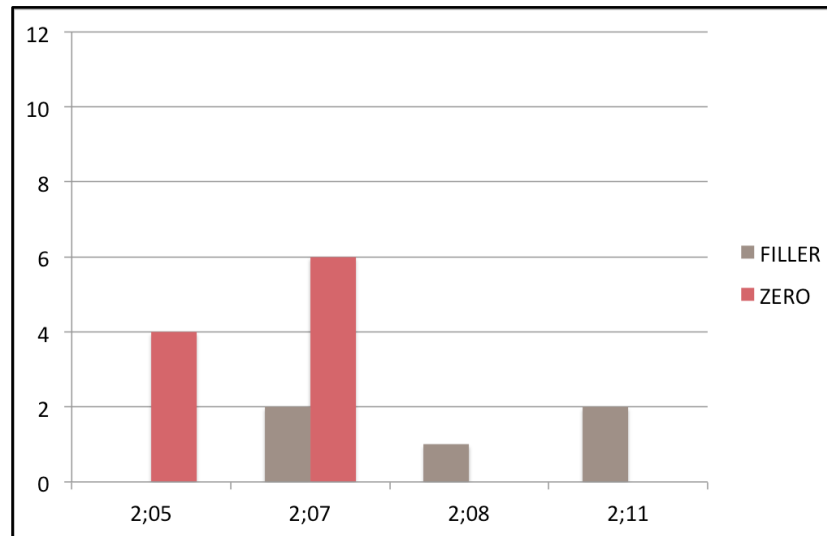
*Graphique 26 - Noms animés*



*Graphique 27 - Noms inanimés*



*Graphique 28 - Noms de procès*



## 2.4 Distribution des formes devant les noms concrets et abstraits

Les mêmes croisements ont été opérés pour les noms concrets et abstraits que pour les noms animés, inanimés et les noms d'actions et d'événements. Toutefois, une représentation graphique des résultats ne serait pas pertinente ici au vu de la quasi-absence des noms abstraits dans nos données. La distribution des omissions, fillers et déterminants devant les noms concrets ne donne rien de significatif à voir du point de vue sémantique et ne fait que refléter la répartition globale des formes chez chaque enfant, telle que nous l'avons présentée au chapitre V. De même que pour les noms de procès, il est difficile d'établir de véritables constats concernant la distribution des formes devant les noms abstraits. Les quelques cas observés dans nos données seront donc seulement discutés ici, sans chercher à définir des associations spécifiques entre formes et noms abstraits.

Dans le corpus de Lubin, seuls quatre noms considérés comme abstraits ont été relevés (*cf.* exemples 4 ci-dessous), l'un d'entre eux à 2;8 et les trois autres à 2;9.

### Exemples 4 - Noms abstraits produits dans le corpus de Lubin

- (a) [wi zε bəzwẽ dɛd] 'oui j'ai besoin d'aide.' (2;8)
- (b) [si sɛ lœɛ (.) lœɛ də gute] 'si c'est l'heure (.) l'heure de goûter.' (2;9)
- (c) [ɛskø y a dɛwa] 'est ce que y a droit ?' (2;9)
- (d) [ʃ pø fɛɛ de pat a modle ø de zitoiø ø de msjø madam ø de kɥizin] 'on peut faire des pâtes à modeler euh des histoires des Monsieur Madame euh des cuisines.' (2;9).

Le premier '-aide-' est produit suite à une question de l'adulte ('tu veux un p(e)tit peu d'aide ?') et est précédé d'un déterminant, identique à celui qui est produit dans l'énoncé précédent du père. Dans le deuxième cas de figure, l'énoncé de l'enfant constitue également une reprise partielle de l'énoncé précédent de la mère ('nan c'est pas l'heure') et le déterminant produit par Lubin est là aussi le même que celui que produit son interlocuteur. Dans l'énoncé *c*, l'enfant reprend une partie de la production précédente de l'observatrice ('va d(e)mander à maman si t'as l(e) droit') pour l'adresser à la mère mais il le fait en omettant la forme pré-nominale. Enfin, dans l'exemple *d*, Lubin imagine une liste de tout ce qu'il est possible de représenter avec le nouveau jeu qu'il vient d'avoir et dans ce contexte, il produit le nom 'histoire' que nous avons caractérisé comme un nom abstrait. Ce dernier est précédé du déterminant 'des'.

Trois noms abstraits sur quatre sont donc précédés de déterminants mais ce constat ne peut en aucun cas nous permettre de tirer des conclusions sur des associations éventuelles entre noms abstraits et formes pré-nominales dans le corpus de Lubin. Outre le fait évident que les données ne sont pas suffisantes pour cela, ces noms sont tous produits au cours des deux dernières séances, pendant lesquelles les déterminants sont majoritaires. L'intérêt de ces exemples repose donc probablement davantage sur l'éclairage qu'ils nous apportent sur la façon dont l'enfant s'approprie les productions de ses interlocuteurs que sur son maniement des noms abstraits.

Dans le corpus de Naël, sept occurrences de noms ont été caractérisées comme abstraites (cf. exemples 5). Deux d'entre elles concernent le nom 'bêtise' qui est produit à la fois dans la première séance (cf. exemple *a*) et dans la dernière séance (cf. exemple *g*). A 1;6, le nom est produit en reprise immédiate de la mère ('ça c'est une bêtise') et un an plus tard, à 2;6, c'est l'enfant qui le produit pour la première fois dans la séance. Dans le premier cas, nous avons analysé le [e] produit devant le nom comme un proto-présentatif et non comme un filler pré-nominal. De fait, le nom est associé à une omission, contrairement à l'occurrence à 2;6, qui est précédée de l'indéfini 'un'. Notons que le père reprend le même syntagme au tour suivant et reformule l'énoncé de l'enfant : 'i(l) pense que c'est une bêtise'.

### Exemples 5 - Noms abstraits produits dans le corpus de Naël

- (a) [e betiz] 'yyy bêtise ?' (1;6)
- (b) [a z ε yn ide] 'ab@i F ai une idée.' (2;4)
- (c) [X ʒε yn ide] 'xxx j'ai une idée.' (2;4)
- (d) [a wi ʒε yn ide] 'ab@i oui j'ai une idée !' (2;4)
- (e) [nã ʒə se pa liʁ le zistwaʁ] 'nan je sais pas lire les histoires.' (2;6)
- (f) [mwa ʃ se pa ʁakɔ̃te le zistwaʁ] 'moi j(e) sais pas raconter les histoires.' (2;6)
- (g) [i ʁuv sa yn betiz] '+, i(l) (t)rouve ça une bêtise.' (2;6)

Les trois énoncés produits à 2;4 (cf. *b*, *c* et *d*) contiennent le nom abstrait 'idée' que l'enfant est le seul à produire dans cette séance. Le nom est à chaque fois associé au déterminant 'un', à un âge où rappelons le, l'enfant ne produit plus que des déterminants. Enfin, dans les deux énoncés qui suivent (*e* et *f*), Naël produit le nom 'histoire', et comme attendu à 2;6, il est dans les deux cas précédé du déterminant 'les'. Cette forme n'est pas reprise des syntagmes nominaux offerts par l'adulte puisque le père produit dans un premier temps 'tu veux lui lire une histoire à Marine ?', puis 'tu lui racontes à Marine l'histoire de Kirikou ?'. Si ces deux exemples apparaissent trop tardivement pour nous renseigner sur le maniement des différents types de formes devant les

noms abstraits, il pourra être intéressant d'y revenir au chapitre VII, lorsque nous ferons, entre autres, la description des formes pré-nominales produites par l'enfant suite à des offres de l'adulte.

Nous n'avons caractérisé comme abstraits que deux des noms produits par Antonin, 'avis' et 'âge' (cf. exemples 6). Ces deux noms sont produits au cours de la dernière séance, à 3;2, et pour la première fois dans la séance.

### Exemples 6 - Noms abstraits produits dans le corpus d'Antonin

- (a) [a mɔ̃ avi i va fɛʁ kɔ̃m X] 'à mon avis i(l) va faire comme xxx.' (3;2)
- (b) [e a mɔ̃ naʒ] 'et à mon âge ?' (3;2)

Bien que l'enfant produise d'autres possessifs dans cette séance, on peut s'interroger sur la valeur de celui qui est produit dans l'exemple *a*. On pourrait faire l'hypothèse ici d'un usage relativement figé de 'mon', lié à la nature même de cette expression, le plus souvent utilisée comme une locution adverbiale invariable. Dans cette perspective, la caractérisation de 'avis' comme un nom à part entière et abstrait est d'ailleurs largement discutable.

Enfin, six occurrences de noms abstraits ont pu être observées dans le corpus d'Adrien (à 2;7) mais toutes concernent le même nom 'bêtise', d'abord repris d'un premier énoncé de la mère ('il fait des bêtises'), puis repris ensuite à la fois par l'enfant et par la mère.

### Exemples 7 - Noms abstraits produits dans le corpus d'Adrien

- (a) [betiç] 'bêtises.' (2;6)
- (b) [betiti beti:] 'bêtises [l] bêtises.' (2;6)
- (c) [avi betiçi] 'yyy bêtises.' (2;6)
- (d) [betiœ] 'bêtises.' (2;6)
- (e) [avi betisi] 'yyy bêtises.' (2;6)

Les dix énoncés de la mère contenant le nom 'bêtise' sont précédés dans cette séance de l'indéfini 'des'. Parallèlement, les productions d'Adrien sont toutes précédées d'omissions. Celles-ci représentant encore plus de 30% des formes pré-nominales à cet âge, ces occurrences du nom 'bêtise' n'apparaissent pas comme spécifiques au statut sémantique du nom, et en tout état de cause, les données sont ici encore insuffisantes pour qu'on cherche à les interpréter d'un point de vue sémantique.

## 2.5 Des combinaisons préférentielles entre formes pré-nominales et statuts sémantiques des noms ?

Après avoir croisé les formes pré-nominales produites et le statut sémantique des noms séance par séance, nous avons cherché à observer si certaines tendances générales se dégagèrent chez chaque enfant et d'un enfant à l'autre. Pour ce faire, nous avons compilé certaines séances et conservé deux périodes distinctes pour chaque enfant, puis identifié la forme la plus fréquemment produite dans deux contextes : devant les noms animés (par défaut concrets) et devant les noms inanimés concrets. Si notre idée initiale était de discuter aussi des noms référant à des procès et à des entités abstraites, ceux-ci ont *in fine* - en raison du peu d'occurrences de ces noms dans notre corpus - été exclus de cette analyse de synthèse.

**Tableau 3 - Formes pré-nominales produites le plus fréquemment devant les noms animés et inanimés concrets (pour chaque enfant et par période)**

Enfant	Période	Statut sémantique du nom	Omissions	Fillers	Déterminants	Total (V.A.)
Lubin	1	Animés	77,11%	13,25%	9,64%	83
		Inanimés concrets	60,00%	26,09%	13,91%	116
	2	Animés	14,69%	16,08%	69,23%	145
		Inanimés concrets	8,60%	15,59%	75,81%	187
Naël	1	Animés	55,19%	27,27%	17,53%	164
		Inanimés concrets	44,71%	23,56%	31,73%	216
	2	Animés	3,70%	7,41%	88,89%	82
		Inanimés concrets	5,13%	2,56%	92,31%	161
Antonin	1	Animés	47,68%	41,35%	10,97%	237
		Inanimés concrets	47,78%	36,95%	15,27%	203
	2	Animés	14,29%	17,86%	67,86%	56
		Inanimés concrets	17,46%	11,24%	71,30%	341
Adrien	1	Animés	54,12%	35,29%	10,59%	85
		Inanimés concrets	42,33%	38,23%	19,44%	463
	2	Animés	3,57%	25,00%	71,43%	28
		Inanimés concrets	11,72%	28,91%	59,38%	140

Le tableau 3 ci-dessus présente donc pour chaque enfant les formes les plus souvent produites devant les animés et les inanimés concrets. Les pourcentages que nous avons faits figurer dans ce tableau sont calculés sur la part totale des noms produits par chaque enfant dans chaque contexte. En suivant les constats effectués au chapitre V et dans la première section de ce chapitre VI, nous avons considéré une première période équivalant, dans le corpus de Lubin et d'Antonin, aux quatre premières séances, et une seconde période, aux quatre suivantes. Dans le corpus de Naël, la première période correspond aux cinq premières sessions et la seconde rassemble les trois séances de 2;2 à 2;6. Dans le corpus d'Adrien, la période 1 compile les six premières sessions et la période 2, les deux dernières, à 3;3 et 3;5.

Ces résultats compilés en deux périodes nous permettent de confirmer certains des constats effectués plus tôt mais donnent également à voir des similitudes plus grandes entre les enfants que ce que nous pouvions observer séance par séance. Dans les données de Lubin, Naël et Antonin, les noms inanimés concrets semblent favoriser la production de déterminants, en période 1 comme en période 2. Au contraire, les omissions et/ou fillers semblent être plus fréquents devant les animés que les inanimés, et là aussi, durant les deux périodes. Dans le corpus d'Adrien, les déterminants sont aussi plus fréquents avec les inanimés qu'avec les animés en période 1. En revanche, en période 2, ils sont plus largement produits avec les animés. Nos données confirment donc en partie les résultats de Bassano (*i.a.* 2005) et Nashawati (2010) concernant la présence plus marquée de déterminants devant les noms renvoyant à des entités inanimées mais elles ne confirment pas (ou n'apparaissent pas en nombre suffisant pour le faire) la préférence pour les omissions - plus que pour les fillers - avec les animés et pour les fillers avec les abstraits. Ces derniers constats, effectués par Bassano, n'avaient pas non plus été confirmés dans la thèse de Nashawati.

Bien que les déterminants ne soient pas considérablement plus fréquents lorsqu'ils sont associés à des noms inanimés, nous pouvons tout de même nous interroger sur le rôle joué par la fréquence de ces noms dans le corpus. Chez les quatre enfants, les inanimés sont beaucoup plus nombreux que les animés. Outre la sensibilité de l'enfant à certaines caractéristiques sémantiques, on peut faire l'hypothèse que leur fréquence d'apparition puisse influencer la nature des formes pré-nominales les précédant. En d'autres termes, il nous semble que ces résultats seraient à interpréter en tenant compte aussi de la représentation plus large et plus précoce de la catégorie



des inanimés - et peut-être également des noms de cette catégorie pris séparément - dans les productions de l'enfant (et probablement dans celles de l'adulte). L'enfant étant plus familier des noms inanimés qu'animés, les opportunités qu'il y associe une forme pré-nominale adulte pourraient également être plus nombreuses.

La présentation de ces résultats est l'occasion de signaler également la difficulté que nous avons rencontrée lors du maniement de ces catégories sémantiques. Il nous a à plusieurs reprises semblé que si les catégories retenues pour confronter nos résultats à ceux de précédents travaux pouvaient être pertinentes pour l'analyse de productions adultes, elles étaient moins adaptées à des productions enfantines. La caractérisation sémantique du nom impliquant de pouvoir avoir accès à l'entité visée par ce lexème, celle-ci peut s'avérer complexe et discutable dans les premières productions des enfants. Nous avons par exemple codé tous les noms renvoyant à des animaux comme des noms de la catégorie des animés mais ceux-ci étant la plupart du temps absents de la situation et seulement représentés sur les supports manipulés par les participants (puzzle, cube, livre, ...), nous pouvons nous demander si dans certains contextes, l'enfant réfère à l'animal animé ou à son illustration inanimée.

Nous avons aussi eu des difficultés parfois à distinguer l'objet visé par l'enfant de l'action impliquée par l'objet et par conséquent, à faire un choix entre 'inanimé' et 'procès'. C'est le cas par exemple dans l'extrait (8) ci-dessous. Naël est sur sa chaise haute. Sa mère et sa tante sont présentes. Suite à quelques échanges entre les trois participants (*cf.* encadré 1) pour désambiguïser la production de l'enfant, la mère et la tante comprennent que Naël veut dessiner (et non descendre de sa chaise). Pendant cette séquence, nous pourrions ainsi analyser 'dessin' comme une action ('faire du dessin').

**Exemple 8 - Cas de nom pouvant référer à une entité inanimé et à un procès (corpus de Naël, 1;8)**

1  
 Naël : [vø desẽ] '*veux dessin.*' (sur sa chaise haute d'où confusion possible pour la mère de Naël entre 'dessin' et 'descendre')  
 Mère : tu veux descendre ?  
 Naël : [desa] '*dessin.*'  
 Tante : dessin.  
 Naël : [ø desa] '*veux dessin.*'  
 Mère : tu veux un dessin ?  
 Mère : ou tu veux descendre ? (insiste sur la dernière syllabe de <descendre>

Naël : [desa] 'dessin.'  
Mère : faire du dessin ?  
Tante : 0. (montre sa boîte de surligneurs à l'enfant)  
Naël : 0. (se tourne vers sa tante et tend la main vers ses surligneurs)  
Mère : ouais c'est ça.  
Naël : [X vø desẽ] 'xxx veux dessin.' (saisit boîte de feutres de sa tante)  
Mère : nan mais on va t(e) donner les tiens Naël ça évitera que +/. (se dirige vers l'endroit où sont rangés les feutres de l'enfant)  
Naël : [mwa ʒə vø desa] 'moi je veux dessin.'

Dans la suite de la séquence (cf. encadré 2), l'enfant signale à ses interlocuteurs qu'il veut dessiner avec les surligneurs de sa tante et non avec ses feutres à lui ([pa desa desa] 'pas dessin (.) dessin'). On peut ici s'interroger sur ce qui est visé par Naël au travers du lexème 'dessin'. Il est possible d'envisager que celui-ci renvoie cette fois aussi bien à l'action de dessiner (avec ses feutres ou avec les surligneurs de sa tante) qu'au matériel employé pour dessiner, et donc aux feutres et aux surligneurs. Cette interprétation impliquerait donc que contrairement au début de la séquence, on analyse ici 'dessin' comme un inanimé.

**2**

Mère : 0. (sort boîtes de feutres de l'enfant du tiroir)  
Naël : [nã pa desa] 'nan pas dessin !'  
Mère : dessin ? (montre boîte de feutres et cahier de l'enfant)  
Naël : [pa desa desa] 'pas dessin (.) dessin.' (montre la boîte de feutres de sa tante)  
Naël : [nã] 'nan.' (regarde vers sa mère qui tient ses feutres et le cahier)  
Mère : nan tu dessines avec tes feutres.  
Mère : pas avec les feutres de tatie.

D'autres exemples de ce type, mêlant objet et action ont pu être observés dans les données et peuvent, de même que les cas impliquant l'identification d'entités animées (et en particulier, d'animaux), remettre en question la pertinence de l'emploi de catégories sémantiques lorsque l'on s'intéresse à de très jeunes enfants. Les productions verbales de l'enfant n'étant pas conformes à la cible adulte, et le plus souvent constituées d'un ou deux terme(s) seulement, elles peuvent donner lieu à des interprétations ambiguës ou multiples, et ne sont donc pas toujours compatibles avec une classification sémantique stricte.

### 3. Impact du contexte lexical et syntaxique

Cette troisième et dernière partie du chapitre VI sera consacrée à la question de l'influence éventuelle du contexte lexical et syntaxique sur la production des formes devant les noms. Nous l'annonçons en introduction, les travaux de Pine et Lieven (*cf.* par exemple Pine et Lieven, 1997 ; Lieven *et al.*, 1997 ; Pine *et al.*, 2013), menés dans une approche basée sur l'usage, ont montré la façon dont se construit progressivement la catégorie des déterminants en anglais. Les formes adultes seraient, d'après les auteurs, d'abord construites au travers d'usages figés et dépendants des contextes lexico-syntaxiques (*cf.* par exemple cas de 'de l'eau' ou 'un autre' déjà évoqués plus tôt et récurrents dans nos données), avant que la catégorie soit pleinement acquise et que l'enfant accède à des usages généralisés des formes.

Deux questions principales découlent de la lecture de ces travaux : i) peut-on établir que dans notre corpus d'enfants francophones, les noms sont d'abord associés à des formes pré-nominales spécifiques (omissions, fillers ou déterminants), ou à des déterminants particuliers ? ii) de même, si l'on se focalise sur les cadres lexico-syntaxiques les plus fréquemment produits dans nos données, peut-on observer des combinaisons préférentielles entre cadres et formes pré-nominales ou de façon plus précise, entre cadres et déterminants ?

Nous passerons en revue l'impact du contexte lexical et syntaxique dans nos données au travers de quatre sections principales. Dans une première section (3.1), nous nous arrêterons brièvement sur la méthodologie qui nous a permis de mesurer le taux d'alternance ('overlap') d'une part et de caractériser les structures syntaxiques d'autre part. Au point 3.2, nous commenterons les résultats des analyses portant sur l'alternance des formes devant les noms et au point suivant (3.3), nous ferons un état des lieux des cadres lexico-syntaxiques les plus fréquemment produits dans notre corpus et des formes pré-nominales avec lesquelles ces constructions sont associées. En dernier lieu (point 3.4), nous tenterons de faire la synthèse de nos observations afin de définir si la catégorie des déterminants se construit, chez les quatre enfants de notre corpus, autour d'un continuum allant des usages figés aux usages généralisés ?

### 3.1 Mesure de l'overlap et caractérisation des structures syntaxiques

#### 3.1.1 Mesure de l'overlap

La mesure de l'alternance des formes devant les noms n'a pas fait l'objet d'un repérage préalable, lors de l'analyse des données, mais d'une analyse *ad hoc*. Deux éléments codés dans notre grille ont été nécessaires pour effectuer ces mesures : les noms produits par chaque enfant à chaque séance et les formes pré-nominales associées à chaque lexème. L'analyse s'est faite en plusieurs étapes :

- a. Nous avons initialement distingué les noms produits une seule fois par séance des noms produits à plusieurs reprises.
- b. Nous avons ensuite croisé chacun des noms produits à plusieurs reprises avec les formes pré-nominales (omission, filler et déterminant) précédentes afin de savoir si ces noms étaient systématiquement produits avec la même forme, ou si au contraire, ils étaient précédés de plusieurs formes différentes.
- c. Parmi les noms produits avec plusieurs formes différentes, deux types d'oppositions ont alors été définies :
  - les alternances entre deux formes seulement : omission et filler, omission et déterminant ou filler et déterminant ;
  - les alternances entre les trois formes observées : omission, filler et déterminant.
- d. Les noms produits plusieurs fois avec un déterminant au sein de la même séance ont ensuite fait l'objet d'une analyse spécifique, croisant noms et types de déterminants, afin d'observer si le syntagme nominal était toujours identique ou si l'enfant alternait le type de déterminant devant le nom.
- e. Une analyse semblable a été conduite sur les noms produits à plusieurs reprises dans tout le corpus (au sein de plusieurs séances distinctes) et précédés d'un déterminant, afin de mesurer si ceux-ci étaient préférentiellement associés à un type de déterminant ou si, toutes séances confondues, ceux-ci étaient combinés avec des déterminants distincts.
- f. En dernier lieu, nous nous sommes intéressée aux noms qui n'étaient jamais repris au sein de la même séance mais qui pouvaient par ailleurs être produits dans une autre séance, afin d'établir si ces derniers étaient d'une séance à l'autre accompagnés ou non de la même forme.

Tous les résultats de ces analyses seront présentés au point 3.2.

### 3.1.2 Caractérisation des structures syntaxiques

En parallèle, nous nous sommes intéressée aux cadres lexico-syntaxiques au sein desquels ont été produits les noms en position post-verbale (et ayant une fonction d'objet direct ou indirect, d'attribut et de circonstant participant à la prédication). Dans un premier temps, nous avons distingué les énoncés à noyau non verbal (ne contenant donc pas de verbe) des énoncés à noyau verbal (que celui-ci soit conjugué ou non), et dans un second temps, nous avons caractérisé les structures lexico-syntaxiques des énoncés verbaux. Pour cela, nous avons eu recours à une grille d'analyse élaborée par G. Fox dans le cadre du projet DIAREF, seule grille existant - à notre connaissance - pour l'acquisition du français. Dans le cas où la structure produite correspondait à l'un des 16 cadres identifiés dans cette grille comme les cadres les plus fréquents dans le langage de l'enfant et dans le langage qui lui est adressé<sup>1</sup>, nous l'avons explicitement codée sous la forme LEMME VERBAL+X (cf. liste présentée dans l'encadré 1).

#### Encadré 1 - Liste des cadres lexico-syntaxiques spécifiés

ALLER+X, AVOIR+X, CEST+X, DEVOIR+X, DIRE+X, DONNER+X, ETRE+X, FAIRE+X, ILYA+X, METTRE+X, POUVOIR+X, PRENDRE+X, REGARDER+X, SAVOIR+X, TIENS+X, VOULOIR+X

Précisons que nous avons inclus dans cette catégorie aussi bien des cadres conformes à la structure adulte que des structures incomplètes ou simplifiées mais qui pouvaient d'après le contexte être identifiées comme certains des cadres spécifiés que nous avons listés ci-dessus. Ainsi, des formes comme [a] (ex (Antonin, 3;2) : [me la a ply d plas] '*mais là (y) a plus d(e) place*') ont pu être analysées comme des cadres 'ILYA+X'. De même, [e] ou [te] (ex (Naël, 1;10) : [te a ɔti] '*c'est F rôti*') ont pu être interprétés comme des structures de type 'CEST+X' et des éléments comme [ø] (ex (Naël, 1;6) : [ø də lo] '*vieux de l'eau*'), comme des cadres 'VOULOIR+X'.

Lorsqu'en revanche, les cadres de production des noms ne correspondaient pas à l'un de ceux que nous avons préalablement listés, nous les avons regroupés de façon générique sous l'étiquette V(erbe)+X. Nous nous sommes néanmoins assurée qu'aucun cadre très fréquemment produit par l'enfant ne figure parmi les cadres non spécifiés.

<sup>1</sup> Cette analyse préalable avait été menée sur le corpus transversal de 13 enfants francophones de 2;2 à 2;4 utilisé dans le cadre de ce projet (cf. par exemple rapport final disponible sur le site de la recherche DIAREF : <http://www.univ-paris3.fr/anr-diaref-37421.kjsp>).

Si de par leur fonction syntaxique, une partie des noms produits par l'enfant ont dû être exclus de cette analyse, nous les avons toutefois repérés et regroupés en cinq ensembles que nous avons faits figurer dans le tableau 4 ci-dessous.

**Tableau 4 - Modalités supplémentaires pour le codage lexico-syntaxique**

<b>N.A.</b>	Le nom n'a pas une des fonctions prises en compte (pour les sujets et compléments du nom).
<b>CIRC</b>	Le nom est dans un groupe circonstant qui ne participe pas à la prédication.
<b>DG</b>	Le nom est en périphérie gauche de la proposition noyau (dislocation gauche).
<b>DD</b>	Le nom est en périphérie droite de la proposition noyau (dislocation droite).
<b>NINT</b>	Le cadre est non identifiable.

Les enfants de notre corpus produisant en période 1 (au stade des énoncés à un ou deux terme(s)) une majorité d'énoncés non verbaux, nous avons dans une étape ultérieure décidé de coder également la structure de l'énoncé immédiatement précédent de l'interlocuteur (lorsque cet énoncé contenait un verbe, cf. exemple 9a et b) ou très rarement, de l'enfant (lorsque l'énoncé précédent était produit par l'enfant et contenait un verbe, cf. exemple 9c). Nous avons en effet pu observer que souvent, le tour de parole précédent pouvait prédéterminer la production de l'enfant, particulièrement dans le cas des réponses de l'enfant aux questions de l'adulte. Il nous a donc paru pertinent de pouvoir également en tenir compte dans notre travail.

**Exemples 9 - Cas de nom pouvant référer à une entité inanimé et à un procès (corpus de Naël, 1;8)**

- (a) Mère : ben assieds toi mais tout à l'heure tu as donné ton gâteau au chien et maintenant c'est trop tard pour manger.  
Naël (1;9) : [gato ə f]ɛ̃] 'gâteau F chien.' (DONNER+X)
- (b) Mère : Tristan i(l) fait des puzzles de bébé alors.  
Naël (1;9) : [e pœlz ə bɛbɛ ] 'F puzzle F bébé ?' (FAIRE+X)
- (c) Lubin (2;5) : [X le poʃɔ̃ i ã na pa] 'xxx F cochon i(l) en pas ?'  
Père : comment ?  
Lubin : [ʁa i ã na pa dy mais] 'yyy i(l) en a pas du maïs ?'  
Lubin : [de gʁɛn] 'des graines.' (AVOIR+X)

Ce codage a toutefois été limité aux cas où l'énoncé précédent était en continuité formelle avec l'énoncé analysé. Par exemple, en 9a, nous avons considéré que la réponse de l'enfant était en continuité avec la demande de dénomination de l'adulte, et donc avec le cadre 'CEST+X'. A

l'inverse, en 9b, l'enfant fait une reprise partielle d'un énoncé qu'il a produit quelques tours de parole plus tôt ([mɛ o fuʁ] 'mets au four') et nous avons considéré ici que sa production n'était pas en continuité avec l'énoncé de l'adulte et par conséquent, pas associée au cadre V+X produit par le père. Les cas pour lesquels il n'a pas été possible de considérer avec certitude si il y avait continuité ou non ont été codés comme non analysables.

### Exemples 10 - Cas de continuité et non continuité entre deux énoncés (Lubin, 2;0)

- (a) Père : c'est quoi ça ?  
Lubin : [apɛ̃] '(l)apins.'
- (b) Père : allez prépare de la soupe .  
[ʃ fuʁ] 'F four' (enfourne la casserole)

## 3.2 Alternance des formes devant les noms

Les résultats de nos analyses concernant l'alternance des formes devant les noms seront présentées comme suit : nous nous arrêterons d'abord brièvement sur la distribution des noms à occurrences uniques et multiples au sein de chaque corpus (3.2.1) ; nous aborderons ensuite la question de l'alternance entre omission, filler et déterminant devant les noms produits plusieurs fois au sein de la même séance (3.2.2) ; dans une troisième partie, nous nous focaliserons spécifiquement sur l'alternance des déterminants (3.2.3), et enfin, dans une dernière section, sur l'alternance des formes toutes séances confondues (3.2.4).

### 3.2.1 Distribution des noms à occurrences uniques ou multiples

Nous l'annonçons plus tôt, la première étape du travail a consisté en un inventaire - pour chaque enfant - des différents noms produits à chaque séance et d'une sélection des noms pour lesquels plusieurs occurrences étaient repérables. Seuls les noms (ou les formes assimilées à des noms) pour lesquels un déterminant était attendu ont été ciblés pour cette analyse.

Dans le tableau qui suit nous avons fait figurer pour chaque enfant la part moyenne représentée par les noms qui ne sont produits qu'une fois, et celle représentée par ceux qui sont utilisés à deux reprises au moins. Cette répartition a été calculée en fonction du nombre de noms différents produits par l'enfant à chaque séance. A titre indicatif, nous avons aussi précisé dans ce tableau le nombre moyen de noms distincts qui ont pu être observés par enfant et par séance.

Ces chiffres ne représentent donc pas les résultats de toutes les séances compilées mais bien les moyennes de chacune des séances.

**Tableau 5 - Part moyenne des noms produits une fois et plusieurs fois par séance (chez les quatre enfants)**

Enfant	Part moyenne des noms produits une seule fois	Part moyenne des noms produits plusieurs fois	Nombre moyen de noms différents
Lubin	55,81%	44,19%	34
Naël	45,18%	54,82%	26
Antonin	50,41%	49,59%	32
Adrien	38,98%	61,02%	25

Le sous-ensemble sur lequel va reposer notre travail sur la dépendance au lexème est constitué de la moitié environ des noms présents dans les données du corpus de Lubin (44,19%), de Naël (54,82%) et d'Antonin (49,59%), et d'un peu plus de six noms sur dix en moyenne dans le corpus d'Adrien.

Les autres noms, qui seront donc exclus de notre analyse (mais conservés toutefois dans les graphiques 29 à 32 qui suivent) sont donc produits une seule fois, soit parce que l'enfant renvoie par ailleurs au référent avec un ou des pronom(s), soit parce qu'il le fait par le biais d'un autre nom ou d'un nom propre.

On observe de plus (*cf.* colonne 'nombre moyen de noms différents') que les quatre enfants produisent une trentaine de noms différents en moyenne par séance. Cette diversité lexicale est plus grande dans les corpus de Lubin et d'Antonin que dans ceux de Naël et Adrien. La variation individuelle est d'autant plus grande si l'on considère le fait que les chiffres sont calculés sur la base de séances de trente minutes dans les trois premiers corpus et de soixante minutes dans le dernier. Ce constat coïncide avec ce que nous avons pu observer en ce qui concerne l'allongement des énoncés et le développement des formes pré-nominales, un peu plus tardif dans le corpus d'Adrien.

### **3.2.2 Alternance entre omission, filler et déterminant**

Après avoir repéré les noms produits plusieurs fois par séance, nous avons cherché à observer si ceux-ci étaient précédés de formes distinctes (omission, filler ou déterminant) ou utilisés dans des constructions dont le type de forme ne varie pas (et donc toujours précédés d'une omission, d'un



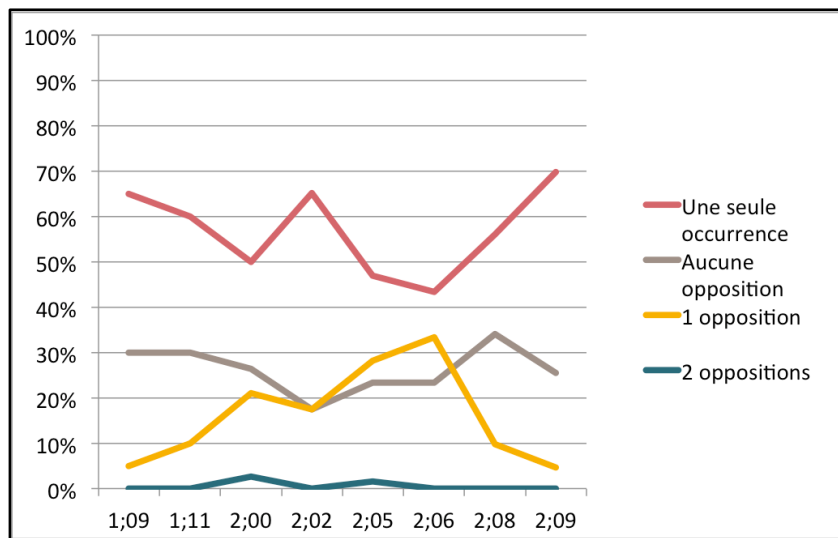
filler ou bien d'un déterminant). L'évolution de l'alternance (ainsi que de l'absence d'alternance) chez les quatre enfants est décrite au moyen des quatre graphiques (29 à 32) ci-dessous. La courbe rouge ('une seule occurrence') représente la part des noms qui ne sont pas repris et les trois autres courbes représentent les absences d'alternance ('aucune opposition'), les alternances entre deux types de formes ('1 opposition') et les alternances entre les trois formes observées ('2 oppositions'). Rappelons que les pourcentages ayant servi à créer ces graphiques sont calculés sur le total des noms produits à chaque séance et devant lesquels un déterminant pouvait être attendu.

En premier lieu, nous pouvons constater que la part des noms produits une seule fois ('une seule occurrence') suit une évolution - bien que non linéaire - assez comparable d'un enfant à l'autre : leur proportion tend à diminuer entre la première et la dernière séance de la première période puis à augmenter de nouveau durant la seconde période. Si la diminution de la part de ces noms peut être expliquée par plusieurs facteurs - parmi lesquels l'augmentation de la part des répétitions - leur augmentation durant la période 2 peut s'expliquer en très grande partie par le développement de la position pré-verbale et l'augmentation de la fréquence des pronoms d'une part, et d'autre part, par la présence de davantage d'énoncés en continuité référentielle, contenant proportionnellement moins de noms.

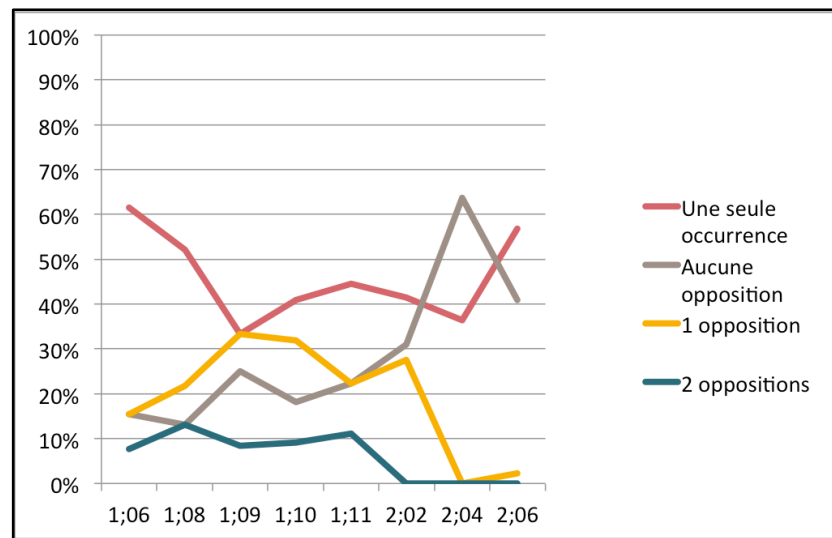
Les noms repris sans opposition du type de forme pré-nominale sont plus fréquents dans les dernières séances que dans les premières dans les corpus d'Antonin et de Naël. Les déterminants devenant à cette période la forme majoritaire, la plupart (ou la totalité) des noms qui sont produits à plusieurs reprises sont précédés de déterminants. Signalons à ce propos que la nette augmentation observée chez Naël à 2;4 est liée à une part plus importante d'auto-répétitions dans une séquence où la mère est occupée et ne participe que peu à l'activité de l'enfant. Dans le corpus d'Adrien, la proportion des noms produits sans opposition diminue avant d'augmenter dans les deux dernières séances, là aussi en raison d'une proportion plus importante de déterminants. La part des noms repris sans opposition est assez stable dans les données de Lubin (entre 20 et 30% des noms). Durant une première période, ces noms sont surtout précédés d'omissions ou de fillers et dans une seconde période, ils sont combinés à un même déterminant ou avec des déterminants différents.

**Graphiques 29 à 32 - Alternance entre omission, filler et déterminant chez les quatre enfants**

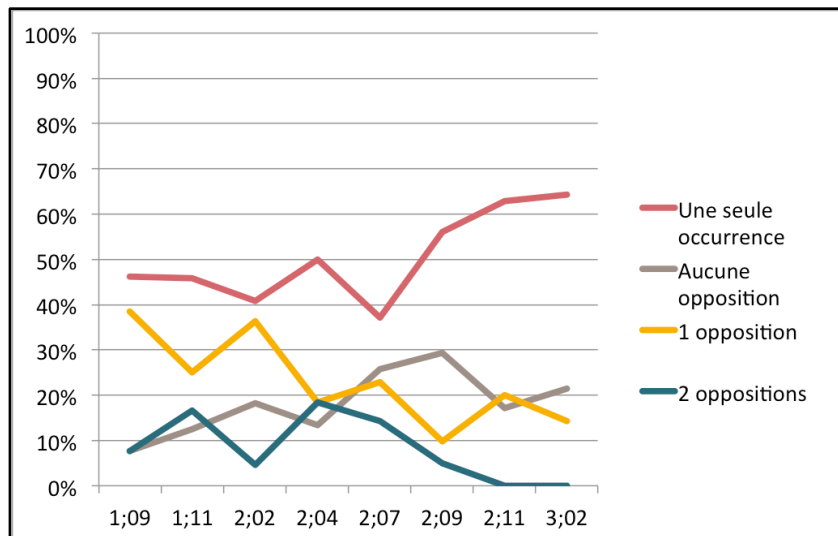
*Graphique 29 - Corpus de Lubin*



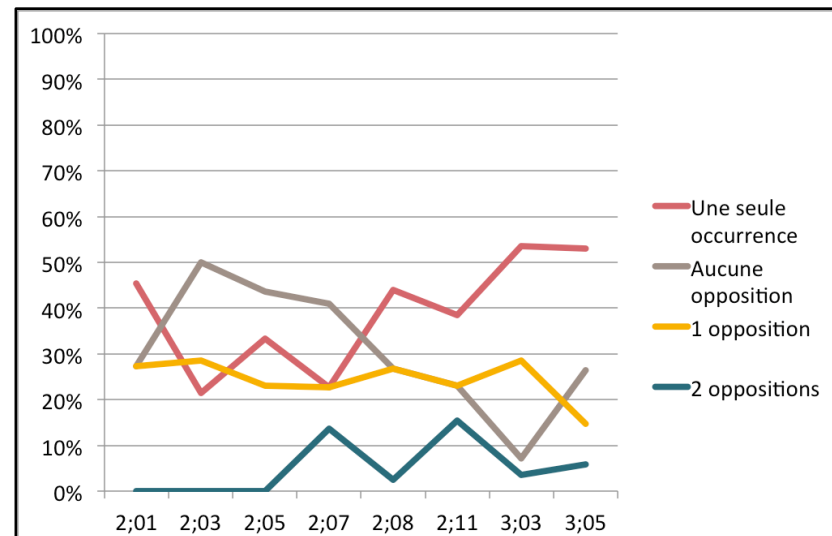
*Graphique 30 - Corpus de Naël*



*Graphique 31 - Corpus d'Antonin*



*Graphique 32 - Corpus d'Adrien*



Dans les quatre corpus, la proportion des oppositions de type 1 diminue nettement dans les premières séances. La fréquence d'apparition de ces oppositions lors de la séance 8 traduit la part que représentent encore les formes de transition dans les productions des enfants. Et en effet, on remarque qu'elles apparaissent plus souvent chez Adrien et Antonin que chez Lubin et Naël dans cette dernière séance.

On n'observe que quelques occurrences d'oppositions entre les trois types de formes chez Lubin (à 2;0 et 2;5). Elles sont plus fréquentes dans les données de Naël et Antonin et diminuent pour disparaître complètement au cours de la période 2. Dans le corpus d'Adrien, ce type d'opposition n'est visible qu'entre 2;7 et 3;5. On observe encore deux cas de variation entre omission, filler et déterminant dans cette dernière séance à 3;5. Tous les noms n'étant pas repris plus de deux fois, la part des cas d'opposition 2 est nécessairement moindre que la part des cas d'opposition 1, et par ailleurs, même lorsque les noms sont produits plus de deux fois au sein de la même séance, les enfants tendent à opposer deux formes et non trois.

Par ailleurs, dans les cas où seulement deux formes alternent, nous avons cherché à connaître la nature de ces formes. La part de chacun des contrastes identifiés (omission et filler, filler et déterminant, omission et déterminant) a été évaluée proportionnellement au total des alternances entre deux formes, pour chaque séance et chaque enfant. Les résultats sont présentés dans les quatre tableaux ci-dessous.

**Tableau 6 - Nature des alternances entre deux formes (1 opposition) dans le corpus de Lubin**

Age	Omission/Filler	Filler/Déterminant	Omission/Déterminant
1;9	100,00%	0,00%	0,00%
1;11	50,00%	50,00%	0,00%
2;0	0,00%	37,50%	62,50%
2;2	0,00%	25,00%	75,00%
2;5	27,78%	61,11%	11,11%
2;6	10,00%	70,00%	20,00%
2;8	0,00%	25,00%	75,00%
2;9	0,00%	50,00%	50,00%

Les trois types d'alternance peuvent être observées dans le corpus de Lubin. Toutefois, elles n'apparaissent pas dans les mêmes proportions. Durant les deux premières séances, c'est l'alternance entre omission et filler ainsi qu'entre filler et déterminant qui est privilégiée. L'alternance entre omission et déterminant est ensuite nettement majoritaire dans les deux

séances à 2;0 et 2;2. Durant les deux premières séances de la seconde période, toutes les alternances sont visibles mais c'est l'alternance entre filler et déterminant qui est la plus fréquente. Dans les deux dernières séances, les oppositions entre omission et filler disparaissent et le plus souvent, lorsqu'il y a alternance entre deux formes, c'est entre l'omission et le déterminant.

Ces résultats montrent donc qu'à l'exception des deux premières séances, Lubin n'a que peu recours à l'opposition entre omission et filler et qu'il privilégie les oppositions entre filler et déterminant et entre omission et déterminant. On peut faire l'hypothèse que dès lors que l'enfant commence à produire davantage de formes adultes, les formes de transition - omission et filler - alternent avec ces formes adultes et de fait, le filler n'apparaît plus comme une forme contrastant avec l'absence de forme.

**Tableau 7 - Nature des alternances entre deux formes (1 opposition) dans le corpus de Naël**

Age	Omission/Filler	Filler/Déterminant	Omission/Déterminant
1;6	100,00%	0,00%	0,00%
1;8	60,00%	40,00%	0,00%
1;9	37,50%	25,00%	37,50%
1;10	28,57%	28,57%	42,86%
1;11	33,33%	33,33%	33,33%
2;2	0,00%	37,50%	62,50%
2;4	---	---	---
2;6	0,00%	100,00%	0,00%

Dans le corpus de Naël, les oppositions constatées à 1;6 et 1;8 sont surtout des oppositions de type 'omission/filler'. Les contrastes entre omission et déterminant n'apparaissent pas du tout au cours de ces deux séances. Durant les trois séances qui suivent, la distribution entre les trois types de contrastes est assez comparable, avec une préférence toutefois (toutes séances confondues) pour l'alternance entre omission et déterminant. A partir de 2;2, les oppositions entre omission et filler disparaissent. Durant cette séance, les oppositions de type 'omission/déterminant' sont majoritaires. A 2;4, l'enfant ne produit que des déterminants donc aucune opposition n'est constatée. Enfin, à 2;6, les déterminants alternent seulement avec des fillers mais rappelons qu'il ne s'agit ici que de trois occurrences.

Outre le fait que le développement global des formes se reflète dans la nature des alternances observées, on peut signaler que Naël, comme Lubin, ne semble pas privilégier l'opposition entre

omission et filler après les deux premières séances. Il tend à l'inverse à produire plus souvent des noms précédés en alternance d'omissions et de déterminants. Cependant, les trois séances de 1;9 à 1;11 étant très semblables, ces remarques sont à apprécier avec prudence.

**Tableau 8 - Nature des alternances entre deux formes (1 opposition) dans le corpus d'Antonin**

Age	Omission/Filler	Filler/Déterminant	Omission/Déterminant
1;9	80,00%	20,00%	0,00%
1;11	100,00%	0,00%	0,00%
2;2	75,00%	25,00%	0,00%
2;4	63,64%	36,36%	0,00%
2;7	12,50%	12,50%	75,00%
2;9	0,00%	75,00%	25,00%
2;11	0,00%	42,86%	57,14%
3;2	0,00%	25,00%	75,00%

Dans le corpus d'Antonin, les deux périodes observées plus tôt se dessinent aussi très distinctement dans la nature des alternances entre les formes. De 1;9 à 2;4, avant que le déterminant apparaisse comme majoritaire devant les noms, l'opposition entre omission et filler est de loin le contraste le plus observé. Ensuite, de 2;7 à 3;2, les cas d'opposition entre filler et déterminant d'une part et omission et déterminant d'autre part sont les plus fréquents. Les deux étant tour à tour privilégiés, il semblerait que l'enfant produise, en alternance avec les déterminants, aussi fréquemment des omissions que des fillers.

**Tableau 9 - Nature des alternances entre deux formes (1 opposition) dans le corpus d'Adrien**

Age	Omission/Filler	Filler/Déterminant	Omission/Déterminant
2;1	66,67%	33,33%	0,00%
2;3	75,00%	25,00%	0,00%
2;5	55,56%	22,22%	22,22%
2;7	40,00%	0,00%	60,00%
2;8	45,45%	45,45%	9,09%
2;11	33,33%	66,67%	0,00%
3;3	12,50%	50,00%	37,50%
3;5	0,00%	100,00%	0,00%

Dans les données d'Adrien, on remarque qu'à l'exception de la séance à 2;7, lorsque deux types de formes alternent, il s'agit plus souvent d'une alternance entre omission et filler et filler et déterminant qu'entre omission et déterminant. Ces résultats suggèrent que l'enfant privilégierait

l'usage du filler pour contraster avec l'omission d'une part et avec le déterminant d'autre part, et utiliserait donc cette forme de transition comme une passerelle entre l'absence de forme et le déterminant. On peut relever par ailleurs que le contraste concerne proportionnellement plus souvent l'omission et le filler dans un premier temps (de 2;1 à 2;5), et à partir de 2;11, le contraste entre filler et déterminant devient majoritaire.

Si ces derniers graphiques et tableaux laissent à voir une part non négligeable d'alternance entre les formes pré-nominales (omission/filler dans un premier temps puis omission/déterminant ou filler/déterminant dans un second temps), l'observation qualitative des données nous a permis de constater que certains lexèmes sont toutefois combinés beaucoup plus fréquemment avec une des deux formes (ou trois dans les contextes de type 'opposition 2') qu'avec la deuxième (ou la troisième). Nous avons en effet considéré qu'il y avait alternance à partir du moment où deux occurrences (ou trois selon le type d'opposition) de formes pré-nominales différentes pouvaient être observées, mais aucun élément dans les chiffres présentés précédemment ne nous permet de savoir si l'enfant tend - même dans les cas d'alternance - à produire des associations préférentielles entre une forme pré-nominale et un nom spécifiques.

Nous avons donc cherché à définir, dans les cas d'alternance entre deux et trois formes, la part des noms produits avec une répartition équivalente des formes (autrement dit, les noms produits aussi fréquemment avec les deux ou trois formes) et ceux produits avec une répartition non-équivalente des formes (c'est-à-dire, produits davantage avec une - ou deux - forme(s) qu'avec l'autre). Une fois cette analyse effectuée, nous avons mesuré la fréquence de production de la forme majoritaire lorsque la répartition n'était pas équivalente. Ces résultats sont présentés dans les tableaux 10 et 11 ci-dessous. Pour chaque enfant, les chiffres proposés représentent la moyenne des résultats obtenus à chaque séance.

**Tableau 10 - Part moyenne des noms avec répartition équivalente et non-équivalente des formes pré-nominales dans les cas de 1 opposition et 2 oppositions (pour tous les enfants)**

Enfant	Répartition équivalente		Répartition non-équivalente	
	1 opposition	2 oppositions	1 opposition	2 oppositions
<b>Lubin</b>	43,00%	0,00%	57,00%	100,00%
<b>Naël</b>	31,79%	18,10%	68,29%	81,90%
<b>Antonin</b>	37,29%	1,52%	62,71%	98,48%
<b>Adrien</b>	37,54%	0,00%	62,46%	100,00%

Ces résultats nous montrent que lorsque les noms sont précédés de plusieurs types de formes, la plupart d'entre elles ne sont pas produites dans les mêmes proportions ('répartition non-équivalente'). En d'autres termes, lorsque l'on observe la présence de formes différentes devant un nom repris, cela ne signifie pas pour autant que l'une d'entre elles ne puisse pas être privilégiée, voire largement privilégiée, et cela qu'il s'agisse des cas de '1 opposition' ou de '2 oppositions'. Par exemple, à 1;8, Naël produit 17 fois le lexème 'lapin', 14 fois précédé d'une omission et 3 fois d'un filler. Deux types de formes - omission et filler - étant observées devant ce nom, il a été identifié comme un cas de type '1 opposition'. Toutefois, bien qu'un contraste existe effectivement entre deux formes, on relève que l'une d'entre elles, en l'occurrence l'omission, est nettement plus fréquente devant ce lexème.

**Tableau 11 - Fréquence de production de la forme majoritaire dans les cas de répartition non-équivalente des formes pré-nominales dans les contextes d'opposition 1 et 2 (pour tous les enfants)**

Enfant	1 opposition	2 oppositions
Lubin	70,50%	57,50%
Naël	74,18%	55,73%
Antonin	70,88%	53,54%
Adrien	74,83%	63,36%

Si l'on se concentre spécifiquement sur les contextes de type '1 opposition' dans lesquels la répartition n'est pas équivalente, on s'aperçoit que les résultats confirment ce que nous avons pu envisager lors du travail préalable sur les données : chez tous les enfants, la forme majoritaire apparaît en moyenne plus de deux fois sur trois. De même, si l'on se penche sur les résultats des cas avec deux oppositions, on voit que l'une des trois formes produites est majoritaire. En moyenne, lorsque la répartition entre omission, filler et déterminant n'est pas équivalente, la forme préférentielle est observée plus d'une fois sur deux, les deux autres formes occupant à elles deux la position pré-nominale dans un peu moins de 50% des cas.

Malgré la variation observée en position pré-nominale (omission / filler, filler / déterminant, omission / déterminant ou omission / filler / déterminant), ces chiffres tendent donc à montrer que le plus souvent, l'enfant n'emploie pas ces trois formes dans les mêmes proportions et que l'une d'entre elles est privilégiée. Dans la mesure où le recours à la même catégorie 'filler' ou 'déterminant' n'implique pas nécessairement que l'enfant produise des formes identiques (la forme phonologique des fillers peut être variable, de même que la nature des déterminants), ceci

ne signifie pas pour autant que les syntagmes nominaux soient les mêmes d'une mention à l'autre. Et quand bien même les fillers ou les déterminants seraient identiques, on ne peut pas conclure à un usage figé des syntagmes sans considérer par ailleurs l'influence du développement global des formes. La forme préférentielle sera de toute évidence le plus souvent une omission ou un filler dans une première période et un déterminant dans une période ultérieure.

### 3.2.3 Alternance entre différents types de déterminants

Après l'analyse de l'alternance entre déterminants et proto-formes, nous nous sommes spécifiquement concentrée sur l'alternance entre différents types de déterminants. Pour cela, nous nous sommes donc arrêtée sur les noms produits à plusieurs reprises au sein d'une même séance et précédés de plusieurs déterminants (identiques ou distincts).

**Tableau 12 - Part des noms repris avec plusieurs déterminants par séance et par enfant**

Séance	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8	Moyenne
<b>Lubin</b>	14,29%	12,50%	15,79%	0,00%	44,12%	58,82%	83,33%	92,31%	38,31%
<b>Naël</b>	40,00%	15,38%	23,53%	30,77%	26,67%	100,00%	100,00%	100,00%	54,54%
<b>Antonin</b>	0,00%	30,77%	7,69%	20,00%	59,09%	88,89%	69,23%	90,00%	45,71%
<b>Adrien</b>	0,00%	9,09%	3,85%	41,18%	21,74%	25,00%	61,54%	81,25%	30,46%

Si l'on observe le tableau 12 ci-dessus, on constate qu'en moyenne, sur les huit séances analysées, ces noms représentent 38,31% des noms repris chez Lubin, 54,54% chez Naël, 45,71% chez Antonin et enfin, 30,46% chez Adrien. Conformément à ce que nous pouvions attendre, la part des noms repris avec un déterminant augmente au fil des séances. S'ils représentent une part non négligeable des noms repris, rappelons néanmoins que leur proportion est calculée sur seulement la moitié environ des noms produits par chaque enfant (l'autre moitié correspondant aux occurrences uniques). Le nombre d'occurrences étant donc *in fine* limité, les résultats seront présentés dans les pages qui suivent en valeurs absolues.

Suite à cette sélection, les noms ont été analysés en distinguant ceux précédés de déterminants du même type (sans alternance) et ceux précédés de déterminants de types différents (avec alternance). Les résultats sont reproduits dans les tableaux ci-dessous pour chaque enfant.



Tableau 13 - Alternance des déterminants chez Lubin

Age	Noms produits plusieurs fois avec des déterminants	Noms sans alternance du type de déterminant	Noms avec alternance du type de déterminant
1;9	1	1	0
1;11	1	0	1
2;0	3	3	0
2;2	0	---	---
2;5	15	12	3
2;6	10	9	1
2;8	15	9	6
2;9	12	7	5

Dans le corpus de Lubin, si l'on s'intéresse aux cinq séances au cours desquelles plusieurs noms sont produits avec plusieurs déterminants (à 2;0 puis de 2;5 à 2;9), on remarque que jusqu'à 2;6, les noms sont surtout combinés à un seul type de déterminant, et plus spécifiquement au même déterminant. A 2;0, 2;6 et 2;8, les noms produits sans alternance du type de déterminant sont tous précédés d'une forme identique. Par exemple, à 2;6, les cinq occurrences du nom 'téléphone' sont précédés d'un défini et dans chaque cas il s'agit du défini 'le'. A 2;8, 'dessins animés' est produit 11 fois avec un défini et toutes les occurrences sont de type 'les dessins animés'. A 2;5, seul un nom sur les douze produits sans alternance du type de déterminant est précédé de deux indéfinis distincts ('une poire'/'des poires'), et à 2;9, deux noms sont produits avec des définis différents ('au goûter'/'le goûter', 'le canapé'/'les canapés').

Dans les deux dernières séances, les alternances sont plus nombreuses, mais impliquent le plus souvent seulement des définis et des indéfinis. Par ailleurs, comme dans le cas de l'alternance entre omissions, fillers et déterminants, on constate que le plus souvent, une forme est privilégiée sur les deux qui sont produites. A 2;8 par exemple, sur les 5 occurrences de 'tortue', la première est associée à un indéfini et les quatre autres mentions sont combinées à un défini (cf. exemple 11 ci-dessous).

#### Exemple 11 - Cas d'alternance avec une forme majoritaire (corpus de Lubin, 2;8)

Lubin : [se yn tɔty] 'c'est une tortue.'

[...]

Lubin : [ty kaf la tɔty menɑ̃] 'tu caches la tortue maintenant.'

[...]

Lubin : [nɑ̃ i fo mɛt la tɔty] 'nan i(l) faut met(tre) la tortue.'

[...]

Père : qu'est ce que j'ai caché ?

Lubin : [ə̃m la tɔty] 'hum la tortue !'

[...]

Lubin : [ty pø læ kaʃe la tɔty] 'tu peux le cacher la to(r)tue ?'

Si ces cas de non alternance peuvent s'expliquer par le fait que l'enfant tende à faire des associations préférentielles entre déterminants et noms, ils peuvent aussi être analysés à la lumière du statut pragmatique du référent porté par le nom. Dans l'exemple que nous venons de citer, la première mention de la tortue fait suite à une demande de dénomination du père, contexte dans lequel l'enfant privilégie l'indéfini. Suite à l'introduction de cet élément dans le discours (et dans l'activité) par le biais d'un syntagme indéfini, l'enfant comme l'adulte y font référence soit au travers d'un syntagme défini, soit par un pronom. Le contexte pragmatique pourrait donc autant que la dépendance au lexème justifier la production d'un déterminant identique.

**Tableau 14 - Alternance des déterminants chez Naël**

Age	Noms produits plusieurs fois avec des déterminants	Noms sans alternance du type de déterminant	Noms avec alternance du type de déterminant
1;6	2	1	1
1;8	2	2	0
1;9	4	3	1
1;10	4	1	3
1;11	4	4	0
2;2	17	8	9
2;4	14	6	8
2;6	18	13	4

Dans les données de Naël, si l'on se focalise sur les trois dernières séances, qui comptent le plus de noms repris avec plusieurs déterminants, on remarque que les noms avec et sans alternance sont produits dans des proportions comparables à 2;2 et 2;4, et contrairement à ce que nous avons pu voir chez Naël, la dernière séance est marquée par une prédominance des noms sans alternance du déterminant. Signalons par ailleurs que les 13 noms repris avec le même type de déterminant (défini, indéfini ou possessif) sont produits avec une forme identique. Toutefois, 10 noms sur 13 n'étant produits que deux fois, il paraît difficile de dire si ces associations sont récurrentes ou seulement liées au contexte, et d'autant plus si l'on considère que lorsqu'on les

confronte à des critères pragmatiques (entre autres, type de référence ou statut attentionnel du référent), les formes auxquelles l'enfant a recours devant les noms sont pertinentes.

Enfin, il nous semble que si dans les premières étapes du développement des déterminants, le nom peut influencer la forme produite, c'est probablement dans un stade antérieur à celui que nous avons pu observer chez Naël, dont les productions sont presque exclusivement constituées de déterminants en période 2. Le même constat peut d'ailleurs être fait au sujet des corpus de Lubin et Antonin, mais le nombre de noms produits avec plusieurs déterminants en période 1 étant très limité, nos observations ne peuvent être fondées que sur la période 2. De ces remarques découle donc un commentaire méthodologique important : si l'analyse de l'alternance/non alternance des déterminants produits devant les noms peut nous éclairer sur la façon dont l'enfant construit progressivement sa maîtrise des formes, elle doit être effectuée sur un corpus dense, contenant un nombre de déterminants bien plus conséquent que celui sur lequel nous nous appuyons ici.

**Tableau 15 - Alternance des déterminants chez Antonin**

Age	Noms produits plusieurs fois avec des déterminants	Noms sans alternance du type de déterminant	Noms avec alternance du type de déterminant
1;9	0	---	---
1;11	4	4	0
2;2	1	1	0
2;4	6	6	0
2;7	13	8	5
2;9	16	11	5
2;11	9	3	6
3;2	9	4	5

Durant la première période, de 1;11 à 2;4, les noms sont tous produits dans le corpus d'Antonin sans alternance ni du type de déterminant ni de la forme. Lorsqu'ils apparaissent à plusieurs reprises, les noms sont donc toujours produits avec le même déterminant. Néanmoins, il apparaît que dans ce corpus également, l'analyse des combinaisons 'déterminant + nom' gagnerait à être complétée par une analyse de leur contexte de production. Par exemple, à 2;4, 'biberon', 'cheval' et 'chat' sont systématiquement accompagnés du défini 'le' mais ces trois noms étant produits au sein d'une activité de lecture, ils renvoient à des référents déjà connus de l'enfant et de son interlocuteur et qui - s'ils ne sont pas repris par des pronoms - pourraient difficilement être

verbalisés avec d'autres types de syntagmes nominaux (et d'autant moins si l'on tient compte du fait qu'à cet âge, les indéfinis sont les seuls à contraster avec les définis).

Dans les deux séances qui suivent, les noms sans alternance restent majoritaires mais en parallèle, certains des noms repris le sont avec des déterminants différents (5 à chaque séance). Enfin, à 2;11 et 3;2, les noms avec alternance deviennent majoritaires. Comme chez Lubin et Naël, cette seconde période ne semble donc plus suffisamment adaptée pour notre analyse.

**Tableau 16 - Alternance des déterminants chez Adrien**

Age	Noms produits plusieurs fois avec des déterminants	Noms sans alternance du type de déterminant	Noms avec alternance du type de déterminant
2;1	0	---	---
2;3	1	1	0
2;5	1	1	0
2;7	7	5	2
2;8	5	5	0
2;11	2	1	1
3;3	8	7	1
3;5	13	7	6

Dans le corpus d'Adrien, lorsque les mêmes noms sont produits avec plusieurs déterminants, ils sont repris plus souvent sans alternance qu'avec une modification du type de déterminant produit. Les noms avec alternance n'apparaissent à part presque égale des noms produits avec le même type de forme que dans la dernière séance. Par ailleurs, il nous semble intéressant de relever que lorsque les noms sont produits sans alternance du type de déterminant, l'enfant produit des formes qui sont identiques. A titre d'exemple, les 13 occurrences du syntagme 'déterminant + pronom *autre*' relevées à 2;3 (et déjà discutées au chapitre V) sont toutes précédées de l'indéfini 'un' et les six occurrences de 'bouteille' produites à 2;6 sont associées au défini 'la'. Sur tout le corpus, seul un nom est produit avec plusieurs déterminants de la même catégorie mais avec des formes différentes.

Il nous semble que le corpus d'Adrien aurait été plus propice à l'analyse d'associations spécifiques entre déterminant et nom que les autres corpus, dans la mesure où les cas de non alternance restent majoritaires pendant une durée plus longue. Cependant, la question de la quantité de données se posant de nouveau ici, il aurait été préférable de pouvoir travailler sur des séances plus longues et plus rapprochées.

L'analyse des combinaisons de type 'déterminant + nom' montre donc que les cas de non alternance sont d'abord plus fréquents que les cas d'alternance, et que ceux-ci apparaissent ensuite dans des proportions comparables, voire plus souvent que les cas de non alternance. Toutefois, les occurrences que nous avons pu observer se situent surtout dans la période 2 du développement pré-nominal des quatre enfants, et nous avons pu voir qu'à ce stade, des facteurs pragmatiques peuvent aussi expliquer la présence de combinaisons préférentielles entre noms et déterminants spécifiques. Ces constats nous ont donc amenée à reconsidérer la pertinence de nos données pour une analyse de la dépendance au lexème. Il semblerait en effet que si nous avions disposé de davantage de données en période 1 - et en particulier de plus d'occurrences de déterminants avec les noms repris - cette période aurait été plus adaptée que la seconde pour ce type de travail.

### 3.2.4 Alternance des formes entre les séances

En dernier lieu, nous nous sommes intéressée aux noms qui sont produits plusieurs fois dans chaque corpus et qui ne sont pas précédés de formes pré-nominales faisant déjà l'objet d'une ou plusieurs alternance(s) à l'intérieur d'une même séance. Nous avons cherché à voir si ceux-ci étaient toujours produits avec les mêmes formes, quelle que soit la séance, ou bien si au contraire ils étaient associés à des formes pré-nominales différentes. Dans le tableau 17, nous avons considéré trois cas de figure : les cas où aucune opposition n'est constatée, les cas pour lesquels nous avons pu observer une opposition entre deux formes (omission et filler, omission et déterminant ou filler et déterminant), et les cas d'opposition entre omission, filler et déterminant. Les pourcentages sont calculés sur le total des noms concernés par cette analyse. Nous avons en outre fait figurer entre parenthèses les valeurs absolues.

**Tableau 17 - Alternance des formes pré-nominales devant les noms repris au sein de plusieurs séances (pour tous les enfants)**

Enfant	Aucune opposition	1 opposition	2 oppositions
Lubin	---	---	---
Naël	68,42% (13)	31,58% (6)	---
Antonin	40,00% (8)	50,00% (10)	10,00 % (2)
Adrien	35% (7)	50,00% (10)	15,00% (3)

Les résultats montrent en premier lieu qu'une fois encore, cette analyse gagnerait à être menée sur un corpus plus dense. Le corpus de Lubin en est une bonne illustration puisque aucun nom - qui n'est pas déjà repris par ailleurs avec plusieurs formes différentes au sein de la même séance - n'est produit dans plusieurs séances distinctes. Dans le corpus de Naël, nous en avons trouvé 19, dont 13 sont repris avec le même type de forme et 6 avec une opposition de type 1 (omission / filler, filler / déterminant, omission / déterminant). Dans le corpus d'Antonin, ces noms sont majoritairement repris avec une opposition entre deux formes (10 d'entre eux) et dans 40% des cas, sans contraste. Deux noms sont produits au sein de séances différentes avec les trois formes pré-nominales. Enfin, sur les 20 cas observés dans les données d'Adrien, 10 sont précédés de deux formes distinctes, 7 sont toujours accompagnés de la même forme et 3 sont associés à la fois à une omission, un filler et un déterminant.

Conformément à ce que nous pouvions attendre, l'alternance ou la non alternance des types de formes devant les noms s'explique surtout par la période à laquelle ils sont produits. Les noms repris au sein de séances se déroulant pendant la période 2 seront préférentiellement associés à des déterminants et donc produits sans opposition, de même que les noms précédés d'omissions ou de fillers en période 1. En revanche, les formes précédant les noms produits à la fois en période 1 et en période 2 font l'objet d'un plus grand nombre d'oppositions (de type 1 ou 2). Par exemple, dans le corpus d'Antonin, le nom 'route' est toujours précédé d'un déterminant, et ceci s'explique aisément par le fait qu'il n'apparaît que dans des séances de la période 2. En revanche, on observe à la fois des omissions et des déterminants devant 'ambulance' ou 'girafe', qui sont produits dans des séances à un an d'intervalle. À l'exception du syntagme 'de l'eau', aucun nom ne semble dans nos données entraîner la production d'un type de forme spécifique quel que soit l'âge de l'enfant.

Nous l'annonçons en introduction du point 3 de ce chapitre, suite à l'analyse du degré de dépendance entre formes pré-nominales et noms, nous nous sommes focalisée sur l'étude des cadres lexico-syntaxiques dans lesquels les syntagmes nominaux sont produits. Nous avons entre autres cherché à voir si les cadres les plus fréquents dans nos données semblaient entraîner la production d'un type spécifique de forme pré-nominale d'une part, et de déterminant d'autre part. Les résultats de ces analyses sont présentés et discutés au point suivant (3.3).

### 3.3 Cadres lexico-syntaxiques

Dans une première partie de ce point 3.3, nous ferons un état des lieux général sur les cadres les plus produits dans chaque corpus (dans les énoncés de l'enfant mais aussi dans les énoncés qui précèdent les énoncés non verbaux de l'enfant), et dans un second temps, nous nous intéresserons aux formes pré-nominales introduites par les cadres syntaxiques les plus fréquents.

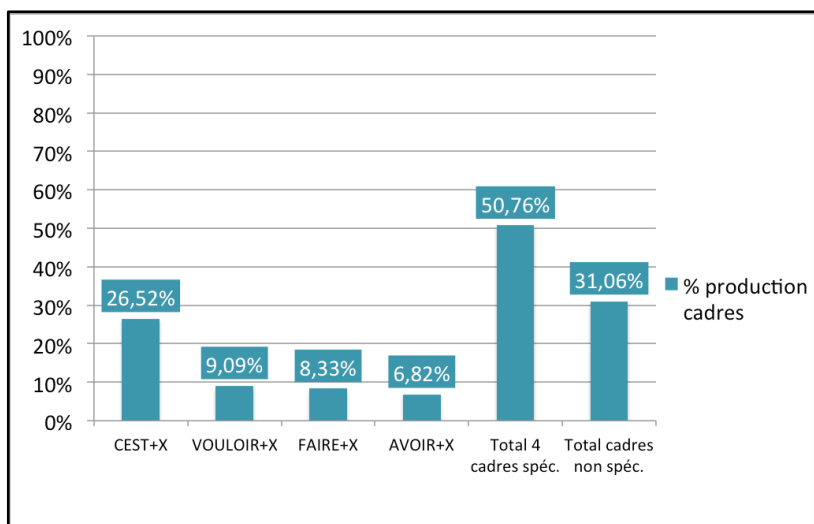
#### 3.3.1 Cadres lexico-syntaxiques les plus fréquemment produits

Sur la base de la liste présentée au point 3.1.2, nous avons identifié les cadres produits par les enfants au sein des énoncés verbaux contenant un nom (devant être précédé d'un déterminant), puis nous avons établi la liste des cadres qui ont été produits le plus souvent dans chaque corpus. Nous avons choisi dans les graphiques 33 à 36 ci-dessous de présenter les quatre cadres les plus fréquents dans les productions de chaque enfant. Les pourcentages mentionnés à l'intérieur des étiquettes sont calculés sur le total des cadres syntaxiques produits, sans prendre en compte les syntagmes nominaux en périphérie gauche ou droite de la structure syntaxique, ainsi que les noms avec des fonctions non analysées pour ce travail (n'ayant donc pas une fonction d'objet direct ou indirect, d'attribut ou de circonstant post-verbal participant à la prédication). Nous avons également fait figurer le total représenté par ces quatre cadres (colonne 'total 4 cadres spécifiés'), ainsi que la proportion représentée par les cadres non spécifiés (colonne 'total cadres non spécifiés').

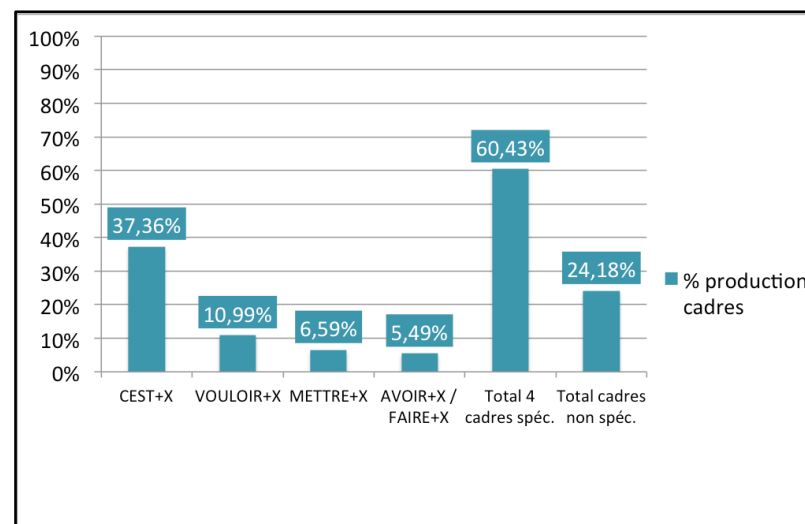
Ces quatre graphiques nous permettent en premier lieu de constater que lorsque les noms sont produits dans des énoncés verbaux et en position post-verbale, plus de la moitié d'entre eux sont associés à seulement quatre cadres différents (*cf.* colonne 'total 4 cadres spécifiés'). En comparaison, les verbes ne figurant pas dans la liste des cadres spécifiés ne représentent à eux tous que 28% en moyenne des contextes de production des noms chez les quatre enfants.

**Graphiques 33 à 36 - Fréquence de production des cinq cadres syntaxiques les plus produits chez les quatre enfants (toutes séances confondues)**

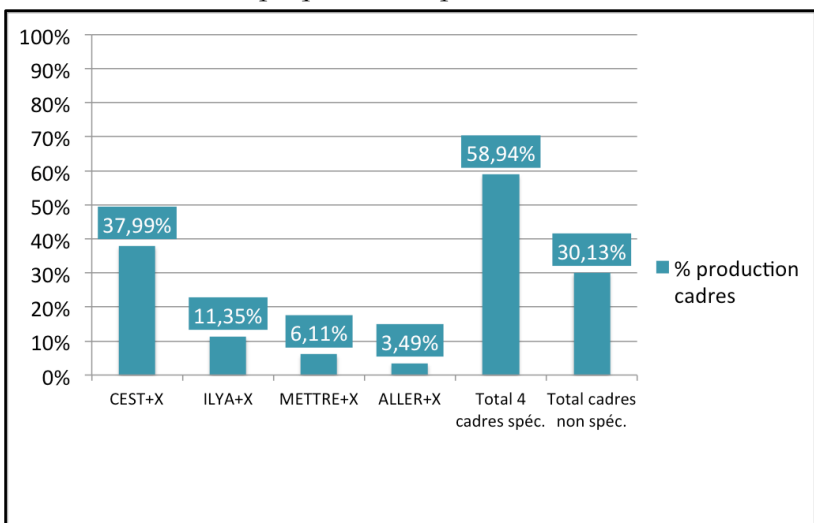
*Graphique 33 - Corpus de Lubin*



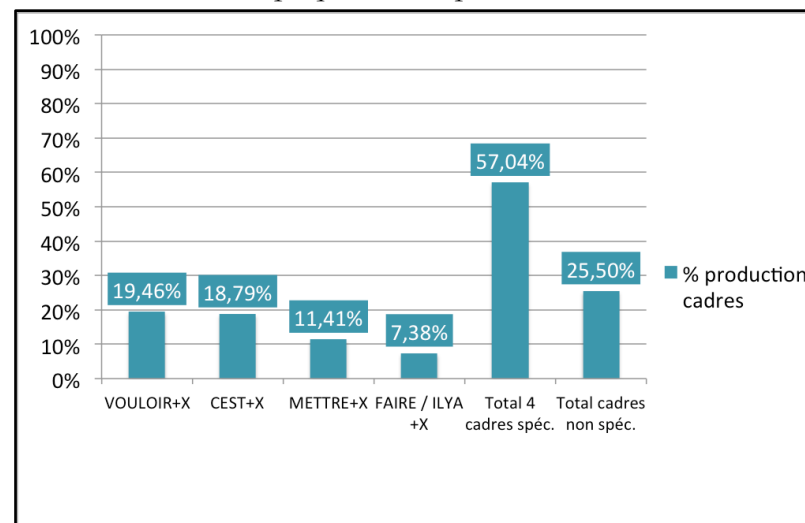
*Graphique 34 - Corpus de Naël*



*Graphique 35 - Corpus d'Antonin*



*Graphique 36 - Corpus d'Adrien*





Si les quatre cadres qui apparaissent comme le plus fréquemment produits ne sont pas identiques dans tous les corpus, ils sont néanmoins largement comparables. Le cadre 'CEST+X' est de beaucoup le plus souvent observé dans les données de Lubin, Naël et Antonin. Il est le deuxième cadre produit chez Adrien, dans des proportions identiques au cadre 'VOULOIR+X' (19,46% pour ce dernier contre 18,79% pour 'CEST+X'). Notons toutefois que lorsque l'on observe toutes les séances séparément, la part représentée par cette structure diminue à mesure que celle des dénominations et plus globalement, des activités générant un grand nombre de dénominations, diminue.

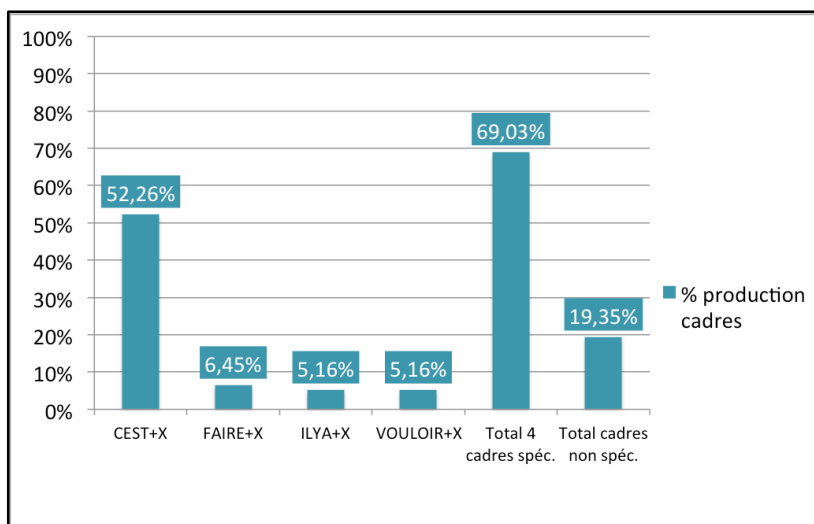
Outre la nette prédominance du présentatif dans nos données, on remarque que d'autres structures sont visibles dans plusieurs corpus. 'VOULOIR+X' apparaît par exemple aussi bien chez Adrien que chez Lubin et Naël, et 'FAIRE+X' dans les données de Lubin, Naël et Adrien.

Enfin, bien que les dislocations ne fassent pas partie de l'analyse systématique que nous avons menée ici, nous avons pu relever que celles-ci sont particulièrement présentes dans le corpus d'Antonin, tant dans le langage de l'enfant que de la mère. Les dislocations en périphérie droite et gauche représentent dans ce corpus 12,75% des contextes syntaxiques de production des noms. Celles-ci seront discutées ultérieurement, en parallèle de la présentation des combinaisons entre cadres lexico-syntaxiques et formes pré-nominales (*cf.* point 3.2.2).

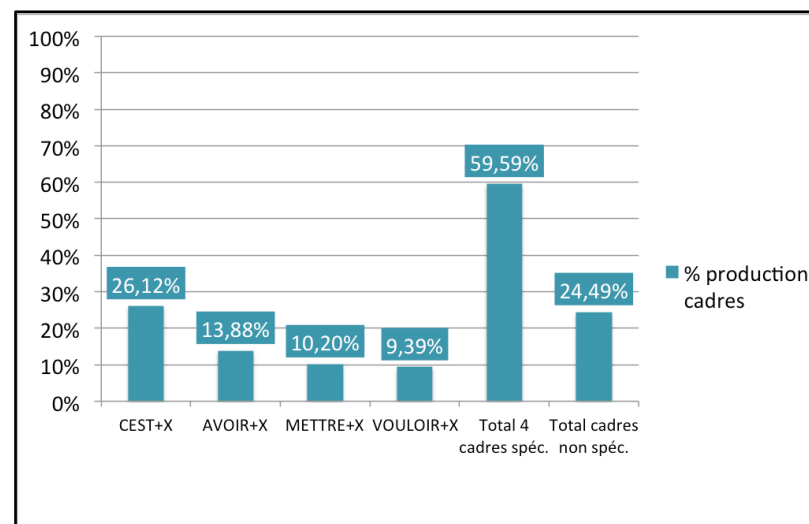
Les mêmes analyses ont été menées pour les énoncés précédant les énoncés non verbaux des enfants. Nous avons reporté dans les quatre graphiques ci-dessous (graphiques 37 à 40), comme nous l'avons fait plus haut, les quatre cadres les plus fréquents dans ces énoncés, majoritairement produits par les interlocuteurs de l'enfant, et en particulier, par ses parents. Rappelons que le codage de ces énoncés nous permettra ultérieurement d'associer les formes pré-nominales contenues dans les énoncés non verbaux des enfants aux cadres syntaxiques des énoncés immédiatement antérieurs.

**Graphiques 37 à 40 - Fréquence de production des cinq cadres syntaxiques les plus produits dans les énoncés précédant les énoncés non verbaux des quatre enfants (toutes séances confondues)**

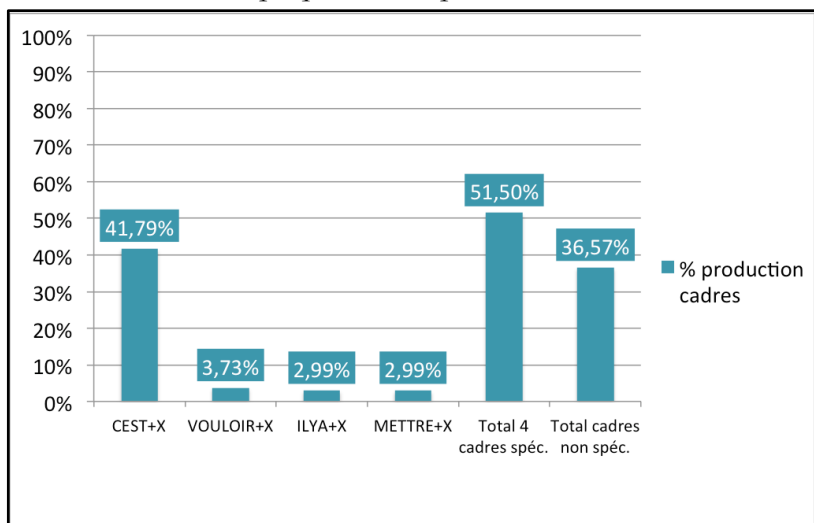
*Graphique 37 - Corpus de Lubin*



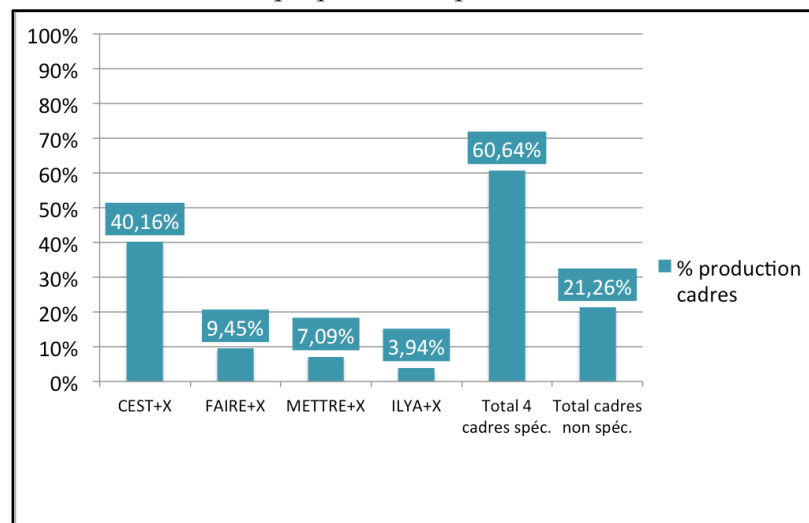
*Graphique 38 - Corpus de Naël*



*Graphique 39 - Corpus d'Antonin*



*Graphique 40 - Corpus d'Adrien*



Lorsque les noms sont produits dans le cadre d'énoncés non verbaux, les énoncés qui les précèdent (produits par l'interlocuteur pour une très large majorité d'entre eux) sont pour beaucoup les mêmes que ceux produits par l'enfant (*cf.* exemples 12 ci-dessous).

**Exemples 12 - Cadres apparaissant parmi les plus fréquents dans les productions des enfants et de leurs interlocuteurs (exemples issus des quatre corpus)**

1. Exemples de 'CEST+X' dans le corpus de Lubin

- (a) Mère : oh mais qu'est-ce que c'est qu(e) ça ?  
Lubin (2;9) : [a se de baʁbi] 'ab c'est des Barbie.'
- (b) Lubin (2;9) : [se la kloʃ] 'c'est la cloche !'
- (c) Mère : c'est un casse tête [/] un casse tête chinois. (*séance à 2;9*)

2. Exemples de 'AVOIR+X' dans le corpus de Naël

- (a) Naël (2;4) : [e il a ẽ bebe] 'et il a un bébé.'
- (b) Naël (2;4) : [i zõ pa d pãtalõ] 'i(l)s ont pas d(e) pantalon.'
- (c) Père : mon bidou t'as des p(e)tites larmes. (*séance à 1;10*)
- (d) Père : t'as le clown ou le papillon. (*séance à 2;6*)

3. Exemples de 'ILYA+X' dans le corpus d'Antonin

- (a) Antonin (3;2) : [ja l kamjõ si tu vø] 'y a l(e) camion si tu veux.'
- (b) Antonin (3;2) : [e j a de zø] 'et y a des jeux ?'
- (c) Mère : tiens r(e)garde y a un accident (*séance à 2;9*)
- (d) Mère : là regarde j(e) pensais qu'il y avait quelque chose et y a rien pour descendre. (*séance à 2;9*)

4. Exemples de 'FAIRE+X' dans le corpus d'Adrien

- (a) Adrien (3;3) : [a fe a lin kə sa] 'a fait F ligne comme ça.'
- (b) Adrien (3;5) : [mwa ə fe əa tesa] 'moi F fais F dessin.'
- (c) Père : qu'est ce qu'elle fait la voiture ? (*séance à 3;5*)
- (d) Mère : tu fais des ronds ? (*séance à 3;5*)

Le cadre 'CEST+X' est très nettement prédominant dans les données de Naël, Antonin et Adrien (45% des contextes de production en moyenne dans ces trois corpus). C'est également le plus fréquent dans le corpus de Naël mais dans une moindre mesure puisqu'il ne représente que 26,12% des cadres syntaxiques observés. Les trois autres cadres sont aussi pour beaucoup les mêmes que ceux qui dominent dans les énoncés verbaux des enfants. Dans le corpus de Lubin, deux d'entre eux sont présents à la fois dans les énoncés de l'enfant et dans les énoncés qui précèdent ('VOULOIR+X' et 'FAIRE+X'). On retrouve trois cadres communs dans le corpus de Naël ('VOULOIR+X', 'METTRE+X' et 'AVOIR+X'), deux dans celui d'Antonin ('METTRE+X' et 'ILYA+X'), et trois dans les données d'Adrien ('METTRE+X', 'FAIRE+X' et 'ILYA+X').

Ces résultats tendent donc à montrer d'une part que les structures les plus fréquentes dans les productions de l'interlocuteur sont reprises par l'enfant lorsque ses énoncés sont verbaux. D'autre part, si l'on s'intéresse aux proportions dans lesquelles les cadres syntaxiques sont produits - et en particulier à la part représentée par le cadre 'CEST+X' - on s'aperçoit qu'elles peuvent également expliquer les proportions dans lesquelles les cadres sont produits chez les enfants.

Bien que notre propos ne soit pas axé dans ce travail sur des questions de développement syntaxique chez l'enfant, on peut toutefois relever que nos données confirment ce qui a pu être avancé par Cameron-Faulkner *et al.* (2003) en anglais (et sur un corpus autrement plus important) au sujet du caractère relativement peu diversifié du langage adressé à l'enfant et de la corrélation évidente entre les premiers énoncés de l'enfant et ceux produits par ses interlocuteurs.

### 3.3.2 Combinaisons entre cadres lexico-syntaxiques et formes pré-nominales

#### 3.3.2.1 Combinaisons entre cadres lexico-syntaxiques et omissions/fillers/déterminants

Nous avons par la suite cherché à observer si les cadres lexico-syntaxiques repérés dans notre corpus entraînaient la production de formes pré-nominales spécifiques ou privilégiées. Pour ce travail, nous avons considéré à la fois les cadres présents dans les productions des quatre enfants et ceux présents dans les énoncés précédant les énoncés non verbaux des enfants. Au sein de cet ensemble, nous avons sélectionné les quatre cadres lexico-syntaxiques les plus fréquents, en excluant évidemment la catégorie 'V+X' regroupant une multiplicité de cadres non spécifiés. Une fois ces cadres présélectionnés, nous les avons confrontés aux catégories de formes pré-nominales auxquelles ils sont associés dans nos données. Les résultats de ces croisements sont présentés dans les quatre tableaux ci-dessous (tableaux 18 à 21). Afin entre autres de faciliter la lecture, nous n'avons pas fait figurer les huit séances de chaque enfant mais seulement les deux périodes (P1 et P2) préalablement repérées. Les quatre cadres retenus pour l'analyse apparaissent dans l'ordre correspondant à leur fréquence de production - toutes séances confondues - dans chacun des corpus (en commençant par le plus fréquent).

Le premier constat que nous pouvons faire à la lecture du tableau 18 (concernant les données de Lubin) - et que nous pourrions d'ailleurs faire avec les quatre enfants - est celui de la symétrie entre les formes produites avec les cadres lexico-syntaxiques préférentiels et le développement

plus général de la position pré-nominale. Durant la première période, les omissions sont majoritaires et *a contrario*, les déterminants sont la forme globalement privilégiée en période 2.

**Tableau 18 - Fréquence de production des formes pré-nominales avec les cadres lexico-syntaxiques les plus fréquents dans le corpus de Lubin (en pourcentages et par période)**

Période	Cadres	Omission	Filler	Déterminant	Total (V.A.)
P1	CEST+X	82,61%	13,04%	4,35%	46
	FAIRE+X	100,00%	---	---	2
	VOULOIR+X	90,91%	9,09%	---	11
	ILYA+X	75,00%	25,00%	---	4
P2	CEST+X	15,71%	17,14%	67,14%	70
	FAIRE+X	0,00%	5,26%	94,74%	19
	VOULOIR+X	---	---	100,00%	9
	ILYA+X	20,00%	30,00%	50,00%	10

Au-delà de ce constat, on note que les omissions, fillers et déterminants ne sont pas produits dans les mêmes proportions avec les quatre cadres sélectionnés. Durant la première période, on constate par exemple que les seuls déterminants produits le sont avec ‘CEST’. Rappelons cependant que les cadres en ‘CEST+X’ étant beaucoup plus fréquents que les trois autres cadres (seuls deux noms apparaissent avec ‘FAIRE’ en P1 et 4 avec ‘ILYA’), les opportunités de voir produites des formes plus hétérogènes, et en particulier des formes adultes, est également plus grande. Dans la période 2, nous pouvons signaler que le cadre ‘VOULOIR+X’ donne lieu à une production exclusive de déterminants. Ceux-ci sont aussi très nettement privilégiés après ‘FAIRE’, mais moins observés en revanche en présence de ‘CEST’ et ‘ILYA’. On pourrait faire l’hypothèse (à confirmer avec davantage de données) que ces deux cadres tendent à favoriser la production d’omissions et de fillers, à une période à laquelle ces formes de transition ne représentent plus qu’environ une forme sur cinq devant les noms.

Dans le corpus de Naël (tableau 19 ci-dessous), les omissions sont moins fréquentes dès la période 1 et tous les cadres sélectionnés entraînent la production des trois types de formes. On note néanmoins que comme dans le corpus de Lubin, elles ne sont pas systématiquement produites dans les mêmes proportions avec tous les cadres. ‘AVOIR’ semble être un contexte plus favorable à l’omission dans la première période et au contraire, les déterminants sont davantage privilégiés avec ‘VOULOIR’ qu’avec les autres cadres. Les fillers sont quant à eux produits plus fréquemment à la suite de ‘METTRE’ et ‘CEST’.

**Tableau 19 - Fréquence de production des formes pré-nominales avec les cadres lexico-syntaxiques les plus fréquents dans le corpus de Naël (en pourcentages et par période)**

Période	Cadres	Omission	Filler	Déterminant	Total (V.A.)
P1	CEST+X	47,76%	25,37%	26,87%	67
	AVOIR+X	71,43%	14,29%	14,29%	7
	VOULOIR+X	43,33%	16,67%	40,00%	30
	METTRE+X	50,00%	33,33%	16,67%	6
P2	CEST+X	9,68%	---	90,32%	31
	AVOIR+X	3,13%	6,25%	90,63%	32
	VOULOIR+X	---	---	100,00%	3
	METTRE+X	4,00%	---	96,00%	25

Au cours de la période 2, les trois occurrences de noms produits avec ‘VOULOIR’ et ‘METTRE’ n’apparaît plus qu’une fois avec une omission, dans le cadre de l’énoncé [tit kuvɛtyʁ] (*(pe)tite couve(r)ture*) produit par l’enfant à 2;2, suite à la question ‘tu lui mets une petite couverture?’ de l’observatrice’. ‘CEST’ est aussi largement suivi de syntagmes nominaux contenant un déterminant (90,32%). Les omissions restent plus représentées qu’avec les autres cadres mais les 9,68% affichés dans le tableau ne représentent que 3 occurrences sur les 31 occurrences de noms produits après ‘CEST’ à cette période (deux omissions suivies du lexème ‘chevalier’ et une devant ‘muffin au chocolat’). De la même façon, ‘AVOIR+X’ semble être un contexte privilégiant plus que les autres la production de fillers mais les 6,25% que représentent ces derniers en période 2 se limitent à deux occurrences, particulièrement proches phonologiquement des déterminants qui pourraient être attendus : [yne kuʁɔ̃n] pour ‘une couronne’ et [t apaxɛj foto] pour ‘P/ton appareil photo’.

Notre première remarque à propos du corpus d’Antonin (tableau 20) concerne le résultat pour le moins surprenant obtenu pour le cadre ‘ALLER+X’ dans la période 1. Il ne s’agit en fait que d’une seule occurrence de défini associé au lexème ‘eau’. Dès lors que le cadre ‘CEST+X’ apparaît comme le cadre largement prédominant dans les productions de la mère et de l’enfant (39,50% des cadres avec lesquels des noms sont produits), les trois autres cadres - et d’autant plus le cadre ‘ALLER+X’ qui n’apparaît qu’en quatrième position des cadres les plus fréquents - sont très peu représentés durant la première période. En période 2, ‘ALLER’ est de nouveau suivi exclusivement de déterminants mais ces 100% affichés dans le tableau ne représentent que 9 occurrences en tout.

**Tableau 20 - Fréquence de production des formes pré-nominales avec les cadres lexico-syntactiques les plus fréquents dans le corpus d'Antonin (en pourcentages et par période)**

Période	Cadres	Omission	Filler	Déterminant	Total (V.A.)
P1	CEST+X	73,91%	19,57%	6,52%	92
	ILYA+X	50,00%	50,00%	---	2
	METTRE+X	75,00%	25,00%	---	4
	ALLER+X	---	---	100,00%	1
P2	CEST+X	15,69%	11,76%	72,55%	51
	ILYA+X	22,22%	14,81%	62,96%	27
	METTRE+X	7,14%	14,29%	78,57%	14
	ALLER+X	---	---	100,00%	9

De même, on ne trouve que deux occurrences de noms produits avec 'ILYA' en période 1 (produit avec une omission pour l'un et avec un filler pour l'autre) et quatre syntagmes seulement produits après 'METTRE', dont trois de type 'omission + nom' et un de type 'filler + nom'. En ce qui concerne le cadre 'CEST+X', on remarque qu'en période 1 il entraîne une production majoritaire d'omissions, presque 20% de fillers et quelques déterminants, mais de par le trop faible nombre d'occurrences de noms produits avec les autres cadres, nous ne pourrions pas dire ici si cette distribution est propre à la structure présentative 'CEST+X' ou si elle ne fait que refléter la distribution attendue pendant cette période.

Durant la période 2, outre l'emploi exclusif du déterminant avec le cadre 'ALLER+X' auquel nous faisons référence plus tôt, on peut signaler que les cadres 'CEST+X' et 'ILYA+X' génèrent une production comparable de déterminants, fillers et omissions. En comparaison, 'METTRE' est plus fréquemment combiné à des syntagmes contenant des déterminants, et parallèlement, moins souvent à des noms précédés d'omissions.

Enfin, nous avons signalé plus tôt que parmi les contextes syntaxiques non retenus pour l'analyse, la dislocation possédait un statut particulièrement privilégié dans les données d'Antonin. En croisant les noms produits dans le cadre de ces dislocations et les formes pré-nominales les précédant, nous avons pu remarquer que celles-ci donnaient lieu à la production de bien plus de formes (fillers et déterminants) dans la période 1 que ce que nous avons pu constater lorsque les noms sont précédés de 'CEST'. Les omissions ne représentent que 37,70% des formes produites, la part des fillers atteint 42,62% et celle des déterminants, 19,67%. En période 2, les mêmes constats peuvent être effectués puisque les déterminants sont produits

devant 82,93% des noms, les fillers devant seulement 7,32% des noms et les omissions ne représentent plus que 9,76% des formes produites devant les noms disloqués. Notons à ce sujet que seules les dislocations en périphérie droite entraînent des omissions en période 2 (ex : [il e mal kamjõ] ‘il est mal Ø camion’ produit à 3;2). On pourrait faire l’hypothèse ici que le référent ayant déjà été introduit par ailleurs par le biais d’un pronom, ce contexte puisse être plus favorable à l’omission du déterminant.

**Tableau 21 - Fréquence de production des formes pré-nominales avec les cadres lexico-syntaxiques les plus fréquents dans le corpus d’Adrien (en pourcentages et par période)**

Période	Cadres	Omission	Filler	Déterminant	Total (V.A.)
P1	CEST+X	65,96%	27,66%	6,38%	47
	VOULOIR+X	81,25%	12,50%	6,25%	16
	METTRE+X	22,73%	63,64%	13,64%	22
	FAIRE+X	41,67%	33,33%	25,00%	12
P2	CEST+X	15,63%	43,75%	40,63%	32
	VOULOIR+X	---	35,29%	64,71%	17
	METTRE+X	---	50,00%	50,00%	4
	FAIRE+X	---	18,18%	81,82%	11

Dans le corpus d’Adrien, ‘METTRE’ et ‘FAIRE’ introduisent une proportion plus grande de syntagmes nominaux avec des déterminants que ‘CEST’ et ‘VOULOIR’ en période 1. Les fillers sont quant à eux largement majoritaires avec ‘METTRE’ (63,64%) et moins produits dans les autres contextes, et les omissions représentent une part plus importante des productions lorsque les syntagmes nominaux sont précédés de ‘CEST’ (65,96%) et ‘VOULOIR’ (81,25%).

Durant la période 2, les déterminants sont majoritaires avec ‘VOULOIR’ et ‘FAIRE’. Ils sont employés dans des proportions identiques aux fillers avec ‘METTRE’ mais relevons qu’il ne s’agit là que de quatre occurrences au total. Il semble de plus que ‘CEST’ génère une production plus importante de formes de transition - omissions et fillers - que les autres cadres. Il est d’ailleurs intéressant de constater que les quelques omissions encore présentes dans ce corpus à 3;3 et 3;5 sont une fois sur trois précédées de ‘CEST’, dans sa version adulte ou sous la forme d’un proto-présentatif [e].

Nous pouvons conclure de ces résultats que pour tous les enfants, l’évolution globale de la position pré-nominale se reflète dans les formes qui sont combinées avec chaque cadre lexico-



syntactique. Outre ce constat attendu, nous avons pu remarquer que certaines formes sont davantage privilégiées avec certains cadres mais à l'exception de quelques rares cas de figure (pouvant être liés à un nombre d'occurrences limité), nous n'avons pas observé de combinaisons exclusives entre des contextes lexico-syntactiques et des formes pré-nominales spécifiques en période 1, marquée contrairement à la période 2 par l'hétérogénéité des formes produites devant les noms. Si l'on s'intéresse aux cadres qui ont été considérés comme préférentiels dans plusieurs corpus, on n'observe pas nécessairement de correspondances d'un enfant à l'autre entre les formes privilégiées pendant les périodes 1 et 2. Ceci pourrait signifier d'une part que les stratégies des quatre enfants ne sont pas identiques et d'autre part, que le cadre lexico-syntactique ne peut pas à lui seul prédéterminer le type de syntagme produit par l'enfant. On relève toutefois quelques similitudes d'un corpus à l'autre : les fillers tendent par exemple à être favorisés avec 'METTRE' ou 'CEST' durant la première période et plus globalement, le cadre 'CEST+X' semble entraîner des productions plus diversifiées que les autres cadres mais il se peut ici que ce phénomène s'explique plus simplement par sa fréquence d'apparition, et donc par une distribution comparable à celle que nous observons plus généralement devant les noms. Toutefois, ce constat d'une plus grande hétérogénéité des formes produites après 'CEST' restant valable en période 2 (marquée chez tous les enfants par une production majoritaire de formes adultes), on peut considérer que ce cadre lexico-syntactique semble favorable à la production de formes de transition.

### *3.3.2.2 Combinaisons entre cadres lexico-syntactiques et types de déterminants*

Dans une étape ultérieure, nous avons croisé de façon spécifique les cadres retenus pour notre analyse et les déterminants produits au sein des syntagmes nominaux introduits par ces structures syntaxiques. Nous avons d'une part cherché à savoir si ces cadres impliquaient la production d'un type de déterminant privilégié, et d'autre part, s'ils pouvaient générer l'emploi d'une forme en particulier.

Les résultats sont présentés pour chaque enfant et toutes séances confondues dans les tableaux 22 à 25. Dans la colonne 'Type de déterminant', nous avons précisé, outre le type de déterminant privilégié, la part représentée par les déterminants de ce type sur la totalité des syntagmes produits avec un déterminant et pour chaque cadre. Dans la colonne suivante, nous avons listé

les formes produites à chaque séance (et appartenant à la catégorie mentionnée dans la colonne précédente), ainsi que le nombre d'occurrences de chaque forme.

**Tableau 22 - Combinaisons entre cadres lexico-syntaxiques et déterminants dans le corpus de Lubin**

Cadre lexico-syntaxique	Type de déterminant privilégié	Formes produites
CEST+X	Indéfini (44,90%)	une (*10), un (*7), des (*5)
FAIRE+X	Indéfini (88,89%)	des (*15), un (*1)
VOULOIR+X	Défini (100,00%)	les (*5), le (*3), la (*1)
ILYA+X	Défini (60,00%)	le (*2), les (1)

Dans le corpus de Lubin, l'indéfini est majoritaire avec 'CEST' (dans des combinaisons impliquant le plus souvent un indéfini singulier 'une' ou 'un'), et avec 'FAIRE' (presque exclusivement dans des combinaisons de type 'FAIRE + 'des' + nom). Au contraire, 'VOULOIR' et 'ILYA' sont davantage suivis du défini. Il semble donc *a priori* que ces cadres soient préférentiellement associés à des déterminants spécifiques.

Plusieurs autres constats s'imposent cependant et viennent modérer ces premières observations. Dans un premier temps, si les types de déterminants produits après 'CEST', 'FAIRE' et 'ILYA' sont les plus fréquents, ils ne sont pas les seuls à être produits dans ces contextes. On peut notamment relever que 'CEST' n'est employé que la moitié du temps avec un indéfini.

Dans un deuxième temps, nous nous devons de mettre en perspective ces premiers constats avec le nombre d'occurrences concernées. Des données supplémentaires devraient en particulier être observées pour pouvoir appréhender les résultats concernant 'ILYA+X' et 'VOULOIR+X' avec plus de certitude.

En dernier lieu, nous avons pu remarquer qu'une association privilégiée entre un cadre syntaxique et un type de déterminant pouvait s'expliquer non seulement par l'influence du cadre syntaxique mais aussi - et probablement surtout - par les usages que font les participants de ces cadres. Par exemple, sur les 22 occurrences de syntagmes indéfinis produits avec 'CEST', 19 d'entre eux le sont dans un contexte de dénomination, dans lequel la structure 'CEST + syntagme indéfini' est largement privilégiée. Si l'on regarde de plus près les indéfinis produits cette fois avec 'FAIRE', on s'aperçoit que sur les 16 occurrences relevées dans le corpus, 13

d'entre elles correspondent à des contextes de référence indéterminée, privilégiant là aussi largement l'emploi de l'indéfini dans ce corpus. Si le type de déterminant produit peut être déterminé par le cadre syntaxique, il semble donc qu'il puisse l'être aussi par le contexte pragmatique dans lequel les combinaisons 'cadre + déterminant + nom' sont produits. Ces questions seront spécifiquement abordées au chapitre VII.

**Tableau 23 - Combinaisons entre cadres lexico-syntaxiques et déterminants dans le corpus de Naël**

Cadre lexico-syntaxique	Type de déterminant privilégié	Formes produites
CEST+X	Défini (36,96%)	la (*9), le (*3), les (*3), l' (*1), du (*1)
AVOIR+X	Indéfini (70,00%)	un (*12), des (*5), une (*4)
VOULOIR+X	Possessif (40,00%)	sa (*2), mon (*2), ma (*1), mes (*1)
METTRE+X	Indéfini (40,00%)	des (*11), une (*1)

Comme chez Lubin, 'AVOIR' est aussi le plus souvent suivi d'un indéfini dans le corpus de Naël (et en particulier, d'un indéfini masculin singulier 'un'). Au contraire, 'CEST' est cette fois associé préférentiellement au défini, et le plus souvent au défini féminin singulier 'la'. 'VOULOIR' entraîne une production plus marquée de possessifs (40,00%) et 'METTRE' est le plus fréquemment associé à un indéfini (40,00% également).

On note que dans ce corpus comme dans le précédent, les déterminants privilégiés ne sont pas les seuls à être observés dans les contextes sélectionnés. Par exemple, si 'CEST' est dans 36,96% des cas suivi du défini, on le trouve également dans 32,61% des cas, et donc à part quasiment égale, suivi de l'indéfini.

Par ailleurs, la remarque que nous avons pu faire quant à la nécessité de la prise en compte du contexte interactionnel est valable aussi dans ce corpus. Par exemple, en nous focalisant sur les 11 occurrences de 'des' relevées après 'METTRE', nous avons vu d'une part que deux d'entre elles (*cf.* exemple 13 ci-dessous), issues de la séance à 2;6, sont produites dans un contexte de référence indéterminée, qui implique le plus souvent la production d'un indéfini, quel que soit le cadre syntaxique introduisant le syntagme.

**Exemples 13 - Indéfinis ‘des’ produits après ‘METTRE’ dans le corpus de Naël à 2;6**

- (a) [i mɛt de kadju dã l dã l panje] ‘i(ls) mettent des cailloux dans l(e) [ / ] dans l(e) panier +...’  
 (b) [ba i zõ mi de piɛk dødã e i i zõ le e i zõ kase le panje] ‘bab i(l)s ont mis des pierres dedans et i(ls) [ / ] i(l)s ont [ / ] les [ / ] et i(l)s ont cassé les paniers.’

Les neuf autres occurrences concernent seulement deux syntagmes comparables produits à 2;6 ([de bɔt də fij] ‘des bottes de fille’ et [de ʃosyɛ də fij] ‘des chaussures de fille’), que l’enfant reprend sur plusieurs tours de parole sans modifier la structure de l’énoncé. L’emploi *a priori* préférentiel de ‘des’ avec METTRE doit donc dans ces données être envisagé au regard de l’influence du contexte pragmatique et du phénomène de répétition. Enfin, si cette question a été posée ici au travers de l’exemple du cadre ‘METTRE+X’, elle aurait aussi pu être soulevée pour ‘CEST+X’ ou ‘AVOIR+X’.

**Tableau 24 - Combinaisons entre cadres lexico-syntaxiques et déterminants dans le corpus d’Antonin**

Cadre lexico-syntaxique	Type de déterminant privilégié	Formes produites
CEST+X	Défini (44,19%)	la (*8), le (*6), l’ (*3), les (*2)
ILYA+X	Partitif (50,00%)	de (*9)
METTRE+X	Défini (90,91%)	la (*4), l’ (*3), le (*2), les (*1)
ALLER+X	Défini (100,00%)	la (*3), l’ (*3), le (*2), au (*2)

Dans les données d’Antonin, le défini est la forme la plus fréquente après ‘CEST’, ‘METTRE’ et ‘ALLER’. Nous avons pu noter que ‘CEST’ est aussi souvent associé à l’indéfini (dans 37,21% des cas). En revanche, ‘ALLER’ et ‘METTRE’ sont exclusivement ou presque exclusivement suivis du défini. Signalons cependant que dans les deux cas, l’emploi du défini peut aussi s’expliquer par le fait que les référents associés aux noms sont pour la plupart particuliers et déjà donnés dans le discours, et pouvant donc être, du point de vue pragmatique, préférentiellement associés à des définis (*cf.* exemples 14).

**Exemples 14 - Définis produits après ‘ALLER’ et ‘METTRE’ dans le corpus d’Antonin**

- (a) [lɔt ɔsi va is ba o gaaʒ ɔsi] ‘l’aut(re) aussi va ic(i) [ / / / ] va au ga(r)age aussi.’ (2;7)  
 (b) [e va lasɛsœk] ‘F va l’ascenseur.’ (2;9)  
 (c) [la pɛtit ɛ va paɛ la ɛut] ‘la petite e(lle) va par la route.’ (2;9)  
 (d) [e lwi la mi bosɪ le bonɔn] ‘et F F a mis yyy les bonhommes.’ (2;7)

Enfin, 'ILYA' est associé dans 50,00% des cas à un partitif, ce qui peut sembler surprenant dans la mesure où ce type de déterminant n'est pas très fréquemment observé dans les données. Et ici encore, la réponse est à chercher dans l'activité en cours lorsque 'ILYA' est produit avec un déterminant, puisque 7 des 9 occurrences de 'de' sont associées au lexème 'place' dans des énoncés de type [a ply d plas] ('(y) a plus d(e) place') produits lorsque l'enfant joue avec ses voitures et cherche à les garer.

**Tableau 25 - Combinaisons entre cadres lexico-syntaxiques et déterminants dans le corpus d'Adrien**

Cadre lexico-syntaxique	Type de déterminant privilégié	Formes produites
CEST+X	Indéfini (50,00%)	des (*4), une (*3), un (*1)
VOULOIR+X	Indéfini (50,00%)	des (*4), un (*2)
METTRE+X	Défini (60,00%)	la (*2), l' (*1)
FAIRE+X	Indéfini (75,00%)	des (*8), une (*1)

Dans le corpus d'Adrien, on peut constater que sur les trois ou quatre types de déterminants produits avec chaque cadre (défini, indéfini, possessif et numéral), l'indéfini est privilégié dans tous les contextes, sauf avec 'METTRE', qui est davantage utilisé avec des définis. Ces constats concordent avec ce que nous avons pu voir avec 'CEST' dans le corpus de Lubin ainsi qu'avec 'METTRE' dans le corpus d'Antonin. Le défini étant par ailleurs beaucoup plus fréquent dans le corpus d'Adrien que l'indéfini, on peut faire l'hypothèse d'une tendance chez l'enfant à préférer certaines formes de déterminants avec certains cadres. Toutefois, les occurrences de déterminants dans des énoncés verbaux ou introduits par des énoncés verbaux de l'interlocuteur sont très peu nombreuses - et cela même si toutes les séances sont confondues - donc il est de toute évidence difficile de tirer ici des conclusions très affirmées sur la question des combinaisons entre cadres et déterminants.

De plus, bien que 'FAIRE' entraîne semble-t-il une production privilégiée de la forme 'des', on remarque en observant ces occurrences que six d'entre elles (listées ci-dessous, cf. exemples 15) le sont dans le cadre de la même séance et de la même activité, et au sein d'énoncés partiellement ou totalement auto-repris. Dès lors que 'FAIRE' n'introduit par ailleurs qu'un seul autre syntagme nominal de type 'des + nom' (syntagme [de gut] 'des gouttes' produit à deux reprises à

2;8), il semble une fois de plus difficile d'attribuer une influence spécifique au cadre syntaxique, plus qu'à l'activité, au contexte pragmatique ou au phénomène des reprises.

### Exemples 15 - Indéfinis 'des' produits après 'FAIRE dans le corpus d'Adrien à 2;6

*Adrien est en train de dessiner et décrit à MOT ce qu'il a dessiné ou ce qu'il est en train de dessiner.*

- (a) [fe de ba] '*fais des b(r)as !*'
- (b) [nɔ̃ ve fa de gɔ̃n] '*non vais fai(re) des gants.*'
- (c) [a vy a fe de gɔ̃n] '*as vu yyy fait des gants.*'
- (d) [a vy a fe de gɔ̃n] '*as vu yyy fait des gants.*'
- (e) [a fe de zø] '*yyy fait des yeux.*'
- (f) [a vy ta fe de putit] '*as vu t'as fait des petites+/.*'

En ce qui concerne l'influence du cadre lexico-syntaxique sur la production des déterminants, nous pouvons dire d'une part que les cadres les plus fréquemment produits par les quatre enfants entraînent une production privilégiée de certains types de déterminants. Toutefois, à l'exception de 'ALLER+X' dans le corpus d'Antonin, aucun des cadres sélectionnés n'entraîne la production exclusive d'un type de déterminant (ni d'une forme spécifique). D'autre part, nos données montrent que d'autres facteurs, liés à l'activité, au contexte pragmatique, ou plus généralement, à l'interaction, peuvent également expliquer l'usage que les enfants font des déterminants avec les cadres lexico-syntaxiques retenus pour l'analyse. Si ces constats ne signifient pas pour autant que le facteur syntaxique est à exclure pour comprendre les productions de déterminants dans nos données, ils invitent néanmoins à ne pas en faire un facteur exclusif mais bien à l'inclure au sein d'un faisceau plus large de facteurs.

### 3.4 La construction de la catégorie 'déterminant' : des usages figés aux usages généralisés ?

Pour clore cette section concernant l'impact du contexte lexical et syntaxique sur les formes pré-nominales produites, nous pouvons avancer d'une part que les usages des syntagmes 'forme pré-nominale + nom' semblent dans nos données être relativement peu figés. En nous intéressant aux noms repris et aux combinaisons de type 'omission/filler/déterminant + nom', nous avons pu relever que les noms ne sont pas exclusivement associés à un type de forme spécifique mais aussi bien à des formes de transition qu'à des formes adultes. Signalons toutefois qu'ils sont combinés plus souvent à l'une de ces trois formes qu'aux deux autres et que les préférences observées dépendent de la période à laquelle les syntagmes sont produits. L'analyse des

combinaisons ‘déterminant + nom’ montre que si la part des cas de non alternance est plus importante que celle des cas d’alternance, c’est aussi en grande partie en raison du statut pragmatique des référents associés aux noms. En outre, nos observations révèlent que l’analyse d’éventuelles associations préférentielles entre noms et déterminants aurait été plus pertinente en période 1, si la proportion des noms repris avec plusieurs déterminants à ce stade nous avait permis de le faire.

Par ailleurs, l’analyse des combinaisons entre cadres lexico-syntaxiques et syntagmes nominaux indique que certaines formes (omission, filler ou déterminant) sont particulièrement - mais pas exclusivement - produites dans certains contextes syntaxiques. Les mêmes constats ont pu être faits au sujet des combinaisons entre cadres et déterminants et nous avons également remarqué que cette fois encore, des facteurs interactionnels pouvaient au même titre que les structures syntaxiques expliquer la production dominante de certains déterminants. L’influence des cadres lexico-syntaxiques n’apparaît donc pas dans notre corpus comme ayant une portée majeure et exclusive sur la nature des formes pré-nominales produites, ce qui suggère que l’acquisition du paradigme des déterminants ne se ferait pas uniquement au travers de combinaisons récurrentes entre cadres et formes pré-nominales. Ces conclusions restent cependant à mettre en perspective avec la quantité de données examinées en période 1.

Nos résultats ne montrent donc pas une préférence claire pour des usages figés et dépendants des contextes lexicaux et syntaxiques, contrairement à ce qui avait pu être observé par Pine et Lieven (1997). Toutefois, nos données et analyses ne sont peut-être pas suffisamment comparables pour que nos résultats puissent être mis en perspective avec ceux de ces auteurs. D’une part, notre étude porte sur une quantité moins importante de données, en particulier dans la période précédant celle où le déterminant est la forme majoritaire devant le nom. D’autre part, nous avons axé notre travail sur l’usage des déterminants mais aussi des formes de transition. Ce choix implique nécessairement une variation en position pré-nominale plus importante que si nous avions seulement observé les contextes dans lesquels des déterminants sont produits. Par ailleurs, la langue que nous avons étudiée n’est pas la même. Les différences morphologiques (masculin/féminin, définis amalgamés ou non, ...) étant plus nombreuses en français qu’en anglais, nous pouvions également nous attendre à ce que les noms soient ici précédés de formes adultes plus variées.

Enfin, il nous paraît important de relever que si les usages ne semblent pas particulièrement figés, nos résultats ne témoignent pas pour autant d'emplois d'emblée généralisés de formes adultes et ne confirment donc en aucun cas les conclusions de Valian *et al.* (1986, 2009). Que l'on s'intéresse aux noms ou aux structures syntaxiques, l'alternance observée implique d'une part des formes de transition et d'autre part, des déterminants dont la nature n'est encore que peu diversifiée en fin de seconde période.



Ce chapitre nous a permis de passer en revue trois des hypothèses avancées dans la littérature pour expliquer la variation observée en position pré-nominale ainsi que la nature des formes produites. Nous avons ainsi soumis nos données à plusieurs interrogations d'ordre phono-prosodique, sémantique et lexico-syntaxique. Nous avons entre autres pu constater que lorsque les productions pré-nominales des enfants sont envisagées en termes de présence ou d'absence de forme, une majorité d'entre elles s'explique par une préférence pour la structure bisyllabique. Nous avons par ailleurs pu relever que les noms renvoyant à des entités inanimées semblent favoriser la production de déterminants et que les noms animés semblent entraîner une production plus fréquente d'omissions et/ou de fillers. Nos données montrent également que lorsque les noms sont produits à plusieurs reprises, ils ne sont pas spécifiquement associés à un type de forme pré-nominale et peuvent être précédés d'omissions, de fillers et de déterminants. Cependant, ils sont la plupart du temps préférentiellement combinés à l'une de ces formes. Lorsque l'on se penche en particulier sur les noms repris avec plusieurs déterminants, on remarque qu'ils apparaissent plus fréquemment associés au même déterminant qu'à des déterminants différents. En outre, en nous intéressant aux cadres lexico-syntaxiques de production des syntagmes nominaux, nous avons pu voir que ceux-ci entraînent la production privilégiée - mais loin d'être exclusive - de certains types de formes pré-nominales ou de déterminants spécifiques.

Toutefois, si ces facteurs peuvent tous - à des degrés différents - expliquer une part du processus d'acquisition des déterminants, la proportion des cas qui ne vont pas dans le sens des hypothèses que nous avons testées nous permet de penser qu'aucun d'entre eux n'explique à lui seul



pourquoi l'enfant peut, au sein de la même séance, produire aussi bien des formes adultes que des omissions, pourquoi il produit un filler plus qu'un déterminant, ou encore pourquoi il produit un type de déterminant plus qu'un autre. D'une part, nous pouvons faire l'hypothèse que le développement des formes pré-nominales gagnerait à être analysé en considérant les facteurs phono-prosodique, sémantique et lexico-syntaxique de façon croisée, et non plus seulement séparément comme nous l'avons fait dans ce chapitre. C'est donc ce que nous tenterons de faire dans le cadre du chapitre VIII. D'autre part, à la suite de certains constats que nous avons pu faire au sein des chapitres V et VI, concernant notamment le statut pragmatique du référent associé au nom ou du cadre interactionnel dans lequel celui-ci est produit, nous pouvons envisager que d'autres facteurs, liés à l'interaction, pourraient venir enrichir l'éclairage apporté par les facteurs distributionnels. Ce sera là l'enjeu de notre prochain chapitre, qui s'intéressera spécifiquement à l'influence des facteurs fonctionnels sur la mise en place des déterminants.



## CHAPITRE VII

*Influence des facteurs fonctionnels sur la mise  
en place des déterminants*

---

Au chapitre précédent, chacune des formes observées en position pré-nominale a été analysée comme un élément constitutif du syntagme produit par l'enfant, ou plus largement, de son énoncé - mais rarement comme un élément inclus dans le cadre plus étendu du discours que l'enfant construit pour et avec l'autre. Or, si l'on considère que l'enfant apprend à parler pour pouvoir interagir avec son environnement (notamment, Bruner, 1987 ; Vygotski, 1997 [1936]) et par conséquent, que l'acquisition de la langue ne se résume pas à l'acquisition des formes mais bien aussi, à celle des combinaisons entre les formes et les fonctions portées par ces formes, il est envisageable que les productions pré-nominales des enfants soient influencées non seulement par des facteurs distributionnels mais aussi par des facteurs fonctionnels tels que les contextes pragmatico-discursifs de production des noms. En nous appuyant sur différents travaux ayant mis en évidence la dimension fonctionnelle de l'acquisition des déterminants (notamment, Gundel *et al.*, 2007 ; Rozendaal, 2008 ; Salazar Orvig *et al.*, 2005, 2013), nous tenterons de compléter les observations que nous avons pu faire au chapitre précédent en évaluant également l'impact des caractéristiques pragmatico-discursives des référents associés aux noms sur les éléments produits en position pré-nominale.

Par ailleurs, dans la mesure où chaque production de l'enfant fait suite à un contexte discursif antérieur et constitue elle-même un fragment du contexte préalable pour les productions qui suivent, nous pouvons nous interroger sur l'influence des mentions précédentes des syntagmes nominaux réutilisés par l'enfant sur la nature des formes qu'il produit devant les noms. A la suite de travaux portant sur le rôle du dialogue et du discours de l'interlocuteur - et notamment des reprises - sur le développement des formes (*cf.* par exemple Veneziano, 1997 ou Bertin, 2011), nous chercherons à voir si les reprises que l'enfant fait du discours de son interlocuteur ou de son propre discours ont un impact sur le type de syntagme produit.

Ces deux thématiques - liées pour l'une au contexte pragmatico-discursif et pour l'autre à la réutilisation des noms dans le discours de l'enfant - seront traitées successivement. Dans un premier temps, nous nous intéresserons à l'influence des caractéristiques pragmatico-discursives des référents, et dans un deuxième temps, nous nous concentrerons sur le rôle des reprises du nom dans l'acquisition des formes. Comme aux chapitres V et VI, tous les résultats et les discussions de ces résultats seront précédés de précisions méthodologiques.

## **1. Influence des caractéristiques pragmatico-discursives des référents**

Nous avons eu l'occasion de constater au travers de certains exemples cités au chapitre VI (et en particulier, au sein de la section 3) que la nature des déterminants produits par l'enfant pouvait s'expliquer par des facteurs tels que le type de référence associée au nom ou le statut attentionnel et discursif du référent. Ces constats concordent avec les conclusions de recherches antérieures sur l'acquisition des fonctions pragmatiques des déterminants. Rozendaal (2008) ou Salazar Orvig *et al.* (2005, 2008, 2013) ont notamment montré que l'enfant fait preuve d'une sensibilité précoce au contexte pragmatico-discursif dans lequel les noms sont produits ainsi qu'à la dimension plurifonctionnelle des déterminants. D'autres auteures comme Karmiloff-Smith ou Hickmann ont également travaillé sur l'interface entre formes et fonctions mais les résultats de leurs travaux (entre autres, Karmiloff-Smith, 1979 ; Hickmann, 1996) témoignent en revanche d'une maîtrise plus tardive de cette plurifonctionnalité. Dans la première partie de ce chapitre, nous reviendrons sur les hypothèses formulées dans ces précédents travaux et nous essaierons de mesurer de façon systématique l'impact de certains facteurs pragmatico-discursifs sur la nature des déterminants produits par les quatre enfants de notre corpus. Mais avant de nous focaliser spécifiquement sur les formes adultes, nous chercherons à voir si selon le statut pragmatico-discursif du référent, l'enfant fait un usage différencié des omissions, fillers et déterminants. Par ailleurs, dès lors que le déterminant n'est pas d'emblée la forme majoritaire, nous essaierons de voir si les différents types de fillers produits par l'enfant peuvent être associés à des contextes spécifiques et/ou utilisés en contraste avec les premières formes adultes.

Après un point méthodologique sur la façon dont ont été caractérisés les référents (section 1.1), nous nous concentrerons sur les manifestations éventuelles de contrastes pragmatico-discursifs dans les productions pré-nominales de chaque enfant. Nous porterons une attention particulière

aux oppositions entre déterminants mais aussi entre formes de transition et déterminants, ainsi qu'entre omissions et fillers. Plusieurs contextes pragmatico-discursifs contrastés ont été considérés pour cette étude : l'opposition entre les syntagmes nominaux utilisés de façon référentielle et ceux utilisés de façon non référentielle (section 1.2.1), l'opposition entre référence particulière et indéterminée (1.2.2), la distinction entre ce qui est nouveau et donné dans le discours (pour les référents particuliers), et en dernier lieu (section 1.2.4), l'influence du degré de connaissance partagée entre les interlocuteurs sur le choix des formes pré-nominales (lorsque le référent est nouveau).

## 1.1 Caractérisation pragmatico-discursive des référents

Avant de traiter la question de l'influence des traits pragmatico-discursifs des référents sur les productions des quatre enfants de notre corpus, nous ferons un point sur la façon dont ces référents ont été caractérisés. Des sections 1.1.1 à 1.1.4, nous précisons quels critères nous ont permis de statuer sur le caractère référentiel ou non du syntagme nominal, et nous indiquerons par ailleurs à quelles catégories nous avons eu recours - lorsque le nom est utilisé de façon référentielle - pour caractériser le type de référence, le statut attentionnel et discursif du référent, ainsi que la familiarité du référent pour l'interlocuteur.<sup>1</sup>

### 1.1.1 Caractère référentiel du syntagme nominal

Dans un premier temps, nous avons cherché à distinguer les usages non référentiels des noms pouvant être précédés d'un déterminant, des usages référentiels. Quatre cas de figure différents ont été identifiés :

- les cas où l'enfant utilise un nom pour dénommer ou catégoriser un objet ou pour répondre à une demande de dénomination de son interlocuteur, comme dans l'exemple 1 ci-dessous.

#### Exemple 1 - Usage dénominatif du nom (Adrien, 2;11)

Mère : c'est quoi ?

Adrien : [ɛip] 'jupe'

Mère : une jupe, ouais !

---

<sup>1</sup> Signalons que la méthodologie décrite dans ces quatre points est très largement inspirée de la méthodologie d'analyse de corpus développée dans le cadre du projet DIAREF déjà mentionné précédemment.

- les contextes dans lesquels l'enfant fait un usage attributif (a) ou métalinguistique (b) du nom (cf. exemple 2).

### Exemple 2 - Usages attributif et métalinguistique du nom

(a) Adrien (3;5)

Adrien : [vø dy zy toladin] 'veux du jus grenadine'

Le nom [toladin] 'grenadine' est ici utilisé de façon attributive.

(b) Adrien (2;1)

Père : tu dis l'eau Adrien ?

Adrien : [lo] 'l'eau'

- les contextes dans lesquels le nom est employé de façon référentielle et où il renvoie donc à une entité extralinguistique (cf. exemple 3).

### Exemple 3 - Usage référentiel du nom (Adrien, 2;11)

Adrien : [t e kase deʒe da pɔt] 't'es caché derrière la porte !'

- et en dernier lieu, les cas pour lesquels il n'a pas été possible de décider si l'enfant faisait un usage dénominatif ou référentiel du nom (le plus souvent, lorsque les noms sont produits dans les énoncés à un terme ou dans les premiers énoncés à deux termes). C'est le cas par exemple dans l'extrait du corpus de Lubin (à 1;9) reproduit ci-dessous.

### Exemple 4 - Incertitude entre usage dénominatif et référentiel (Lubin, 1;9)

Lubin et sa mère font de la peinture. La mère de Lubin est en train de peindre un escargot.

Lubin : [gago ga] '(es)ca(r)got yyy ?' (pointe dessin de sa mère)

Lubin : [gago] '(es)ca(r)got ?'

Lubin : [cago] '(es)ca(r)got ?'

Lubin : [mamã gago] 'maman (es)ca(r)got ? (pointe de nouveau le dessin de sa mère)

Mère : oui.

Lubin : [cago] '(es)ca(r)got ?'

Mère : veux tu qu(e) j(e) t'aide à en dessiner un ?

Mère : tu veux qu(e) je t'aide à dessiner un escargot ?

Lubin : [wɛ] 'ouais.'

*Dans toutes les mentions du référent 'escargot' ci-dessus, nous pourrions imaginer au moins trois paraphrases possibles : 'c'est un escargot', 'tu dessines un escargot' ou encore 'je veux dessiner un escargot'. Dans le premier cas, les productions de l'enfant ne seraient pas référentielles. En revanche, dans les deux suivants, le nom serait utilisé de façon référentielle.*

Signalons par ailleurs que nous n'avons pas considéré pour cette analyse du caractère référentiel du syntagme nominal les noms produits dans le cadre de chants ou de comptines, apprises par cœur ou récitées à la suite de l'adulte. Dans l'exemple 5 qui suit, aucun des noms produits par l'enfant n'a fait l'objet d'une analyse pragmatico-discursive.

### Exemple 5 - Cas d'énoncé exclu de l'analyse (Lubin, 2;6)

Lubin : [o le mɛ̃ po də lapɛ̃ mɛʁɛs a majo də bɛ̃] 'haut les mains peau d(e) lapin maî(t)resse yyy maillot d(e) bain.' (lance la fléchette)

Pour les analyses présentées au point 1.2.1, trois catégories seront considérées : la catégorie des usages non référentiels (pour la plupart dénominatifs, mais comprenant aussi quelques occurrences de noms utilisés de façon attributive ou métalinguistique), celle des usages référentiels et enfin, la catégorie regroupant les cas d'incertitude sur la valeur dénomminative ou référentielle du syntagme nominal. Ces cas pouvant être fréquents, voire très fréquents, dans les premières séances de certains enfants, nous avons souhaité les conserver pour l'analyse, afin de cerner notamment si la distribution des formes produites dans ces contextes se rapproche davantage de celle associée aux usages référentiels ou aux usages non référentiels. En revanche, toutes les analyses qui suivent n'ont été conduites que sur les cas clairement identifiés comme référentiels.

#### 1.1.2 Type de référence

Dans une seconde étape, tous les noms préalablement catégorisés comme référentiels ont fait l'objet d'une analyse concernant le type de référence porté par ces noms. Nous avons distingué ici trois types de référence distincts : la référence particulière, la référence indéterminée et la référence générique.

Il s'agit dans le premier cas de toute référence à des individus particuliers, qu'ils soient plus ou moins spécifiés (cf. usage du nom 'train' dans l'exemple 6 ci-dessous).

### Exemple 6 - Référence particulière (Naël, 1;10)

Naël : [u mɔ̃ kʁɛ̃] 'où mon train ?'

Mère : il est là.

Naël : [i il e la] 'i(l) [l] il est là ?'

Mère : huhum.

Naël : [vø mɔ̃ kʁɛ̃] 'veux mon train !'

Sous l'étiquette 'indéterminé', nous avons classé tous les référents renvoyant à n'importe quel individu d'une catégorie, comme dans l'exemple 7 ci-dessous.

### Exemple 7 - Référence indéterminée (Lubin, 2;5)

Père : qu'est ce que tu veux faire comme jeu ?

Lubin : [a pɔʒɔɛl] 'F puzzle !'

Père : tu veux faire un puzzle ?



Père : oh nan pas de puzzle.

Lubin : [me si] ‘mais si.’

*Lubin verbalise ici son souhait de faire un puzzle, mais il ne précise pas lequel. Nous avons donc considéré ici que le nom ‘puzzle’ renvoyait à un référent indéterminé.*

Tous les noms référant à l’ensemble de la catégorie représentée par le nom ont été catégorisés comme ‘génériques’ (cf. exemple 8).

### Exemple 8 - Référence générique (Naël, 1;11)

Naël : [na sevø maʁin] *‘yyy cheveux Marine.’*

Mère : elle a des cheveux Marine ?

Mère (à OBS) : pa(r)ce qu'on lui a expliqué tout à l'heure que les cheveux c'était des poils et que les poils n'étaient pas des cheveux.

Observatrice : d'accord.

Naël : [se de fwal] ‘c'est des poils.’

Mère : oui [/] oui [/] oui mais on dit des poils pour les chiens.

Naël : [se de fwal də puʁ le ʃɛ̃] ‘c'est des poils yyy pour les ch(i)ens.’

*Dans le dernier énoncé ci-dessus, Naël reprend l'énoncé de sa mère. Le nom ‘chien’ renvoie ici non pas à un référent particulier mais à l'ensemble de la catégorie ‘chien’, partageant pour caractéristique le fait d'avoir des poils (et non des cheveux).*

Les cas dans lesquels l'énoncé ou le contexte ne permettaient pas de statuer sur le type de référence ont été codés comme ‘indécidables’.

### 1.1.3 Statut attentionnel et discursif du référent

Nous avons également souhaité - dans le cas des référents particuliers - analyser la distribution des formes pré-nominales en fonction du statut attentionnel et discursif du référent. Quatre cas de figure différents ont été considérés :

- les noms introduisant un nouvel objet de discours dans le focus d'attention des interlocuteurs ;
- les noms renvoyant à des référents déjà donnés dans le discours, et donc préalablement mentionnés par l'enfant ou par son interlocuteur, y compris par un interrogatif (cf. exemple 9).

### Exemple 9 - Référent donné dans le discours (Naël, 1;11)

Naël : [a] ‘ah ? (fait semblant de manger cheval)’

Mère : tu le manges ?

Naël : [wi] ‘oui.’

Mère : mais si tu manges ton cheval tu pourras plus galoper ?

Naël : [ʃ mɑ̃ʃ ə ʃəval] ‘on mange F cheval.’

- les noms renvoyant à des référents n'ayant pas été verbalisés (et donc pas encore donnés dans le discours) mais se trouvant sous l'attention des interlocuteurs ('focalisés') au moment de la première mention (cf. première mention de 'enfants' dans l'exemple 10 ci-dessous).

**Exemple 10 - Référent focalisé (Lubin, 2;5, activité de lecture de livre avec son père)**

Père : oui ça c'est la femme du fermier.

Lubin : [iʁa le fãfã i vɔ̃ ʁãte la] 'yyy les enfants i(ls) vont rent(r)er là ?'

Père : comment ?

Lubin : [ka le zãfã ə a ʁãte la lə papa] 'yyy les enfants yyy rent(rer) là le papa ?'

Père : les enfants i(ls) vont rentrer là tu crois ?

Lubin : [wi] 'oui.'

Père : ça c'est des en(fants) [//] ça c'est une classe.

*Dans cet extrait, Lubin introduit le référent 'enfants', mais lui et son père étant impliqués dans une activité de lecture (centrée sur la même page depuis plusieurs tours de parole), le référent est déjà sous l'attention des interlocuteurs avant d'être verbalisé. Nous avons donc considéré que le référent associé au nom était ici focalisé et non nouveau.*

Enfin, dans une dernière catégorie, nous avons classé tous les noms que l'enfant produit pour réintroduire un référent après un changement thématique.

Comme pour les analyses précédentes, une catégorie 'indécidable' a été réservée aux référents pour lesquels le statut attentionnel et discursif ne pouvait pas être déterminé avec certitude.

**1.1.4 Familiarité du référent**

Dans les cas où le référent est introduit pour la première fois dans le discours par l'enfant, nous avons souhaité mesurer l'impact de la familiarité du référent pour l'interlocuteur sur les types de formes pré-nominales produites par l'enfant. Ainsi, nous avons différencié pour tous les référents nouveaux les cas où le référent est connu des deux interlocuteurs (cf. mention de 'chaussettes' dans l'exemple 11), et les cas - très rares - où au contraire, le référent est connu du locuteur mais de toute évidence, inconnu pour l'interlocuteur (cf. mention de 'chapeaux' dans l'exemple 12).

**Exemple 11 - Référent connu des deux interlocuteurs (Antonin, 2;11, jeu de construction)**

Mère : tiens vas y mets le dessus.

Mère : de l'aut(re) côté [/] de l'aut(re) côté.

Mère : t'appuies fort ?

Antonin : [ɔ̃m] 'hum.'

Mère : c'est bon ?

Mère : alors ensuite.

Antonin : [i mə ɡʁat də le ʁosɛt] 'yyy me gratte dans les chaussettes.' (se gratte sous le pied)

Mère : sous l(e) pied ?

Antonin : [wi] 'oui.'

### Exemple 12 - Référent inconnu de l'interlocuteur de l'enfant (Antonin, 3;2)

Mère : d'ailleurs t'as joué aux pompiers à la crèche aujourd'hui ?

Antonin : [e u pɔpje] 'yyy où pompier ?'

Mère : le camion de pompier non ?

Antonin : [we avɛ le ʃapo pɔpje] 'ouais ave(c) les chapeaux pompier.'

Mère : t'as mis les chapeaux de pompier ?

Antonin : [kwi] 'oui.'

Antonin : xxx . (inaudible)

Mère : mais sur ta tête ou sur des personnages ?

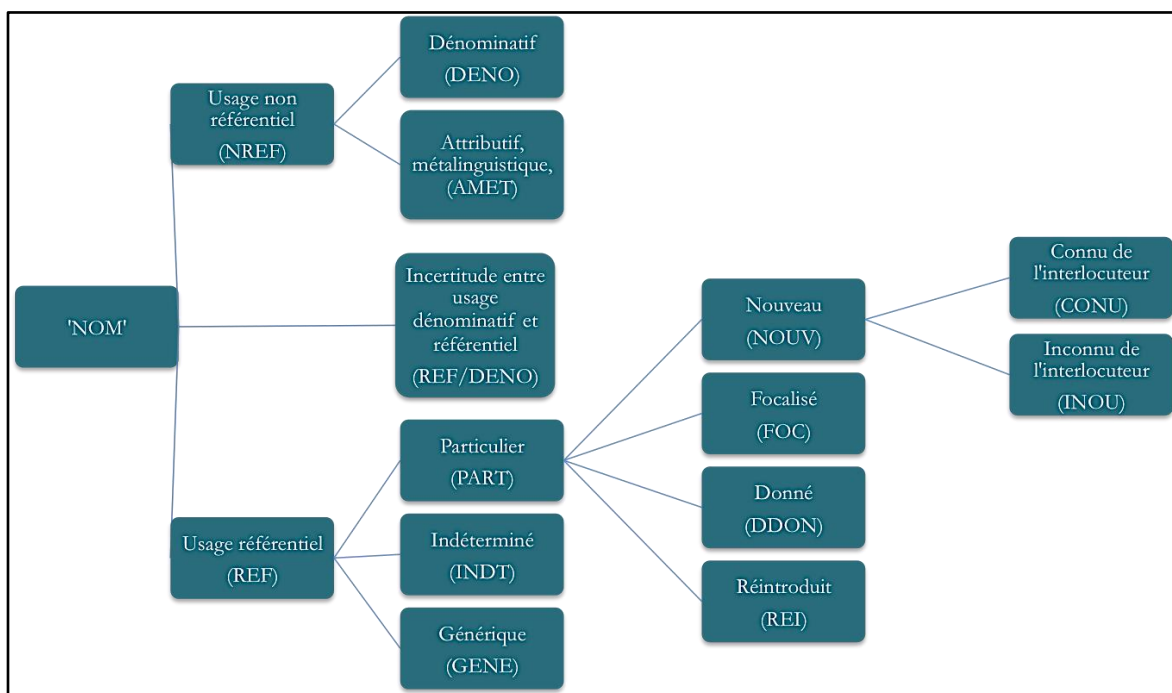
Antonin : [syʁ le syʁ le X e ɑ̃sɥit ʒɛ mi syʁ ma tɛt] 'sur les [ / ] sur les xxx et ensuite j'ai mis sur ma tête.'

*Dans cet exemple, le référent 'chapeaux de pompier' est inférable du contexte discursif. Toutefois, la mère n'étant pas présente lors de l'activité de l'enfant à la crèche, nous avons considéré qu'elle n'avait pas connaissance du référent particulier verbalisé ici par Antonin.*

Les cas pour lesquels le corpus ne nous a pas permis de trancher en faveur d'une analyse plus qu'une autre ont été codés comme des cas indécidables.

Toutes les analyses pragmatico-discursives décrites dans les pages précédentes sont synthétisées à l'aide du schéma 1 ci-dessous.

### Schéma 1 - Récapitulatif des analyses pragmatico-discursives



## 1.2 Gestion des contrastes pragmatico-discursifs chez l'enfant

### 1.2.1 L'opposition entre syntagmes référentiels et non référentiels

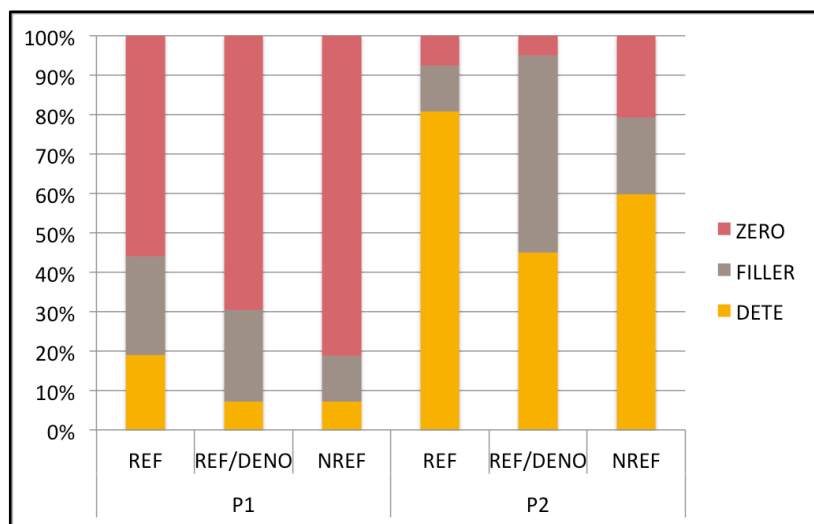
La première partie de notre travail sur la gestion des contrastes pragmatico-discursifs par l'enfant porte sur l'opposition entre les contextes référentiels et non référentiels. Nous avons cherché à observer si l'enfant pouvait avoir recours à des formes différentes selon que le nom était utilisé dans l'un ou l'autre de ces deux contextes. Dans le cadre du premier point du chapitre II, portant sur les caractéristiques sémantico-référentielles des syntagmes nominaux, nous avons eu l'occasion de signaler que dans la littérature (*cf.* par exemple Charolles, 2002), l'indéfini est présenté comme le déterminant privilégié dans les cas d'usages non référentiels des noms (comme les contextes de dénomination), par opposition au défini. Appliqué à la question de l'acquisition des déterminants et de la production concomitante d'omissions, de fillers et de déterminants, ce constat soulève deux interrogations principales : a) peut-on observer dans notre corpus un usage différencié des formes selon que le nom est produit dans un contexte référentiel ou non référentiel ? b) de même, peut-on observer des usages contrastés des premiers déterminants (et en particulier, une opposition entre définis et indéfinis) et des différents types de fillers ?

Pour répondre à ces questions, nous avons commencé par identifier les formes pré-nominales produites lorsque le nom est utilisé de façon référentielle (REF dans les graphiques qui suivent) et de façon non référentielle (NREF). Les usages non référentiels correspondent ici aux dénominations, aux usages attributifs et aux usages métalinguistiques. Par ailleurs, rappelons que les cas d'incertitude sur la valeur dénominate ou référentielle du syntagme nominal ont été regroupés dans une troisième catégorie (REF/DENO) et conservés pour cette première analyse.

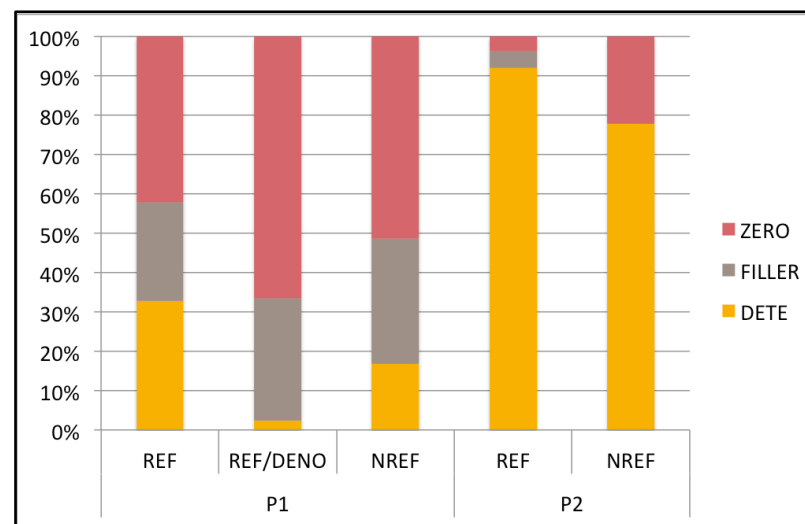
Les quatre graphiques qui suivent présentent la distribution des omissions, fillers et déterminants dans ces trois contextes, pour chaque enfant et par période (P1 et P2). Les pourcentages affichés dans les graphiques représentent la moyenne (par période) des fréquences d'utilisation de chaque forme pré-nominale à chaque âge et dans chacun des contextes.

**Graphiques 1 à 4 - Distribution des formes pré-nominales dans les contextes référentiels, non référentiels et dans les cas d'incertitude entre valeur référentielle et dénomminative (chez les quatre enfants)**

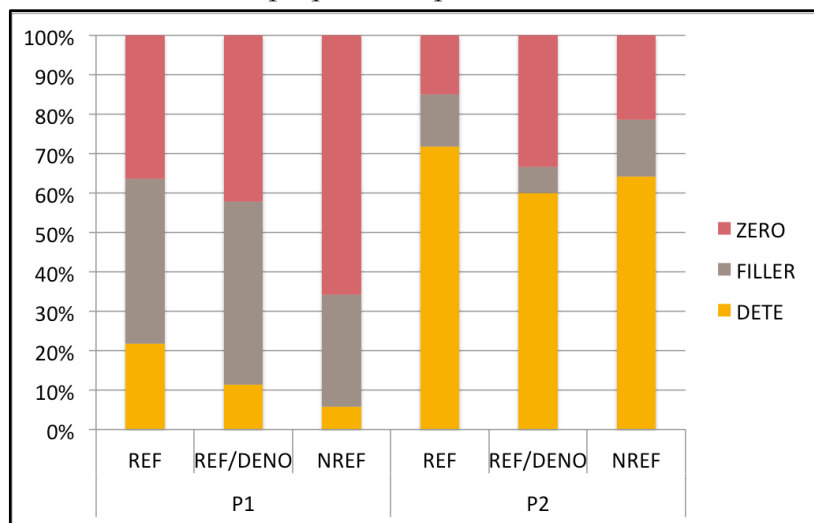
*Graphique 1 - Corpus de Lubin*



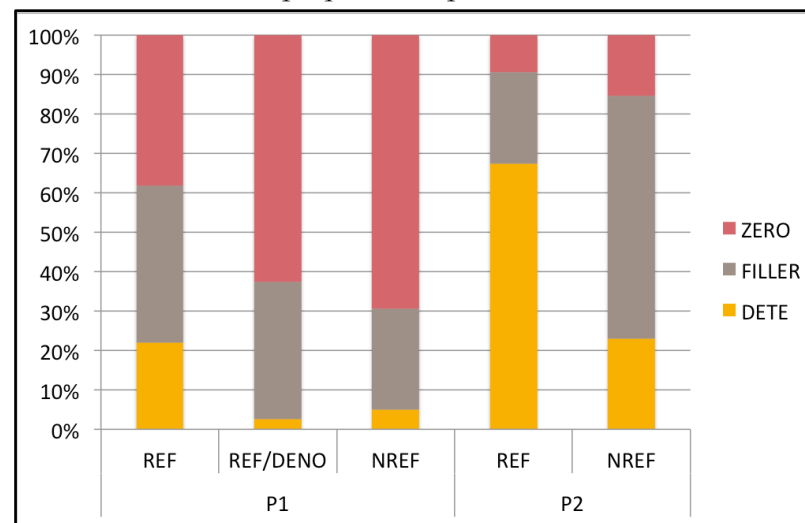
*Graphique 2 - Corpus de Naël*



*Graphique 3 - Corpus d'Antonin*



*Graphique 4 - Corpus d'Adrien*



A la lecture de ces graphiques, nous pouvons constater qu'en période 1, la part des déterminants est plus importante dans les contextes référentiels que dans les deux autres contextes chez les quatre enfants. Les omissions sont quant à elles plus fréquentes dans les contextes non référentiels dans les données de Lubin, Antonin et Adrien. Dans le corpus de Naël, les omissions sont les plus fréquentes dans les contextes où nous n'avons pas pu trancher sur la valeur dénomminative ou référentielle du nom et au contraire, c'est dans ces contextes que les déterminants sont le moins souvent produits. Les omissions restent toutefois plus fréquentes dans ce corpus dans les contextes non référentiels que référentiels.

Dans les corpus de Lubin, Antonin et Adrien, la distribution des formes pré-nominales dans les contextes d'incertitude entre valeur référentielle et dénomminative peut être considérée comme un intermédiaire entre ce que nous observons dans les contextes référentiels et non référentiels. Chez Lubin et Adrien, la proportion d'omissions y est plus importante que dans le cas des usages référentiels et à l'inverse, moins importante que lorsque les noms sont employés de façon non référentielle. Les fillers sont par ailleurs moins fréquents que dans les contextes référentiels et plus fréquents que dans les contextes non référentiels. Chez Antonin, les omissions apparaissent dans des proportions comparables avec les usages référentiels des noms et dans les cas d'incertitude entre usage référentiel et dénomminatif mais elles sont moins représentées dans ce dernier contexte que dans le contexte non référentiel. La proportion de fillers est plus grande avec ces noms pouvant avoir une valeur référentielle ou dénomminative que dans les deux autres contextes.

En période 2, le contexte référentiel entraîne à nouveau une part plus importante de déterminants que les autres contextes chez les quatre enfants. Les omissions sont par ailleurs plus observées avec les usages non référentiels que référentiels. Chez Lubin, Antonin et Adrien, les fillers sont aussi plus représentés dans les contextes non référentiels que référentiels.

Les cas d'incertitude entre usage dénomminatif et référentiel sont bien moins nombreux qu'en période 1. Les productions de l'enfant étant plus longues, il devient en effet plus évident de distinguer les cas où l'enfant emploie un nom pour dénommer des cas où les syntagmes nominaux sont utilisés pour parler d'un référent existant. Seuls quelques cas ont été considérés comme indécidables dans les données de Lubin et Antonin. Dans le corpus de Lubin, ces noms entraînent une proportion de fillers plus importante qu'avec les noms employés de façon

référentielle et non référentielle et dans celui d'Antonin, les omissions sont plus fréquentes avec ces cas indécidables que dans les deux autres contextes mais en revanche, les fillers sont plus représentés dans le cas des usages référentiels et non référentiels. Les déterminants sont moins fréquents chez Lubin devant ces noms dont le statut référentiel n'a pas pu être défini - et seulement légèrement moins fréquents dans le corpus d'Antonin - que lorsque le nom a une valeur référentielle ou non référentielle. Toutefois, ces cas de figure étant très largement minoritaires dans les deux corpus, aucun bilan fiable ne peut être dressé ici.

Ces premiers résultats suggèrent donc que le contexte référentiel serait favorable à la production de formes adultes et *a contrario*, que le contexte non référentiel générerait une part plus importante de formes de transition, en période 1 comme en période 2. Par ailleurs, dans les corpus de Lubin, Antonin et Adrien, la distribution des formes dans les contextes dits 'référentiels/dénommatifs' nous laisse penser que ces cas rassemblent bien des usages référentiels et non référentiels. En revanche, la distribution des formes dans le corpus de Naël nous permet d'envisager que ces noms au statut indécidable sont plus probablement employés par l'enfant de façon non référentielle.

Outre la répartition des trois types de formes pré-nominales, nous avons souhaité observer si la distribution des différents types de déterminants et de fillers semblait dépendre, dans nos données, du caractère référentiel ou non référentiel de la production nominale de l'enfant. Dans la mesure où Lubin, Antonin et Adrien produisent surtout des déterminants définis et indéfinis, et seulement quelques occurrences d'autres déterminants, nous nous sommes uniquement focalisée ici (ainsi que pour les deux sections qui suivent) sur le contraste entre définis et indéfinis. Dans le cas de Naël, nous avons aussi fait figurer les résultats concernant les possessifs, largement représentés dans ce corpus en période 2 (*cf.* section 3.3 du chapitre 5). Dans les quatre tableaux à suivre (correspondant à chacun des corpus), nous avons donc affiché en valeurs absolues et pour chaque période la distribution des déterminants et des fillers dans les contextes référentiels et non référentiels ainsi que dans les cas où le nom a pu être utilisé de façon référentielle ou dénomminative.

Si l'on s'en tient aux observations faites dans la littérature (*cf.* par exemple Salazar Orvig *et al.*, 2013), nous pouvons nous attendre ici à observer une part plus importante d'indéfinis que de

définis dans les contextes non référentiels. Toutefois, dans la mesure où les indéfinis apparaissent surtout en seconde période dans les quatre corpus, nous pouvons imaginer également que les fillers (et peut-être plus particulièrement, les fillers vocaliques, plus fréquents en première période) puissent dans un premier temps se ‘substituer’ aux indéfinis dans ces contextes non référentiels. Dans les contextes référentiels, nous pouvons faire l’hypothèse que les définis - de par leur présence majoritaire dans nos données - seront plus représentés que les indéfinis. Par ailleurs, conformément à ce que nous venons de voir au travers des graphiques précédents, il est envisageable que dans les corpus de Lubin, Antonin et Adrien, la distribution des formes produites dans les contextes d’incertitude entre valeur référentielle et dénomminative soit sur certains aspects comparable au contexte référentiel et sur d’autres, au contexte non référentiel. A l’inverse, nous pouvons penser que celle-ci sera plus proche du contexte non référentiel dans les données de Naël.

Dans le corpus de Lubin (tableau 1), les définis et les fillers vocaliques sont les formes les plus utilisées dans les contextes référentiels en période 1. On note par ailleurs que même si les fillers vocaliques sont plus utilisés dans les contextes référentiels, ils sont proportionnellement plus que les autres formes utilisés dans les contextes non référentiels. Seuls 5 indéfinis ont été observés en période 1, dont 3 dans le contexte non référentiel. La plupart des fillers de type CV (7 sur 12) sont produits dans des contextes d’incertitude entre usage référentiel et dénomminatif. Rappelons par ailleurs qu’aucun filler de type ‘consonne seule’ n’a été relevé dans ce corpus, en période 1 comme en période 2.

**Tableau 1 - Distribution des différents types de déterminants et fillers dans les contextes référentiels, non référentiels et dans les cas d’incertitude entre valeur référentielle et dénomminative (dans le corpus de Lubin)**

Période	Déterminant/ Filler	Référentiel	Incertain référentiel/ dénomination	Non référentiel
1	Défini	11	3	1
	Indéfini	1	1	3
	V	16	4	7
	CV	4	7	1
2	Défini	107	4	8
	Indéfini	50	2	29
	V	16	8	11
	CV	10	2	3



Durant la seconde période, les usages référentiels sont (comme en période 1), plus nombreux que les cas de noms employés de façon non référentielle. Le défini est de nouveau privilégié dans les contextes référentiels mais les indéfinis sont également le plus souvent produits associés à des noms utilisés de façon référentielle. A l'inverse, dans les contextes non référentiels, les indéfinis sont nettement plus visibles que les définis, et les fillers vocaliques sont également plus représentés que les fillers de type CV. Conformément à ce que nous avons vu plus tôt, les cas d'incertitude entre valeur référentielle et dénomminative du nom sont peu nombreux en seconde période. Le plus souvent (8 cas sur 16), ces syntagmes sont de type 'filler V + nom'.

Dans le corpus de Naël (tableau 2), les définis ainsi que les fillers V et CV sont surtout utilisés en période 1 en contexte référentiel, et le possessif est quant à lui seulement produit dans ce contexte. Seul l'indéfini est plus observé en contexte non référentiel (8 occurrences sur 10 en période 1). On remarque par ailleurs que les fillers vocaliques sont majoritaires dans ce contexte mais que les fillers de type CV sont plus nombreux que les indéfinis. Comme dans le corpus de Lubin, les fillers CV sont la forme la plus utilisée dans les contextes que nous avons catégorisés comme des incertitudes entre usages référentiels et dénominatifs.

**Tableau 2 - Distribution des différents types de déterminants et fillers dans les contextes référentiels, non référentiels et dans les cas d'incertitude entre valeur référentielle et dénomminative (dans le corpus de Naël)**

Période	Déterminant/ Filler	Référentiel	Incertain référentiel/ dénomination	Non référentiel
1	Défini	38	---	10
	Indéfini	1	1	8
	Possessif	24	---	---
	V	33	2	19
	C	1	---	---
	CV	24	11	14
2	Défini	74	---	---
	Indéfini	62	---	6
	Possessif	60	---	---
	V	2	---	---
	C	2	---	---
	CV	4	---	---

En période 2, nous n'avons relevé presque aucun syntagme nominal produit en contexte non référentiel. Seulement 6 cas ont été classés comme tels et tous sont associés à un déterminant

indéfini. Les définis (74 occurrences) sont majoritaires dans les contextes référentiels mais les indéfinis et les possessifs y sont presque aussi fréquents (62 indéfinis et 60 possessifs).

Dans le corpus d'Antonin (tableau 3), les noms employés de façon non référentielle sont plus nombreux que dans les autres corpus en période 1, et cela en raison du plus grand nombre d'activités de dénomination. Ces usages sont surtout associés à des fillers vocaliques (36) mais aussi à quelques définis (6). Les fillers V sont par ailleurs largement représentés dans les cas d'incertitude entre usage référentiel et dénominatif, ainsi que dans les contextes référentiels. Les définis et les fillers CV sont ensuite les formes les plus fréquentes dans ce dernier contexte.

**Tableau 3 - Distribution des différents types de déterminants et fillers dans les contextes référentiels, non référentiels et dans les cas d'incertitude entre valeur référentielle et dénominative (dans le corpus d'Antonin)**

Période	Déterminant/ Filler	Référentiel	Incertitude référentiel/ dénomination	Non référentiel
1	Défini	30	13	6
	Indéfini	2	1	2
	V	43	59	36
	C	1	---	1
	CV	15	6	3
2	Défini	175	12	13
	Indéfini	21	6	14
	V	19	2	4
	C	3	---	---
	CV	14	---	1

En période 2, les fillers C et CV sont quasiment tous utilisés dans des contextes référentiels. Une seule occurrence de filler CV est observée en contexte non référentiel. Les fillers vocaliques sont utilisés plus souvent dans un contexte référentiel mais on en observe toutefois quelques occurrences dans les deux autres contextes. Les définis sont plus fréquents dans les contextes référentiels et sont aussi la forme la plus utilisée dans les contextes d'incertitude entre valeur référentielle et dénominative du nom. Les définis et les indéfinis sont produits à part quasiment égale dans les contextes non référentiels. Nous avons en effet pu observer au préalable que les dénominations sont souvent faites au travers du défini, chez Antonin comme chez sa mère.

Dans les contextes référentiels, Adrien produit en première période (correspondant rappelons le aux six premières séances du corpus) une majorité de fillers CV, et dans un deuxième temps, des

fillers vocaliques (*cf.* tableau 4 ci-dessous). Les définis et les indéfinis sont quant à eux produits dans des proportions comparables lorsque le nom est utilisé de façon référentielle. Les fillers vocaliques sont les plus observés dans les cas où nous n'avons pas pu trancher sur la valeur référentielle ou dénomminative du nom, et contrairement à ce que nous avons pu relever dans les autres corpus, les fillers C sont privilégiés en contexte non référentiel.

**Tableau 4 - Distribution des différents types de déterminants et fillers dans les contextes référentiels, non référentiels et dans les cas d'incertitude entre valeur référentielle et dénomminative (dans le corpus d'Adrien)**

Période	Déterminant/ Filler	Référentiel	Incertain référentiel/ dénomination	Non référentiel
1	Défini	40	---	2
	Indéfini	39	1	---
	V	67	23	---
	C	8	---	11
	CV	85	3	3
2	Défini	63	---	1
	Indéfini	19	---	2
	V	7	---	5
	C	3	---	---
	CV	18	2	3

En deuxième période, les noms sont très peu utilisés de façon non référentielle. Lorsqu'ils le sont, ils sont produits d'abord avec des fillers vocaliques (5 occurrences sur 11) mais également avec des fillers CV, des indéfinis et un défini. Les occurrences ne sont quoi qu'il en soit pas suffisamment nombreuses pour pouvoir dire si une forme est privilégiée dans ce contexte en période 2. Les définis sont les plus utilisés avec les noms employés de façon référentielle et viennent ensuite, dans un peu moins de 20% des cas, les indéfinis et les fillers CV.

Les quatre tableaux précédents nous permettent de constater qu'en période 1, le contexte référentiel favorise la production de définis (et de possessifs chez Naël) et/ou de fillers vocaliques chez les quatre enfants. En contexte non référentiel, les indéfinis et/ou les fillers vocaliques sont privilégiés, sauf dans le corpus d'Adrien qui produit une majorité de fillers de type 'consonne seule'. En période 2, le défini, ainsi que le possessif chez Naël, sont les formes majoritaires lorsque le nom a une valeur référentielle. L'indéfini ou le filler vocalique sont les formes les plus utilisées en contexte non référentiel. Lorsque le nom peut avoir été utilisé de

façon référentielle ou dénominative (cas d'incertitude), nous n'avons pas observé de distribution spécifique et comparable dans tous les corpus, en période 1 comme en période 2. Nous ne pouvons par ailleurs pas définir si les formes produites dans ces contextes tendent à s'approcher davantage des usages référentiels ou non référentiels.

Nous pouvons en conclure que dans les quatre corpus, les enfants utilisent les formes pré-nominales de façon contrastée, selon que le nom est utilisé de façon référentielle ou non référentielle. Toutefois, nous n'observons pas d'associations exclusives entre certaines formes et certains contextes. Nous avons notamment pu constater d'une part que le défini n'est pas réservé aux usages référentiels mais peut aussi être utilisé, au même titre que l'indéfini, pour dénommer. Conformément à ce que nous pouvions attendre, l'indéfini est pour sa part utilisé non seulement pour les usages non référentiels mais aussi dans le cas des usages référentiels. Nous ferons l'analyse de ces emplois d'indéfinis en contexte référentiel dans les pages qui suivent. Par ailleurs, nous avons pu constater que les fillers vocaliques, largement produits dans tous les corpus, sont observés aussi bien lorsque le nom est employé de façon référentielle que non référentielle. Ils restent néanmoins plus présents que les fillers CV dans les contextes non référentiels chez Lubin, Naël et Antonin.

### **1.2.2 Caractère particulier ou indéterminé du référent**

Au-delà de la distinction entre usage référentiel et non référentiel, nous avons également cherché à repérer - pour les usages référentiels des noms - si la forme pré-nominale produite pouvait dépendre du caractère particulier ou indéterminé du référent. Nous avons vu au sein des travaux mentionnés dans la première partie de ce travail (*cf.* par exemple Rozendaal, 2008 ou Rozendaal et Baker, 2008) que le défini serait plus utilisé dans la langue adulte lorsque le locuteur a à l'esprit une entité particulière et qu'au contraire, l'indéfini serait privilégié lorsque celui-ci est quelconque (ou non spécifique). Deux questions principales découlent de ces observations : a) les omissions, fillers et déterminants sont-ils utilisés de façon contrastée dans nos données - tout au long du développement ou à une période donnée - en fonction du caractère particulier ou indéterminé du référent ? Et b) cette opposition est-elle repérable au travers des types de déterminants et de fillers produits par les quatre enfants ?

Nous l'avons signalé plus tôt (*cf.* point 1.1 de ce chapitre), nous avons initialement distingué trois types de référence : la référence particulière, indéterminée et générique. Toutefois, dans la mesure où notre analyse n'a porté que sur les énoncés de l'enfant, contenant très peu d'occurrences de noms employés comme des génériques, nous avons fait le choix de ne pas les prendre en compte dans les graphiques ci-dessous. Nous y ferons référence ultérieurement, lorsque nous commenterons les résultats concernant l'emploi des déterminants et des fillers. Les graphiques 5 à 8 qui suivent représentent donc les usages des formes pré-nominales chez les quatre enfants, pour chaque période et dans les deux contextes de référence que nous venons de citer, à savoir, les contextes de référence particulière et indéterminée.

En premier lieu, nous pouvons constater que les formes pré-nominales ne se distribuent pas dans les contextes de référence particulière et indéterminée de façon équivalente chez tous les enfants, et cela, quelle que soit la période.

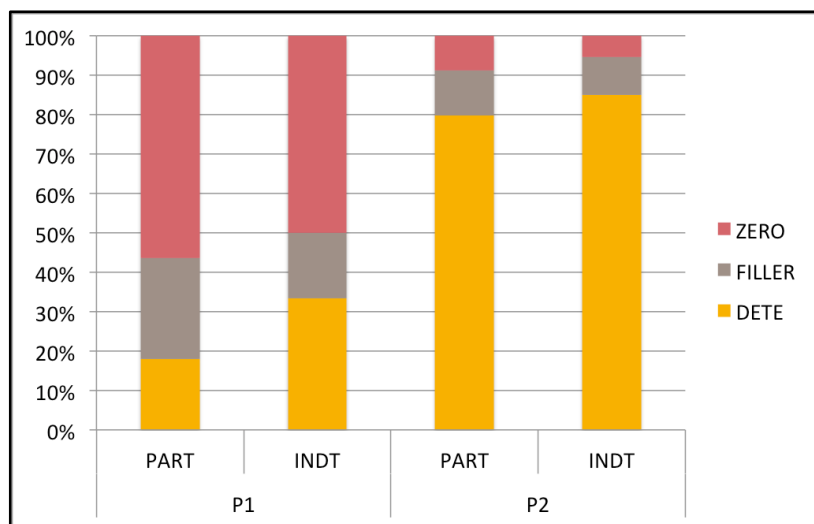
En période 1, les déterminants apparaissent comme plus fréquemment produits lorsque la référence est indéterminée dans les corpus de Lubin et Antonin, et les omissions semblent être privilégiées dans les contextes de référence particulière chez Antonin. Toutefois, chez ces deux enfants, la distribution des formes ne peut pas être véritablement comparée dans les deux contextes avant la période 2 puisque seuls 6 noms sont utilisés pour renvoyer à un référent indéterminé dans les données de Lubin en période 1 et 2 dans celles d'Antonin.

Les résultats concernant le corpus de Naël montrent que la proportion de déterminants est plus importante dans les contextes de référence particulière et les fillers ainsi que les omissions sont quant à eux plus fréquents lorsque la référence portée par le nom est indéterminée. La proportion de formes adultes est très nettement moindre dans ce contexte (moins de 10% des formes produites) que dans le cas des noms renvoyant à des référents particuliers (36% des formes pré-nominales).

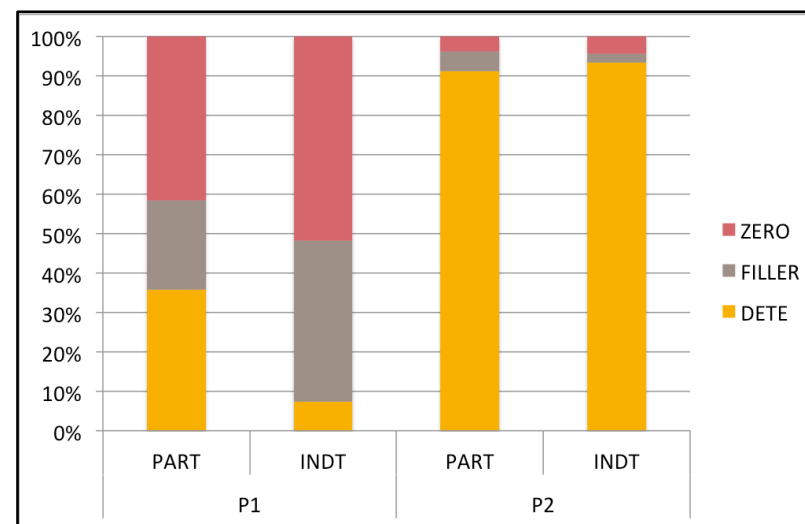
Dans le corpus d'Adrien, les différences entre les deux contextes sont moindres. La proportion de déterminants est comparable dans les cas de référence particulière et indéterminée. Les fillers sont pour leur part plus fréquents lorsque les noms renvoient à des référents indéterminés et les omissions lorsque ceux-ci renvoient à des référents particuliers.

**Graphiques 5 à 8 - Distribution des formes pré-nominales dans les contextes de référence particulière et indéterminée (chez les quatre enfants)**

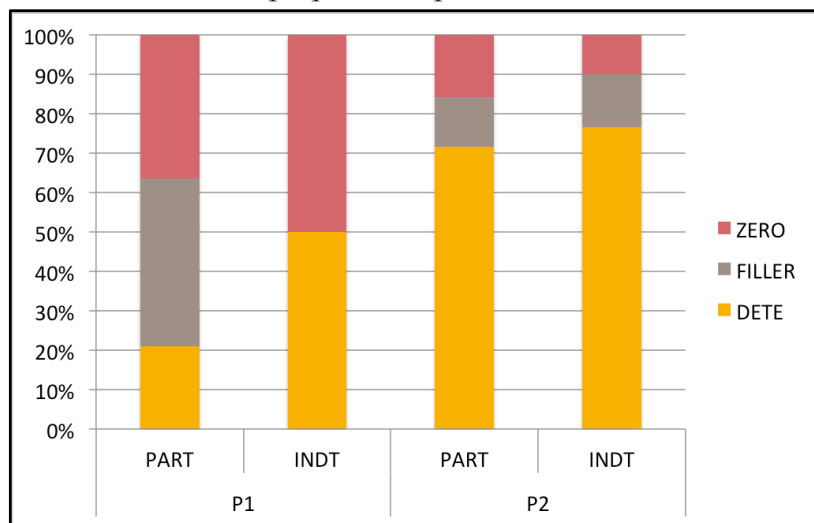
*Graphique 5 - Corpus de Lubin*



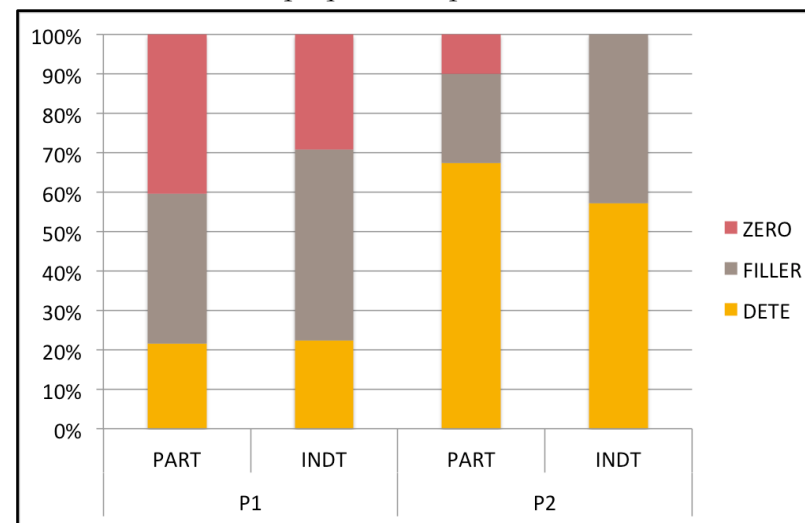
*Graphique 6 - Corpus de Naël*



*Graphique 7 - Corpus d'Antonin*



*Graphique 8 - Corpus d'Adrien*



En période 2, la distribution des formes est équivalente dans les deux contextes dans le corpus de Naël. Dans les données de Lubin et Antonin, les déterminants apparaissent plus fréquemment dans les contextes de référence indéterminée, et les omissions lorsque les noms sont associés à des référents particuliers. Les différences sont en apparence plus marquées dans le corpus d'Adrien. Les déterminants sont moins souvent produits dans les contextes de référence indéterminée, contrairement aux fillers qui sont plus fréquents dans ces contextes, et les omissions n'apparaissent qu'avec les noms renvoyant à des référents particuliers. Cependant, comme dans les corpus de Lubin et Antonin en période 1, seuls quelques noms (7) sont employés ici pour renvoyer à des référents indéterminés. Nos observations peuvent de fait difficilement donner lieu à des conclusions fiables concernant des contrastes éventuels dans l'usage qu'Adrien fait des formes pré-nominales dans les deux contextes.

Contrairement à ce que nous avons pu observer avec l'opposition entre syntagmes référentiels et non référentiels, aucun des deux contextes analysés dans cette section - référence particulière *vs.* indéterminée - ne semble donc entraîner la production privilégiée d'une forme pré-nominale qui soit la même pour tous les enfants. Outre la variabilité individuelle dans la distribution des trois types de formes, nous avons également pu relever que la faible fréquence des noms renvoyant à des référents indéterminés rendait difficile la comparaison des contextes d'une part et des usages des quatre enfants d'autre part.

Dans la continuité de ce que nous avons fait lors de l'analyse de l'opposition entre syntagmes nominaux référentiels et non référentiels, nous avons observé la distribution des différents types de déterminants et de fillers selon le caractère indéterminé ou particulier du référent. Plusieurs hypothèses sont à envisager ici. En premier lieu, nous pouvons penser que lorsque l'enfant en produit, les indéfinis seront plus souvent utilisés que les définis dans les contextes de référence indéterminée. Par ailleurs, dans les contextes de référence particulière, les définis, qui sont plus fréquents dans nos données, pourraient être plus utilisés que les indéfinis. Deux hypothèses divergentes peuvent être considérées en ce qui concerne l'usage des indéfinis dans les contextes de référence particulière. Si l'on s'en tient aux constats de Karmiloff-Smith (1979) ou Hickmann (notamment, 1996), nous pouvons imaginer qu'ils seront moins produits dans ce contexte que dans les cas de référence indéterminée. A l'inverse, les observations de Salazar Orvig (2005,

2013), Rozendaal (2008), ou Gundel (2007) nous permettent d'envisager que l'indéfini pourrait être d'emblée utilisé de façon plurifonctionnelle et ne serait ainsi pas réservé en priorité aux cas de référence indéterminée. Quant aux fillers, nous pourrions notamment faire l'hypothèse - en particulier en période 2 - d'un usage différencié des fillers vocaliques et des fillers contenant des consonnes dans ces deux contextes de référence particulière et indéterminée. Mais nous ne pouvons toutefois pas avancer que les fillers vocaliques (phonologiquement plus proches des indéfinis que les fillers C ou CV) pourraient, à l'inverse des fillers consonantiques, être réservés avant tout au contexte de référence indéterminée, d'une part parce que comme les indéfinis, ils pourraient être associés à des référents dont les caractéristiques pragmatico-discursives sont variées, et d'autre part parce qu'ils pourraient être utilisés à la fois comme des précurseurs de définis et d'indéfinis. Leurs usages dans nos données ne peuvent donc pas *a priori* être comparés davantage aux usages attendus d'indéfinis qu'à ceux attendus pour les définis.

Les premiers résultats (*cf.* tableau 5) montrent que dans le corpus de Lubin, les deux référents indéterminés qui sont en période 1 verbalisés par des noms - sur six au total si l'on inclut ceux produits avec des omissions - sont associés pour l'un à un indéfini (le seul observé ici) et pour l'autre à un filler CV. Dans les cas de référence particulière, les fillers vocaliques apparaissent comme la forme préférentielle. Les seules formes adultes relevées dans ce contexte durant la première période sont des définis (11 occurrences).

**Tableau 5 - Distribution des différents types de déterminants et fillers dans les contextes de référence particulière et indéterminée (dans le corpus de Lubin)**

Période	Déterminant/ Filler	Particulière	Indéterminée
1	Défini	11	---
	Indéfini	---	1
	V	16	---
	CV	3	1
2	Défini	93	14
	Indéfini	14	36
	V	9	7
	CV	8	---

En deuxième période, on remarque que les définis sont beaucoup plus utilisés en référence particulière qu'indéterminée (93 occurrences contre 14). La majorité des occurrences de définis observées en contexte de référence indéterminée (11 sur 14) sont des occurrences du défini



pluriel ‘les’ associées au lexème ‘dessin animé’. En revanche, les indéfinis sont plus utilisés avec les noms qui renvoient à des référents indéterminés (36 occurrences contre 14). Comme les formes adultes, les fillers vocaliques apparaissent dans les deux contextes, et seulement en nombre légèrement supérieur dans les contextes de référence particulière (9 occurrences) qu’indéterminée (7 occurrences). Tous les syntagmes de type ‘filler CV + nom’ relevés lorsque les noms ont une valeur référentielle renvoient à des référents particuliers.

Dans les données de Naël (*cf.* tableau 6), comme dans celles de Lubin, les noms porteurs d’une valeur référentielle sont surtout associés à des référents particuliers. En première période, la référence indéterminée est surtout marquée par des syntagmes de type ‘filler V + nom’ et en deuxième période, par des syntagmes indéfinis. Les indéfinis ne sont cependant pas exclusivement réservés aux cas de référence indéterminée puisque presque la moitié d’entre eux (43%) sont produits lorsque le référent est particulier.

**Tableau 6 - Distribution des différents types de déterminants et fillers dans les contextes de référence particulière et indéterminée (dans le corpus de Naël)**

Période	Déterminant/ Filler	Particulière	Indéterminée
1	Défini	35	1
	Indéfini	1	---
	Possessif	24	---
	V	23	9
	C	1	---
	CV	21	2
2	Défini	66	1
	Indéfini	27	36
	Possessif	59	---
	V	2	---
	C	2	---
	CV	3	1

Le défini est majoritaire dans les contextes de référence particulière, en période 1 comme en période 2. Les possessifs, ainsi que les fillers vocaliques et CV sont fréquents également dans ce contexte (en période 1) puisque chacune de ces catégories représente entre 20 et 23% des formes référencées ici. Conformément à nos précédentes observations, les fillers sont quasiment absents en période 2. Et outre l’emploi dominant du défini, le possessif est exclusivement produit en contexte de référence particulière.

Nous avons vu plus tôt que dans le corpus d'Antonin, seuls deux noms renvoient à des référents indéterminés en première période. L'un d'entre eux est précédé d'une omission, et l'autre d'un déterminant défini (cf. tableau 7 ci-dessous). Les données ne montrent qu'une augmentation légère de ces cas de référence indéterminée en période 2. Les syntagmes indéfinis sont alors privilégiés pour les verbaliser (8 cas sur 13 au total).

**Tableau 7 - Distribution des différents types de déterminants et fillers dans les contextes de référence particulière et indéterminée (dans le corpus d'Antonin)**

Période	Déterminant/ Filler	Particulière	Indéterminée
1	Défini	28	1
	Indéfini	2	---
	V	43	---
	C	1	---
	CV	15	---
2	Défini	173	1
	Indéfini	13	8
	V	17	1
	C	3	---
	CV	11	3

Dans les cas de référence particulière, les fillers vocaliques sont majoritaires en première période (devant les définis et les fillers CV), et les définis sont largement privilégiés au cours de la période 2. Les fillers vocaliques, de même que les indéfinis, sont surtout produits pour renvoyer à des référents particuliers. Comme dans les deux corpus précédents, l'emploi des indéfinis n'est donc pas réservé exclusivement aux cas de référence indéterminée.

Dans le corpus d'Adrien (tableau 8 ci-dessous), les cas de référence indéterminée sont surtout verbalisés en période 1 au travers de syntagmes de type 'filler V + nom'. Dans les contextes de référence particulière, les fillers CV sont privilégiés par l'enfant, devant les fillers vocaliques. Contrairement à ce que nous venons de voir dans les trois premiers corpus, les indéfinis sont plus nombreux dans ce contexte que les définis (15 occurrences contre 6). Ces résultats concordent avec les constats faits au chapitre 5, puisque nous avons vu qu'Adrien produisait une part plus importante d'indéfinis en période 1 que les trois autres enfants.

**Tableau 8 - Distribution des différents types de déterminants et fillers dans les contextes de référence particulière et indéterminée (dans le corpus d'Adrien)**

Période	Déterminant/ Filler	Particulière	Indéterminée
1	Défini	6	1
	Indéfini	15	4
	V	34	33
	C	8	---
	CV	75	9
2	Défini	38	---
	Indéfini	19	20
	V	5	2
	C	3	---
	CV	18	---

En période 2, les référents indéterminés sont surtout verbalisés par des syntagmes indéfinis. À l'inverse, le défini est majoritaire dans les contextes de référence particulière, devant les indéfinis et les fillers CV. Ces deux types de formes représentent chacune presque un quart des productions en contexte de référence indéterminée (respectivement, 23% et 22% des formes produites).

Pour faire la synthèse des résultats que nous venons de présenter, nous pouvons dire d'une part qu'en période 1 comme en période 2, les définis (et les possessifs chez Naël) sont surtout produits dans des contextes de référence particulière. Les indéfinis sont quant à eux privilégiés en contexte de référence indéterminée. Toutefois, ce contraste n'a pu être clairement observé qu'en période 2, lorsque l'indéfini devient plus fréquent dans les productions des quatre enfants. Par ailleurs, si l'indéfini est la forme préférentielle pour marquer la référence indéterminée en période 2, nous avons pu voir que celui-ci était aussi largement employé dans les contextes de référence particulière. Ces résultats suggèrent que les premiers indéfinis ne seraient pas exclusivement associés aux cas de référence indéterminée mais pourraient être au contraire d'emblée utilisés de façon plurifonctionnelle. Cette question sera développée dans les deux sections qui suivent. Comme les indéfinis, les fillers vocaliques sont représentés aussi bien dans les contextes de référence particulière qu'indéterminée, ce qui pourrait confirmer que les enfants y ont recours dans nos données aussi bien comme précurseurs de définis que d'indéfinis. En revanche, les fillers C et CV sont produits plus fréquemment en contexte de référence particulière et

apparaissent moins souvent en contexte de référence indéterminée que les fillers vocaliques. Ces fillers consonantiques pourraient donc être considérés davantage comme des proto-formes de définis que d'indéfinis.

Enfin, signalons que seuls 15 référents ont été identifiés comme génériques dans nos données. Le seul cas observé dans le corpus de Lubin et Antonin est verbalisé par un syntagme de type 'filler + nom'. Les sept cas relevés chez Naël sont tous des syntagmes définis. Dans les données d'Adrien, les six référents caractérisés comme génériques sont mentionnés dans le discours au travers de quatre syntagmes définis et deux noms précédés d'omissions. Le peu de cas observés ne nous permet pas de faire de constats solides concernant le maniement des formes pré-nominales dans les contextes de référence générique. On pourra toutefois souligner que lorsque les enfants verbalisent ces référents au moyen de syntagmes adultes, ils ont recours au défini, forme qui semble être privilégiée par les adultes dans ce contexte (*cf.* exemples 13 ci-dessous).

### **Exemple 13 - Cas de référents génériques verbalisés par l'adulte au moyen de syntagmes définis**

*Extrait du corpus de Lubin à 2;5*

Père : c'est pour (.) pouvoir bouger la paille.

Lubin : [apa buge də paj] *ʒʒʒ bouger də@fs paille ?*

Père : oui la fourche c'est pour euh pour pouvoir attraper la paille.

*Extrait du corpus de Naël à 1;11*

Naël : [se de fwal] *'c'est des poils.'*

Mère : oui [/] oui [/] oui mais on dit des poils pour les chiens.

Naël : [se de fwal də puʁ le ʃɛ̃] *'c'est des poils ʒʒʒ pour les ch(i)ens.'*

*Extrait du corpus d'Antonin à 2;7*

Mère : c'est pour monter sur quoi les chevaliers ?

Antonin (2;7) : [e puʁ əl ʃəal le] *ʒʒʒ pour ək@fs che(v)al ʒʒʒ.'*

Mère : eh oui pour le cheval.

Antonin : [e si tu] *'et c'est tout.'*

Mère : et c'est tout oui.

### **1.2.3 Distinction entre référent nouveau et donné dans le discours**

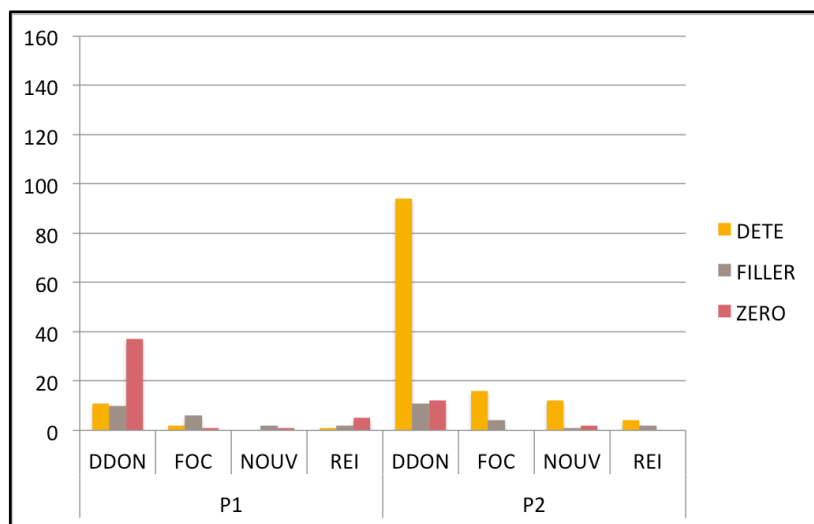
Suite à l'analyse des oppositions entre syntagmes référentiels et non référentiels, ainsi qu'entre contextes de référence particulière et indéterminée, nous avons cherché à évaluer si les noms renvoyant à des référents particuliers pouvaient être plus fréquemment associés à une forme pré-

nominale qu'à une autre selon le statut attentionnel et discursif du référent. Nous avons vu plus tôt (*cf.* section 1.3.3 du chapitre II, sur le modèle adulte de combinaisons forme-fonction) que la distinction entre un référent nouveau et donné pouvait être marquée à l'oral par un emploi contrasté des déterminants. On sait en effet que lorsque le référent est verbalisé par un syntagme nominal, le syntagme défini est privilégié pour les référents déjà mentionnés, donnés dans le discours, et l'indéfini pour les référents qui n'ont pas encore été verbalisés et sont donc nouveaux dans le discours. En nous fondant sur cette observation dans la langue adulte, nous nous sommes interrogée sur l'existence de telles oppositions entre les formes que l'enfant produit avant de maîtriser pleinement toutes les formes de déterminants et leur plurifonctionnalité. Autrement dit, est-ce qu'omissions, fillers et déterminants semblent être dans nos données préférentiellement associés à un statut attentionnel et discursif spécifique ? Par ailleurs, si des contrastes existent, s'agit-il des mêmes contrastes au cours des deux périodes ?

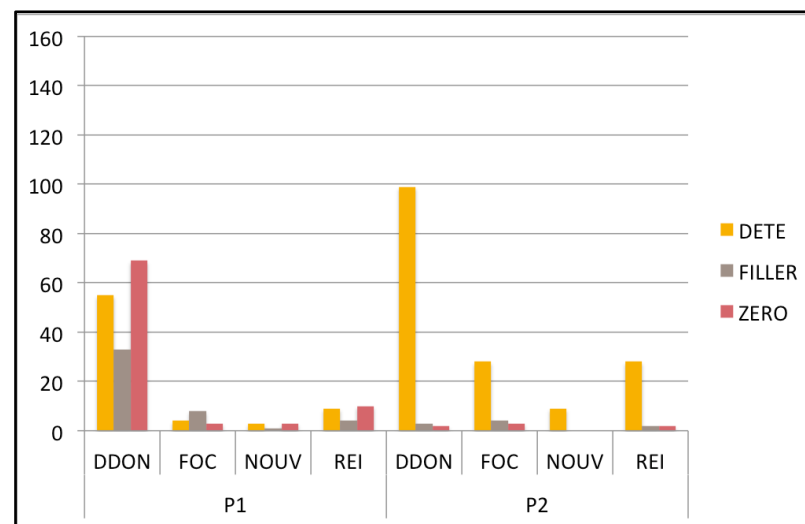
Rappelons que pour répondre aux questions que nous venons de soulever, nous avons distingué les cas où le référent est nouveau dans le discours (NOUV) et ceux où il a déjà été mentionné (DDON) mais aussi les cas où il est dit 'focalisé' (FOC) - donc non mentionné mais sous l'attention des interlocuteurs - et ceux dans lesquels le référent est réintroduit après un changement thématique (REI). Les graphiques ci-dessous (9 à 12) nous permettront de discuter de la distribution des omissions, fillers et déterminants dans ces quatre contextes chez les quatre enfants. Dans notre corpus, les référents particuliers sont pour une grande majorité des référents déjà donnés dans le discours. Les référents particuliers et nouveaux, focalisés ou réintroduits sont en parallèle très peu représentés. Pour éviter de baser nos observations sur des pourcentages ne représentant que quelques occurrences de noms, nous avons préféré faire figurer dans ces graphiques les valeurs absolues.

**Graphiques 9 à 12 - Distribution des formes pré-nominales selon le statut attentionnel et discursif des référents particuliers (dans les quatre corpus)**

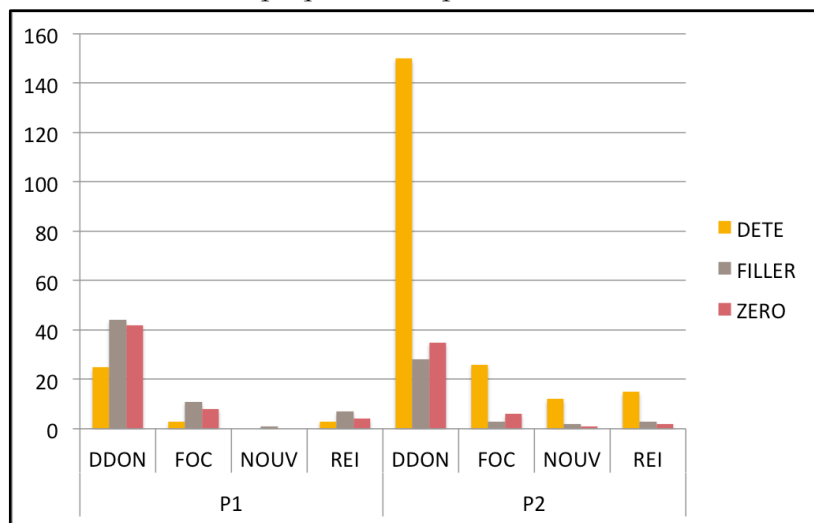
*Graphique 9 - Corpus de Lubin*



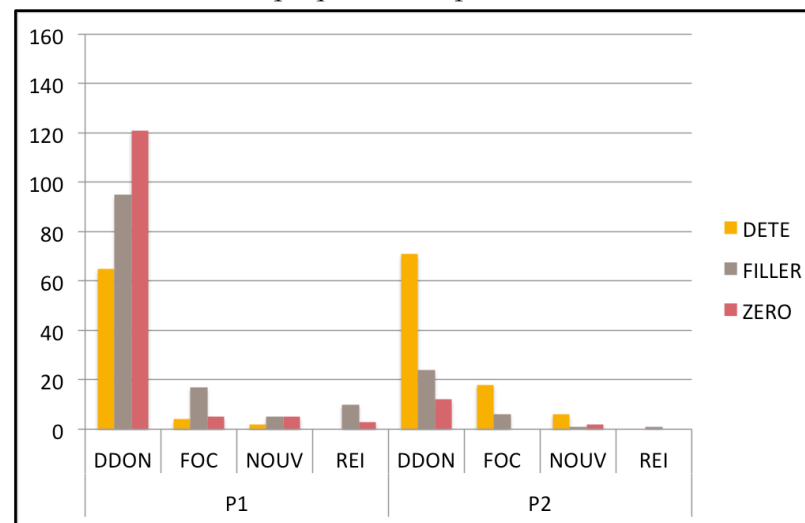
*Graphique 10 - Corpus de Naël*



*Graphique 11 - Corpus d'Antonin*



*Graphique 12 - Corpus d'Adrien*



Outre la prédominance des noms renvoyant à des référents donnés dans le discours, les quatre graphiques ci-dessus font apparaître qu'en P1 comme en P2, la répartition des formes dans ces contextes où le référent est donné coïncide avec le développement global de la position pré-nominale. Les formes de transition sont majoritaires durant la première période et les déterminants dans la seconde. Signalons cependant que comme nos analyses reposent sur le sous-ensemble des noms produits dans des contextes référentiels et associés à des référents particuliers, la part des fillers et des déterminants est plus importante chez Naël, Antonin et Adrien que lorsque l'on observe la distribution de toutes les formes produites durant cette période.

Dans les rares cas où le référent est focalisé, les quatre enfants ont davantage recours aux fillers qu'aux deux autres formes. Dans le corpus de Lubin, 6 noms sur 9 sont précédés d'un filler, 8 sur 15 dans le corpus de Naël, 11 sur 22 dans celui d'Antonin et 17 sur 26 dans les données d'Adrien. Bien que le nombre limité d'occurrences ne permette pas de l'affirmer avec certitude, il semblerait donc que les contextes où le référent est focalisé favorisent la production des fillers.

Les quelques occurrences de noms renvoyant à des référents nouveaux nous ont seulement permis de constater que conformément à ce qui peut être attendu en période 1, les formes de transition sont dans ce contexte plus fréquentes que les déterminants. La forme privilégiée - omission ou filler - est par ailleurs variable d'un corpus à l'autre.

En ce qui concerne les référents réintroduits, les fillers sont privilégiés dans les corpus d'Antonin et d'Adrien, et les omissions dans ceux de Lubin et de Naël. Cette préférence concorde avec celle observée dans les cas où le référent est donné, sauf dans le cas d'Adrien qui produit proportionnellement plus de fillers avec les noms associés à des référents réintroduits que donnés (pour lesquels l'omission est privilégiée).

Conformément à ce que nous pouvions attendre, les déterminants sont majoritaires dans tous les contextes en période 2. Dans le corpus d'Antonin, les trois formes sont produites (en quantité variable) dans tous les contextes. Chez Naël, les noms associés à des référents nouveaux (9 au total en période 2) sont tous précédés de déterminants et quelques occurrences de fillers et d'omissions sont encore observables dans les trois autres contextes. Dans les données de Lubin, de même que dans celles d'Adrien, les omissions n'apparaissent qu'avec les noms associés aux référents donnés et nouveaux.

Pour conclure sur ce point, nous pouvons dire qu'à l'exception de l'emploi majoritaire de fillers que nous avons pu observé en période 1 devant les noms renvoyant à des référents focalisés, aucune association claire et identique dans les quatre corpus ne se dessine entre formes et statuts attentionnels et discursifs. Ce constat reste toutefois à mettre en perspective avec le peu de données que nous avons pu analyser dans les contextes autres que ceux où le référent est donné.

Nous nous sommes par la suite focalisée spécifiquement sur les productions de déterminants et de fillers dans chacun des contextes que nous venons d'observer. Au vu de la très nette dominance des référents donnés, nous pouvons imaginer que toutes les formes seront proportionnellement plus fréquentes dans ce contexte que dans les trois autres. Toutefois, si l'on s'appuie sur les résultats de certains travaux déjà cités au sein de ce chapitre (*cf.* par exemple Rozendaal, 2008 ; Rozendaal et Baker, 2008 ; Salazar Orvig *et al.*, 2013), nous pouvons nous attendre à observer plus d'indéfinis que de définis lorsque le référent est nouveau et à l'inverse, plus de définis que d'indéfinis lorsque le référent a déjà été donné dans le discours. Bien que les référents dits 'focalisés' n'aient pas encore été verbalisés, ils sont sous l'attention des interlocuteurs donc on peut imaginer qu'ils seront plus associés à des syntagmes nominaux définis que les référents nouveaux. Enfin, les référents réintroduits ayant déjà été verbalisés, il serait envisageable que les noms y renvoyant soient davantage précédés de définis que d'indéfinis. Selon le corpus et les précédentes observations faites au sujet du statut des fillers vocaliques, nous pourrions nous attendre à ce que ceux-ci soient utilisés soit en suivant des tendances comparables aux indéfinis (ou à ce qui serait attendu des indéfinis), soit de façon indifférenciée, dans des contextes en théorie privilégiés pour les indéfinis mais aussi dans ceux susceptibles d'être privilégiés pour les définis. Si l'on part de l'idée que les fillers contenant des consonnes cibleraient - au moins en période 2 - plutôt des définis que des indéfinis, nous pouvons faire l'hypothèse qu'ils seront produits en contraste avec les indéfinis et/ou les fillers vocaliques.

Dans le corpus de Lubin, nous pouvons constater (*cf.* tableau 9 ci-dessous) qu'en période 1, le défini et le filler vocalique sont utilisés à part égale lorsque le référent est donné dans le discours. Le filler vocalique est par ailleurs la forme la plus utilisée lorsque le référent est focalisé mais le nombre de cas de référents identifiés comme focalisés étant trop limité, nous ne pouvons pas en



tirer de conclusions fiables. Les seuls cas de référents nouveaux que nous avons pu observer sont associés pour l'un à un nom précédé d'un filler vocalique, et pour l'autre à un syntagme de type 'filler CV + nom'. Lorsque le référent est réintroduit, nous avons relevé deux occurrences de fillers vocaliques et une occurrence de défini. En dernier lieu, rappelons que dans ce corpus, aucun indéfini n'a été produit en référence particulière au cours de la période 1, d'où l'absence de ligne consacrée à l'indéfini dans le tableau.

**Tableau 9 - Distribution des déterminants définis et indéfinis selon le statut attentionnel et discursif du référent (dans le corpus de Lubin)**

Période	Type de déterminant/ Filler	Donné	Focalisé	Nouveau	Réintroduit
1	Défini	8	2	---	1
	V	8	5	1	2
	CV	2	---	1	---
2	Défini	67	12	10	3
	Indéfini	12	2	---	---
	V	7	2	---	---
	CV	4	2	---	2

En période 2, le défini est privilégié dans tous les contextes, y compris lorsque le référent est nouveau. Ce résultat qui peut sembler de prime abord surprenant sera discuté plus en détail dans la section suivante (1.2.4). De même, il peut sembler inattendu d'observer des indéfinis lorsque le référent a été donné dans le discours, mais ces cas s'expliquent surtout par les choix que nous avons faits lors de la catégorisation du statut attentionnel et discursif des référents. Ces choix sont illustrés au travers des deux exemples ci-dessous.

**Exemples 14 - Emplois d'indéfinis dans des contextes où le référent est donné (dans le corpus de Lubin)**

- (a) Père : c'est un ver de terre !  
Lubin (2;5) : [ɛ̃ vɛ də tɛʁ i mɑ̃ʒ] 'un ver de terre i(l) mange ?'
- (b) Mère : qu'est ce qu'i(l) y a dans l'enveloppe ?  
Lubin (2;9) : [de putɔ̃] 'des boutons.'

Dans l'exemple *a*, Lubin et son père sont en train de lire un livre. Suite à une dénomination du père ('c'est un ver de terre'), Lubin reprend partiellement l'énoncé du père mais le syntagme 'un ver de terre' est cette fois utilisé de façon référentielle. Dans ces cas de reprise d'un nom d'abord produit avec une valeur non référentielle, nous avons considéré que le référent avait déjà été

donné dans le discours. Par ailleurs, dès lors que le référent est individualisé (ici, le ver de terre représenté sur le livre), nous avons considéré que quand bien même il était peu spécifié, le référent n'était pas pour autant indéterminé. C'est également le cas dans l'exemple *b*. La mère demande à l'enfant d'identifier les éléments du jeu dans l'enveloppe et Lubin répond par le syntagme indéfini 'des boutons'. Nous avons considéré qu'il s'agissait ici de l'ensemble particulier des pièces du jeu et que la référence n'était donc pas indéterminée. En outre, rappelons que le référent a été identifié comme donné dans le discours dans la mesure où il est introduit par l'interrogatif produit par la mère dans l'énoncé précédent. Si la présence d'indéfinis dans les contextes où le référent a déjà été donné dans le discours peut paraître inattendue, elle n'en est donc pas moins conforme à ce qui pourrait être attendu dans la langue adulte.

Dans le corpus de Naël (*cf.* tableau 10 ci-dessous), le défini et le possessif sont privilégiés en période 1 dans les contextes où le référent est donné. Toutefois, les fillers (V et CV) représentent également une part notable des formes observées dans ce contexte (environ 20% des productions). Conformément à ce que nous avons vu plus tôt, les fillers (de nouveau, V et CV) sont majoritaires lorsque le référent est sous l'attention des interlocuteurs (focalisé) mais n'a pas encore été mentionné. Nous n'avons relevé que quatre référents nouveaux verbalisés par des syntagmes nominaux, deux d'entre eux par des syntagmes contenant des possessifs et les deux autres, par des syntagmes de type 'filler + nom' (un filler V et un filler CV).

**Tableau 10 - Distribution des déterminants définis et indéfinis selon le statut attentionnel et discursif du référent (dans le corpus de Naël)**

Période	Type de déterminant/ Filler	Donné	Focalisé	Nouveau	Réintroduit
1	Défini	23	2	1	7
	Indéfini	1	---	---	---
	Possessif	21	---	2	1
	V	15	4	---	4
	C	1	---	---	---
	CV	17	3	1	---
2	Défini	39	8	5	13
	Indéfini	21	4	2	1
	Possessif	35	11	1	13
	V	1	1	---	---
	C		2	---	---
	CV	2	---	---	---

En période 2, nous pouvons noter que comme dans le corpus de Lubin, le défini est non seulement majoritaire lorsque le référent est donné mais également lorsqu'il est nouveau. Le possessif est la forme la plus observée - devant les définis - dans les contextes où le référent est focalisé. Les référents sont presque exclusivement réintroduits avec des syntagmes contenant des définis et des possessifs. Enfin, on remarque que dans ce corpus également, l'indéfini est surtout produit lorsque le référent est donné. L'observation des énoncés concernés nous a permis de constater qu'il s'agissait dans la plupart des cas de reprises - partielles ou totales - par l'enfant d'énoncés précédents produits par l'adulte et introduisant des référents nouveaux ou focalisés. Ces cas de figure sont illustrés au travers des deux exemples ci-dessous.

**Exemples 15 - Emplois d'indéfinis dans des contextes où le référent est donné (dans le corpus de Naël)**

- (a) Père : regarde Naël là y a une p(e)tite mousse par dessus.  
 Naël (1;10) : [t a yn mu dəs] 'yy a une mou(sse) des(sus) +/.'
- (b) Père : il lui rend [///] il lui ramène un panier.  
 Naël (2;6) : [ẽ tu pti panje] 'un tout p(e)tit panier ?'

Dans ces exemples *a* et *b*, les référents 'mousse' et 'panier' sont verbalisés pour la première fois par le père et ces énoncés sont ensuite repris par Naël au tour suivant, avec la même intention communicative.

Dans le corpus d'Antonin (*cf.* tableau 11 ci-dessous), les référents donnés dans le discours sont verbalisés en période 1 d'abord par des syntagmes de type 'filler V + nom', puis par des syntagmes définis. Le filler vocalique est la forme privilégiée dans des contextes où le référent est focalisé et réintroduit. Par ailleurs, dans la mesure où un seul référent a été identifié comme nouveau, il nous est impossible de dire ici si l'enfant privilégie une forme plus qu'une autre dans les contextes où le référent est nouveau.

Comme dans le corpus de Lubin, le défini est majoritaire dans tous les contextes en période 2. S'il peut sembler attendu que le défini soit la forme la plus utilisée lorsque le référent est donné, focalisé ou réintroduit, il peut ici aussi paraître surprenant que l'enfant n'ait pas davantage recours à des syntagmes indéfinis pour verbaliser les référents nouveaux, et qu'il utilise par ailleurs des syntagmes indéfinis surtout dans les cas où le référent a déjà été mentionné dans le discours.

**Tableau 11 - Distribution des déterminants définis et indéfinis selon le statut attentionnel et discursif du référent (dans le corpus d'Antonin)**

Période	Type de déterminant/ Filler	Donné	Focalisé	Nouveau	Réintroduit
1	Défini	23	2	---	3
	Indéfini	1	1	---	
	V	29	8	---	6
	C	1		---	
	CV	11	2	1	1
2	Défini	126	23	8	16
	Indéfini	10	1	1	---
	V	12	2	1	2
	C	3	---	---	---
	CV	9	1	1	---

Rappelons que les emplois de définis dans les contextes où le référent est nouveau seront traités dans la section qui suit. En ce qui concerne les emplois de syntagmes indéfinis pour verbaliser des référents nouveaux, nous avons pu relever en observant les dix énoncés concernés qu'il ne s'agissait pas ici non plus d'usages inadéquats sur le plan pragmatique, mais plutôt de productions liées à des contextes discursifs spécifiques ou du résultat de certains choix méthodologiques de notre part. Dans l'exemple 15a ci-dessous, Antonin et sa mère sont en train de construire une caserne de pompiers. L'enfant cherche une pièce représentée sur la notice de montage. Le référent a été considéré ici comme introduit par l'énoncé de la mère dans la question 'qu'est-ce que tu cherches?', puis mentionné une première fois par l'enfant dans sa réponse 'je cherche le truc là le noir'. La mère demande ensuite confirmation à Antonin ('le noir?') et l'enfant confirme en répondant par le syntagme indéfini 'un truc noir là'. Nous pouvons imaginer ici que bien que le référent ait été donné dans les énoncés précédents, l'enfant tienne compte du fait que la mère n'ait pas d'emblée compris à quelle pièce il faisait référence et qu'ainsi, il reformule son énoncé en réintroduisant le référent par un syntagme indéfini.

**Exemples 16 - Emplois d'indéfinis dans des contextes où le référent est donné (dans le corpus d'Antonin)**

- (a) Mère : qu'est ce que tu cherches ?  
 Antonin (2;11) : [zə sɛʁsə lə tʁyk la lə nwaʁ] 'je cherche le truc là le noir.'  
 Mère : le noir ?  
 Antonin [wi] 'oui.'  
 Antonin : [ɛ̃ tʁyk nwaʁ la] 'un truc noir là.'

- (b) Antonin (3;2) : [e mɛ la ð̃ a de plas la] *'eh mais là on a des places là.'*  
 Antonin : [e ã pø de plas la] *'et un peu des places là.'*

Dans le deuxième exemple (15b), Antonin joue avec ses voitures et cherche à les ranger dans le garage. Les référents associés au syntagme 'des places' produit par l'enfant dans le deuxième énoncé n'ont pas été mentionnés au préalable, ni comme un tout ni de façon individuelle. L'enfant fait référence à d'autres places que celles qu'ils vient de cibler dans l'énoncé précédent, et nous aurions donc pu considérer le syntagme 'des places' comme renvoyant à des référents nouveaux dans le discours. Toutefois, dans la mesure où d'autres référents du même ensemble général 'places' sont verbalisés dans l'énoncé précédent, nous avons caractérisé les référents du deuxième énoncé comme donnés dans le discours. Bien que ces cas ne soient pas très fréquents, il convient ici de rappeler que les productions de l'enfant ne sont pas inadéquates et que les résultats présentés sont à interpréter à la lumière de certaines de nos décisions méthodologiques.

Chez Adrien comme chez les trois autres enfants, toutes les formes sont majoritairement produites dans des contextes où le référent est donné, en période 1 comme en période 2 (cf. tableau 12 ci-dessous). Durant la première période, le filler CV est la forme la plus fréquente dans ce contexte, devant les définis puis les fillers vocaliques. Les syntagmes indéfinis sont donc ici aussi employés surtout pour verbaliser des référents déjà mentionnés (14 occurrences au total en période 1). On note par ailleurs que le filler vocalique est majoritaire lorsque les référents sont dits 'focalisés'.

**Tableau 12 - Distribution des déterminants définis et indéfinis selon le statut attentionnel et discursif du référent (dans le corpus d'Adrien)**

Période	Type de déterminant/ Filler	Donné	Focalisé	Nouveau	Réintroduit
1	Défini	36	1	1	---
	Indéfini	14	3	1	---
	V	21	10	1	2
	C	7	---	---	1
	CV	61	6	3	5
2	Défini	48	9	4	---
	Indéfini	8	6	1	---
	V	3	2	---	---
	C	2	1	---	---
	CV	14	3	1	---

En période 2, les définis sont surtout employés dans les contextes où le référent est donné, mais ils apparaissent également comme majoritaires dans les cas où le référent est focalisé et nouveau. Les indéfinis sont peu nombreux et produits non pas dans des contextes où le référent est nouveau mais lorsqu'il est donné ou focalisé. Cette fois encore, nous nous sommes intéressée aux syntagmes indéfinis verbalisant des référents déjà mentionnés et nous avons pu constater que ces résultats s'expliquaient en majeure partie par des reprises par l'enfant de son propre discours, comme dans les exemples 16a et 16b ci-dessous.

**Exemples 17 - Emplois d'indéfinis dans des contextes où le référent est donné (dans le corpus d'Adrien)**

- (a) Adrien (2;11) : [d a de su] *'yyy a des sous !'*  
Adrien : [a de su] *'a des sous !'*
- (b) Adrien (3;5) : [a vy a fe de gɔn] *'as vu yyy fait des gants !'*  
Adrien : [a vy a fe de gɔn] *'as vu yyy fait des gants !'*

Dans les deux cas, les référents mentionnés dans le deuxième énoncé ('sous' et 'gants') ont été analysés comme donnés mais ils sont verbalisés dans le cadre de productions reprenant à l'identique l'intention communicative de l'enfant dans l'énoncé précédent.

Nos résultats concernant les usages que les enfants font des fillers et des déterminants en fonction du statut attentionnel et discursif du référent ne semblent donc pas aller dans le sens d'un usage contrasté des formes. Les référents sont très majoritairement donnés dans le discours et toutes les formes pré-nominales sont plus souvent produites dans ce contexte que dans les trois autres. Conformément à nos hypothèses, le défini est la forme privilégiée lorsque le référent a déjà été mentionné dans le discours mais de façon plus inattendue, l'enfant a aussi recours au défini plus qu'à l'indéfini dans les cas où le référent est nouveau. Nous tenterons de voir au point suivant si ces emplois peuvent être expliqués par le fait que les interlocuteurs soient tous deux familiers du référent mentionné par l'enfant. Nous avons par ailleurs remarqué que comme les autres formes pré-nominales, l'indéfini apparaît surtout dans des contextes où le référent a été donné dans le discours. L'observation des énoncés concernés par ces productions nous a permis de constater que ces résultats inattendus n'étaient pas pour autant pragmatiquement inadéquats. Nous avons notamment vu qu'ils pouvaient s'expliquer d'une part par nos choix de codage, et d'autre part, par le fait que l'enfant reprenne - de son propre discours ou de celui de son

interlocuteur - des syntagmes indéfinis introduisant au tour ou dans l'énoncé précédents de nouveaux référents.

#### 1.2.4 Impact du degré de connaissance partagée entre les interlocuteurs

En dernier lieu, nous nous sommes intéressée à la relation entre la nature des formes produites devant les noms et la connaissance que l'interlocuteur pouvait avoir du référent verbalisé par l'enfant. D'une part, nous avons souhaité observer si ce facteur pouvait influencer le type de forme produite, et d'autre part, nous avons cherché à voir si nos données confirmaient les conclusions des travaux ayant fait le constat d'une prise en compte tardive de la connaissance de l'interlocuteur par l'enfant (*i.a* Karmiloff-Smith, 1979 ; Kail et Hickmann, 1992 ; Hickmann, 2003). Pour cela, nous nous sommes focalisée sur les syntagmes nominaux associés à des référents particuliers et nouveaux et nous avons relevé les formes produites lorsque l'interlocuteur avait connaissance du référent mentionné dans le discours de l'enfant et lorsqu'au contraire, il n'en avait pas connaissance.

Dans la mesure où les syntagmes associés à des référents nouveaux (et particuliers) ne représentent qu'une part très faible du total des syntagmes produits dans notre corpus, nous avons préféré dans le tableau 13 ci-dessous discuter sur la base des valeurs absolues et par ailleurs, regrouper la présentation des types de formes produites et celle des formes spécifiques de déterminants et de fillers observés dans chacun des contextes. Les résultats sont présentés par période et pour chacun des enfants.

Dans un premier temps, nous pouvons préciser que nous n'avons relevé que très peu de cas où le référent est connu de l'enfant mais pas de son interlocuteur, et cela, quelle que soit la période observée. Ceci est lié en grande partie au fait que l'enfant interagisse avec son parent (et donc avec un interlocuteur privilégié pour le savoir partagé), et que par ailleurs, les référents verbalisés par l'enfant soient le plus souvent présents dans la situation, et par conséquent, nécessairement connus de l'interlocuteur. En période 1, le seul cas relevé (dans le corpus d'Adrien) est verbalisé par un syntagme de type 'filler + nom'. En période 2, seuls deux noms (produits pour l'un par Naël et pour l'autre par Antonin) renvoient à des référents nouveaux et identifiés comme inconnus de l'interlocuteur. Ils sont tous les deux précédés d'un déterminant défini.

**Tableau 13 - Formes pré-nominales produites par les quatre enfants dans les contextes de connaissance partagée ou non du référent entre les interlocuteurs (pour les référents nouveaux)**

Période	Enfant	Familiarité du référent (pour l'interlocuteur)	Déterminants	Fillers	Omissions
1	Lubin	Connu	---	2 (1 V, 1 CV)	1
		Inconnu	---	---	---
	Naël	Connu	3 (1 DEF, 2 POS)	1 (CV)	---
		Inconnu	---	---	---
	Antonin	Connu	---	1 (CV)	---
		Inconnu	---	---	---
	Adrien	Connu	2 (1 DEF, 1 IND)	4 (3 CV, 1 V)	5
		Inconnu	---	1 (AUTRES)	---
2	Lubin	Connu	12 (10 DEF, 2 POS)	1 (AUTRES)	2
		Inconnu	---	---	---
	Naël	Connu	7 (4 DEF, 2 IND, 1 POS)	---	---
		Inconnu	1 (DEF)	---	---
	Antonin	Connu	9 (7 DEF, 2 POS, 1 IND)	2 (1 V, 1 CV)	1
		Inconnu	1 (DEF)	---	---
	Adrien	Connu	6 (4 DEF, 1 IND, 1 POSS)	1 (CV)	2
		Inconnu	---	---	---

Ces deux cas sont présentés ci-dessous (exemples 18). Dans le premier exemple, l'enfant introduit un référent fictif ('la fille') au travers d'un syntagme défini. Ce référent n'ayant pas été mentionné au préalable dans l'échange et n'étant pas connu du père, nous aurions pu nous attendre à ce qu'il soit introduit dans la langue adulte par un syntagme indéfini. De même, dans l'exemple issu du corpus d'Antonin (et déjà été cité au préalable dans la section concernée à notre méthodologie d'analyse), la mère peut comprendre à quel référent renvoie le discours de l'enfant mais l'emploi d'un syntagme indéfini aurait été pragmatiquement plus adéquat pour introduire le référent 'chapeaux de pompier' que l'enfant est le seul à connaître.

**Exemples 18 - Syntagmes nominaux renvoyant à des référents nouveaux (dans le corpus de Naël et Antonin)**

*Extrait du corpus de Naël à 2;6*

Père : c'est ton copain ?

Naël : [nã] 'nan.'



Père : non ?

Naël : [sɛ sɛ] 'c'est [//] c'est +...'

Père : et pourtant il est gentil.

Naël : [sɛ sɛ la fij sɔ̃ kopɛ̃] '+, c'est [//] c'est la fille son copain.'

Père : ah d'acc(ord) +/.

Père : copine oui.

*Extrait du corpus d'Antonin à 3;2*

Mère : d'ailleurs t'as joué aux pompiers à la crèche aujourd'hui ?

Antonin : [e u pɔpje] 'yyy où pompier ?'

Mère : le camion de pompier non ?

Antonin : [we avɛ le fapo pɔ̃pje] 'ouais ave(c) les chapeaux pompier.'

Mère : t'as mis les chapeaux de pompier ?

Antonin : [kwi] 'oui.'

Antonin : xxx . (inaudible)

Mère : mais sur ta tête ou sur des personnages ?

Antonin : [syɁ le syɁ le X e ʔsɥit ʒɛ mi syɁ ma tɛt] 'sur les [//] sur les xxx et ensuite j'ai mis sur ma tête.'

Le peu de cas où le référent n'est pas connu de l'interlocuteur rend donc impossible notre analyse des éventuels contrastes dans les usages que les enfants font des formes lorsque le référent est connu et inconnu de l'interlocuteur. Il nous a cependant paru intéressant de relever qu'à l'exception des deux cas que nous venons de mentionner, les syntagmes définis produits dans les contextes où le référent est nouveau renvoient à des référents connus de l'interlocuteur. Les usages *a priori* inattendus que nous avons relevés au point précédent peuvent donc s'expliquer par le fait que l'adulte comme l'enfant soient familiers du référent dont il est question dans le discours. Dans ce contexte, l'emploi du défini est donc conforme à ce qui pourrait être attendu d'un point de vue pragmatique, et même plus adéquat que l'emploi d'un indéfini. Et outre le degré de connaissance partagée entre les interlocuteurs, l'observation des cas concernés a fait apparaître que le caractère unique du référent pouvait lui aussi favoriser l'emploi d'un défini dans un contexte où le référent est nouveau. C'est le cas par exemple dans l'extrait du corpus d'Antonin ci-dessous.

**Exemple 19 - Syntagme défini renvoyant à un référent nouveau (dans le corpus d'Antonin à 2;11)**

Mère : tu veux qu'on les lise ?

Antonin : [wi] 'oui.'

Mère : allez viens !

Mère : Ernest et Célestine.

Antonin : [ɔ ba wi] 'oh bab oui !'

Antonin : [e lɛ dø] 'et les deux +/.'

Antonin : [nã dã l kanapɛ] 'nan dans l(e) canapé.'

Le référent 'canapé' est introduit pour la première fois par l'enfant par un syntagme défini ('le canapé') et ceci peut s'expliquer d'une part par le fait que la mère connaisse le référent ciblé par l'enfant, et d'autre part, par le fait que l'enfant fasse référence au seul grand canapé du salon où il a l'habitude de lire et jouer avec sa mère.

### 1.3 Une sensibilité de l'enfant aux facteurs pragmatico-discursifs ?

Les résultats que nous venons de présenter suggèrent que les enfants de notre corpus semblent être sensibles à certaines caractéristiques pragmatico-discursives des référents. Toutefois, toutes les oppositions considérées dans cette première partie du chapitre ne semblent pas influencer de la même façon la nature des formes produites par l'enfant. En ce qui concerne la distribution globale des omissions, fillers et déterminants, nous avons vu par exemple que les déterminants sont plus fréquents lorsque le nom est utilisé de façon référentielle que lorsqu'il possède une valeur non référentielle, et qu'au contraire, le contexte non référentiel favorise davantage la production de formes de transition que le contexte référentiel. En revanche, nous n'avons pas constaté que les contextes de référence particulière et indéterminée généraient la production d'un type de forme plus que d'un autre. A l'exception d'une préférence pour le filler dans les contextes où le référent est focalisé, nous n'avons pas non plus relevé que la distribution des formes pré-nominales semblait être clairement déterminée par le statut attentionnel et discursif du référent. Si des contrastes existent dans l'usage que l'enfant fait des trois types de formes pré-nominales en fonction du type de référence et du statut attentionnel et discursif du référent, ceux-ci se manifesteraient différemment chez les quatre enfants. Par ailleurs, dès lors que les cas de référents nouveaux et inconnus de l'interlocuteur sont quasiment absents de nos données, nous n'avons pas pu établir si les quatre enfants avaient recours à des formes distinctes selon que le référent était connu ou non de leur interlocuteur. Cette analyse de la prise en compte de la connaissance de l'autre aurait probablement nécessité que l'enfant interagisse avec un interlocuteur moins familier que son parent.

En ce qui concerne cette fois la nature des fillers et des déterminants produits par l'enfant, nous avons pu relever que les contextes référentiels favorisaient la production de définis et de fillers

vocaliques. Dans les contextes non référentiels, les indéfinis et les fillers vocaliques sont plus fréquents que les définis et les fillers consonantiques. Ces associations ne sont toutefois pas exclusives puisque toutes les formes peuvent apparaître à la fois en contexte référentiel et non référentiel. Nous avons également constaté que le syntagme défini est privilégié pour verbaliser les référents particuliers et le syntagme indéfini pour renvoyer aux référents indéterminés. L'indéfini n'est néanmoins pas uniquement associé à la référence indéterminée puisqu'il est produit aussi dans les cas de référence particulière. Le peu de cas de référents indéterminés en première période ne nous a pas permis d'évaluer si des contrastes entre fillers et définis pouvaient déjà être observables avant l'apparition des indéfinis. En période 2, les fillers vocaliques sont produits aussi bien lorsque le référent est particulier que lorsqu'il est indéterminé. Ils sont cependant plus représentés que les fillers consonantiques dans les contextes de référence indéterminée. Les fillers C et CV semblent quant à eux être plutôt réservés aux cas de référence particulière. Si nos résultats concernant l'emploi des définis et indéfinis en contexte référentiel/non référentiel et particulier/indéterminé sont conformes à ceux de recherches antérieures (*cf.* par exemple, Gundel, 2007 ; Rozendaal, 2008 ; Salazar Orvig *et al.*, 2013), nous n'avons pas - contrairement à ces travaux - observé de contrastes dans l'usage des formes selon le statut attentionnel et discursif du référent. En première période, la grande majorité des référents sont donnés dans le discours et tous les déterminants et fillers sont davantage produits dans ce contexte que lorsque le référent est focalisé, nouveau ou réintroduit. En outre, les syntagmes définis sont plus fréquents que les indéfinis lorsque les référents sont nouveaux et les indéfinis sont produits le plus souvent dans des contextes où le référent est nouveau. Ces résultats *a priori* inattendus s'expliquent pour les uns par la connaissance partagée (entre l'enfant et l'interlocuteur) du référent, et pour les autres, par des biais méthodologiques ainsi que par des cas de reprise par l'enfant de syntagmes indéfinis introduisant de nouveaux référents.

## 2. Le rôle des reprises du nom dans l'acquisition des formes

Si la première partie du chapitre a mis en lumière l'influence de certains facteurs pragmatiques sur les usages que les enfants font des formes pré-nominales, nous avons constaté par ailleurs que parmi les exemples analysés, certains pouvaient soulever la question de la relation entre les

syntagmes nominaux produits par l'enfant et ceux apparaissant dans le discours de son interlocuteur. Considérons par exemple cet extrait du corpus d'Antonin à 2;9 :

**Exemple 20 - Reprises de syntagmes nominaux dans le corpus d'Antonin à 2;9**

Mère : oh bah la petite elle a eu un accident hein. ('la petite' renvoie ici à une voiture)

Antonin : [wɛj] 'ouais.'

Antonin : [la pətɪt ɛ va paʁ la ʁut] 'la petite e(lle) va par la route.'

Mère : e(lle) va sur la route ?

Antonin : [wi] 'oui.'

Antonin : [sa la ʁut] 'sur / ça la route.' (fait emprunter à la voiture la route dessinée sur le tapis de sol)

Dans son premier énoncé, la mère réfère à la plus petite des voitures de l'enfant par le syntagme 'la petite'. C'est la première fois que ce syntagme est observé dans la séance mais le référent a déjà été mentionné dans le discours. Au tour suivant, l'enfant reprend ce syntagme 'la petite' (ainsi que la structure disloquée) de l'énoncé de la mère et y associe une nouvelle prédication : 'e(lle) va par la route'. Le syntagme 'la route' a déjà été produit à l'identique par l'enfant quelques tours plus tôt, sans reprise de la mère. Dans cette séquence, la mère reprend et reformule l'énoncé de l'enfant, en lui offrant comme alternative au syntagme prépositionnel 'par la route' un autre syntagme prépositionnel : 'sur la route'. Dans l'énoncé qui suit, Antonin reprend le segment 'la route', précédé d'une forme qui aurait pu être transcrite comme un démonstratif 'ça', mais qui au vu de l'énoncé précédent pourrait également avoir pour cible la préposition 'sur'. Les échanges suivants ne nous ont pas permis de trancher sur le statut de cette forme.

Si cet échange entre Antonin et sa mère illustre bien ce qui vient d'être traité dans la première partie de ce chapitre, puisque les lexèmes 'petite' et 'route', associés à des référents donnés dans le discours, sont combinés à des définis, il témoigne également d'une continuité ou d'un alignement réciproque - relatifs aux noms mais aussi aux déterminants ou aux structures syntaxiques - entre le discours de l'enfant et celui de l'adulte.

Ce phénomène de continuité a d'ailleurs été largement discuté dans la littérature, tant pour ses effets sur le développement lexical, morphologique ou syntaxique de l'enfant que pour la construction de sa capacité à dialoguer avec l'autre. Dans une perspective dialogique, Salazar Orvig (2000) a par exemple montré qu'en créant des liens entre les énoncés des participants (immédiats ou à distance), les reprises aident l'enfant à se situer et à s'insérer dans le dialogue. Au

travers des reprises qu'il fait des productions de son interlocuteur, l'enfant manifesterait non seulement sa faculté à imiter mais aussi sa capacité croissante à « prendre en compte ce que l'autre a dit, à se saisir des mots d'autrui et les réutiliser dans des situations similaires » (2000 : 89). Dans leur travail de 2006, Bernicot, Salazar Orvig et Veneziano soulignent également l'importance des reprises dans la création d'un 'espace dialogique intersubjectif' et mettent par ailleurs en évidence le caractère complexe de ces reprises. Celles-ci peuvent avoir quatre fonctions majeures pour l'adulte, étiquetées comme suit par les auteures : 'demander', 'préciser la description du monde', 's'accorder sur la forme des énoncés' et 'ajuster les états mentaux des interlocuteurs'. Chez l'enfant, les reprises ont des fonctions moins variées et la dernière est largement dominante. De par ces différentes fonctions qu'elles sont susceptibles de remplir dans le discours de l'adulte et de l'enfant, les reprises joueraient donc un rôle fondamental au sein des interactions adulte-enfant.

Dans une perspective plus formelle, le travail de Bloom *et al.* (1974) sur l'imitation (basé sur un corpus d'enfants dont la LME se situe entre 1 et 2), a d'une part montré que tous les enfants n'imitent pas le discours entendu, et d'autre part, que lorsque le jeune enfant imite, il ne reprend pas ce qu'il ne comprend pas. L'imitation faciliterait l'acquisition progressive de mots et/ou de structures syntaxiques encore rarement produites mais pas absentes des contextes de production spontanée. Les auteures précisent par ailleurs que si l'enfant n'imité pas ce qui est totalement nouveau pour lui, il ne reprend pas pour autant des segments lexicaux ou syntaxiques avec lesquels il est déjà très familiarisé. En s'appuyant sur le corpus d'une enfant de 1;5.23 à 1;8.15 en interaction avec sa mère, Veneziano *et al.* (1990) ont pour leur part mis en évidence l'importance des reprises de l'enfant dans le développement des énoncés combinant deux termes reliés sémantiquement. Et en se penchant à la fois sur les reprises de l'enfant (ou énoncés imitatifs) et sur les reprises de l'adulte (interprétations, reformulations), Veneziano (notamment, 1997b) a pu montrer que les échanges de type imitatif sont favorables au développement lexical, et mêmes corrélés au rythme d'acquisition du lexique partagé entre interlocuteurs lorsqu'ils sont réciproques ('mère-enfant-mère' ou 'enfant-mère-enfant'). En ce qui concerne plus particulièrement l'acquisition des morphèmes grammaticaux, Bertin (2011) a entre autres relevé, dans le cadre de sa thèse, que les reprises adaptées et immédiates que l'adulte fait du discours et

des tâtonnements de l'enfant soutiennent le processus d'appropriation progressive des déterminants et des clitiques sujets.

Nous nous intéresserons pour notre part, dans cette deuxième partie du chapitre, aux reprises par l'enfant de syntagmes nominaux déjà mentionnés dans l'échange, ainsi qu'à l'influence de ces reprises sur la production des formes pré-nominales<sup>1</sup>. Quatre questions principales - abordant une partie des thématiques que nous venons d'évoquer - guideront notre réflexion : a) Lorsque les noms produits par l'enfant sont repris du discours de l'interlocuteur, sont-ils plus fréquemment précédés de formes (fillers ou déterminants) que les noms qui apparaissent pour la première fois dans l'échange ? En d'autres termes, l'hétéro-reprise est-elle, dans nos données, un contexte favorable à la production des premières formes de déterminants ou des fillers ? b) Lorsque l'on s'intéresse non plus seulement aux hétéro-reprises mais aussi aux reprises que l'enfant fait de son propre discours, peut-on observer des régularités entre le rang d'apparition du nom ( $n$  et  $n-1$ ) et le type de forme produite ? c) Quelles différences peut-on relever dans la nature des formes pré-nominales produites si l'on met en perspective les noms pour lesquels la dernière mention est une mention de l'enfant, les noms repris du discours de l'interlocuteur ainsi que les noms qui ne sont pas précédemment mentionnés dans le discours ? Les productions de l'enfant semblent-elles dépendre davantage du rang d'apparition du nom ou du type de forme pré-nominale associé au nom dans la mention précédente ? d) Enfin, dans les contextes d'auto- et d'hétéro-reprise, la distance entre le nom analysé et la mention précédente peut-elle jouer un rôle sur la nature des formes produites par l'enfant ?

Trois sections principales composent ce point 2. Dans un premier temps (section 2.1), nous présenterons les grandes lignes de la méthodologie que nous avons adoptée pour décrire les reprises de l'enfant : identification de la reprise, source, distance de cette reprise et caractérisation de la mention précédente du nom. Dans une deuxième section (2.2), nous ferons un état des lieux des contextes dans lesquels les noms sont produits dans notre corpus (auto/hétéro-reprises ou premières mentions), ainsi que des différents types de reprises (immédiates ou à distance) que nous avons pu observer. Enfin, dans un troisième temps (section 2.3), nous comparerons la

---

<sup>1</sup> Nous aurons recours ici au terme 'reprise' pour désigner le processus de réutilisation des noms par l'enfant mais les phénomènes que nous observons dans cette partie du chapitre ne se limitent pas aux reprises en tant que telles et peuvent correspondre à ce que des recherches antérieures ont caractérisé comme des répétitions, des imitations ou encore, des reprises avec reformulation.

distribution des formes pré-nominales dans les trois contextes préalablement listés - première mention du nom, auto-reprise et hétéro-reprise - et nous tenterons d'évaluer l'impact de la distance entre deux mentions sur les formes produites par l'enfant.

## 2.1 Description des reprises de l'enfant

### 2.1.1 Identification des reprises

Pour identifier les reprises présentes dans notre corpus, nous avons parcouru tous les noms préalablement relevés (*cf.* point 1 du chapitre V) et pouvant être précédés d'un déterminant, et pour chacun d'entre eux, nous avons noté si le lexème était produit pour la première fois dans la séance en cours (auquel cas, nous avons codé une absence de reprise) ou bien au contraire, si la production de l'enfant reprenait une forme déjà produite en amont de l'échange. Dès lors que notre intérêt s'est porté sur le suivi de la forme et non plus sur le suivi du référent, nous avons signalé la présence d'une reprise, qu'il y ait continuité référentielle ou non, autrement dit, que la forme produite et la forme précédente renvoient ou non au même référent.

### 2.1.2 Source et distance de la reprise

Afin de distinguer les auto-reprises des hétéro-reprises, nous nous sommes également intéressée à la source de la reprise de l'enfant. Nous avons ainsi précisé dans chaque cas le statut du participant (enfant ou interlocuteur) à qui nous pouvions attribuer la dernière mention du nom.

Une fois la source de la reprise repérée, nous avons également codé la distance entre la production de l'enfant et la dernière mention. Deux cas de figure ont été distingués : les cas de reprises immédiates (pour lesquelles la mention précédente se situe à moins de quatre tours de parole de distance) et les reprises à distance, comptant plus de quatre tours de parole entre la mention analysée et la forme précédente. Ces deux cas de figure sont illustrés dans les exemples ci-dessous.

#### Exemples 21 - Cas de reprises immédiate et à distance

(a) Reprise immédiate (Lubin, 2;8)

Père : y a pas d(e) locomotive.

Lubin : [mɛ si sa lokomotiv] '*mais c'est ça locomotive.*'

(b) Reprise à distance (Naël, 2;6)

Père : là où y avait tous les enfants qui jouaient dans l'eau mais il ne sont plus là.

[16 tours de parole]

Naël : [pas kə le zãfã i zõ kase se panje] ‘pa(r)ce que les enfants i(l)s ont cassé ses paniers.’

### 2.1.3 Caractérisation de la forme pré-nominale de la mention précédente

Dans les cas d’une reprise du nom, qu’il s’agisse d’une reprise par l’enfant de son propre discours ou du discours de son interlocuteur, immédiate ou à distance, nous avons identifié et caractérisé la forme pré-nominale précédant la dernière mention du nom. Cette caractérisation s’est faite en trois temps. Nous avons dans un premier temps distingué les cas où le nom repris et son antécédent sont précédés des mêmes formes - omissions, fillers ou déterminants (identiques) - ou de formes différentes - omission/filler, omission/déterminant, filler/déterminant ou deux déterminants distincts - et nous avons dans un deuxième temps relevé la catégorie de la forme pré-nominale de la mention précédente. Nous avons ici réutilisé les mêmes catégories que celles exposées au point 1 du chapitre V, en précisant toutefois le type de déterminant produit plutôt que de n’avoir recours qu’à la macro-catégorie ‘déterminant’. En dernier lieu, nous avons spécifié la forme phonétique du filler ou orthographique du déterminant lorsque la dernière mention du nom était précédée de l’un de ces deux éléments.

Ces analyses nous seront particulièrement utiles pour discuter de l’ordre d’apparition des formes devant les noms lorsque celles-ci alternent dans les contextes d’auto-reprise.

## 2.2 Inventaire des contextes de production des noms et des types de reprises

Avant de nous pencher spécifiquement sur la question du rôle joué par les reprises dans la production des formes pré-nominales, nous ferons d’abord l’état des lieux au sein de ce point 2.2 des différents contextes de production des noms dans nos données. Pour ce faire, nous avons évalué pour chaque enfant la proportion représentée par les syntagmes nominaux repris, de son propre discours ou du discours de son interlocuteur, ainsi que la proportion des syntagmes nominaux non repris et donc introduits pour la première fois par l’enfant. Cette analyse a été menée pour tous les noms pouvant être précédés d’un déterminant, et pour toutes les séances de chaque corpus. Rappelons toutefois que dans le cas où le nom analysé a déjà été produit dans la séance en cours, seule la dernière mention a été considérée pour ce travail.

Les résultats de ces analyses sont présentés dans les quatre graphiques qui suivent. La part des noms pour lesquels la dernière mention est une mention de l’enfant (REP ENF) et ceux pour



lesquels la dernière mention est produite par l'interlocuteur (REP INT) est affichée dans la partie inférieure des graphiques, et la part des noms qui ne font pas l'objet d'une reprise (SREP), dans la partie supérieure.

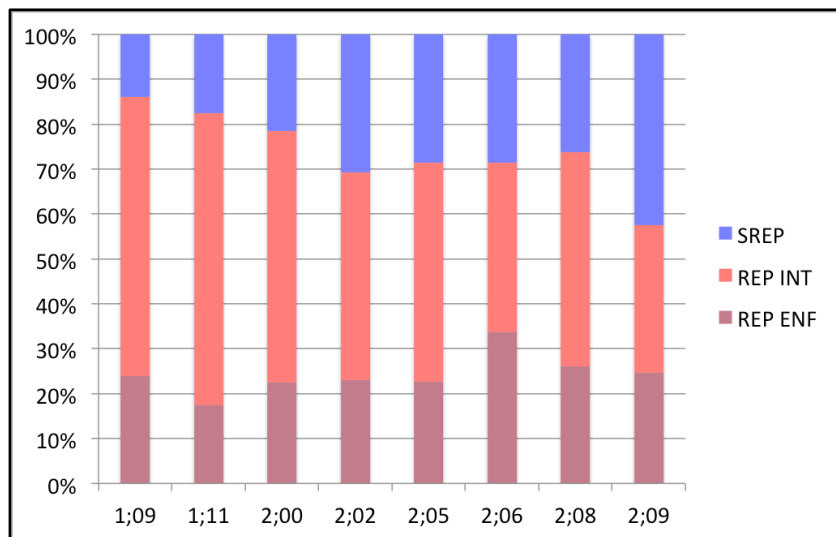
Ces graphiques 13 à 16 ci-dessous nous montrent dans un premier temps que les syntagmes nominaux produits par les quatre enfants sont en très grande majorité repris, de leur propre discours ou du discours de leur interlocuteur. Au contraire, les cas de noms introduits dans l'échange par l'enfant sont minoritaires, quels que soient le corpus ou la séance observés.

On peut néanmoins remarquer que la part de ces noms introduits pour la première fois évolue dans le corpus de Lubin puisqu'elle passe de 14% des productions de l'enfant dans la première séance à 43% dans la dernière. Autrement dit, presque la moitié des noms que l'enfant produit au cours de la séance à 2;9 ne font pas partie des lexèmes disponibles dans les échanges qui précèdent. Leur proportion augmente également chez Naël mais de façon nettement moins importante (12% à 1;6 contre 23% un an plus tard). Dans le corpus d'Adrien, seule la dernière séance se démarque par une augmentation de la part des noms non repris. Contrairement à ce que nous observons chez Lubin, aucune évolution notable n'est constatée dans les données d'Antonin, dans la mesure où la part représentée par ces noms augmente et diminue en alternance au fil des séances.

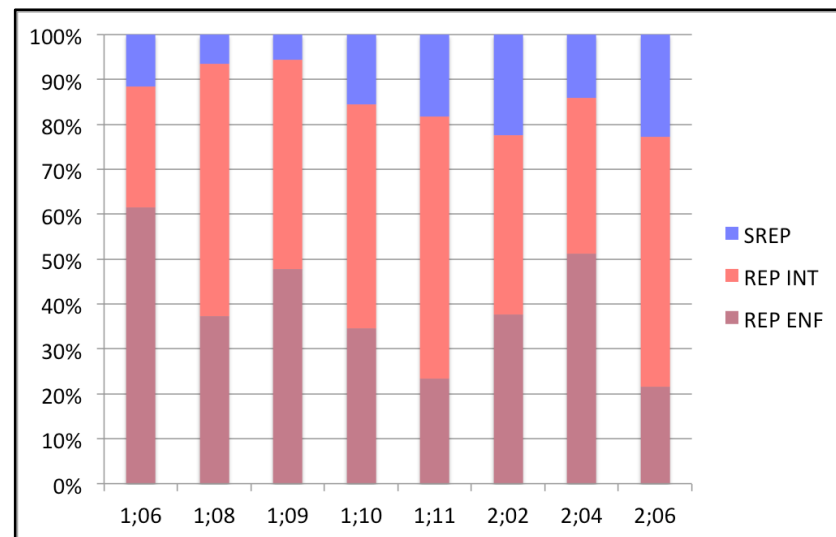
En parallèle de l'augmentation des noms introduits par l'enfant, la part des reprises de noms produits dans la mention précédente par l'interlocuteur diminue de moitié dans le corpus de Lubin entre la première (62%) et la dernière séance (33%). On ne relève pas une telle progression (ou diminution) dans les autres corpus, au sein desquels la proportion de ces noms est beaucoup plus fluctuante d'une séance à l'autre.

**Graphiques 13 à 16 - Distribution des syntagmes nominaux repris et non repris par l'enfant (dans les quatre corpus)**

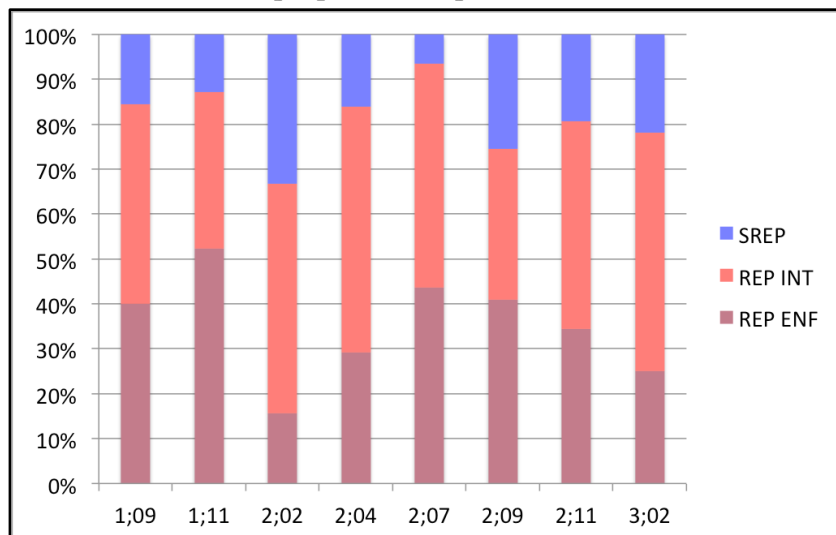
*Graphique 13 - Corpus de Lubin*



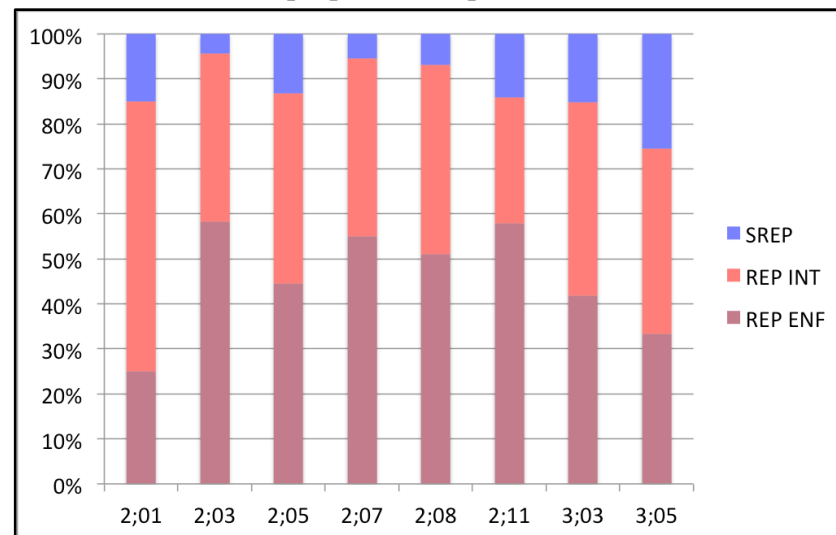
*Graphique 14 - Corpus de Naël*



*Graphique 15 - Corpus d'Antonin*



*Graphique 16 - Corpus d'Adrien*



La proportion des auto-reprises est assez constante au fil des séances chez tous les enfants. On peut certes observer une diminution de leur proportion entre la séance 1 et 8 dans le corpus de Naël mais la diminution n'est pas régulière entre ces séances. On relève par ailleurs que tous les enfants n'y ont pas recours dans les mêmes proportions. On trouve environ 46% de noms dont la mention précédente a été produite par l'enfant dans le corpus d'Adrien, et seulement 25% dans celui de Lubin. Ces reprises sont présentes dans les données de Naël et d'Antonin dans des proportions intermédiaires : elles représentent 39% des syntagmes identifiés chez Naël et 37% chez Antonin. Cette variation dans les proportions observées peut témoigner entre autres de stratégies différentes chez les quatre enfants, en ce qui concerne le maniement des noms, ou de styles interactionnels différents d'une dyade à l'autre.

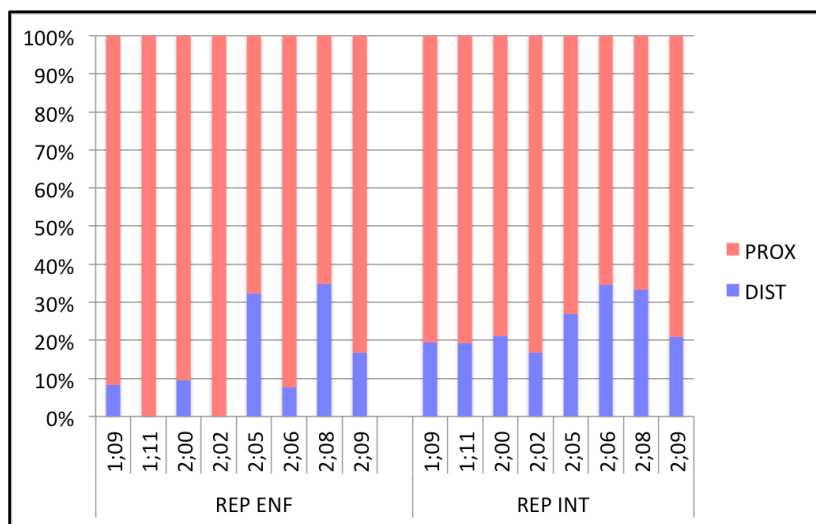
Afin de poursuivre notre état des lieux des reprises au sein des quatre corpus étudiés, nous avons par ailleurs cherché à observer la distance séparant le nom analysé de la dernière mention. Rappelons que deux cas de figure différents ont été considérés : les reprises immédiates (pour lesquelles le nom analysé et son antécédent sont séparés par moins de quatre tours de parole), et les reprises à distance (pour lesquelles les deux mentions sont séparées de plus de quatre tours de parole).

Nous pouvons faire l'hypothèse ici les reprises immédiates (que l'enfant fait de son propre discours ou de celui de son interlocuteur) seront plus fréquentes que les reprises à distance. Dès lors que les noms tendent à être repris plutôt dans le cadre d'échanges autour du même thème, nous pouvons en effet nous attendre ce que l'enfant s'appuie davantage sur des lexèmes produits dans les énoncés immédiatement précédents que dans des énoncés produits à plus de quatre tours de parole de distance. Il est cependant envisageable que les résultats soient variables d'un corpus à l'autre et cela, entre autres, en raison du cadre interactionnel dans lequel les noms sont produits.

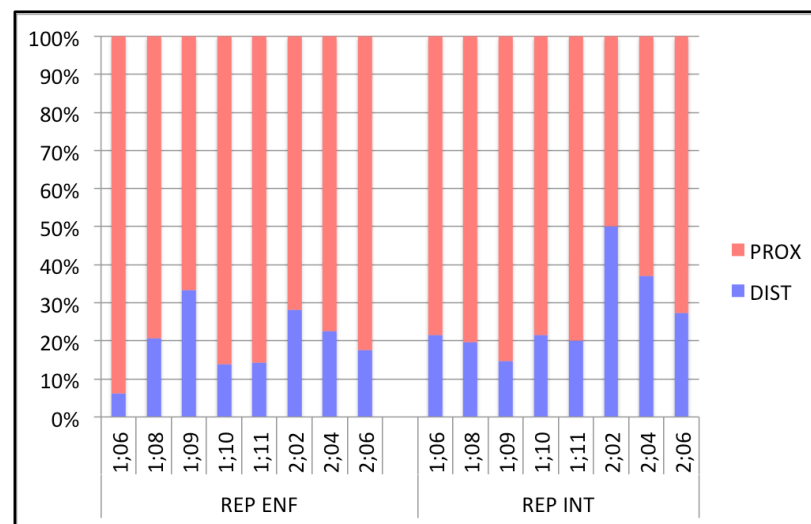
Les résultats affichés dans les quatre graphiques ci-dessous (17 à 20) présentent la part des reprises immédiates (PROX) et à distance (DIST) dans les contextes d'auto-reprise (REP ENF) et d'hétéro-reprise (REP INT).

**Graphiques 17 à 20 - Répartition des reprises immédiates et à distance dans les contextes d'auto- et d'hétéro-reprise (dans les quatre corpus)**

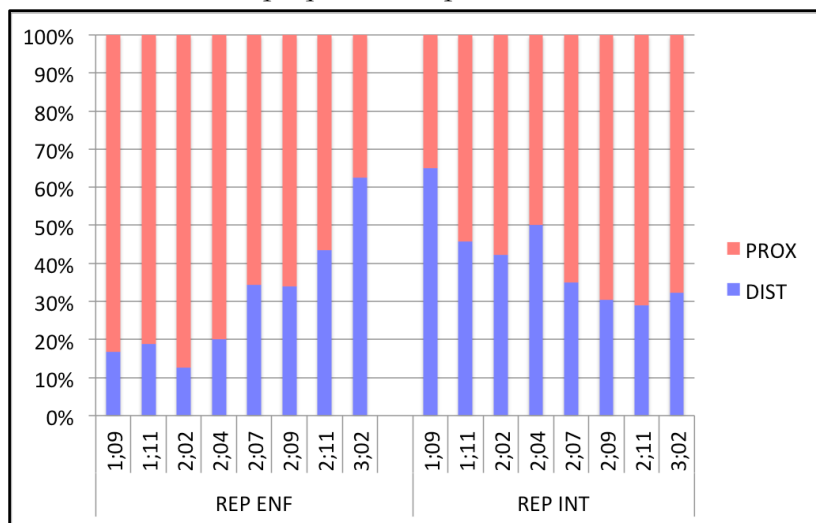
*Graphique 17 - Corpus de Lubin*



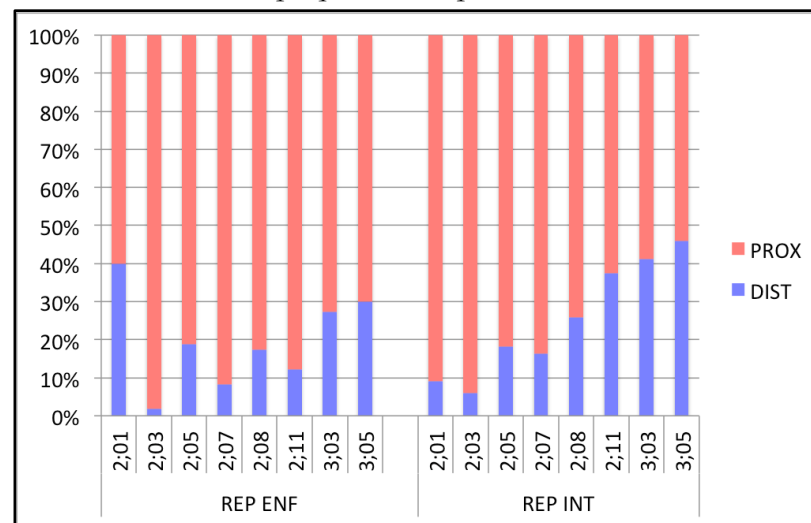
*Graphique 18 - Corpus de Naël*



*Graphique 19 - Corpus d'Antonin*



*Graphique 20 - Corpus d'Adrien*



Si l'on s'intéresse aux cas d'auto-reprises, on constate d'une part que les reprises à distance apparaissent - conformément à ce que nous pouvions envisager - dans des proportions nettement moins importantes que les reprises immédiates. Elles sont quasiment absentes dans le corpus de Lubin en première période et tendent à augmenter en seconde période (23% en moyenne si l'on compile les quatre séances). Elles augmentent de façon notable dans les données d'Antonin entre la première et la dernière séance. Elles représentent moins de 20% des auto-reprises à 1;9 et plus de 60% à 3;2. Dans le corpus d'Adrien, la proportion de ces auto-reprises diminue entre la séance 1 et les suivantes puis augmente de nouveau dans les deux dernières séances. Toutefois, la part représentée par les noms que l'enfant reprend d'un syntagme qu'il a lui-même produit (immédiatement ou à distance) ne correspond à 2;1 qu'à 5 occurrences (contre 53 par exemple dans la séance qui suit), et les auto-reprises à distance, à 2 occurrences (du même lexème 'nez'). Au vu de ces chiffres, il apparaît préférable de considérer que la part des auto-reprises à distance est minime de 2;1 à 2;11 et tend à augmenter à 3;3 et 3;5 (séances correspondant à la seconde période chez Adrien). Dans le corpus de Naël, aucune évolution significative n'est visible en ce qui concerne ces auto-reprises.

En parallèle, la proportion de reprises immédiates diminue dans le corpus d'Antonin (83% en séance 1 contre 37% en séance 8). De même, elles diminuent dans celui de Lubin si l'on compare la part qu'elles représentent en période 1 et en période 2, ainsi que dans le corpus d'Adrien si l'on exclut la première séance pour les raisons évoquées plus tôt. Dans la continuité de ce que nous venons de signaler au sujet des auto-reprises à distance dans le corpus de Naël, aucune tendance particulière ne se dessine lorsque l'on se focalise sur les auto-reprises immédiates (représentant en moyenne 79% des auto-reprises).

L'observation des noms pour lesquels la dernière mention est une mention de l'interlocuteur nous montre d'une part que, comme pour les auto-reprises, les reprises à distance sont bien moins fréquentes que les reprises immédiates. Leur proportion augmente en période 2 dans les corpus de Naël et Adrien. Ce constat est également valable pour celui de Lubin mais dans une moindre mesure (19% en période 1 en moyenne et 29% en période 2). Les reprises immédiates sont par conséquent moins fréquentes durant la période 2 au sein de ces trois corpus. En revanche, dans le corpus d'Antonin, les reprises à distance apparaissent dans des proportions

plus importantes dans la première période que dans la seconde et de fait, les reprises immédiates sont plus fréquentes en période 2 qu'en période 1.

Si nous avons initialement pensé que les augmentations de la proportion des reprises à distance observées dans les contextes d'auto- ou d'hétéro-reprise pouvaient peut-être être interprétées comme un signe de la capacité croissante de l'enfant à s'extraire du répertoire lexical immédiat et à réutiliser des noms autres que ceux qui viennent d'être produits, il nous semble que cette conclusion n'est *in fine* pas pertinente dans nos données. Les différences individuelles relevées dans les graphiques que nous venons de présenter nous ont amenée à nous pencher en particulier sur les noms repris à distance et à faire le constat de l'interaction entre la présence de reprises et d'autres facteurs, tels que la situation ou l'activité et les ressources linguistiques (autres que les noms) dont dispose l'enfant pour parler des référents.

D'une part, les noms sont plus souvent repris parce que l'on continue à échanger à propos d'un référent et/ou d'un autre référent du même type (une voiture et une autre voiture par exemple) dans la situation en cours. De fait, les reprises sont largement corrélées à ce que font les participants. Plus leurs activités sont longues et focalisées sur des référents spécifiques, plus les chances d'observer des reprises - et notamment, à distance - sont nombreuses. Par exemple, en nous intéressant aux noms concernés par les reprises à distance dans le corpus d'Antonin, nous avons pu relever qu'une grande partie d'entre eux sont associés à la même activité de jeu avec ses voitures ('camion', 'garage', 'voiture', 'accident', 'quatre-quatre', etc.). Bien que plusieurs autres activités puissent se dérouler au sein de chacune des séances de ce corpus, l'intérêt de l'enfant se porte régulièrement sur son circuit automobile, installé dans la pièce où se déroulent sept des huit séances retenues pour l'analyse. Dans plusieurs séances, Antonin joue longuement avec ses voitures et le fait que l'activité se prolonge durant presque toute la séance pourrait par exemple se manifester par une part plus importante de reprises à distance des noms renvoyant aux objets qui sont régulièrement - mais pas simultanément - manipulés. D'autre part, en nous penchant cette fois plus spécifiquement sur la séance à 2;2 chez Naël, nous avons pu voir que, outre le rôle joué par la durée et la focalisation de l'activité, la façon dont l'enfant réfère aux objets de discours, la nature des expressions référentielles auxquelles il a recours influence nécessairement l'emploi qu'il fait des noms. Dans cette séance, nous avons pu observer que les noms repris à distance sont bien souvent les mêmes ('château fort', 'chevalier', 'cheval') et qu'entre deux mentions, les

référents auxquels ils renvoient peuvent être verbalisés par des pronoms. Nous pouvons donc faire l'hypothèse que plus l'enfant emploie des pronoms, plus la part des noms utilisés dans des contextes de continuité référentielle diminue, ceux-ci tendant alors à être destinés plutôt aux contextes d'introduction, de réactivation suite à un changement thématique ou encore, d'oppositions entre deux référents. Si cette hypothèse était vérifiée dans nos données, elle pourrait en partie expliquer l'augmentation de la part des noms repris à distance. La distribution des différents types de reprises pourrait donc peut-être s'expliquer davantage par des facteurs situationnels ou par l'interaction entre différents instruments langagiers que par l'âge de l'enfant ou par son degré de maîtrise des noms de sa langue.

## **2.3 Emploi des formes pré-nominales devant les noms repris**

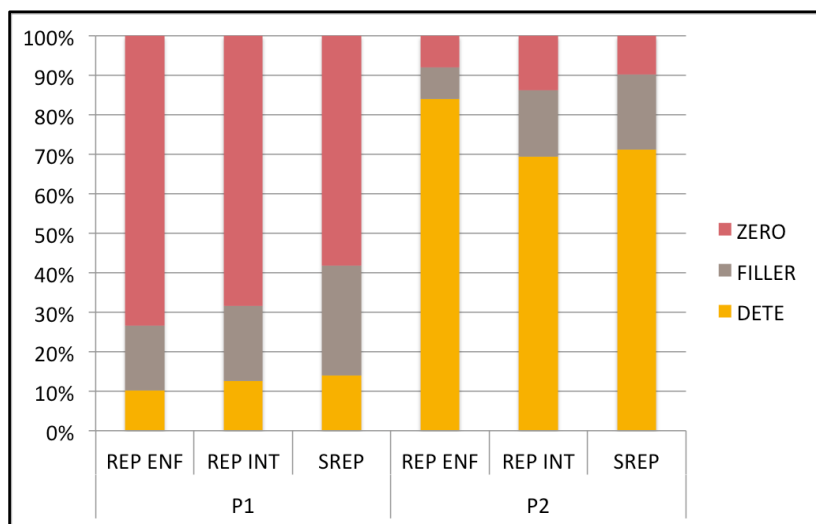
### **2.3.1 Comparaison de la distribution des formes pré-nominales devant les noms introduits dans le discours et repris**

A la suite de cet état des lieux, nous avons cherché à évaluer l'influence des reprises sur la production des formes pré-nominales. Pour cela, nous avons comparé la distribution des formes devant les noms repris et les noms que l'enfant introduit pour la première fois dans la séance. Afin de mesurer l'impact éventuel de la source de la reprise, nous avons distingué la distribution devant les noms repris du discours de l'enfant et ceux pour lesquels la dernière mention est produite par l'interlocuteur.

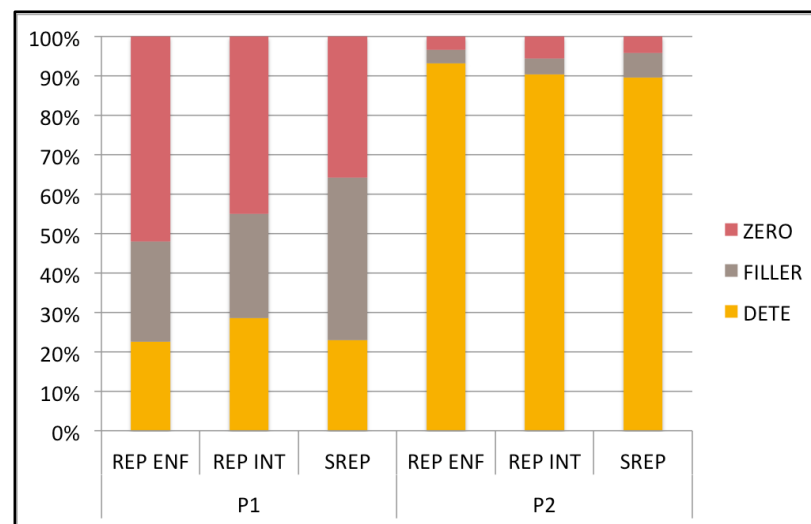
Les résultats sont présentés pour chaque enfant et par période dans les quatre graphiques qui suivent. Nous avons affiché la proportion représentée par les omissions, les fillers et les déterminants dans chaque contexte : noms repris du discours de l'enfant (REP ENF), noms repris du discours de l'interlocuteur (REP INT) et noms non repris (SREP).

**Graphiques 21 à 24 - Distribution des formes pré-nominales devant les noms introduits dans le discours et repris (dans les quatre corpus)**

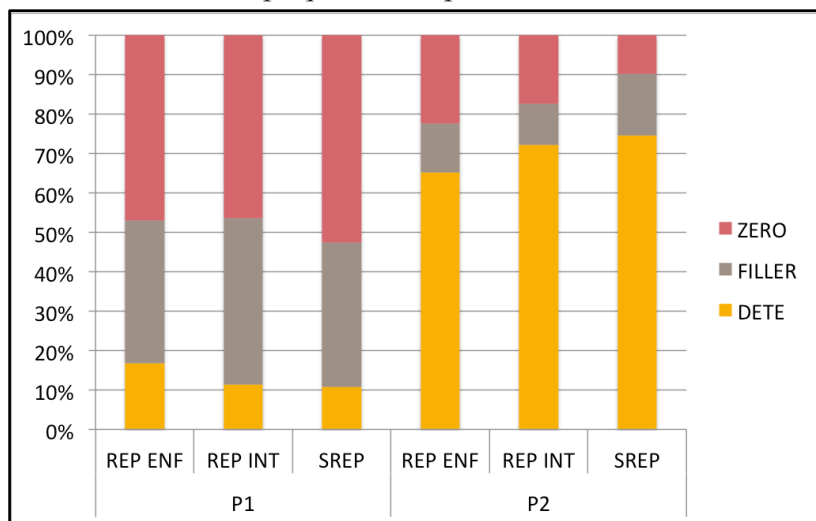
*Graphique 21 - Corpus de Lubin*



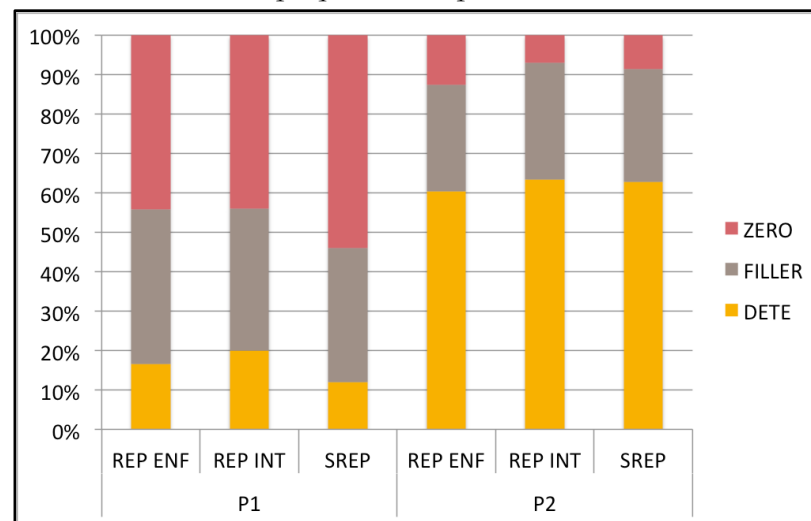
*Graphique 22 - Corpus de Naël*



*Graphique 23 - Corpus d'Antonin*



*Graphique 24 - Corpus d'Adrien*





On remarque tout d'abord que quels que soient le corpus ou la période, la distribution des formes pré-nominales apparaît comme très comparable dans les trois contextes. On relève toutefois quelques différences - minimes - d'un contexte à l'autre.

Dans le corpus de Lubin, la part des fillers devant les noms qui ne sont pas repris est plus importante que devant les noms repris (en auto- et en hétéro-reprise) en période 1, et *a contrario*, la part des omissions est moindre dans les contextes où le nom est introduit par l'enfant. En période 2, les fillers et les omissions sont plus fréquents devant les noms repris du discours de l'interlocuteur, ainsi que devant les noms qui ne sont pas repris, que devant ceux dont la dernière mention est une mention de l'enfant.

Dans le corpus de Naël, les omissions semblent être favorisées dans le contexte des auto-reprises, les déterminants dans celui des hétéro-reprises et les fillers devant les noms qui ne sont pas repris. En période 2, omissions, fillers et déterminants sont produits dans ces trois contextes dans des proportions quasiment identiques.

Dans les données d'Antonin, les déterminants sont légèrement plus fréquents dans le cas des auto-reprises en période 1, de même que les fillers dans le contexte des reprises du discours de l'interlocuteur et les omissions avec les noms introduits par l'enfant. Durant la seconde période, la tendance s'inverse puisque les déterminants apparaissent moins fréquemment devant les noms produits en auto-reprise que dans les autres contextes, et les fillers devant les noms qui ne sont pas repris.

Enfin, dans le corpus d'Adrien, les omissions sont plus fréquentes devant les noms apparaissant pour la première fois dans l'échange, et en parallèle, c'est dans ce contexte que les déterminants sont le moins souvent produits. La distribution est par ailleurs équivalente quelle que soit la source de la reprise dans cette première période. Durant la seconde, les trois formes sont produites dans les mêmes proportions dans les trois contextes.

A l'exception de ces quelques différences - variables d'un enfant à l'autre et d'une période à l'autre - nos résultats ne font pas apparaître d'associations nettement privilégiées entre certaines formes et certains contextes. Ceci manifeste d'une part l'existence de styles individuels dans le traitement des reprises et pourrait suggérer d'autre part que la reprise n'est pas dans nos données

un facteur susceptible de déterminer véritablement la nature des formes pré-nominales produites par l'enfant.

Dans la mesure où les cas dans lesquels l'adulte ne produit pas de déterminant - dans les énoncés métalinguistiques de type 'vas-y dis ! piscine !' (mère Adrien, séance à 2;5) - ont été exclus de nos analyses, toutes les productions de l'enfant en hétéro-reprise font suite à un syntagme 'déterminant + nom'. La distribution des formes pré-nominales dans ces contextes correspond donc à la distribution que nous venons d'observer dans les quatre graphiques qui précèdent (cas de 'REP INT'). En revanche, les formes observées devant les noms pour lesquels la dernière mention est une mention de l'enfant peuvent avoir été produites après une forme adulte ou après une forme de transition, omission ou filler. Nous nous sommes donc interrogée dans un deuxième temps sur la forme produite devant la mention précédente (pouvant elle-même faire partie des noms introduits par l'enfant ou des noms repris) et sur le degré de correspondance entre les formes pré-nominales relevées dans les mentions  $n$  et  $n-1$ .

Dans le tableau qui suit, nous avons affiché la répartition des formes en position pré-nominale dans les contextes d'auto-reprise, en précisant cette fois la forme produite devant la mention précédente ( $n-1$ ) du nom (déterminant, omission, filler). Les résultats sont présentés pour chaque enfant et par période.

Le premier constat que nous pouvons faire est celui de la reprise du type de forme produite d'une mention à l'autre durant la première période. En nous focalisant sur les formes majoritaires dans chaque contexte (*cf.* cellules grisées), nous avons pu noter que pendant cette période, les formes produites dans les mentions  $n$  et  $n-1$  sont le plus souvent les mêmes. Seul Antonin produit plus d'omissions suite à un déterminant durant cette première période, mais les chiffres étant très faibles et assez comparables d'un contexte à l'autre, nous n'en tirerons pas de conclusion spécifique.

Tableau 14 - Nature des alternances en position pré-nominale dans les contextes d'auto-reprise (pour les quatre corpus)

Mention <i>n-1</i>		Déterminant			Filler			Omission			Total
Mention <i>n</i>		Dét.	Fill.	Ø	Dét.	Fill.	Ø	Dét.	Fill.	Ø	
P1	Lubin	11,54%	1,92%	1,92%	1,92%	9,62%	3,85%	1,92%	3,85%	63,46%	100,00%
	Naël	14,29%	3,90%	6,49%	3,90%	12,34%	10,39%	3,25%	9,09%	36,36%	100,00%
	Antonin	3,29%	3,95%	4,61%	6,58%	17,11%	13,16%	7,24%	14,45%	29,61%	100,00%
	Adrien	12,15%	3,47%	0,35%	3,82%	31,60%	4,51%	0,69%	4,17%	39,24%	100,00%
P2	Lubin	72,84%	4,94%	1,23%	8,64%	3,70%	2,47%	2,47%	0,00%	3,70%	100,00%
	Naël	89,77%	1,14%	1,14%	0,00%	2,27%	0,00%	3,41%	0,00%	2,27%	100,00%
	Antonin	54,14%	6,37%	7,64%	5,73%	2,55%	2,55%	5,74%	3,18%	12,10%	100,00%
	Adrien	47,62%	12,7%	6,35%	6,35%	14,29%	0,00%	6,35%	0,00%	6,35%	100,00%

On relève par ailleurs que lorsque les deux formes ne sont pas les mêmes, l'omission dans la mention *n-1* entraîne une production plus importante de fillers que de déterminants dans la mention suivante, et de même, les syntagmes 'filler + nom' sont davantage suivis de noms associés à des omissions qu'à des déterminants. Les résultats sont plus diversifiés en ce qui concerne les déterminants. Lubin a recours presque exclusivement au déterminant si c'est un déterminant qu'il a produit dans la mention précédente. Dans le cas de Naël, la mention qui suit un syntagme 'déterminant + nom' peut aussi être précédée d'abord d'omissions, puis de fillers. Dans les données d'Adrien, les quelques noms qui ne sont pas précédés de déterminants sont associés à des fillers.

Au cours de la deuxième période, les syntagmes 'déterminant + nom' dans la mention *n-1* entraînent une production très majoritaire de déterminants devant les noms de la mention suivante. On observe néanmoins une part plus importante de fillers et d'omissions dans les corpus d'Antonin et d'Adrien que dans ceux de Lubin et Naël. Ces observations concordent avec la distribution globale des formes puisque ces deux corpus sont également ceux dans lesquels la proportion de déterminants est la moins importante en période 2. Lorsque la mention précédente du nom est combinée à un filler, elle donne lieu le plus souvent à un syntagme 'déterminant + nom' dans la mention *n* chez Lubin et Antonin. Dans le corpus de Naël, on observe que les fillers ne sont suivis que par des fillers, mais signalons tout de même qu'il ne s'agit là que de 2,27% des syntagmes produits dans le cadre d'auto-reprises (soit 2 occurrences). Les fillers apparaissent aussi plus fréquemment à la suite d'un syntagme 'filler + nom' dans le corpus d'Adrien. Les occurrences de syntagmes 'omission + nom' (dans la mention *n-1*) entraînent une production plus fréquente d'omissions dans la mention suivante dans les corpus de Lubin et d'Antonin, de déterminants dans celui de Naël, et à part égale, de déterminants et d'omissions dans les données d'Adrien. Nous nous devons toutefois de signaler que la part des omissions étant globalement très minoritaire durant cette seconde période dans les données de Lubin et de Naël, ces observations peuvent difficilement donner lieu à des conclusions solides.

Les résultats concernant la première période suggèrent donc que dans les contextes où l'enfant reprend un nom qu'il vient lui-même de produire, il reprend le plus souvent non seulement le lexème mais également la forme qui le précède. Si l'on s'intéresse aux cas qui diffèrent de ce schéma majoritaire, on voit que conformément au développement global de la position pré-

nominales, les omissions alternent plus souvent avec des fillers et les fillers avec des omissions, qu'avec des déterminants. En période 2, la part représentée par les cas où les deux noms produits sont précédés des mêmes formes reste majoritaire mais on constate toutefois que les syntagmes 'omission + nom' et 'filler + nom' donnent lieu à une alternance plus fréquente qu'en période 1 et à de la variation d'un enfant à l'autre. Il nous paraît intéressant de relever que lorsque l'alternance est plus importante que la non-alternance des formes pré-nominales, c'est au profit du déterminant dans la mention  $n$ , et non de l'omission pour les fillers ou du filler pour les cas d'omission. Nous pouvons donc faire l'hypothèse que dans le cas d'un tâtonnement sur la forme pré-nominale dans la mention  $n-1$ , l'auto-reprise pourrait être un contexte favorable en période 2 à la production de formes adultes dans la mention suivante.

### **2.3.2 Comparaison de la distribution des formes pré-nominales dans les contextes de reprises immédiates et à distance**

Au travers des graphiques 21 à 24, nous avons vu d'une part que, dans nos données, la distribution des formes pré-nominales est très largement comparable dans les contextes où le nom est repris et dans ceux où il est introduit par l'enfant. D'autre part, nous n'avons pas observé de différences importantes dans les proportions représentées par chacune des formes (en période 1 comme en période 2) lorsque la mention précédente est produite par l'enfant et lorsqu'elle est produite par l'adulte.

Nous avons toutefois cherché à observer si ce dernier constat se confirmait quelle que soit la distance entre le nom analysé et la mention précédente. Nous nous sommes donc intéressée en premier lieu à la distribution des omissions, fillers et déterminants dans les contextes de reprise immédiate et à distance, d'abord dans le cas des auto-reprises, puis dans celui des hétéro-reprises. Dans un deuxième temps, nous avons comparé l'influence de la distance entre la mention  $n$  et  $n-1$  dans les contextes où le nom est repris du discours de l'enfant et lorsqu'il est repris du discours de l'interlocuteur.

Les quatre graphiques qui suivent (25 à 28) présentent la distribution des formes pré-nominales dans les contextes d'auto-reprise immédiate et à distance dans les quatre corpus. Les résultats sont affichés par période. La répartition des formes produites dans les cas de reprises immédiates

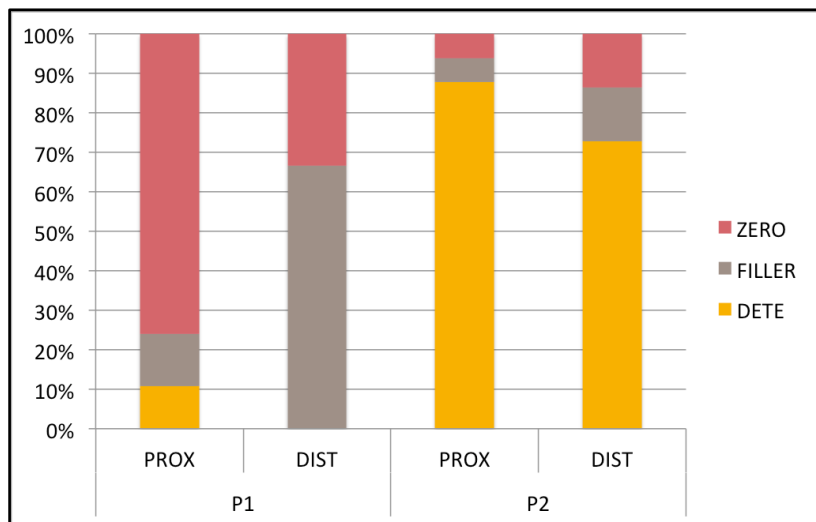
(PROX) est représentée pour chaque période par la colonne de gauche et la répartition des formes dans les reprises à distance (DIST), par celle de droite.

Ces résultats montrent d'une part que la distribution des formes pré-nominales est très comparable dans les contextes d'auto-reprise immédiate et à distance en période 1. Si les proportions paraissent plus diversifiées dans le corpus de Lubin, signalons que les reprises à distance sont très rares dans ce corpus en période 1 (et même absentes dans deux séances sur quatre), et que l'analyse ne porte ici que sur trois noms. Il aurait d'ailleurs été préférable de n'afficher que les valeurs absolues mais par souci de cohérence avec les choix de représentation que nous avons faits pour les trois autres enfants, nous avons conservé un affichage en pourcentages. Dans les données de Naël, Antonin et Adrien, la part des omissions, fillers et déterminants est semblable lorsque les reprises sont immédiates et à distance. On relève néanmoins que les formes de transition - omissions et fillers - sont un peu plus fréquentes chez Naël et Antonin lorsque les noms sont produits dans des contextes de reprise à distance, et en parallèle, les reprises immédiates tendent à être légèrement plus favorables à la production d'un déterminant. Ces tendances sont en revanche inversées dans le corpus d'Antonin.

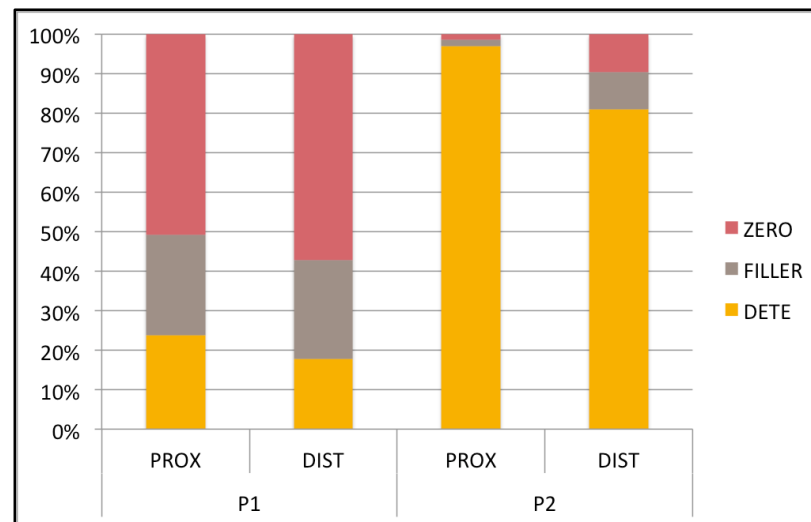
D'autre part, les différences que nous observons en période 2 sont plus marquées que celles visibles en période 1. Dans les données de Lubin et Naël, les omissions et les fillers sont plus représentés dans le cas des reprises à distance. De nouveau, la tendance est inverse dans le corpus d'Antonin puisque l'on remarque que les omissions et les fillers sont plus fréquemment produits lorsque les noms sont repris du contexte immédiat. Dans le corpus d'Antonin, les déterminants apparaissent aussi souvent dans les deux contextes. Les omissions sont plus fréquentes dans le cas des reprises immédiates et les fillers dans le cas des reprises à distance. Si certaines formes tendent à être privilégiées dans un contexte plus que l'autre en période 2, celles-ci sont donc variables d'un enfant à l'autre.

**Graphiques 25 à 28 - Distribution des formes pré-nominales dans les contextes d'auto-reprise immédiate et à distance (dans les quatre corpus)**

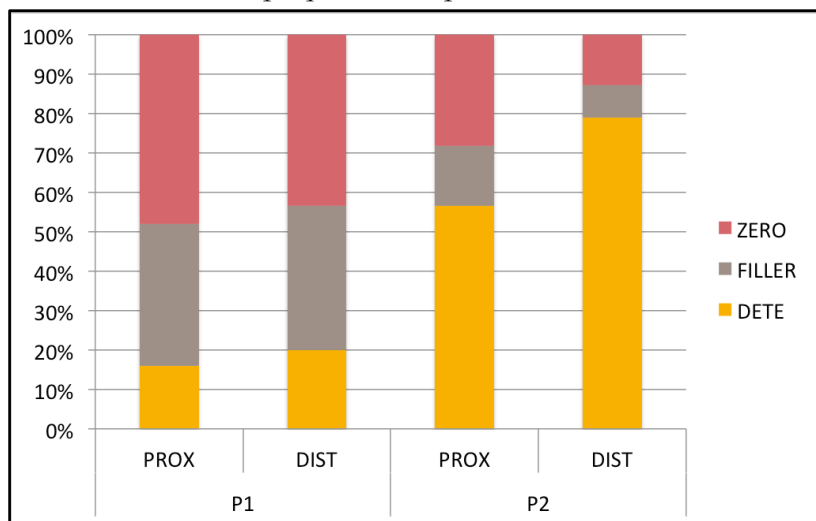
*Graphique 25 - Corpus de Lubin*



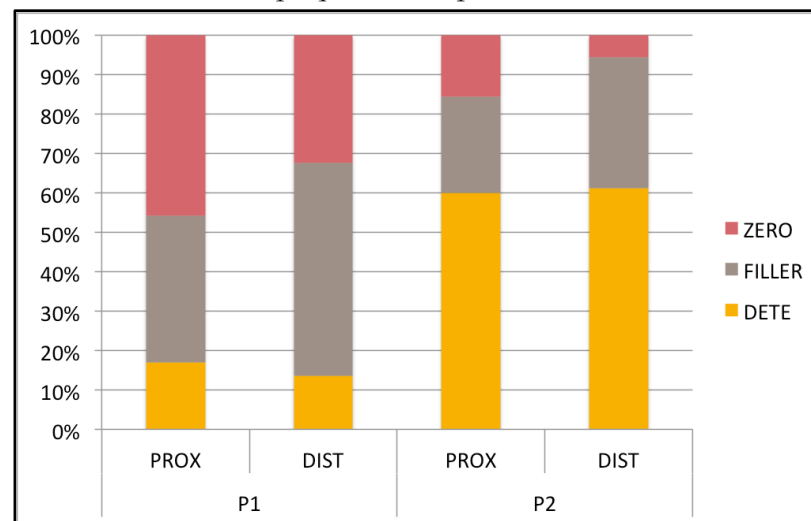
*Graphique 26 - Corpus de Naël*



*Graphique 27 - Corpus d'Antonin*



*Graphique 28 - Corpus d'Adrien*



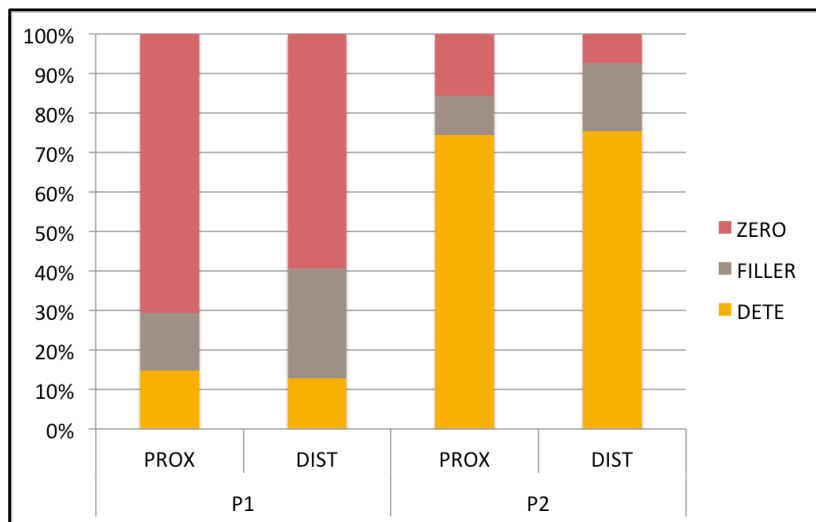
Les graphiques 29 à 32 représentant quant à eux la distribution des formes pré-nominales dans les contextes d'hétéro-reprise immédiate et à distance, montre qu'en période 1, les reprises à distance semblent être un contexte favorable à la production de fillers dans les corpus de Lubin, Antonin et Adrien. Ce constat est particulièrement valable pour Antonin qui produit 53% (51 occurrences) de fillers dans les contextes de reprise à distance contre 37% (38 occurrences) dans les contextes de reprise immédiate. En parallèle, la proportion de déterminants est plus importante chez Antonin dans le cas des reprises immédiates. Dans les données de Lubin et Adrien, les omissions sont plus fréquentes avec les reprises immédiates que les reprises à distance. Dans le corpus de Naël, l'enfant a recours aux trois formes dans des proportions tout à fait comparables quelle que soit la distance entre les deux mentions du nom. Le même constat peut d'ailleurs être fait en ce qui concerne la période 2. Durant cette seconde période, les noms repris à distance entraînent une production de fillers légèrement plus importante dans les données de Lubin et Adrien et les noms repris du contexte immédiat sont un peu plus souvent précédés d'omissions dans ces deux corpus, ainsi que dans celui d'Antonin. On relève par ailleurs qu'Antonin produit 15% de plus de déterminants devant les noms repris à distance que devant les noms apparaissant en reprise immédiate.

On peut donc conclure de l'observation des formes pré-nominales produites dans les contextes d'hétéro-reprise immédiate et à distance que l'influence de la distance entre la mention  $n$  et  $n-1$  est variable d'un enfant à l'autre. Elle ne se manifeste pas dans la distribution des formes chez Naël et lorsque des différences peuvent être pointées dans les trois autres corpus, elles ne concernent pas systématiquement les mêmes formes.

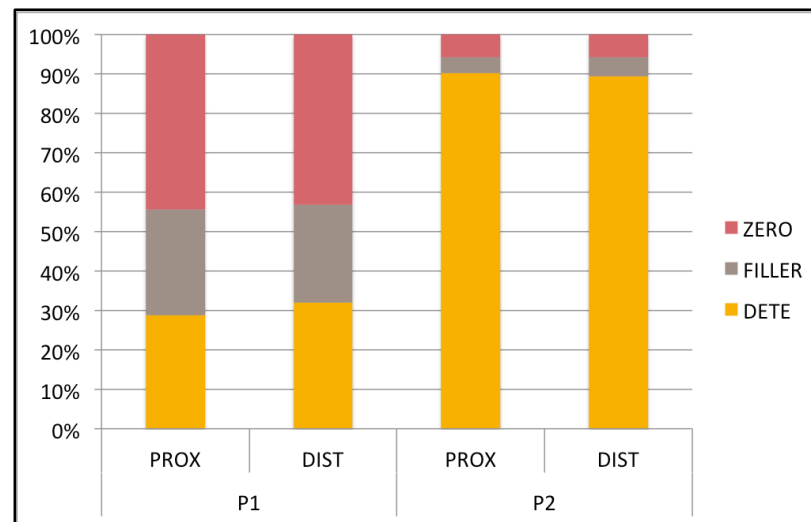


**Graphiques 29 à 32 - Distribution des formes pré-nominales dans les contextes d'hétéro-reprise immédiate et à distance (dans les quatre corpus)**

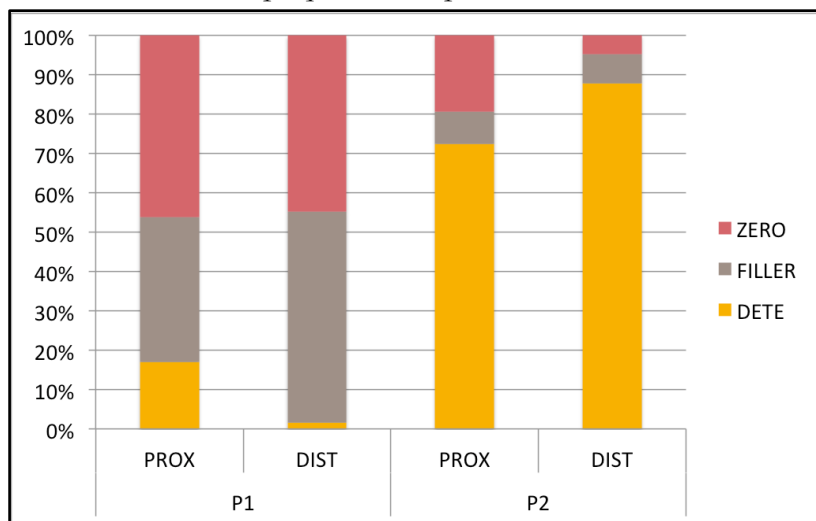
*Graphique 29 - Corpus de Lubin*



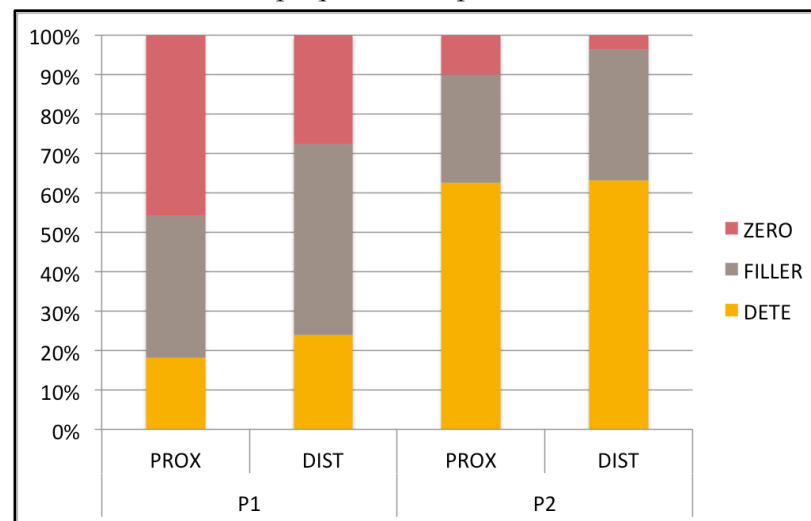
*Graphique 30 - Corpus de Naël*



*Graphique 31 - Corpus d'Antonin*



*Graphique 32 - Corpus d'Adrien*



Si l'on compare les contextes d'auto-reprise et d'hétéro-reprise, on remarque que la répartition des formes dans les cas de reprises immédiates et à distance est très semblable. Qu'il s'agisse des reprises du discours de l'enfant ou de l'interlocuteur, la distance entre le nom analysé et la mention précédente ne semble pas influencer le type de forme que l'enfant produit. Par ailleurs, les quelques différences observées entre reprises immédiates et à distance ne sont pas systématiquement confirmées dans les deux contextes (auto- et hétéro-reprises). Par exemple, dans la période 2 du corpus d'Antonin, la moindre proportion de déterminants que nous avons pu observée pour les reprises immédiates lorsque l'enfant reprend son propre discours est confirmée lorsqu'il reprend le discours de son interlocuteur. Le même constat peut être fait en ce qui concerne les omissions, plus représentées dans le cas des reprises immédiates. En revanche, les différences observées chez Naël en période 2 dans les contextes d'auto-reprise (plus de fillers et d'omissions avec les reprises à distance) ne sont pas confirmées par la distribution dans les contextes d'hétéro-reprise. Ceci peut donc suggérer que si les reprises peuvent à un degré limité influencer le type de forme que l'enfant produit, il s'agit probablement plutôt d'une influence combinée de la distance et de la source de la reprise que de l'influence privilégiée de l'un de ces deux facteurs.

#### **2.4 Une influence des reprises sur les formes pré-nominales produites par les enfants ?**

Les résultats présentés dans cette deuxième partie du chapitre montrent que les noms analysés dans notre corpus sont davantage repris - du discours de l'enfant ou du discours de l'interlocuteur - qu'introduits dans l'échange par l'enfant. Nous avons également pu observer que le plus souvent, ces reprises s'inscrivent en continuité avec des énoncés immédiatement antérieurs (plus qu'avec des énoncés produits à distance). Bien que ces constats soient valables pour les quatre enfants, nous avons pu relever toutefois que tous n'ont pas recours aux reprises ainsi qu'aux différents types de reprises dans les mêmes proportions. La part des syntagmes introduits par l'enfant, ainsi que celle des noms repris dans des contextes d'auto- ou d'hétéro-reprise ne sont pas équivalentes dans tous les corpus. Des différences ont aussi pu être relevées dans la distribution entre auto-reprises et hétéro-reprises d'une part, et entre reprises immédiates et à distance d'autre part. Outre l'existence de styles individuels dans le traitement des reprises,

nous avons fait l'hypothèse que ces variations d'un corpus à l'autre pourraient s'expliquer par le format de l'activité en cours et par la nature des formes employées pour verbaliser les référents.

Nous avons par ailleurs pu constater que les reprises de l'enfant ne semblent pas, dans nos données, avoir une influence majeure sur les formes pré-nominales produites. La comparaison des contextes dans lesquels les noms sont produits (repris/non repris) n'a pas fait apparaître de différences importantes dans la distribution des formes devant les noms. De même, les omissions, fillers et déterminants sont produits dans des proportions assez comparables, et cela, quelles que soient la distance entre les mentions  $n$  et  $n-1$  du nom ou la période observée. Lorsque les noms sont produits dans des contextes d'auto-reprise, les deux mentions du nom sont le plus fréquemment précédées des mêmes formes. Nous avons cependant pu relever qu'en période 2, l'omission ou le filler ne sont pas privilégiés par tous les enfants après un syntagme de type 'omission/filler + nom' dans la mention précédente. Dans ces cas où l'alternance entre les deux mentions est plus fréquente que la non-alternance, le déterminant est la forme la plus observée dans la mention  $n$ .

Notre constat principal - du faible impact des reprises de l'enfant sur les formes produites en position pré-nominale - reste cependant à envisager à l'aune des choix méthodologiques que nous avons faits. En comparant la distribution des omissions, fillers et déterminants devant les noms introduits par l'enfant, repris de son propre discours et de celui de son interlocuteur, nous n'avons observé l'influence des reprises que sous un angle synchronique. Une des étapes ultérieures de notre travail pourrait consister en une étude du degré de corrélation entre la proportion et la nature des reprises au début de la période 1 et le développement de la position pré-nominale au début et à la fin de la période 2. Si nous n'avons pas observé d'effet immédiat des reprises sur les productions des enfants, les résultats de travaux précédents (*cf.* par exemple Bertin, 2011) nous permettent de penser que la reprise de syntagmes nominaux par l'enfant est un processus favorable à la mise en place ultérieure des formes devant les noms.

Rappelons par ailleurs que nos analyses s'appuient seulement sur le lien entre un nom produit par l'enfant et sa mention précédente. Or, dans certains cas, le nom est produit plus d'une fois par l'enfant et son interlocuteur, et notamment dans le cadre de ce que Veneziano a appelé des 'échanges de type imitatif réciproques' (notamment, Veneziano, 1997b). Par exemple, dans l'extrait ci-dessous (exemple 21), nous avons analysé trois paires d'énoncés (1-2, 3-4 et 4-5), mais

notre analyse ne nous a pas permis de mesurer de façon quantitative l'influence combinée des énoncés 1 à 5 sur la dernière mention de l'enfant ([la tɛ mamã] 'la tête(m) maman'). Le dernier énoncé de la mère (6) n'a pas non plus été pris en compte dans ce travail dans la mesure où cette séquence apparaît en fin de séance et que l'enfant ne reprend pas ultérieurement les noms produits par la mère dans cet énoncé 6.

**Exemple 22 - Cas de reprise 'mixte' ou 'réciproque' dans le corpus de Lubin à 2;2**

*Lubin et sa mère sont en train de faire un puzzle.*

Mère : alors regarde comment il est le dessin ?

Mère : comme ça. (montre à Lubin le modèle du puzzle)

Mère : ça c'est la tête de la maman regarde. (1)

Lubin : 0. (regarde le modèle)

Lubin : [ɛt mamã] '(t)ête maman.' (2)

Mère : oui la tête de la maman là regarde. (3)

Lubin : [tɛt mamã] 'tête maman.' (4)

Lubin : [a la tɛ mamã] 'yyy la tête(m) maman.' (5)

Mère : la tête de la maman. (6)

Il serait donc judicieux de compléter notre travail par l'observation systématique et non segmentée de l'intégralité des chaînes de production des noms, en prenant en compte aussi bien les reprises de l'enfant que celles de l'adulte. Nous pourrions ensuite comparer les formes produites dans les contextes où le nom analysé apparaît dans des échanges simples (de type 'Interlocuteur-Enfant' ou 'Enfant-Interlocuteur') et lorsqu'il est produit au sein d'échanges réciproques (de type 'Interlocuteur-Enfant-Interlocuteur' ou 'Enfant-Interlocuteur-Enfant'), afin de mesurer l'impact de ces deux types de reprises sur le développement des déterminants.

En outre, il nous paraît important de souligner que nos résultats n'engagent pas l'intérêt plus général des reprises - de l'enfant mais aussi de l'adulte - dans la construction d'un cadre dialogique favorable aux acquisitions langagières, et en particulier, à l'acquisition du nom. Toutes les études mentionnées en introduction de ce point 2 s'accordent sur le fait que les reprises contribuent à l'acquisition progressive des formes conventionnelles de la langue. Dans la lignée de ces travaux, nous pouvons envisager que dans nos données, les reprises de l'adulte et de l'enfant encourageraient l'acquisition des noms, tant sur le plan de la forme que du sens et par extension, offriraient davantage d'opportunités à l'enfant de produire et d'entendre des noms et des formes pré-nominales. Le maniement du nom 'ambulance' dans le corpus d'Antonin nous

semble en être une bonne illustration. Nous avons sélectionné ici quatre extraits de ce corpus, à 2;4, 2;9 et 3;2.

**Exemple 23 - Reprises du lexème ‘ambulance’ dans le corpus d’Antonin (à 2;4, 2;9 et 3;2)**

(a) *Extrait à 2;4*

Antonin : [papɔ̃ pɔ̃pje] *‘yyy pompier.’*<sup>1</sup> (regarde pièce du puzzle qui représente une ambulance)

Mère : nan c'est pas un camion de pompier [=! rires].

Mère : c'est une ambulance.

Antonin : [mylãs] *‘(am)bulance.’*

Mère : ambulance.

(b) *Extrait un peu plus tard dans la même séance à 2;4*

Antonin : [mylãʃ] *‘(am)bulance.’* (encastre l’ambulance sur la planche du jeu)

Mère : l’ambulance très bien !

(c) *Extrait à 2;9*

Mère : et là c'est quoi ? (pointe pièce de puzzle représentant l’ambulance)

\*CHI: [ãbylãs] *‘ambulance.’*

\*MOT: l’ambulance.

(d) *Extrait à 3;2*

Mère : tu veux l’ambulance ?

Antonin : [wi] *‘oui.’*

Mère : ça se gare l’ambulance ?

Antonin : [wi] *‘oui.’*

Antonin : [lãbylãs] *‘l’ambulance.’*

Dans le premier énoncé d’Antonin à 2;4, la verbalisation [papɔ̃ pɔ̃pje] n’est pas adéquate d’un point de vue lexical puisque l’enfant est en train de regarder une pièce de puzzle représentant une ambulance. La mère reformule et offre à l’enfant le syntagme ‘une ambulance’. Suite à cet énoncé, l’enfant fait une reprise partielle de l’offre de la mère ([mylãs]). Au tour suivant, la mère reprend le nom ‘ambulance’ (sans déterminant), et par ce biais, elle apporte une deuxième fois à l’enfant la forme phonologique attendue. Un peu plus tard dans l’échange, Antonin manipule à nouveau la pièce du puzzle qui représente l’ambulance et reprend à distance le lexème précédemment introduit par la mère. Le nom est, comme à 2;4, précédé d’une omission mais la production de l’enfant est phonologiquement conforme au lexème adulte. La mère complète ensuite la production de l’enfant en la combinant à un défini (‘l’ambulance’) et signale explicitement que la forme produite est attendue dans ce contexte (‘très bien’). Enfin, à 3;2, la

<sup>1</sup> [papɔ̃] est régulièrement produit dans le corpus d’Antonin en période 1, associé à [pɔ̃ pje]. Cette combinaison est probablement à interpréter en considérant à la fois la cible ‘camion (de) pompier’ et le bruit émis par la sirène du véhicule.

mère pose deux questions à l'enfant contenant le syntagme 'l'ambulance', mais contrairement à ce que nous avons pu voir plus tôt, les deux énoncés ne sont ni dénominatifs ni métalinguistiques. L'enfant répond d'abord par deux phrases 'oui' puis reprend le syntagme complet [lãbylãs]. Dans cette séance, le nom est donc non seulement conforme à la cible adulte sur le plan phonologique mais également précédé d'un défini, identique à celui que la mère vient de produire.

Deux éléments nous ont particulièrement intéressée dans ces extraits : la double fonction des reprises de la mère et l'évolution dans le maniement du lexème par l'enfant. D'une part, on note ici que conformément à ce qui a pu être largement décrit dans la littérature (notamment, Clark et Chouinard, 2000, Bernicot *et al.*, 2006), les énoncés de la mère servent à ratifier les productions de l'enfant mais aussi à apporter des reformulations phonologiques et/ou morphologiques implicites. D'autre part, on voit que la progression de l'enfant se situe elle aussi à deux niveaux puisque dans la séance à 3;2, soit dix mois après le premier extrait, le nom est phonologiquement adéquat et en outre, produit au sein d'un syntagme 'déterminant + nom'. Si dans les deux premières séances, le nom n'est précédé que d'omissions, les reprises de la mère permettent à Antonin d'entendre le nom produit dans le cadre de syntagmes nominaux complets. La progression semble être d'abord d'ordre phonologique (passage d'une production bisyllabique à une production trisyllabique), mais nous pouvons faire l'hypothèse que les offres de la mère puissent dès ce premier stade avoir un impact sur la grammaticalisation du nom par l'enfant quelques mois plus tard.

Enfin, nous nous sommes appuyée pour notre étude sur l'analyse des reprises d'un même nom, mais les résultats de travaux relatifs à l'amorçage syntaxique nous permettent d'envisager que les productions pré-nominales de l'enfant pourraient être influencées par les énoncés antérieurs même dans les cas où l'on n'observe pas de continuité lexicale entre l'énoncé cible et l'énoncé précédent. Par exemple, l'étude de Kemp *et al.* (2005) portant principalement sur l'appréhension des catégories 'déterminant' et 'adjectif' par l'enfant, a particulièrement retenu notre attention pour l'éclairage qu'elle apporte sur le lien entre les énoncés entendus précédemment par l'enfant et la nature des formes pré-nominales auxquelles il a recours ultérieurement. Dans cette étude expérimentale, les auteurs ont comparé la production de syntagmes nominaux au sein de deux groupes d'enfants, l'un ayant bénéficié d'un amorçage syntaxique préalable, et l'autre non, et ont

relevé que le premier groupe (avec amorçage syntaxique) manifestait une plus grande capacité à comprendre et produire des déterminants et des adjectifs que le second. Dans l'expérience menée sans amorçage syntaxique (à laquelle il est fait référence comme l'expérience 1 dans l'article), chacun des 60 enfants (âgés de 2 à 4 ans) a entendu deux mots nouveaux ('toma' et 'modi') associés à l'indéfini 'a' ou au défini 'the', et renvoyant à deux peluches utilisées au cours du jeu avec l'expérimentateur. Ces deux noms ont ensuite été repris par l'enfant, spontanément ou parce que le contexte impliquait qu'il y fasse référence. Sur les 50% d'enfants de 2 ans ayant ensuite réutilisé ces noms précédés d'un déterminant, la moitié d'entre eux les ont produits combinés au même déterminant que celui offert dans le modèle de l'expérimentateur. L'autre moitié a eu recours à d'autres combinaisons mais impliquant à chaque fois l'indéfini 'a', en opposition au défini 'the'. Dans le cadre de la seconde expérience, basée sur de la description d'images, 192 enfants âgés de 2 à 7 ans environ ont entendu l'expérimentateur décrire une série d'images ('prime pictures') à l'aide de syntagmes nominaux variés et ont été en alternance eux-mêmes sollicités pour décrire les images cibles de l'expérience ('target pictures'). Les résultats montrent que dans ce contexte, les enfants ont plus fréquemment produit des syntagmes de type 'déterminant + nom' (le plus souvent de type 'a + nom') que dans la première expérience. Les auteurs signalent toutefois que ces résultats sont à interpréter en tenant compte du plus grand degré de difficulté de l'expérience 1 par rapport à l'expérience 2. Les noms utilisés dans l'expérience 1 n'existant pas dans la langue, ils étaient nécessairement nouveaux pour tous les enfants. En revanche, dans le deuxième cas, les enfants étaient familiers des noms utilisés pour décrire les images. Au-delà de ce biais lié au lexique utilisé, nous pourrions aussi penser que l'enfant est plus sensible au fait d'entendre à plusieurs reprises des énoncés contenant différents types de syntagmes nominaux (et dont le but - ici, décrire une image - correspond à celui de son énoncé) qu'au fait d'entendre au préalable les syntagmes cibles associés à un seul déterminant. Appliquées à nos données, ces interprétations pourraient nous amener à nous interroger sur les interactions entre les syntagmes nominaux produits par l'adulte et ceux produits par l'enfant, y compris lorsque ceux-ci impliquent des noms différents. Suite à ces observations, nous pourrions évaluer plus globalement l'influence de la structure 'déterminant + nom' sur les usages que les enfants de notre corpus font des formes pré-nominales.

Avant de conclure ce chapitre, nous pouvons donc rappeler que nos résultats ne semblent pas aller dans le sens d'une influence notable des reprises de l'enfant sur la nature des formes produites. Toutefois, ces résultats sont à interpréter à la lumière des choix méthodologiques que nous avons faits et ne représentent en aucun cas un bilan général sur le rôle des reprises dans le développement langagier des quatre enfants de notre corpus.



Au sein de ce chapitre VII, nous avons cherché à compléter les observations faites au chapitre VI en analysant également l'influence de certains facteurs fonctionnels sur la production des formes pré-nominales. Nous nous sommes intéressée d'une part aux contextes pragmatico-discursifs dans lesquels sont produits les syntagmes nominaux, et d'autre part, aux reprises que l'enfant fait de son propre discours ou de celui de son interlocuteur. Les résultats ont montré que la distribution globale des formes pré-nominales, ainsi que la distribution plus spécifique des types de fillers et de déterminants produits par l'enfant pouvaient en partie dépendre de facteurs pragmatico-discursifs. En revanche, nous n'avons pas constaté que les reprises semblaient - dans une perspective synchronique - influencer de façon notable la nature des syntagmes produits par l'enfant.

Le dernier objectif que nous nous étions fixé dans le cadre de cette thèse était de mettre en perspective les facteurs distributionnels et fonctionnels impliqués dans la mise en place des déterminants. Au chapitre suivant, nous observerons donc conjointement tous les facteurs considérés au sein des chapitres VI et VII, et nous tenterons de mesurer statistiquement l'influence respective de chacun d'entre eux, en période 1 et en période 2, dans le processus d'acquisition des déterminants chez les quatre enfants de notre corpus.



# CHAPITRE VIII

*Observation croisée des influences  
distributionnelles et fonctionnelles*

---



Au sein du chapitre VI, nous avons relevé que les facteurs phono-prosodique, sémantique et lexico-syntaxique pouvaient - à des degrés différents - nous renseigner sur les contextes facilitant le développement progressif des déterminants. Nous avons toutefois signalé qu'aucun d'entre eux ne semblait rendre compte à lui seul de l'intégralité des productions de l'enfant, quelle que soit la période observée. Ces premiers constats nous ont ensuite conduit à considérer également, au chapitre VII, l'impact de certains facteurs fonctionnels. Les résultats de nos analyses nous ont entre autres montré que si des facteurs pragmatico-discursifs pouvaient expliquer une grande partie des usages que les enfants font des déterminants, ceux-ci ne nous éclairaient pas de façon systématique sur la variation entre omissions, fillers et déterminants, particulièrement observée en première période. Il semblerait donc que, comme les facteurs distributionnels, les facteurs fonctionnels ne suffiraient pas à expliquer tout le processus d'acquisition des déterminants.

Dans ces deux chapitres VI et VII, chacun des facteurs étudiés a fait l'objet d'une analyse spécifique et a donc été traité indépendamment des autres facteurs. Or, au même titre que l'acquisition du langage ne repose pas sur le développement successif de plusieurs niveaux langagiers mais plutôt sur la mise en place concomitante de compétences phonologiques, lexicales, grammaticales et pragmatiques, nous pouvons faire l'hypothèse que l'acquisition du paradigme des déterminants reflèterait la sensibilité de l'enfant à plusieurs dimensions langagières. Dans cette perspective, il semblerait qu'une étude multidimensionnelle, proposant une observation conjointe de l'ensemble des facteurs distributionnels et fonctionnels examinés dans cette thèse, pourrait compléter et affiner notre compréhension du phénomène d'acquisition des déterminants.

En s'appuyant sur cette idée d'une interaction entre différents niveaux langagiers, des chercheurs ont examiné certaines des caractéristiques du langage de l'enfant en combinant l'analyse de plusieurs dimensions. On trouve notamment dans la littérature un certain nombre de travaux sur l'interface entre syntaxe et pragmatique, comme ceux de Serratrice *et al.* (2004) et d'Allen (*i.a.* 2006). Dans la première étude, les auteures ont comparé les productions d'un enfant bilingue anglais-italien à celles d'enfants monolingues anglais et italiens et ont entre autres constaté que dans le corpus de l'enfant bilingue, l'influence de la structure syntaxique de la langue *a* ou *b* sur les formes produites dans l'autre langue se manifestait non pas de façon aléatoire mais dans des contextes pragmatico-discursifs spécifiques. L'analyse de facteurs fonctionnels a donc permis de compléter et d'affiner la compréhension des processus à l'œuvre dans le développement syntaxique des enfants bilingues. Dans son travail, Allen s'est elle aussi intéressée aux positions sujet et objet, et en particulier aux omissions, et a pu établir que les omissions étaient liées à la fois à des facteurs syntaxiques et pragmatiques, et non pas à l'un ou l'autre de ces facteurs. Elle a par ailleurs souligné qu'en faisant appel à plusieurs perspectives différentes (« 'cocktail' view »), la description et l'explication des cas d'omissions étaient plus riches que dans le cas d'une focalisation sur l'une d'entre elles, supposée répondre à elle seule à toutes les questions (« 'magic bullet' view »). Plus récemment, les travaux menés dans le cadre du projet de recherche DIAREF (Salazar Orvig, de Weck, Hassan et Rialland, en préparation) ont établi d'une part que le choix et l'usage des expressions référentielles pouvait être déterminé à la fois par des facteurs distributionnels (alternance des formes, cadres syntaxiques) et fonctionnels (statut attentionnel et discursif du référent, place dans le dialogue), et d'autre part, que certains aspects formels tels que la prosodie pouvaient être influencés par des critères pragmatiques et discursifs. Enfin, on peut citer également, dans une perspective plus globale, les travaux de Bates et MacWhinney (par exemple, 1989) autour du modèle dit 'de compétition'. En mettant en évidence la concurrence ou au contraire la convergence entre des indices prosodique, syntaxique, sémantique et pragmatique, les auteurs ont élaboré un modèle permettant d'après eux de prédire la façon dont l'enfant acquiert sa langue et l'ordre dans lequel sont faites les acquisitions. Bien que ces quelques exemples ne portent pas tous spécifiquement sur la question de l'acquisition des déterminants, ils nous permettent cependant de confirmer l'intérêt que peut représenter une approche multidimensionnelle, et en particulier, une approche tenant compte à la fois de facteurs formels et fonctionnels.

Dans la lignée des travaux que nous venons de citer, nous nous concentrerons dans ce dernier chapitre sur l'apport conjoint des facteurs formels et fonctionnels, et de façon plus globale, nous mettrons en perspective toutes les contraintes analysées dans cette thèse, afin d'une part d'essayer expliquer une plus grande partie des productions de l'enfant, et d'autre part, de répondre à certaines de nos questions initiales : a) l'un des facteurs considérés prévaut-il tout au long du processus d'acquisition ? b) est-ce que plusieurs d'entre eux prédominent de façon successive ou bien au contraire, est-ce qu'ils affectent la mise en place des déterminants durant la même période et de façon complémentaire ? c) les trois formes pré-nominales (omissions, fillers et déterminants) sont-elles soumises à l'influence des mêmes facteurs ?

Les résultats présentés aux chapitres précédents nous permettent d'envisager que tous les facteurs ne seront pas aussi pertinents les uns que les autres pour expliquer la production des formes pré-nominales. On peut par exemple penser que le facteur phono-prosodique apparaîtra comme un facteur prédominant en période 1 pour toutes les formes, et qu'en période 2, les facteurs pragmatico-discursifs seront les plus significatifs. En revanche, il peut être attendu qu'aucune des formes n'apparaisse comme déterminée en priorité par le statut sémantique du nom ou par le fait que celui-ci soit repris. En ce qui concerne le facteur lexico-syntaxique, on peut faire l'hypothèse que celui-ci ne prédira pas de façon très significative la production des trois formes. Toutefois, dans la mesure où tous les facteurs sont ici analysés les uns par rapport aux autres, il se peut aussi qu'ils apparaissent comme ayant une portée différente (et donc plus ou moins prononcée) de celle observée lorsque ces facteurs sont considérés de façon indépendante.

Ce chapitre est organisé en deux parties. La première sera consacrée à quelques préalables méthodologiques, concernant notamment le choix et l'élaboration du test statistique que nous avons utilisé. Dans la seconde partie, nous ferons la présentation des résultats de notre analyse pour chacune des formes pré-nominales, ainsi qu'une synthèse générale de ces résultats.

## **1. Traitement statistique des données**

Afin de répondre aux questions que nous venons de soulever, nous avons dans un premier temps mesuré le degré de significativité des facteurs grâce à un modèle mixte (Generalized Linear Mixed Model). Le choix de cet outil sera justifié au point 1.1 ci-dessous. Nous décrirons ensuite au point 1.2 la façon dont nous avons procédé pour établir le modèle. En dernier lieu, nous

avons représenté sous la forme d'arbres de partition la façon dont se distribuent les omissions, fillers et déterminants en fonction des facteurs conservés au sein de notre modèle. Quelques précisions nécessaires à la lecture de ces arbres seront apportées au lecteur au point 1.3.

### **1.1 Intérêts du modèle mixte pour notre étude**

Le modèle mixte nous a semblé être le plus adapté à notre étude. D'une part, ce test permet d'évaluer le degré de significativité de plusieurs variables (considérées comme les 'effets fixes'), mais aussi de tenir compte de certains facteurs dits 'aléatoires', tels que les enfants ou les observations. Nous pouvions donc au travers de ce modèle prendre en considération la question primordiale de la variabilité individuelle. D'autre part, ce test ne nécessite pas que l'analyse repose sur un nombre d'occurrences identiques pour tous les enfants et pour chaque session. Dès lors que nos corpus et nos séances ne sont pas équivalent(e)s les un(e)s par rapport aux autres et contiennent un nombre variable de formes pré-nominales, ce test paraissait particulièrement approprié au format de données que nous observons. Par ailleurs, signalons que dans la mesure où nos données ne suivent pas une distribution dite 'normale', nous ne pouvions pas envisager d'avoir recours à des analyses de la variance (ANOVA) ou à des régressions linéaires simples.

### **1.2 Elaboration du modèle**

Tous nos tests ont été réalisés avec la fonction `glmer` du programme de statistiques R. Dans un premier temps, nous avons testé un modèle comprenant tous les facteurs analysés pour chaque axe (phono-prosodique, sémantique, lexico-syntaxique, pragmatico-discursif et pour les reprises). Seuls deux facteurs ont d'emblée été exclus de l'analyse : le statut concret/abstrait du nom (du fait du peu d'occurrences de noms abstraits observés), ainsi que la dépendance au lexème, dans la mesure où il nous a été impossible de convertir l'analyse qualitative de l'overlap (présentée au point 3.2 du chapitre VI) en système binaire. Nous avons ensuite sélectionné le modèle le plus adapté, par élimination successive des variables les moins significatives. Nous avons commencé par éliminer celles pour lesquelles une autre variable relative à la même catégorie de facteurs était déjà représentée dans le modèle et apparaissait comme significative. C'est le cas par exemple de la variable 'référence particulière/indéterminée' qui ne semblait pas être une variable fonctionnelle significative alors qu'en parallèle, le caractère référentiel du nom apparaissait lui comme une variable fonctionnelle pertinente.

Nous le verrons plus tard, ce ne sont pas les mêmes facteurs qui *in fine* apparaissent comme les plus significatifs pour les trois formes pré-nominales, mais pour pouvoir comparer les résultats obtenus, le modèle final que nous avons établi est identique pour les omissions, les fillers et les déterminants. A chaque analyse, nous avons codé 1 pour la variable dépendante (respectivement filler, omission, déterminant) et 0 pour toutes les autres formes. Ce modèle comporte une variable d'ordre phono-prosodique, une variable relative au statut sémantique du nom, une variable lexico-syntaxique et trois variables fonctionnelles, deux concernant le statut pragmatico-discursif du référent et une définissant si le nom a été repris ou non du discours antérieur. Au vu de la grande différence dans la distribution des formes entre la période 1 et 2, nous avons aussi considéré la période observée comme un facteur fixe (1 ou 2). Enfin, nous avons retenu pour les effets aléatoires les enfants et les séances. Toutes les variables pour le calcul des effets fixes sont présentées et décrites dans le tableau 1 ci-dessous.

**Tableau 1 - Présentation des variables retenues pour le calcul des effets fixes**

Facteurs	Variables retenues	Description
<b>Distributionnels</b>		
Phono-prosodique	IMPAIRSYLL	contextes où le nom n'est pas bisyllabique (1), par opposition à ceux où il est bisyllabique ou pair (0)
Sémantique	INANIMÉ	cas où le nom est inanimé (1), par opposition à ceux où il est inanimé ou renvoie à un procès (0)
Lexico-syntaxique	CADREPRIV	contextes dans lesquels les noms sont employés en combinaison avec l'un des quatre cadres considérés comme privilégiés dans chaque corpus (1) <i>vs.</i> cas de noms utilisés hors cadre ou avec des cadres non privilégiés (0)
<b>Fonctionnels</b>		
Pragmatico-discursif	RÉFÉRENTIEL	nom possédant une valeur référentielle (1) <i>vs.</i> noms utilisés de façon non référentielle et incertitudes entre valeur référentielle et dénominative du nom (0)
	DONNÉ	cas où le référent a été donné dans le discours (1) en contraste avec les cas où il est nouveau, focalisé, réintroduit et non référentiel (0)
Reprise	REPRISE	nom repris du discours antérieur (1) ou non repris (0)
<b>Autres</b>		
Période	PÉRIODE	Première période ( $\leq 1$ ) ou deuxième période ( $> 1$ )

Les résultats de ces analyses sont présentés pour chacune des formes pré-nominales dans les trois premiers points de la section 2.

### 1.3 Arbres de partition

Nous avons par ailleurs voulu comprendre comment se distribuent les facteurs présentés ci-dessus et de quelle façon ils interagissent. Nous avons eu recours pour cela à des arbres de partition (conditional inference trees), réalisés à l'aide de la fonction `ctree` du Package `Tree` du logiciel R. Ces arbres permettent de visualiser d'une part la distribution des variables considérées et d'autre part, l'interaction entre ces différentes variables (Hothorn *et al.*, 2016).

Nous avons donc fait figurer dans la section 2 qui suit trois arbres, correspondant aux résultats pour chacune des trois formes pré-nominales et élaborés sur la base des facteurs considérés au préalable dans notre modèle final (IMPAIRSYLL, INANIMÉ, CADREPRIV, RÉFÉRENTIEL, DONNÉ et REPRISE), y compris ceux n'apparaissant pas comme significatifs. Nous verrons ultérieurement, lors de la présentation des résultats pour les fillers, que quelques différences ont pu être relevées entre les résultats obtenus avec le modèle et ceux représentés par les arbres. Ces différences peuvent être liées à la fois au fait que les arbres de partition ne prennent pas en compte les effets aléatoires et au fait que ceux-ci représentent non seulement la distribution des formes en fonction des facteurs mais aussi leur interaction.

Tous les facteurs sont sur les arbres représentés par des nœuds. Plus ceux-ci se situent en haut de l'arbre et plus ils rendent compte de la production de la forme pré-nominale étudiée. Les histogrammes figurant au bas des arbres correspondent à la proportion d'utilisation des formes pour chaque combinaison de facteurs. Chaque arbre sera présenté et discuté à la suite des résultats du modèle, dans la seconde partie des sections 2.1, 2.2 et 2.3.

## 2. Prévalence - constante ou périodique - d'un facteur ou complémentarité des facteurs ?

Comme préalablement annoncé, les résultats seront présentés pour chacune des formes pré-nominales. Nous commencerons par la présentation des résultats pour les omissions (2.1) puis nous poursuivrons avec la discussion de ceux concernant les fillers (2.1) et les déterminants (2.3). Dans chaque section, nous commenterons d'abord les résultats du test et ensuite, l'arbre de partition. Pour rappel, nous ferons au point 2.4 une synthèse générale des résultats obtenus.



## 2.1 Omissions

Le tableau 2 ci-dessous, comme les tableaux 3 et 4 qui suivent, nous montre d'une part quelles variables ('effets fixes') apparaissent comme les plus significatives (*cf.* colonne 'p') et d'autre part, quel type de relation - positive ou négative - entretiennent les variables avec la forme pré-nominale étudiée (*cf.* colonne 'z').

**Tableau 2 - Présentation des résultats du test pour les omissions**

Effets aléatoires	Variance	SD		
Session	0.506262	0.71152		
Enfant	0.002358	0.04856		
Effets fixes	Estimate	SE	z	p
(Intercept)	3.03294	0.45593	6.652	2.89e-11 ***
IMPAIRSYLL	-1.11475	-1.11475	-11.183	< 2e-16 ***
INANIMÉ	-0.12401	0.10899	-1.138	0.2552
CADREPRIV	0.48953	0.11407	4.291	1.78e-05 ***
RÉFÉRENTIEL	-0.80013	0.16029	-4.992	5.99e-07 ***
DONNÉ	0.30652	0.15790	1.941	0.0522 .
REPRISE	0.18680	0.14681	1.272	0.2032
PÉRIODE	-2.35421	0.30187	-7.799	6.26e-15 ***

Note : SD=Standard Deviation ; SE=Standard Error ; Nombre d'observations : 2716 ; groupes sessions : 32 ; enfants : 4 ; C : 0.8240282

Quatre variables apparaissent comme particulièrement significatives pour les omissions : le nombre de syllabes du nom, la présence d'un cadre lexico-syntaxique privilégié, le caractère référentiel du nom et la période. Nous pouvons constater que les omissions ne sont pas favorisées dans les cas où le nom n'est pas bisyllabique (et donc impair) puisque la relation entre la variable 'IMPAIRSYLL' et la forme est négative (-11.183). Ces résultats sont conformes à ceux obtenus au chapitre VI puisque nous avons pu constater que les noms bisyllabiques généraient une part plus importante d'omissions que de fillers ou de déterminants. La relation entre la variable 'CADREPRIV' et la forme est positive, ce qui signifie que les cadres lexico-syntaxiques privilégiés tendent à favoriser la production d'omissions. Si ces résultats n'étaient *a priori* pas

attendus, nous verrons ultérieurement qu'ils s'expliquent notamment par le lien entretenu entre cette variable et d'autres facteurs.

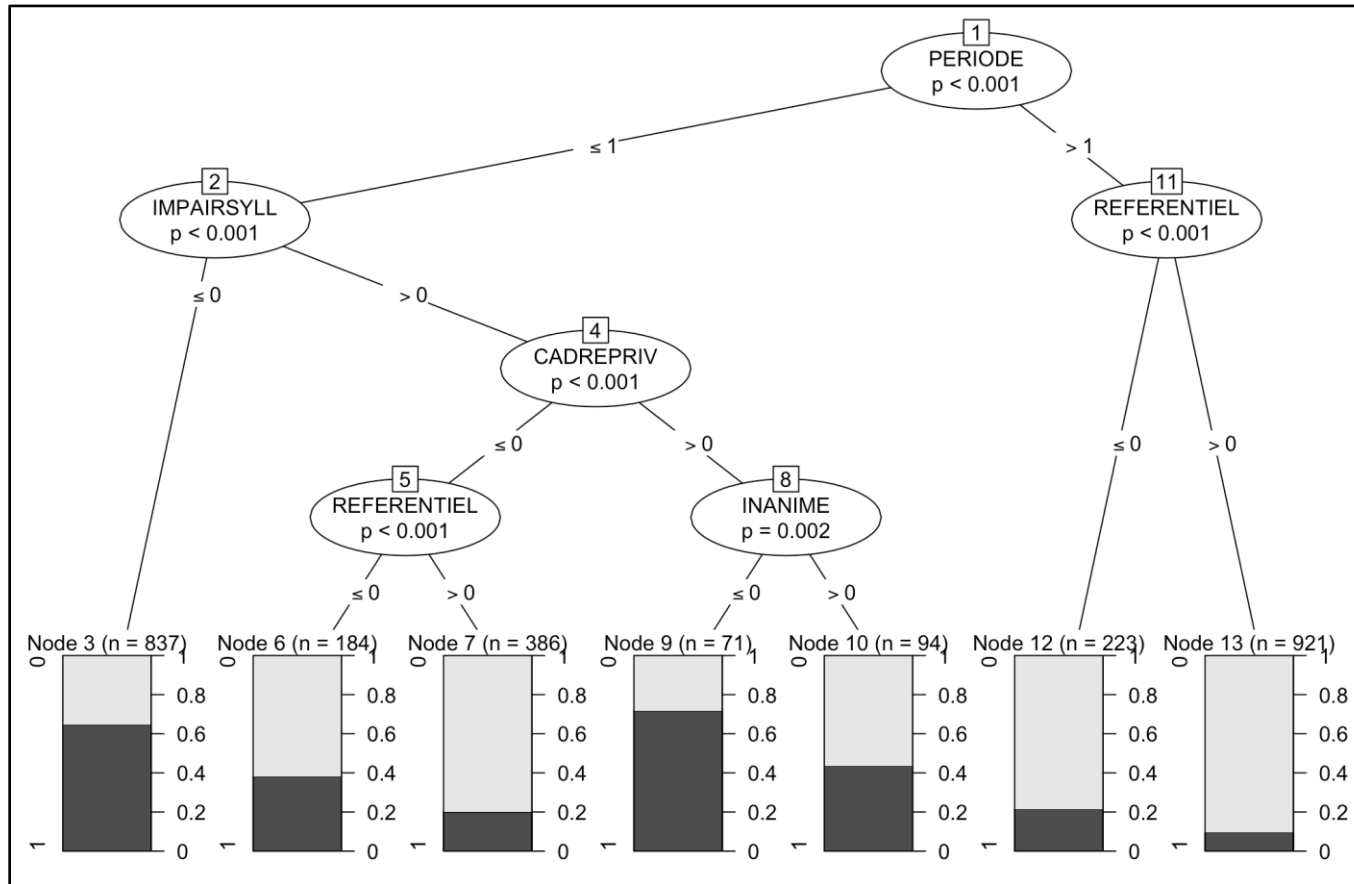
Le contexte référentiel apparaît quant à lui comme défavorable ( $z=-4.992$ ) à la production d'omissions. Ce résultat ne paraît pas surprenant dans la mesure où au chapitre précédent, nous avons vu que ce contexte favorisait la production de déterminants et le contexte non référentiel, la production de formes de transition. Conformément à ce que nous pouvions attendre, les résultats confirment que la période 1 est favorable à l'omission, contrairement à la période 2.

Enfin, nous pouvons observer une tendance significative avec la variable 'INANIMÉ', et une relation négative entre cette variable et l'omission, ce qui suggère que les omissions tendraient à être moins produites dans les contextes où le nom est inanimé. Ici aussi, les résultats sont conformes aux observations antérieures puisque nous avons pu relever que les animés étaient légèrement plus fréquemment associés à des formes de transition que les inanimés. Le statut attentionnel 'DONNÉ' et le fait que le nom soit réutilisé n'apparaissent quant à eux pas comme des variables significatives, ce qui confirme également les résultats présentés au sein du chapitre VII.

Plusieurs constats peuvent être faits à la lecture de l'arbre ci-dessous. En premier lieu, nous pouvons remarquer que la période est le facteur le plus significatif pour prédire la production des omissions.

Nous pouvons noter ensuite qu'en période 1, la structure syllabique apparaît comme le deuxième facteur susceptible d'entraîner (ou non) la production des omissions. Celles-ci sont plus fréquentes lorsque le nom est bisyllabique (et donc 'non impair') que lorsqu'il n'est pas bisyllabique. Dans ces contextes où le nom n'a pas une structure paire, le cadre syntaxique est le troisième facteur le plus à même de prédire la production d'omissions. On observe par ailleurs une interaction avec le caractère référentiel dans le cas où le cadre n'est pas un cadre privilégié, et avec le statut inanimé lorsqu'au contraire, le cadre est privilégié.

Arbre 1 - Présentation des résultats pour les omissions



Les omissions sont plus fréquentes dans les contextes où le nom est non référentiel que dans les cas où il est produit avec une valeur référentielle. Le contexte non référentiel favorise donc la production d'omissions, conformément à ce que nous avons pu voir à la fois au chapitre VII et dans les résultats qui précèdent. En outre, lorsque le nom est produit dans un cadre privilégié, les omissions sont plus nombreuses avec les noms qui ne sont pas inanimés qu'avec les noms inanimés.

En période 2, le facteur référentiel est le seul facteur apparaissant comme significatif (après la période) pour pouvoir prédire la production d'un déterminant, et nous pouvons remarquer que les omissions sont plus souvent observées dans les contextes non référentiels que référentiels.

## 2.2 Fillers

En ce qui concerne les fillers, nous pouvons voir (*cf.* tableau 3 ci-dessous) que comme pour les omissions, la structure syllabique apparaît comme une variable significative, mais contrairement à ce que nous venons de voir, la relation entre cette variable et le filler est positive. Dans les contextes où le nom possède un nombre impair de syllabes (donc non bisyllabique), le filler est plus produit que lorsque le nom est bisyllabique. Ces résultats corroborent donc les observations faites au chapitre VI.

Le cadre lexico-syntaxique est lui aussi un facteur significatif mais à l'inverse des résultats concernant les omissions, la relation entre cette variable et la forme pré-nominale est négative. Les fillers tendraient donc à être produits davantage lorsque le nom n'est pas associé à un cadre lexico-syntaxique privilégié. Par ailleurs, le fait que le référent ait déjà été mentionné ne favorise pas la production de fillers. Ceci peut s'expliquer d'une part par le constat que nous avons fait d'une présence plus importante de fillers que d'autres formes dans les contextes où le référent est focalisé, et d'autre part, par le fait que par défaut, tous les noms qui n'ont pas une valeur référentielle ont été codés 0 alors que les noms renvoyant à un référent donné dans le discours ont été codés 1. Ces résultats peuvent donc signifier à la fois que le filler est plus fréquemment associé aux cas où le référent n'a pas été mentionné mais aussi plus largement, qu'il serait privilégié dans les contextes où le nom n'est pas référentiel, ce qui confirmerait ce que nous avons pu noter au chapitre précédent. Conformément à nos hypothèses, pour les fillers comme pour les omissions, la période est une variable qui apparaît comme très significative.

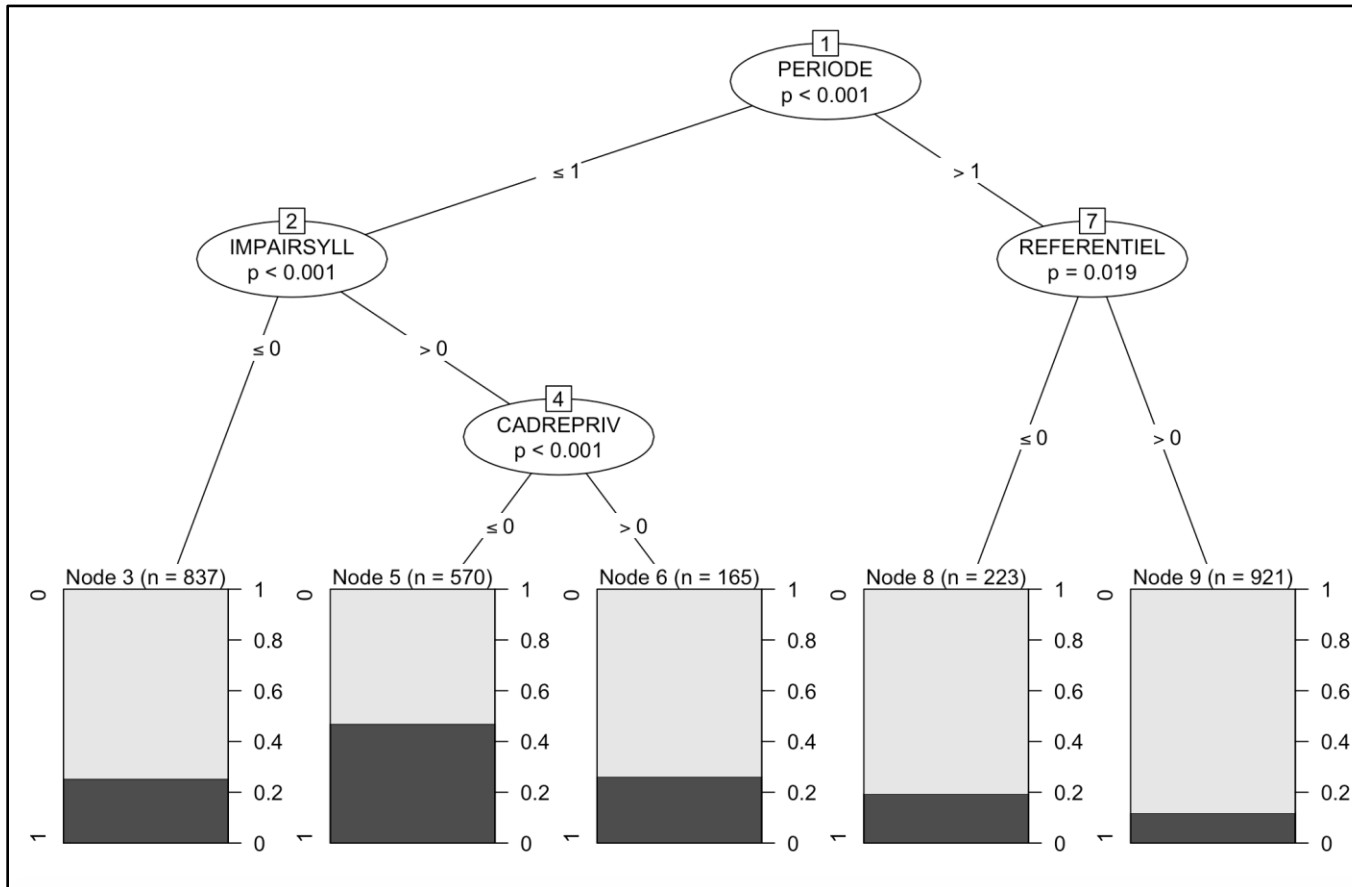
**Tableau 3 - Présentation des résultats du test pour les fillers**

Effets aléatoires	Variance	SD		
Séance	0.5414	0.7358		
Enfant	0.1187	0.3445		
Effets fixes	Estimate	SE	z	p
(Intercept)	0.68475	0.49509	1.383	0.16664
IMPAIRSYLL	0.52970	0.09776	5.418	6.02e-08 ***
INANIMÉ	-0.15740	0.10959	-1.436	0.15094
CADREPRIV	-0.34378	0.11834	-2.905	0.00367 **
RÉFÉRENTIEL	0.20106	0.15203	1.322	0.18601
DONNÉ	-0.45606	0.14804	-3.081	0.00206 **
REPRISE	-0.11603	0.14001	-0.829	0.40725
PÉRIODE	-1.46072	0.30977	-4.715	2.41e-06 ***

Note : SD=Standard Deviation ; SE=Standard Error ; Nombre d'observations : 2716 ; groupes sessions : 32 ; enfants : 4 ; C : 0.7445376

En revanche, le statut sémantique et le caractère référentiel du nom, ainsi que le fait que celui-ci soit repris (ou non repris) ne sont pas des variables significatives. Les résultats de l'analyse concernant les reprises concordent avec les observations que nous avons faites dans la seconde partie du chapitre VII. De même, les inanimés ne générant qu'une part légèrement moins importante de fillers et d'omissions que les animés, les résultats relatifs au statut sémantique sont conformes à ceux des analyses menées plus tôt pour chaque enfant. Enfin, bien que la variable 'RÉFÉRENTIEL' n'apparaisse pas comme significative, nous pouvons tout de même faire l'hypothèse que l'influence de cette variable (et en particulier, du statut non référentiel du nom) puisse se manifester au travers des résultats concernant le statut attentionnel et discursif du référent (variable 'DONNÉ').

Arbre 2 - Présentation des résultats pour les fillers



A la lecture de l'arbre ci-dessus (arbre 2), on voit que de la même façon que pour l'omission, la période est le facteur prédominant, et qu'en période 1, la structure syllabique est la variable qui prédit le mieux la présence ou l'absence de fillers. Ceux-ci sont plus nombreux lorsque le nom a un nombre de syllabes impair que pair. Lorsque le nom est bisyllabique, le facteur phono-prosodique interagit aussi avec le cadre syntaxique et les fillers sont plus observés dans les cas où le nom n'est pas inclus dans un des cadres préférentiels. Ces résultats concordent avec les résultats que nous venons de présenter.

En période 2, le caractère référentiel est la seule variable qui apparaît comme significative et les résultats nous montrent que le contexte non référentiel est plus favorable aux fillers que le contexte référentiel. Cette variable ne figurait pas parmi les facteurs significatifs dans les résultats du test présentés ci-dessus. En revanche, le statut attentionnel apparaissait comme une variable significative (et reliée négativement aux fillers) et nous pouvons à nouveau préciser que dans la mesure où les noms non référentiels sont tous identifiés comme des noms 'non donnés', l'influence du facteur 'non référentiel' a pu se traduire dans le test statistique précédent par une valeur significative de la variable relative au statut attentionnel.

### 2.3 Déterminants

Comme pour les fillers et les omissions, la structure syllabique est une variable significative pour prédire la production des déterminants. Conformément à ce que nous pouvions attendre, les déterminants sont plus associés au contexte impair que bisyllabique ( $z=6.758$ ). Le statut sémantique apparaît lui aussi comme une variable significative. Rappelons que nous avons vu au chapitre VI que les inanimés sont plus fréquemment associés à des déterminants que les animés. C'est aussi ce que les résultats présentés ici nous montrent puisque la relation entre la variable 'INANIMÉ' et les déterminants est positive.

Le caractère référentiel du nom semble également influencer de façon significative la production des déterminants. Contrairement à ce que nous venons de relever pour les fillers et les omissions - mais conformément aux résultats obtenus pour chaque enfant au chapitre VII - les déterminants sont plus associés aux noms référentiels que non référentiels.

**Tableau 5 - Présentation des résultats du test pour les déterminants**

Effets aléatoires	Variance	SD		
Séance	0.4978	0.7055		
Enfant	0.1847	0.4298		
Effets fixes	Estimate	SE	z	p
(Intercept)	-5.85048	0.52164	-11.216	< 2e-16 ***
IMPAIRSYLL	0.71886	0.10638	6.758	1.40e-11 ***
INANIMÉ	0.35671	0.12149	2.936	0.00332 **
CADREPRIV	-0.15303	0.12268	-1.247	0.21224
RÉFÉRENTIEL	0.66814	0.16159	4.135	3.55e-05 ***
DONNÉ	0.29311	0.14896	1.968	0.04911 *
REPRISE	-0.06466	0.15332	-0.422	0.67321
PÉRIODE	3.09389	0.29650	10.435	< 2e-16 ***

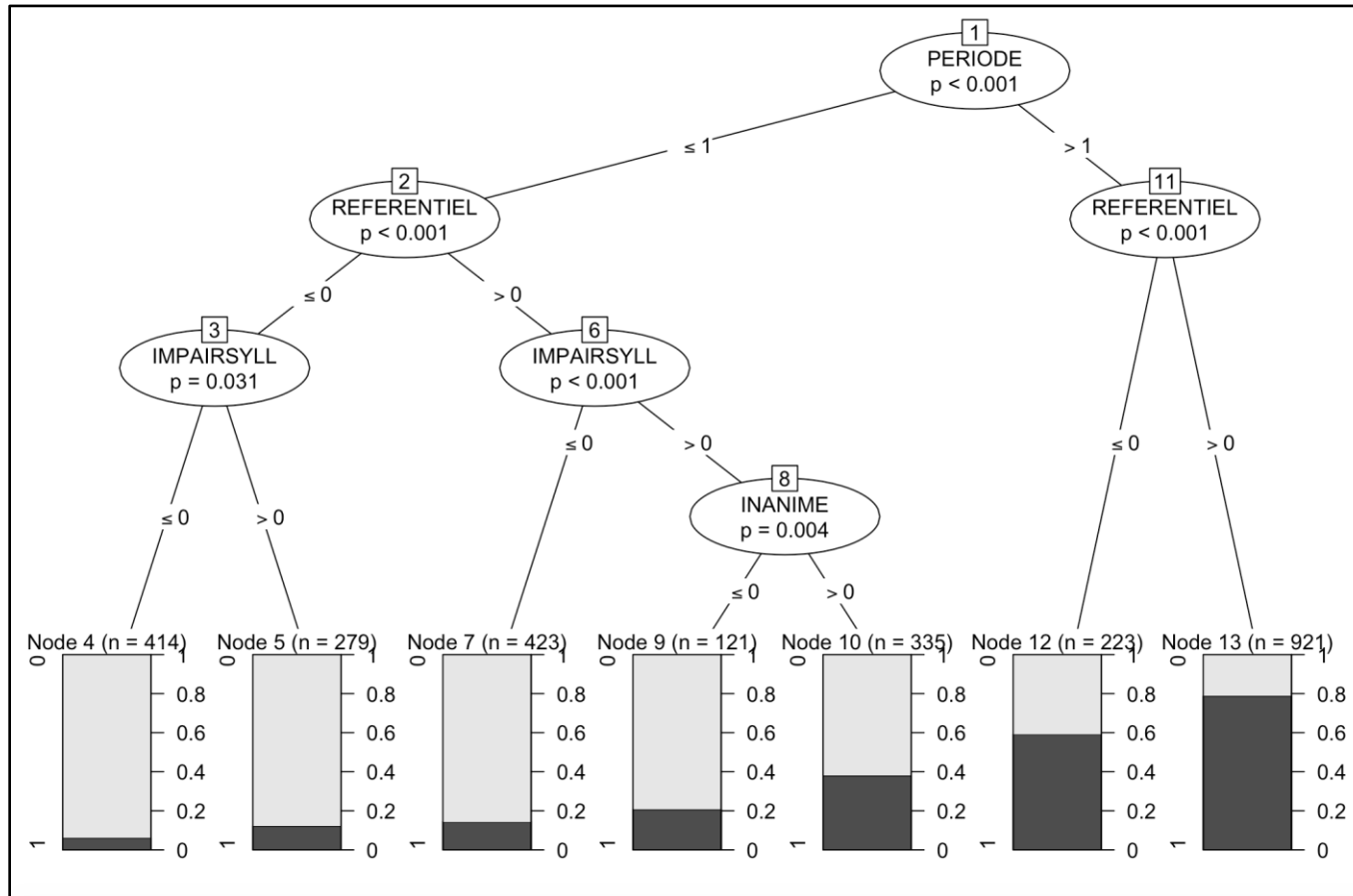
Note : SD=Standard Deviation ; SE=Standard Error ; Nombre d'observations : 2716 ; groupes sessions : 32 ; enfants : 4 ; C : 0.8752159

Les mêmes constats peuvent être faits concernant le statut attentionnel. Les déterminants sont davantage produits lorsque le référent est donné que non donné (et donc, rappelons le, focalisé, nouveau ou réintroduit mais aussi non référentiel). La période apparaît elle aussi comme un facteur pertinent pour prédire la production de déterminants, ceux-ci étant de toute évidence observés davantage en période 2 qu'en période 1.

En revanche, le cadre syntaxique et la réutilisation du nom ne sont pas des facteurs prédisant de façon significative la production de déterminants. Si le cadre influence la production des déterminants, nous avons vu qu'il s'agissait davantage d'une influence sur la nature du déterminant produit que sur la présence ou l'absence de déterminant, donc ces résultats, comme ceux portant sur les reprises n'invalident pas ce que nous avons pu voir plus tôt dans ce travail.



Arbre 3 - Présentation des résultats pour les déterminants



En ce qui concerne la distribution et l'interaction entre les facteurs, on peut noter en observant l'arbre ci-dessous que comme pour les fillers et les omissions, la période est le facteur prépondérant pour prédire la production des déterminants. En revanche, la structure syllabique n'apparaît plus comme le facteur le plus significatif en période 1. Nous remarquons qu'ici, c'est le caractère référentiel qui prime en période 1 comme en période 2. En période 2, il s'agit d'ailleurs du seul facteur observé et on relève que les déterminants sont plus fréquents en contexte référentiel que non référentiel.

En période 1, lorsque le nom n'est pas référentiel, les déterminants sont plus souvent produits dans les contextes impairs que bisyllabiques. Dans les cas où le nom a une valeur référentielle, les déterminants sont également moins nombreux dans les contextes bisyllabiques que pairs et par ailleurs, plus représentés avec les noms inanimés qu'animés. En période 1, les déterminants sont donc privilégiés dans les contextes où le nom est référentiel, impair et inanimé.

## 2.4 Synthèse des résultats

Pour conclure, nous pouvons dire que les résultats concernant le degré de significativité des différentes variables considérées dans notre modèle corroborent les observations que nous avons pu faire aux chapitres V, VI et VII. Les résultats que nous venons de présenter montrent également que la production des omissions, fillers et déterminants semble être conditionnée par un ensemble de facteurs plus que par un facteur unique. En nous intéressant à la façon dont se distribuent et interagissent les variables, nous avons pu voir, outre la place prépondérante de la période, que la production des trois formes s'explique dans un premier temps par une combinaison de facteurs (en période 1), et dans un second temps (en période 2), par le caractère référentiel ou non référentiel du nom. Dans le cas des omissions et des fillers, les facteurs distributionnels (et notamment prosodique et syntaxique) sont prédominants durant la première période, et dans le cas des déterminants, le facteur référentiel apparaît déjà à ce stade comme le plus à même de prédire la présence ou l'absence de déterminants.

Ces résultats concernant les formes adultes suggèrent que la sensibilité de l'enfant aux facteurs fonctionnels se manifesterait dès la première période dans les usages contrastés que celui-ci fait des formes (déterminant *vs.* omission ou filler). Bien que l'on observe que le facteur référentiel apparaisse aussi, en combinaison avec des facteurs distributionnels, comme un facteur

susceptible d'expliquer la production des omissions en période 1, il semble toutefois qu'il ne devienne particulièrement pertinent pour les omissions comme pour les fillers qu'en période 2. Ces deux formes de transition pourraient donc être utilisées dans un premier temps comme un ensemble que l'enfant oppose aux déterminants, avant que chacune d'entre elles ne devienne spécifiquement déterminée par le statut référentiel du nom, et que les facteurs phono-prosodique et syntaxique ne deviennent moins signifiants. Cette évolution des facteurs prédominants suggère que l'acquisition des déterminants est d'emblée multi-déterminée.



# CONCLUSION

En nous appuyant sur les observations des nombreuses études existant sur la mise en place des déterminants chez le jeune enfant, nous nous étions fixé comme objectif principal dans cette thèse de rendre compte de l'acquisition de ce paradigme dans une perspective multidimensionnelle. Les chapitres V à VIII nous ont permis de dresser un certain nombre de constats et de répondre aux questions dont nous avons fait l'inventaire au chapitre III. Cette conclusion sera l'occasion d'en faire la synthèse générale (section 1) mais aussi, et peut-être surtout, de faire un état des lieux des questions restant encore à ce stade sans réponse, et que nous souhaiterions explorer dans des travaux à venir (section 2).

## 1. Synthèse générale

Outre notre objectif général d'une description multidimensionnelle du processus d'acquisition des déterminants, nous avons, en amont de nos chapitres d'analyse, passé en revue les interrogations déjà abordées dans la littérature et sur lesquelles nous souhaitions revenir dans notre étude, et celles n'ayant - à notre connaissance - pas encore été soulevées dans des travaux antérieurs. Nous avons organisé ces différentes questions en quatre points principaux, à l'origine de l'agencement des quatre chapitres précédents : la distribution et la description des formes pré-nominales, l'influence des facteurs distributionnels, l'influence des facteurs fonctionnels, et enfin, la mise en perspective de ces deux dernières dimensions. Ces quatre aspects sont repris ci-dessous afin de faire le bilan non exhaustif mais synthétique des principaux résultats de notre travail.

## 1.1 Description des formes

Au chapitre V, nous avons en premier lieu fait le constat d'une mise en place progressive mais non régulière des déterminants. Conformément à ce qui a déjà pu être observé dans la littérature, nous avons vu que la part des formes de transition - fillers et omissions - diminue à mesure que celle des formes adultes augmente mais nous avons relevé par ailleurs que ces diminution et augmentation ne suivent pas une évolution dont le rythme est strictement régulier.

Nous avons en outre remarqué que le développement de la position pré-nominale n'est pas identique chez les quatre enfants de notre corpus, tant en termes de rythme que de style. Le déterminant devient la forme majoritaire à un âge qui peut être variable d'un enfant à l'autre et tous n'ont pas recours aux fillers pendant la même période et dans les mêmes proportions.

En ce qui concerne plus spécifiquement la nature des déterminants produits, nous avons pu relever que le défini est le déterminant le plus fréquent dans nos données, avant que les formes ne deviennent, en période 2, plus diversifiées. Enfin, suite à l'observation de la forme phonologique des fillers produits dans notre corpus, nous avons pu mettre en évidence le caractère à la fois évolutif et hétérogène des fillers.

## 1.2 Influence des facteurs distributionnels

Au sein du chapitre VI, nous avons tenté de mesurer l'influence des facteurs phono-prosodique, sémantique et lexico-syntaxique sur la mise en place des déterminants. Nous avons vu d'une part que les productions pré-nominales des quatre enfants sont le plus souvent produites en période 1 dans des contextes bisyllabiques ('omission + nom de deux syllabes' et 'filler/déterminant + nom d'une syllabe'), et que cette sensibilité manifeste à la structure syllabique tend à être moins marquée en période 2.

Nous avons d'autre part relevé que les noms renvoyant à des entités inanimées sont plus fréquemment associés à des déterminants que les noms animés. En parallèle, les noms animés semblent être plus favorables à la production de formes de transition que les noms inanimés.

Nous avons ensuite essayé d'évaluer si les formes pré-nominales pouvaient être préférentiellement associées à des noms ou des cadres syntaxiques spécifiques. Nos résultats montrent que lorsque les noms sont produits plusieurs fois, ils tendent à être préférentiellement

- mais pas exclusivement - combinés à l'une des trois formes pré-nominales. En nous penchant plus particulièrement sur les noms produits plusieurs fois avec un déterminant, nous avons pu relever que ceux-ci sont, dans nos données, plus souvent combinés à un seul déterminant qu'à plusieurs déterminants distincts. L'analyse des cadres lexico-syntaxiques dans lesquels les noms sont produits a mis en évidence que ceux-ci peuvent être favorables à la production d'un type de forme ou de déterminant spécifiques mais qu'aucun d'entre eux n'entraînent la production exclusive d'une seule forme ou d'un seul déterminant.

Les résultats discutés au chapitre VI ont donc globalement pu montrer que les trois facteurs distributionnels étudiés semblaient influencer à des degrés divers la production des formes pré-nominales. Ils ont toutefois montré par ailleurs qu'aucun de ces facteurs ne pouvait expliquer toutes les formes produites par l'enfant et/ou tout le processus d'acquisition des déterminants, et que dans certains cas, les facteurs distributionnels semblaient interagir avec certains facteurs fonctionnels.

### **1.3 Influence des facteurs fonctionnels**

Dans le cadre du chapitre VII, nous avons mesuré l'impact de certains facteurs fonctionnels sur le développement du paradigme des déterminants. Nous avons pu noter d'une part que les quatre enfants de notre corpus semblent être sensibles à des facteurs pragmatico-discursifs mais que tous les facteurs étudiés ne semblent pas aussi pertinents les uns que les autres pour expliquer la distribution des formes pré-nominales. En considérant l'ensemble des formes pré-nominales (omissions, fillers et déterminants), nous avons pu voir que seul le contraste entre valeur référentielle et non référentielle donne lieu à des différences claires dans les usages que les enfants font des formes. Les déterminants apparaissent plus fréquemment dans les contextes référentiels que non référentiels et en parallèle, les formes de transition sont plus fréquentes dans les contextes non référentiels que référentiels. En nous focalisant sur la production des déterminants et des fillers, nous avons vu que les définis et les fillers consonantiques apparaissent plus fréquemment que les indéfinis dans les contextes référentiels et particuliers que les indéfinis et les fillers vocaliques. Au contraire, les indéfinis et les fillers vocaliques sont majoritaires en contexte non référentiel et indéterminé. L'enfant semble donc faire des usages contrastés des formes, et aussi bien des déterminants que des fillers.

## CONCLUSION

Nos résultats n'ont par ailleurs pas révélé de différences importantes dans la distribution des formes selon que le nom est repris ou non du discours antérieur, et cela quelle que soit la distance entre les mentions  $n$  et  $n-1$  du nom ou la période observée. Nous avons seulement pu relever qu'en période 2 et dans les cas d'auto-reprise, les syntagmes de type 'omission/filler + nom' peuvent être chez certains enfants plus souvent repris par un syntagme de type 'déterminant + nom' que par un syntagme identique. Enfin, nous avons signalé que ces résultats relatifs aux reprises doivent être mis en perspective avec les choix méthodologiques que nous avons faits (et en particulier, celui d'une analyse synchronique) et qu'ils ne constituent pas un bilan global sur le rôle des reprises dans le développement langagier des quatre enfants.

### **1.4 L'acquisition du déterminant : un phénomène au cœur de plusieurs dimensions**

En dernier lieu, nous avons fait l'analyse conjointe de l'ensemble des facteurs distributionnels et fonctionnels étudiés au sein des chapitres VI et VII. Les résultats de notre analyse statistique ont d'une part confirmé les tendances qui avaient pu être dégagées au préalable, et d'autre part, mis en évidence l'influence complémentaire de plusieurs facteurs, et en particulier au cours de la première période. Pendant cette période, les omissions et les fillers semblent être déterminés surtout par des facteurs d'ordre phono-prosodique et syntaxique, et en période 2, par la valeur référentielle ou non référentielle du nom. La production des formes adultes apparaît quant à elle comme conditionnée en priorité par le facteur référentiel, en période 1 comme en période 2, et suivi durant la première période par le facteur phono-prosodique. Ces résultats témoignent donc de la sensibilité de l'enfant à la fois au contexte pragmatique et à des facteurs formels tels que la structure syllabique du nom. Plus globalement, ils permettent de souligner et de confirmer l'intérêt de considérer la multiplicité des facteurs en jeu dans le processus d'acquisition des déterminants.

## **2. Un tableau à compléter**

Si nous avons essayé, nous semble-t-il, de rendre compte dans cette thèse de la multiplicité des facteurs à considérer pour pouvoir décrire plus précisément la mise en place des déterminants chez le tout jeune enfant, d'autres aspects seraient toutefois à considérer pour compléter le tableau que nous avons dépeint dans ce travail. Il apparaît notamment assez clairement que la



catégorie des fillers pourrait être explorée de façon plus précise et que le développement des déterminants devrait être, dans une étape ultérieure, appréhendé dans le cadre global de celui des expressions référentielles. Il nous semblerait également pertinent de considérer le rôle joué par les activités dans lesquelles les enfants sont engagés, ainsi que les types de discours qui y sont associés. Nos analyses pourraient aussi être étendues à d'autres langues, afin de mesurer entre autres quels facteurs sont spécifiques à chaque langue et lesquels sont à envisager comme universels, ou du moins, communs à des langues dont les structures sont comparables. Enfin, augmenter la quantité de données analysées (langage produit par l'enfant mais aussi par l'adulte) permettrait d'affiner les descriptions que nous avons effectuées et de prendre en compte de façon systématique l'influence du langage adressé à l'enfant sur la production des formes. Toutes ces perspectives très diverses de prolongement du travail mené dans cette thèse seront développées dans les cinq sections à suivre.

## **2.1 La catégorie transitionnelle des fillers : que reste-t-il à explorer ?**

Dans la mesure où notre travail s'est intéressé au développement des formes en position pré-nominale, seuls les fillers pré-nominaux ont été analysés dans cette thèse. Ces fillers nous ont renseignée sur la façon dont l'enfant s'approprie la forme et la fonction des déterminants mais pour en dire davantage sur l'usage que l'enfant fait des fillers et couvrir toutes les dimensions de cette catégorie transitionnelle, il nous faudrait également dans un second temps les étudier dans le contexte de l'acquisition des formes pré-verbales. Explorer la catégorie transitionnelle des fillers dans son ensemble, en ciblant donc aussi bien les fillers produits devant les noms que devant les verbes, nous permettrait par exemple, comme a pu le faire Veneziano (2003), d'appréhender plus précisément leur évolution et leur éventuelle spécification phonologique et fonctionnelle selon leur position dans l'énoncé.

Un autre aspect concernant l'étude des fillers nous semblerait particulièrement intéressant à considérer plus largement : celui de la dimension formelle ou phonologique de ces éléments. Dans un travail en cours mené conjointement avec N. Yamaguchi, nous avons fait l'analyse de deux corpus : celui d'Adrien (utilisé dans cette thèse), pour lequel nous avons retenu six séances de 2;5 à 3;5, et celui d'une fille, Madeleine, pour qui nous avons sélectionné six séances

également, de 1;7 à 2;2<sup>1</sup>. Nous avons notamment pu constater que les macro-catégories utilisées pour décrire la forme des fillers pré-nominaux (consonne, voyelle et consonne-voyelle) n'étaient pas suffisantes pour travailler sur l'interface phonologie-pragmatique. A la suite d'un travail initié par Yamaguchi, Rialland et Salazar Orvig (2014), il nous a paru préférable de nous appuyer sur le contexte pragmatique pour faire des hypothèses sur le type de déterminant qui pourrait être attendu dans la langue adulte, et ensuite, de décrire les fillers produits notamment en fonction de la conformité ou non de leurs consonnes (lorsqu'une consonne était produite) avec des déterminants de la langue adulte contenant des consonnes. Ce travail nous a entre autres permis de repérer que Madeleine produit des fillers dont la consonne est conforme aux déterminants définis (et contenant donc la consonne /l/) dans des contextes exclusivement référentiels, alors que dans une étape antérieure, nous avons seulement pu constater que l'enfant semblait utiliser indistinctement des fillers vocaliques et consonantiques dans des usages référentiels et non référentiels. Affiner notre description des fillers nous permettra donc d'approfondir la question des contrastes pragmatiques présents dans le langage de l'enfant avant même la mise en place des contrastes entre formes adultes. De plus, dans la perspective d'un travail global sur les fillers pré-nominaux et pré-verbales, l'analyse détaillée des formes phonologiques permettra de dire avec plus de précision si ces deux ensembles tendent à une certaine période à ressembler davantage à des formes adultes de déterminants d'une part et à des formes adultes de pronoms d'autre part et si donc, elles se spécifient progressivement.

## **2.2 Appréhender le développement des déterminants en lien avec le développement global des expressions référentielles**

De même que la catégorie transitionnelle des fillers pourrait être explorée davantage en prenant en compte à la fois les formes pré-nominales et pré-verbales, il nous semblerait intéressant, dans un travail futur, d'essayer d'explorer et de comprendre le développement des syntagmes nominaux, et donc celui des déterminants, en lien avec le développement global des expressions référentielles. Si les enfants réfèrent avec des syntagmes nominaux, ils réfèrent aussi (et entre autres) avec des pronoms. Comment et à quelle période se mettent en place ces formes dans notre corpus ? Peut-on observer certaines 'interférences' entre le développement des pronoms et

---

<sup>1</sup> Corpus recueilli dans le cadre du corpus CoLaJe et disponible sur la base de données CHILDES (<http://childes.psy.cmu.edu/browser/index.php?url=French/Paris/Madeleine/>)

celui des déterminants ? Est-ce que certaines des ‘régressions’ constatées dans nos corpus pourraient être imputées au développement parallèle de la position pré-verbale ? Est-ce que la grammaticalisation (à la fois formelle et fonctionnelle) de l’une de ces positions peut freiner ou au contraire étayer la grammaticalisation de l’autre ? Les contrastes dans l’usage alterné des syntagmes nominaux et des pronoms sont-ils significatifs du point de vue pragmatique ? Peut-on imaginer que la maîtrise progressive des fonctions pragmatico-discursives des déterminants puisse aider l’enfant à appréhender le maniement général des expressions référentielles ? S’il semble difficile de pouvoir analyser chaque aspect du développement de l’enfant en regard de tous les autres aspects de son développement, il semblerait toutefois bénéfique pour la connaissance du contexte dans lequel se met en place le déterminant, de pouvoir répondre à ces différentes questions, en considérant donc la mise en place des noms et des déterminants, non seulement comme un phénomène au centre d’un faisceau de facteurs mais aussi au sein du développement général du langage.

Une part des analyses sur la position pré-verbale ayant déjà été effectuée pour le corpus d’Adrien, ce travail comparatif entre développement pré-nominal et pré-verbal pourra être initié prochainement.

### **2.3 Influence de l’activité et des types de discours**

Certains auteurs (*i.a.* De Weck et Rosat, 2003 ; Kern et Chenu, 2010) ont montré que les facteurs - formels ou fonctionnels - influençant la production des formes chez l’enfant et chez son interlocuteur ne peuvent pas être observés sans considérer les activités dans lesquelles les participants sont engagés et les types d’énoncés ou de discours qui leur sont associés. Dans la perspective de ces travaux et suite aux premiers constats que nous avons pu faire dans cette thèse (influence de l’activité sur la proportion des noms produits, ainsi que sur leur nature sémantique ou sur le choix des types de cadres syntaxiques et de formes pré-nominales), un travail ultérieur pourra donc avoir pour objectif d’explorer l’impact des activités et des types de discours sur la production des formes pré-nominales produites par les enfants de notre corpus.

Nous pourrions ainsi analyser chaque occurrence de nom et chaque forme pré-nominale selon le type d’activité en cours (jeu symbolique, lecture, goûter, ...) et le type de discours (narration, description, dénomination, ...). A la suite de certains travaux (Salazar Orvig *et al.*, à paraître), nous

## CONCLUSION

pouvons dans un premier temps faire l'hypothèse que certaines activités comme les activités du quotidien (goûter ou bain par exemple) pourraient entraîner un usage plus important de noms que celles qui s'appuient sur des supports iconiques. Nous pourrions par ailleurs supposer que parmi ces activités générant une production plus fréquente de noms, certaines formes pré-nominales - dont le type serait variable d'un enfant à l'autre et d'une période à l'autre - seraient associées de façon privilégiée à certains types de discours. Nous pouvons envisager que la maîtrise du paradigme des déterminants puisse se manifester, comme nous l'avons vu dans cette thèse, par une augmentation de la proportion des formes adultes et des usages adéquats de ces formes dans le contexte interactionnel, mais aussi par une diversification des usages qu'en font les enfants. Dans la lignée de ce qui a pu être observé dans le domaine de l'acquisition des constructions et des catégories grammaticales (*cf.* Pine et Lieven, 1997 ou Tomasello, 2000a), nous pouvons imaginer que l'enfant puisse passer de la production de syntagmes nominaux très figés et dépendant des activités et des types de discours à des usages des formes plus décontextualisés et abstraits.

### 2.4 Comparaisons inter-langues

A la lecture des travaux portant sur des comparaisons inter-langues (*cf.* par exemple, Rozendaal, 2008 ou Lleò et Demuth, 1999), nous avons pu mesurer l'importance du rôle joué par les spécificités de la langue sur la production des formes pré-nominales. En nous limitant dans un premier temps à un périmètre de langues pour lesquelles les déterminants sont, comme pour le français, requis la plupart du temps devant les noms (l'anglais et l'allemand par exemple), nous aimerions, en collaboration avec des collègues locuteurs de ces langues mesurer entre autres l'impact de la langue en cours d'acquisition sur le rythme et les stratégies d'acquisition des déterminants, ainsi que sur les caractéristiques des fillers et des premières formes adultes. Identifier des fillers et des déterminants ainsi que leurs conditions d'apparition dans d'autres langues nous permettrait par ailleurs de pouvoir observer si les enfants peuvent être sensibles aux mêmes facteurs, formels et fonctionnels, que ceux examinés dans cette thèse ou si au contraire, les spécificités (notamment phonologiques, morphologiques, syntaxiques) de chaque langue entraînent un traitement très différent des formes, lié de façon privilégiée à certains facteurs plus qu'à d'autres.

## 2.5 Extension du corpus analysé

Notre travail soulève donc des interrogations auxquelles nous n'avons pas pu répondre ici mais qui constitueront autant de perspectives ultérieures de travail. Toutefois, sans même aborder des questions différentes de celles que nous avons pu poser dans cette étude, le prochain défi à relever sera d'augmenter la quantité de données analysées. Cette extension se fera à trois niveaux différents. Nous aimerions dans un premier temps faire l'analyse de toutes les séances enregistrées pour les corpus que nous avons collectés et examiner des séances supplémentaires pour Adrien, afin, dans les deux cas, de pouvoir baser notre étude sur un corpus d'environ une séance par mois pendant tout le temps du processus d'acquisition des déterminants. Une fois ces séances supplémentaires analysées, il sera envisageable de compléter les séances qui n'ont été qu'en partie traitées. Nous pourrions ainsi étoffer nos codages, affiner nos descriptions, et notamment celles des étapes par lesquelles passe l'enfant pendant la construction du paradigme.

Il nous semblerait par ailleurs particulièrement judicieux de faire l'analyse des productions des interlocuteurs de l'enfant au sein du corpus actuellement analysé. Nous avons essayé dans ce travail de prendre en compte le contexte discursif, et plus largement, le contexte interactionnel, au travers des analyses des cadres syntaxiques, du statut pragmatico-discursif des référents et de l'étude des reprises. Néanmoins, nous n'avons pas directement examiné le langage adressé à l'enfant. Nous aimerions dans une prochaine étude pouvoir comparer systématiquement les productions de l'enfant à celles de l'adulte pour traiter de façon plus approfondie la relation entre le développement du langage et l'expérience communicative de l'enfant. Nous pourrions ainsi par exemple mesurer l'impact des formes utilisées par les adultes ou les aînés devant les noms sur les productions de l'enfant, et définir non pas des profils individuels de gestion des noms et des déterminants mais des profils de dyades. Ceci impliquerait peut-être à terme de reconsidérer certains des aspects de l'émergence des formes que nous avons initialement attribués à des questions de variation individuelle et qui nécessiteraient d'être requalifiés comme des phénomènes liés à des variations inter-dyadiques plus qu'inter-individuelles.



## BIBLIOGRAPHIE

- Allen, S.E.M. (2006), « Formalism and functionalism working together ? Exploring roles for complementary contributions in the domain of child null arguments ». In R. Slabakova, S. Montrul et P. Prévost (eds.), *Inquiries in linguistic development : In honor of Lydia White*, pp. 233-255
- Ambridge, B. (sous presse), « Syntactic categories in child language acquisition : innate, induced or illusory ? ». In H. Cohen et C. Lefebvre (eds.), *Handbook of categorization in cognitive science*, Amsterdam : Elsevier
- Ambridge, B. et Lieven, E.V.M. (2015), « A Constructivist account of child language acquisition ». In B. MacWhinney and W. O'Grady (eds.), *Handbook of Language Emergence*, Hoboken, NJ: Wiley Blackwell, pp. 478-510
- Austin, J.L. (1969), *Quand dire, c'est faire*, Paris : Editions du Seuil
- Bailey, J. (2008), « First steps in qualitative data analysis : transcribing », *Family practice*, 25, pp. 127-131
- Baldwin, D.A. (1991), « Infants' contribution to the achievement of joint reference », *Child Development*, 63, pp. 875-890
- Bassano, D. (2015), « The acquisition of nominal determiners: Evidence from crosslinguistic approaches ». In L. Serratrice et S. E. M. Allen (eds.), *The acquisition of Reference*, Amsterdam / Philadelphia : John Benjamins Publishing Company, pp. 25-49
- Bassano, D. (2010), « L'acquisition du déterminant nominal en français : une construction progressive et interactive de la grammaire », *CogniTextes*, 5, [En ligne], consulté le 02 mars 2015. URL : <http://cognitextes.revues.org/315>
- Bassano, D. (2008), « Acquisition du langage et grammaticalisation: le développement des noms et des verbes en français ». In F. Labrell et G. Chasseigne (eds.), *Cognition, Santé et Vie Quotidienne*, Numéro spécial : *Le développement conceptuel*, Paris : Publibook Université, collection Psychologie Cognitive, pp. 17-50
- Bassano, D., Maillochon I., et Mottet S. (2008), « Noun grammaticalization and determiner use in French children's speech: a gradual development with prosodic and lexical influences », *Journal of Child Language*, 35, pp. 403-438
- Bassano, D. (2005), « Production naturelle précoce et acquisition du langage : l'exemple du développement des noms ». *LIDIL*, 31, [En ligne], consulté le 21 août 2014. URL : <http://lidil.revues.org/136>
- Bassano, D., et Eme, P.E. (2001), « Variabilités développementale et inter-individuelle dans la grammaticalisation des noms en français ». In A. Flieller, C. Bocerean, J.L. Kopp, E. Thiebaut, A.M. Toniolo, & J. Tournois, *Questions de Psychologie Différentielle*, Rennes : Presses Universitaires de Rennes, pp. 215-222

## BIBLIOGRAPHIE

- Bassano, D. (2000), « Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar », *Journal of Child Language*, 27, pp. 521-559
- Bates, E., Dale, P. et Thal, D. (1995), « Individual differences and their implications for theories of language development ». In P. Fletcher et B. MacWhinney (eds.), *Handbook of child language*, Oxford : Blackwell, pp. 96-151
- Bates, E. et MacWhinney, B. (1989), « Functionalism and the Competition Model ». In B. MacWhinney et E. Bates (eds.), *The cross-linguistics study of sentence processing*, New-York : Cambridge University Press, pp. 3-73
- Bates, E., Bretherton, I. et Snyder, L. (1988), *From first words to grammar : individual differences and dissociable mechanisms*, Cambridge : Cambridge University Press
- Belletti, A. et Guasti, M.T. (2015), *The acquisition of Italian, Morphosyntax and its interfaces in different modes of acquisition*, Amsterdam / Philadelphia : John Benjamins Publishing Company, pp. 25-49
- Bernicot, J., Salazar Orvig, A. et Veneziano, E. (2006), « Les reprises : dialogue, formes, fonctions et ontogenèse », *La Linguistique*, 42, pp. 29-50
- Bernicot, J. (1998), « L'acquisition du langage : étapes et théories ». In R. Ghiglione et J. F. Richard (eds.), *Développement et intégration des fonctions cognitives, Cours de psychologie* (vol. 3), Paris : Dunod, pp. 420-439
- Bernicot, J. et Roux, M. (1998), « La structure et l'usage des énoncés : comparaison d'enfants uniques et d'enfants seconds nés ». In J. Bernicot, H. Marcos, C. Day, M. Guidetti, V. Laval, J. Rabain-Jamin et G. Babelot (eds), *De l'usage des gestes et des mots chez l'enfant*, Paris : A. Colin, pp. 157-178
- Bernicot, J. (1992), *Les actes de langage chez l'enfant*, Paris : PUF
- Bertin, T. (2011), *Grammaticalisation du langage de l'enfant: processus interactionnel d'appropriation des articles et des clitiques sujets chez des enfants francophones entre 1 et 3 ans*, Thèse de doctorat, Université de Nancy
- Blanche-Benveniste, C. (2010), *Approches de la langue parlée en français*, Paris : Ophrys
- Blanche-Benveniste, C. et Jeanjean, C. (1987), *Le français parlé*, Paris : Didier Erudition
- Bloom, L., Hood, L. et Lightbown, P. (1974), « Imitation in language development : if, when and why », *Cognitive Psychology*, 6, pp. 380-420
- Bloom, L. (1970), *Language development : form and function in emerging grammars*, Cambridge, MA : MIT Press
- Bloom, P. (1990), « Subjectless sentences in child language », *Linguistic Inquiry*, 2, pp. 491-504
- Bogny Yapo, J. (2007), « Le Modèle Chomskyen de la description linguistique : des Principes et Paramètres au Programme Minimaliste », *Séminaire sur la Grammaire Générative*, Université Félix-Houphouët-Boigny, août 2007, Abidjan



- Bottari, P., Cipriani, P. et Chilosi, A.M. (1993/94), « Protosyntactic devices in the acquisition of Italian free morphology », *Language Acquisition*, 3, pp. 327-369
- Bourdieu, P. (1993), *La misère du monde*, Paris : Seuil
- Boysson-Bardies, B. de (1996), *Comment la parole vient aux enfants*, Paris : Editions Odile Jacob
- Braine, M.D.S. (1992), « What sort of innate structure is needed to “bootstrap” into syntax ? », *Cognition*, 45, pp. 77-100
- Braine, M.D.S. (1971), « The acquisition of language in infant and child ». In C. E. Reed (ed.), *The learning of language*, New York : Appleton-Century-Crofts, pp. 7-95
- Braine, M. D. S. (1963), « The ontogeny of English phrase structure », *Language*, 39, pp. 1-13
- Bronckart, J.-P. (1977), « Acquisition du langage et développement cognitif ». In *La genèse de la parole, Symposium de l'Association de psychologie scientifique de langue française*, Paris : PUF, pp. 137-159
- Brogère, G. (2006), « Les enfants, les adultes et l'observateur ». In R. Hess et G. Weigand (dir), *L'observation participante dans les situations interculturelles*, Paris : Anthropos, pp. 207-223
- Brown, R. (1973), *A first language : the early stages*, Cambridge, MA : Harvard University Press
- Bruner, J.S. (1987), *Comment les enfants apprennent à parler*, Paris : Retz
- Bruner, J.S. (1983), *Le développement de l'enfant : Savoir faire, savoir dire*, Paris : PUF
- Bruner, J.S. (1975), « From communication to language: a psychological perspective », *Cognition*, 3, pp. 255-287
- Caët, S. (2013), *Référence à soi et à l'interlocuteur chez des enfants francophones et anglophones et leurs parents*, Thèse de doctorat, Université Sorbonne Nouvelle-Paris 3
- Cameron-Faulkner, T., Lieven, E.M.V et Tomasello, M. (2003), « A construction based analysis of child directed speech », *Cognitive Science*, 27, pp. 843-873
- Caronia, L. (2014), « Talking objects : the embarrassing epistemic work of things in the research setting », communication orale, *Séminaire de l'équipe Sesylia*, (université Sorbonne Nouvelle-Paris 3), Paris, novembre 2014
- Charolles, M. (2002), *La référence et les expressions référentielles en français*, Paris : Ophrys
- Choi, S. et Gopnik, A. (1995), « Early acquisition of verbs in Korean : a cross-linguistic study », *Journal of child language*, 22, pp. 497-529
- Chomsky, N. (2006)[1968], *Language and Mind*, New-York : Cambridge University Press
- Chomsky, N. (2002)[1966], *Cartesian Linguistics, A chapter in the history of rationalist thought*, Christchurch, New Zealand : Cybereditions
- Chomsky, N. (2002)[1957], *Syntactic structures*, Berlin/New-York : Mouton de Gruyter
- Chomsky, N. (1993), *Language and Thought*, Wakefield, RI: Moyer Bell
- Chomsky, N. (1988), *Language and Problems of Knowledge : The Managua Lectures*, Cambridge, MA : The MIT Press

## BIBLIOGRAPHIE

- Chomsky, N. (1981), *Lectures on government and binding*, Dordrecht : Foris
- Chomsky, N. (1965), *Aspects of the theory of syntax*, Cambridge, MA, MIT Press
- Chomsky, N. (1964), *Current issues in linguistic theory*, The Hague : Mouton
- Clark, E.V. et Chouinard, M.M. (2000), « Enoncés enfantins et reformulations adultes dans l'acquisition du langage », *Langages*, 140, pp. 9-23
- Coene, M. (2005), « On the acquisition of the indefinite article : A cross-linguistic study of french, italian, romanian and spanish child speech », *Belgian Journal of Linguistics*, 19, pp. 121-143
- Dannequin, C. (1989), « Le corpus oral dans l'observation du langage de l'enfant », *Linx*, 20, pp. 21-30
- Darwin, C. (1877), « A biographical sketch of an infant », *Mind*, 2, pp. 285-294
- De Cat, C. (2011), « Information tracking and encoding in early L1: linguistic competence vs. cognitive limitations », *Journal of Child Language*, 38, pp. 828-860
- Demuth, K., Patroliu, M., Song, J.Y. et Masapollo, M. (2012), « The development of articles in children's early Spanish : prosodic interactions between lexical and grammatical form ». In J. Rothman et Pedro Guijarro-Fuentes (eds.), *Linguistic interfaces and language acquisition in childhood*, *First Language*, 32, pp. 17-37.
- Demuth, K. et Tremblay, A. (2008), « Prosodically-conditioned variability in children's production of French determiners », *Journal of Child Language*, 35, pp. 99-127
- Demuth, K. (2001), « A prosodic approach to filler syllables ». *Journal of Child Language*, 28, pp. 246-249
- Demuth, K. (2000), « Bantu noun class systems : Loan word and acquisition evidence of semantic productivity ». In G. Senft (ed.), *Classification Systems*, Cambridge : Cambridge University Press, pp. 270-292
- Demuth, K. (1996), « The prosodic structure of early words ». In J. Morgan et K. Demuth (eds.), *Signal to syntax : Bootstrapping from speech to grammar in early acquisition*, Mahwah, NJ : Lawrence Erlbaum Associates, pp. 171-184
- Demuth, K. (1996), « Stages in the acquisition of prosodic structure ». In E. Clark (ed.), *Proceedings of the 27th Child Language Research Forum*, Stanford University : CSLI, pp. 39-48
- Demuth, K. (1992), « Accessing functional categories in Sesotho: Interactions at the morpho-syntax interface ». In J. Meisel (ed.), *The Acquisition of Verb Placement : Functional Categories and V2 Phenomena in Language Development*, Kluwer Academic Publishers : Dordrecht, pp. 83-107
- Dore, J., Franklin, M.B., Miller, R.T. et Ramer, A.L.H. (1976), « Transitional phenomena in early language acquisition », *Journal of Child Language*, 3, pp. 13-28
- Dubois-Charlier, F. et Vautherin, B. (2008), « La grammaire générative et transformationnelle : bref historique », *La Clé des Langues*, [En ligne], consulté le 3 avril 2015. URL :

- <http://cle.ens-lyon.fr/plurilingues/la-grammaire-generative-et-transformationnelle-bref-historique-48418>
- Ducrot, O. et Schaeffer, J.-M. (1995), *Nouveau Dictionnaire Encyclopédique des Sciences du Langage*, Paris : Seuil
- Ducrot, O. (1972), *Dire et ne pas dire*, Paris : Hermann
- Egré, P. (2011), « Philosophie de la linguistique ». In A. Barberousse, D. Bonnay et M. Cozic (eds), *Précis de philosophie des sciences*, [En ligne], consulté le 2 avril 2015. URL : [paulegre.free.fr/Papers/philo\\_ling7.doc](http://paulegre.free.fr/Papers/philo_ling7.doc)
- Ferguson, C.A. (1977), « Baby talk as a simplified register ». In C.E. Snow et C.A. Ferguson (eds.), *Talking to children*, New York : Cambridge University Press, pp. 209-235
- Ferguson, C.A. (1964), « Baby talk in six languages », *American Anthropologist*, 66, pp. 103-114
- Fikkert, P. (1994), *On the acquisition of prosodic structure*, Thèse de doctorat, Université de Leiden
- Fortis, J. M. (2008), « Le langage est-il un instinct? Une critique du nativisme linguistique, de Chomsky à Pinker », *Texte !*, 13, [En ligne], consulté le 10 avril 2015. URL : <http://www.revue-texto.net/index.php?id=1870>
- Gadet, F. (2000), « Derrière les problèmes méthodologiques du recueil des données ». In M. Bilger (dir.), *Linguistique sur corpus*, Perpignan : Presses Universitaires de Perpignan, pp. 59-75
- Gerken, L. (1996), « Prosodic structure in young children's language production », *Language*, 72, pp. 683-712
- Gerken, L. (1994), « Young children's representation of prosodic phonology: Evidence from English-speakers' weak syllable productions », *Journal of Memory and Language*, 33, pp. 19-38
- Gibson, B.E. (2005), « Co-producing video diaries : the presence of the "absent" researcher », *International Journal of Qualitative Methods*, 4, pp. 34-43
- Glanzer, P.D. et Dodd, D.H. (1975), « Developmental changes in the language spoken to children », communication orale, *Biennial meeting of the Society for Research in Child Development*, Denver, Avril 1975
- Gleitman, L.R. (1990), « The structural sources of verb meaning », *Language Acquisition*, 1, pp. 3-55
- Gleitman, L.R. et Wanner, E. (1982), « Language acquisition: The state of the state of the art ». In E. Wanner et L. R. Gleitman (eds.), *Language acquisition: State of the art*, New York: Cambridge University Press, pp. 3-48
- Granfeldt, J. (2003), *L'acquisition des catégories fonctionnelles. Etude comparative du développement du DP français chez des enfants et des adultes*, Thèse de doctorat, Université de Lund
- Grégoire, A. (1937). *L'apprentissage du langage, les deux premières années*, Paris : Alcan

## BIBLIOGRAPHIE

- Green, J., Franquiz, M. et Dixon, C. (1997), The myth of the objective transcript : transcribing as a situated act », *TESOL Quarterly*, 31, pp. 172-176
- Gubrium, J.F. et Holstein, J.A. (1997), *The new language of qualitative method*, New-York : Oxford University Press
- Gumperz, J. (1982), *Discourse strategies*, Cambridge : Cambridge University Press
- Gundel, J.K., Ntelitheos, D. et Kowalsky, M. (2007), « Children's use of referring expressions : Some implications for Theory of Mind », *ZAS Papers In Linguistics*, 48, pp. 1-21
- Gundel, J.K., Sera, M., Kowalsky, M., et Page, S. (2001), « Cognitive status, implicature, and children's use of referring forms in English and Spanish », communication affichée, *Society for Research in Child Development Annual Meeting*, Minneapolis, mai 2001
- Gundel, J., Sera, M. et Page, S. (1999), « The acquisition of referring expressions in English and Spanish », Communication orale, *Child Language Research Forum*, Stanford, avril 1999
- Gundel, J.K. et Page, S. (1998), « The givenness hierarchy and children's use of referring expressions ». Communication orale, *Annual Meeting of the Linguistic Society of America*, New-York, janvier 1998
- Halliday, M.A.K. (1973), « The functional basis of language ». In B. Bernstein (ed.), *Applied studies towards a sociology of language (vol. 2), Class, codes and control*, Londres : Routledge et Kegan Paul, pp. 343-366
- Heurdier, J. (2009), « Fillers en position pré-nominale en français : éléments entre phonologie, morphologie et pragmatique dans le cadre de la grammaticalisation du paradigme des déterminants », *Actes des XI<sup>èmes</sup> RJC de l'ED 268*, Paris, pp. 61-64
- Hickmann, M., Schimke, S. et Colonna, S. (2015), « From early to late mastery of reference : Multifunctionality and linguistic diversity ». In L. Serratrice et S. E. M. Allen (eds.), *The acquisition of Reference*, Amsterdam / Philadelphia : John Benjamins Publishing Company, pp. 181-211
- Hickmann, M. (2003), *Children's Discourse: Person, Time and Space Accross Languages*, Cambridge, MA : Cambridge University Press
- Hickmann, M., Hendricks, H., Roland, F., et Liang, J. (1996), « The marking of new information in children's narratives: A comparison of English, French, German and Mandarin Chinese ». *Journal of Child Language*, 3, pp. 591-619
- Hilaire-Debove, G., Régol, V. et Jisa, H. (2002), « Développement morpho-phonologique de deux enfants en train d'acquérir le français après un implant cochléaire », *Actes des XXIV<sup>èmes</sup> Journées d'Etude sur la Parole*, Nancy, pp. 202-205
- Hirsh-Pasek, K. et Golinkoff R.M. (1996), *The origins of grammar. Evidence from early language comprehension*, Cambridge, MA : MIT Press
- Hoekstra, T. et Hyams, N. (1995), « Missing heads in child language ». In F. Wijnen et C. Koster (eds.), *Proceedings of GALA*, Groningen : Center for Language and Cognition, pp. 251-260

- Höhle, B. (2009), « Bootstrapping mechanisms in first language acquisition », *Linguistics*, 47, pp. 359-382
- Hothorn, T., Kornik, K., Strobl, C. et Zeileis, A. (2016), « Package party. A laboratory for recursive partytioning », Texte publié sur le site du projet CRAN-R, [En ligne], consulté le 2 décembre 2016. URL : <https://cran.r-project.org/web/packages/party/party.pdf>
- Hyams, N.M. (1986), *Language acquisition and the theory of parameters*, Dordrecht : Reidel Publishing Company
- Ingram, D. (1989), *First Language Acquisition. Method, Description and Explanation*, Cambridge, MA : Cambridge University Press
- Itard, J. (1801-1807), *Mémoire et Rapport sur Victor de l'Aveyron*. Publiés à la suite de l'ouvrage Lucien Malson, *Les enfants sauvages*, Paris : Bibliothèque 10/18, octobre 2002
- Jisa, H. et Richaud, F. (1994), « Quelques sources de variation chez les enfants », *Acquisition et Interaction en Langue Etrangère*, 4, pp. 7-51
- Joly-Rissoan, O. (2004), « Une méthode de traitement sociologique de données filmées », *Bulletin de méthodologie sociologique*, 82, pp. 27-41
- Judd, C.M., Smith, E.R. et Kidder, L.H. (1991), *Research Methods in Social Relations*, 6th edition, London: Harcourt Brace Jovanovich
- Kail, M. (2000), « Perspectives sur l'acquisition du langage ». In M. Kail et M. Fayol (eds.), *L'acquisition du langage 1*, Paris : PUF, pp. 9-27
- Kail, M. et Bassano, D. (2000), « Méthodes d'investigation et démarches heuristiques ». In M. Kail et M. Fayol (eds.), *L'acquisition du langage 1*, Paris : PUF, pp. 29-60
- Kail, M., et Hickmann, M. (1992), « French children's ability to introduce referents in narratives as a function of mutual knowledge », *First Language*, 12, pp. 73-94
- Karmiloff-Smith, A. (1979), *A functional approach to child language: a study of determiners and reference*. Cambridge : Cambridge University Press
- Kemp, N., Lieven, E.V.M et Tomasello, M. (2005), « Young children's knowledge of the "determiner" and "adjective" categories », *Journal of Speech, Language and Hearing Research*, 48, pp. 592-609
- Kern, S. et Chenu, F. (2010), « Contextes de production et LAE en français : caractéristiques pragmatiques et syntaxiques », *Actes du Congrès Mondial de Linguistique Française (CMLF)*, Nouvelle-Orléans, USA, juillet 2010
- Kilani-Schoch, M. et Dressler, W.U. (2000), « Are fillers as precursors of morphemes relevant for morphological theory ? A case story from the acquisition of French ». In W.U. Dressler, O.E. Pfeiffer, M.A. Pochtrager et J.R. Rennison (eds.), *Morphological Analysis in Comparison*, Amsterdam/Philadelphia : John Benjamins, pp. 89-111
- Kilani-Schoch, M. (1998), « Continuité des catégories dans l'acquisition L1 : l'exemple de fillers », *Cahiers de l'ILSL*, 11, pp. 155-173

## BIBLIOGRAPHIE

- Kupisch, T. (2006), « The emergence of article forms and functions in a German-Italian bilingual child ». In C. Lleó (ed.), *Interfaces in multilingualism: Acquisition, representation and processing*, Amsterdam : John Benjamins, pp. 45-109
- Labov, W. (1972a), *Sociolinguistic patterns*, Philadelphia : University of Pennsylvania Press
- Labov, W. (1972b), « Some principles of linguistic methodology », *Language in Society*, 1, pp. 97-120
- Lallier, C. (2011/2), « L'observation filmante. Une catégorie de l'enquête ethnographique », *L'Homme*, 198-199, pp. 106-130
- Lave, J. et Wenger, E. (1991), *Situated Learning, Legitimate Peripheral Participation*, Cambridge, MA : Cambridge University Press
- Le Mené, M., Yamaguchi, N., Fox, G., Jullien, S. et Salazar Orvig, A. (en préparation), « Acquiring a paradigm : the case of determiners ». In Salazar Orvig, A., de Weck, G., Hassan, R. et Rialland, A. (eds), *Constructing reference in dialogue*, Amsterdam / Philadelphia : John Benjamins Publishing Company
- Le Mené, M. et Yamaguchi, N. (2016), « Determiner use in early French at the crossroads of phonology and pragmatics », *Communication orale, VIIIth International Conference of Language Acquisition*, Palma (Mallorca), septembre 2016
- Leroy-Collombel, M. (2009), « La reformulation dans les interactions adulte-enfant : une analyse longitudinale de 1;06 à 2;08 », *Cahiers de praxématique*, 52, pp. 59-80
- Lleó, C. et Demuth, K. (1999), « Prosodic constraints on the emergence of grammatical morphemes : crosslinguistic evidence from germanic and romance languages », In A. Greenhill, H. Littlefield, et C. Tano (eds), *Proceedings of the 23rd Annual Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press, pp. 407-418.
- Lieven, E.V.M., Pine, J.M. et Baldwin, G. (1997), « Lexically-based learning and early grammatical development », *Journal of child language*, 24, pp. 187-219
- López Ornat, S. (2003), « Learning earliest grammar: evidence of grammar variations in speech before 22 months ». In S. Montrul et F. Ordóñez (eds.), *Linguistic Theory and Language Development in Hispanic languages*, Somerville, MA: Cascadilla Press, pp. 254-274
- Lyons, C. (1999), *Definiteness*, Cambridge : Cambridge University Press
- MacWhinney, B. (2001), « The Competition Model : the input, the context, and the brain ». In P. Robinson (ed.), *Cognition and second language instruction*, Cambridge : Cambridge University Press, pp. 69-90
- MacWhinney, B. (2000), *The CHILDES Project: Tools for Analyzing Talk. Volume 1: Transcription format and programs. Volume 2: The Database*. Mahwah, NJ : Lawrence Erlbaum Associates
- MacWhinney, B. (2010), « Introduction to CHILDES and TalkBank », *Présentation powerpoint extraite du site de CHILDES*, [en ligne], consulté le 5 mai 2015. URL : <http://childes.psy.cmu.edu/intro/>

- Maratsos, M.P. (1976), *The use of definite and indefinite reference in young children : an experimental study of semantic acquisition*, Cambridge, MA : Cambridge University Press
- Marchman, V. et Bates, E. (1994), « Continuity in lexical and morphological development ; a test of the critical mass hypothesis », *Journal of Child Language*, 21, pp. 339-366
- Marcos, H. et Bernicot, J. (1997), « How do young children reformulate assertions ? A comparison with requests », *Journal of Pragmatics*, 27, pp. 781-798
- Matthews, D., Lieven, E.V.M., Theakston, A. et Tomasello, M. (2006), « The effect of perceptual availability and prior discourse on young children's use of referring expressions », *Applied Psycholinguistics*, 27, pp. 403-422
- McNeill, D. (1970), *The Acquisition of Language: The Study of Developmental Psycholinguistics*. New York : Harper & Row
- McNeill, D. (1966), « Developmental psycholinguistics ». In F. Smith et G.A. Miller (eds), *The genesis of language, a psycholinguistic approach*, Cambridge, MA : MIT Press, pp. 15-84
- Mitterand, H. (1980), « La nomenclature grammaticale (version 1975) », *Langue française*, 47, pp. 90-99
- Mondada, L. (1998), « Technologies et interactions dans la fabrication du terrain du linguiste », *Cahiers de l'ILSL*, 10, pp. 39-68
- Morgan, J. L. et Demuth, K. (1996), *Signal to syntax : bootstrapping from speech to grammar in early acquisition*, Hillsdale, NJ : Lawrence Erlbaum Associates
- Morgenstern, A. et Parisse, C. (2012), « The Paris Corpus », *French Language Studies*, 22, pp. 7-12
- Morgenstern, A. (2009a), *L'enfant dans la langue*, Paris : Presses de la Sorbonne Nouvelle
- Morgenstern, A. (2009b), « Ecllosion de la morphosyntaxe : le rôle des fillers pré-nominaux dans un corpus longitudinal », *Rééducation orthophonique, la morphosyntaxe*, pp. 63-82
- Morgenstern, A. et Parisse, C. (2007), « Codage et interprétation du langage spontané d'enfants de 1 à 3 ans », *Corpus*, 6, *Interprétation, contextes, codage*, pp. 55-78
- Müller, N. et Kupisch, T. (2007), « Acquisition des déterminants et des clitiques objets chez des enfants bilingues (français-allemand) », *Acquisition et interaction en langue étrangère*, 25, [en ligne], consulté le 3 avril 2016. URL : <http://aile.revues.org/2333>
- Nashawati, S. (2010), *Le développement des expressions référentielles chez le jeune enfant : noms et pronoms dans des dialogues mère-enfant*, Thèse de doctorat, Université Sorbonne Nouvelle-Paris 3
- Nelson, K. (1981), « Individual differences in language development : implications for development and language », *Developmental Psychology*, 17, pp. 170-187
- Nelson, K. (1973), « Structure and strategy in learning to talk », *Monographs of the Society for Research in Child Language Development*, 38, pp. 1-135
- Nespor, M. et Vogel, I. (1986), *Prosodic phonology*, Dordrecht : Foris

- Newport, E.L., Gleitman, H. et Gleitman, L.R. (1977), « Mother, I'd rather do it myself : Some effects and non-effects of maternal speech style ». In C. E. Snow et C. A. Ferguson (eds.), *Talking to children*, New York : Cambridge University Press, pp. 109-149
- Ochs, E. (1979), « Transcription as theory ». In E. Ochs et B. Schieffelin (eds.), *Developmental Pragmatics*, New York, NY : Academic Press, pp. 43-72
- O'Neill, D. (1996), « Two-year-old children's sensitivity to a parent's knowledge state when making requests », *Child Development*, 67, pp. 659-677
- Parisse, C. et Morgenstern, A. (2010), « Transcrire et analyser les corpus d'interactions adulte-enfant ». In E. Veneziano, A. Salazar Orvig, J. Bernicot (eds), *Acquisition du langage et interaction*, Paris : L'Harmattan, pp. 201-222
- Pepinsky, T., Demuth, K. et Roark, B. (2001), The status of filler syllables in childrens early speech, *Proceedings of the 25<sup>th</sup> Annual Boston University Conference on Language Development*, pp. 575-586
- Peters, A. M. (2001), « Filler syllables : what is their status in emerging grammar ? », *Journal of Child Language*, 28, pp. 229-242
- Peters, A.M. (1995), « Strategies in the acquisition of syntax ». In P. Fletcher et B. MacWhinney (eds.), *Handbook of Language Acquisition*, Oxford : Blackwell, pp. 462-482
- Peters, A.M., et Menn, L. (1993), « False starts and filler syllables: ways to learn grammatical morphemes », *Language*, 4, pp. 742-777
- Peters, A.M. (1983), *The units of language acquisition*. Cambridge, MA : Cambridge University Press
- Peters, A.M. (1977), « Language learning strategies : does the whole equal the sum of the parts ? », *Language*, 53, pp. 560-573
- Piaget, J. (1966)[1923], *Le langage et la pensée chez l'enfant*, Paris : Delachaux et Niestlé
- Piaget, J. (1990)[1937], *La construction du réel chez l'enfant*, Paris : Delachaux et Niestlé
- Pine, J.M., Freudenthal, D., Krajewski, G. et Gobet, F. (2013), « Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner », *Cognition*, 127(3), pp. 345-360
- Pine, J.M. et Lieven, E.V.M. (1997), « Slot and frame patterns and the development of the determiner category », *Applied psycholinguistics*, 19, pp. 123-138
- Pine, J.M. et Martindale, H. (1996), « Syntactic categories in the speech of young children : the case of determiner », *Journal of Child Language*, 23, pp. 369-395
- Pinker, S. (1996)[1984], *Language learnability and language development*, Cambridge, MA : Harvard University Press
- Pullum, G.K. et Scholz, B.C. (2002), « Empirical assessment of stimulus poverty arguments », *The Linguistic Review*, 19, pp. 9-50



- Rabain-Jamin, J. (1998), « Usage du langage et contexte culturel : le langage de la mère adressé à l'enfant ». In J. Bernicot et H. Marcos (eds.), *De l'usage des gestes et des mots chez l'enfant*, Paris : Armand Colin, pp. 179-207
- Radford, A. (1990), *Syntactic theory and the acquisition of English syntax*, Oxford: Blackwell.
- Riegel, M., Pellat, J.C. et Rioul, R. (1994), *Grammaire méthodique du français*, Paris : PUF
- Roberts, C. (1997), « Transcribing talk : issues of representation », *TESOL Quarterly*, 31, pp. 167-172
- Rondal, J.-A. (1983), *L'interaction adulte-enfant et la construction du langage*, Bruxelles : Mardaga
- Rondal, J.-A. (1978), *Langage et éducation*, Bruxelles : Mardaga
- Roodenburg, J. (2004), « French bare arguments are not extinct : The case of coordinated bare nouns », *Linguistic Inquiry*, 35, pp. 301-313
- Rozendaal, M.I. (2008), *The acquisition of reference : a cross-linguistic study*, Utrecht : LOT
- Rozendaal, M.I. et Baker, A.E. (2008), « A cross-linguistic investigation of the acquisition of the pragmatics of indefinite and definite reference in two-year-olds », *Journal of Child Language*, 35, pp. 773-808
- Ryckebusch, C. et Marcos, H. (2000), « La communication du jeune enfant avec le père ou la mère : influence de l'interlocuteur et du type d'activité », *Enfance*, 53, pp. 127-147
- Salazar Orvig, A., de Weck, G., Hassan, R. et Rialland, A. (en préparation), *Constructing reference in dialogue*, Amsterdam / Philadelphia : John Benjamins Publishing Company
- Salazar Orvig A., Marcos H., Heurdier J. et Da Silva C. (à paraître), « Referential features, speech genres and activity types ». In M. Hickmann, E. Veneziano et H. Jisa (eds.), *Sources of variation in first language acquisition: Languages, contexts, and learners*, Amsterdam : John Benjamins
- Salazar Orvig, A., Marcos, H., Caët, S., Corlateanu, C., Da Silva, C., Hassan, R., Heurdier, J., Le Mené, M., Leber-Marin, J. et Morgenstern, A. (2013), « Definite and indefinite determiners in French-speaking toddlers: Distributional features and pragmatic-discursive factors », *Journal of Pragmatics* (56), pp. 88-112
- Salazar Orvig, A. (2013), *Rapport final du projet ANR DIAREF*, [En ligne], consulté le 8 novembre 2015. URL : [http://www.univ-paris3.fr/medias/fichier/rapport-final-diaref-partie-scientifique\\_1477492126483.pdf?INLINE=FALSE](http://www.univ-paris3.fr/medias/fichier/rapport-final-diaref-partie-scientifique_1477492126483.pdf?INLINE=FALSE)
- Salazar Orvig, A., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marin, J. et Parès, J. (2008), « Entre Gramática y Pragmática: Adquisición de los determinantes en francés ». In L. Miranda (ed.), *Actas del V Congreso Nacional de Investigaciones Lingüístico-Filológicas*, Lima : Universidad Ricardo Palma, Facultad de Lenguas Modernas, pp. 55-87
- Salazar Orvig, A., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marin, J. et Pares, J. (2005), « Can first determiners be considered plurifunctional ? A longitudinal and cross-sectional study on French children », communication orale, *Congrès International de l'Association*

## BIBLIOGRAPHIE

- d'Etudes du Langage de l'Enfant (LASCL)*, Berlin, juillet 2005
- Salazar Orvig, A. et Morgenstern, A. (2002), « Can fillers have a discursive function ? », communication orale, *Congrès International de l'Association d'Etudes du Langage de l'Enfant (LASCL)*, Madison, juillet 2002
- Salazar Orvig, A. (2000), « La reprise aux sources de la construction discursive », *Langages*, 140, pp. 68-91
- Schaff, A. et Brendel, C. (1973), « La grammaire générative et la conception des idées innées », *L'Homme et la société*, 28, pp. 3-50
- Schaeffer, J. et Matthewson, L. (2005), « Grammar and pragmatics in the acquisition of article systems », *Natural Language and Linguistic Theory*, 23, pp. 53-101
- Schane, S.A. (1967), « Introduction », *Langages*, 8, pp. 3-12
- Schieffelin, B. et Ochs, E. (1986), *Language socialization across cultures*, Cambridge : Cambridge University Press
- Scollon, R. (1979), « A real early stage: An unzipped condensation of a dissertation on child language ». In E. Ochs et B. Schieffelin (eds.), *Developmental Pragmatics*, New York : Academic Press, pp. 215-227
- Searle, J.R. et Vanderveken, D. (1985), *Foundation of illocutionary logic*, Cambridge : Cambridge University Press
- Searle, J. R. (1982), *Sens et expression*, Paris : Editions de minuit
- Selkirk, E. (1984), *Phonology and syntax: The relation between sound and structure*, Cambridge, MA : MIT Press
- Serratrice, L., Sorace, A. et Paoli, S. (2004), « Cross-linguistic influence at the syntax-pragmatics interface : subjects and objects in English-Italian bilingual and monolingual acquisition », *Bilingualism : Language and Cognition*, 7, pp.183-205
- Serratrice, L. (2000), *The emergence of functional categories in bilingual language acquisition*, Thèse de doctorat, Université d'Edinburgh
- Simonsen, H. G. (1999), « Phonological fillers : data from Norwegian ». In M. Almgren, A. Barrena, M.J. Ezeizabarrena, I. Idiazabal et B. MacWhinney (eds.), *Research on child language acquisition, Proceedings of the 8th Conference of the International Association for the Study of Child Language*, San Sebastian, Spain : Cascadilla Press, pp. 701-712
- Sinclair, H. (1967), *Acquisition du langage et développement de la pensée*, Paris : Dunod
- Skinner, B.F. (1957), *Verbal behaviour*, Acton, MA : Copley Publishing Group
- Slobin, D.I. (2004), « From ontogenesis to phylogenesis : what can child language tell us about language evolution ». In J. Langer, S.T. Parker et C. Milbrath (eds.), *Biology and knowledge revisited : from neurogenesis to psychogenesis*, Mahwah, NJ : Lawrence Erlbaum Associates, pp. 255-285

- Slobin, D.I. (1973), « Cognitive prerequisites for the development of grammar ». In C.A. Ferguson et D.I. Slobin (eds.), *Studies of child language development*, New York : Holt, Rinehart et Winston, pp. 175-208
- Smith, N. (1973), *The acquisition of phonology*, Cambridge, MA : Cambridge University Press
- Snow, C.E. (1977), « Mothers' speech research : from input to interaction ». In C.E. Snow and C.A. Ferguson (eds.), *Talking to children*, New York : Cambridge University Press, pp. 31-51
- Snow, C.E. (1972), « Mothers' speech to children learning language », *Child Development*, 43, pp. 549-565
- Speer, S.A. (2002), « 'Natural' and 'contrived' data : a sustainable distinction ? », *Discourse studies*, 4, pp. 511-525
- Taine, H. (1876), « L'acquisition du langage chez les enfants et dans l'espèce humaine », *Revue Philosophique de la France et de l'Etranger*, 1, pp. 357-395
- Thomas, M. (2002), « Development of the concept of "the poverty of the stimulus" », *The Linguistic Review*, 19, pp. 51-71
- Tomasello, M. (2009), « The usage-based theory of language acquisition ». In E. L. Bavin, *The Cambridge Handbook of Child Language*, Cambridge : Cambridge University Press, pp. 69-88
- Tomasello, M. (2006), « Acquiring linguistic constructions ». In D. Kuhn et R. Siegler (eds.), *Handbook of Child Psychology (vol. 2)*, New-York : Wiley, pp. 255-298
- Tomasello, M. (2004), *Aux origines de la cognition humaine*, Paris : Retz
- Tomasello, M. et Stahl, D. (2004), « Sampling children's spontaneous speech : how much is enough ? », *Journal of Child Language*, 31, pp. 101-121
- Tomasello, M. (2003), *Constructing a language : a usage-based theory of language acquisition*, Cambridge, MA : Harvard University Press
- Tomasello, M. (2000a), « Do young children have adult syntactic competence ? », *Cognition*, 74, pp. 209-253
- Tomasello, M. (2000b), « First steps towards a usage-based theory of language acquisition », *Cognitive Linguistics*, 11, pp. 61-82
- Tomasello, M. (1992), *First verbs, A case study of early grammatical development*, Cambridge, MA : Cambridge University Press
- Tomasello, M. et Farrar, M. J. (1986), « Joint attention and early language », *Child Development*, 57, pp. 1454-1463
- Tomasello, M., Mannle, S. et Kruger, A. (1986), « The linguistic environment of one to two year old twins », *Development Psychology*, 22, pp. 169-176
- Tomasello, M. et Todd, J. (1983), « Joint attention and early lexicon acquisition style », *First Language*, 4, pp. 197-212

## BIBLIOGRAPHIE

- Tremblay, A., et Demuth, K. (2007), « Prosodic licensing of determiners in children's early French ». In A. Belikova, L. Meroni et M. Umeda (eds.), *Proceedings of the 2<sup>nd</sup> Conference on Generative Approaches to Language Acquisition North America (GALANA)*, Somerville, MA: Cascadilla Proceedings Project, pp. 426-436
- Tremblay, A. (2006), « Prosodic constraints on the production of grammatical morphemes in French : The case of determiners ». In K. Deen, J. Nomura, B. Schulz et B.D. Schwartz (eds.), *Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition*, Cambridge, MA : MIT Working Papers in Linguistics, pp. 377-388
- Tremblay, A. (2005), « On the status of determiner fillers in early French : What the child knows ». In A. Brugos, M.R. Clark-Cotton et S. Ha (eds.), *Proceedings of the 29<sup>th</sup> Annual Boston University Conference on Language Development*, Somerville, MA : Cascadilla Press, pp. 604-615
- Valian, V., Solt, S. et Stewart, J. (2009), « Abstract categories or limited-scope formulae ? The case of children's determiners », *Journal of Child Language*, 36, pp. 743-778
- Valian, V. (1986), « Syntactic categories in the speech of young children », *Developmental psychology*, 22, pp. 562-579
- van Hout, A., Harrigan, K. et de Villiers, J. (2009), « Comprehension and production of definite and indefinite noun phrases in English preschoolers. In J. Crawford, K. Otaki et M. Takahash (eds.), *Proceedings of the 3<sup>rd</sup> Conference on Generative Approaches to Language Acquisition North America (GALANA)*, Somerville, MA: Cascadilla Proceedings Project, pp. 76-87
- Veneziano, E. (2014a), « Filler syllables and acquisition of grammatical morphemes », in P.J. Brooks et V. Kempe *Encyclopedia of Language Development*, pp. 215-216
- Veneziano, E. (2014b), « There are fillers and fillers on the way to grammatical morphemes : A three-period account in the acquisition of French », communication orale, *Congrès International de l'Association d'Etudes du Langage de l'Enfant (IASCL)*, Amsterdam, juillet 2014
- Veneziano, E. (2003), « The Emergence of Noun and Verb Categories in the Acquisition of French », *Psychology of Language and Communication* 7, pp. 23-36.
- Veneziano, E. (2001). « The importance of studying filler-producing children », *Journal of Child Language* 28, pp. 275-278
- Veneziano, E., et Sinclair, H. (2000), « The changing status of 'filler syllables' on the way to grammatical morphemes », *Journal of Child Language* 27, pp. 461-500
- Veneziano, E. (2000), « Interaction, conversation, et acquisition du langage dans les trois premières années ». In M. Kail et M. Fayol (eds.), *L'acquisition du langage 1*, Paris : PUF, pp. 231-265
- Veneziano, E. (1999), « Pragmatique et cognition dans les premières acquisitions morphosyntaxiques ». In F. Cordier et J.-E. Tyvaert (eds.), *Actes des journées scientifiques du CIRLEP*, Reims : Presses Universitaires de Reims, pp. 209-227

- Veneziano, E. (1997a), « Echanges conversationnels et premières acquisitions langagières ». In J. Bernicot, J. Caron-Pargue et A. Trognon (eds), *Conversation, Interaction et fonctionnement cognitif*. Nancy : PUN, pp. 191-123
- Veneziano, E. (1997b), « Processus conversationnels et constructivisme dans les débuts du langage », *CAlap*, 15, pp. 93-108
- Veneziano, E. et Sinclair, H. (1997), « From the surface inward : a discontinuous continuity in the emergence of grammatical morphology », *Archives de psychologie*, 65, pp. 107-116
- Veneziano, E., Sinclair, H. et Berthoud, I. (1990), « From one word to two words : repetition patterns on the way to structured speech », *Journal of Child Language*, 17, pp. 633-650
- Voss, J. (1973), « Noam Chomsky et la linguistique cartésienne », *Revue Philosophique de Louvain*, Quatrième série, Tome 71, 11, pp. 512-538
- Vygotski, L.S. (1997)[1936], *Pensée et langage*, Paris : Editions sociales
- Wauquier, S. et Yamaguchi, N. (2013), « Templates in French ». In M. Vihman et T. Keren-Portnoy, *The emergence of phonology : whole word approaches and cross-linguistic evidence*, Cambridge : Cambridge University Press, pp. 317-342
- Weck, G. de et Rosat, M.-C. (2003), *Troubles dysphasiques. Comment raconter, relater, faire agir à l'âge préscolaire*, Paris : Masson
- Weck, G. de (2010), « Interactions adulte-enfant et troubles du développement du langage : bilan des recherches et questions ouvertes ». In J. Bernicot, E. Veneziano, M. Musiol et A. Bert-Erboul (eds.), *Interactions verbales et acquisition du langage*, Paris : L'Harmattan, pp. 151-170
- Wittek, A. et Tomasello, M. (2005a), « Young children's sensitivity to listener's knowledge and perceptual context in choosing referring expressions », *Applied psycholinguistics*, 26, pp. 541-558
- Wittek, A. et Tomasello, M. (2005b), « German-speaking children's productivity with syntactic constructions and case morphology : local cues act locally », *First Language*, 25, pp. 103-125
- Wittgenstein, L. (1953), *Philosophical Investigations*, Oxford : Blackwell
- Wolfson, N. (1976), « Speech events and natural speech : some implications for sociolinguistic methodology », *Language in Society*, 5, pp. 189-209
- Yamaguchi, N., Salazar Orvig, A., Le Mené, M., Caët, S. et Rialland, A. (en préparation), « Emergence of grammatical units : why filler syllables matter ». In Salazar Orvig, A., de Weck, G., Hassan, R. et Rialland, A. (eds), *Constructing reference in dialogue*, Amsterdam / Philadelphia : John Benjamins Publishing Company
- Yamaguchi, N., Rialland, N. et Salazar Orvig, A. (2013), « The intertwining of form and function in the emergent path from fillers to determiners », communication orale, *Congrès International de l'Association d'Etudes du Langage de l'Enfant (IASCL)*, Amsterdam, juillet 2014

## BIBLIOGRAPHIE

- Yamaguchi, N. et Rialland, A. (2013), « Acquiring consonants in lexical words and in grammatical words », *AERef*, 25-26 octobre 2013, Paris
- Yamaguchi, N. (2012), *Parcours d'acquisition des sons du langage chez deux enfants francophones*, Thèse de doctorat, Université Sorbonne Nouvelle-Paris 3
- Yamaguchi, N. et Rialland, A. (2012), « Comparaison de l'acquisition des consonnes dans les mots lexicaux et dans les mots grammaticaux », communication orale, *conférence ELA*, Lyon, décembre 2012

# **PARTIE IV**

## *Annexes*

---





## FORMULAIRE D'AUTORISATION PARENTALE

*Formulaire élaboré sur la base de l'autorisation parentale utilisée dans le cadre du projet ANR CoLaJE (responsable : A. Morgenstern)*

**Autorisation des parents (ou du responsable légal) pour l'enregistrement audio/vidéo et l'archivage, l'exploitation scientifique et la diffusion des données enregistrées (cochez les cases qui vous conviennent).**

Nous, soussignés,  
Monsieur, Madame .....

Agissant en qualité de père et mère (ou responsable légal)

Demeurant à .....

Code Postal .....

Ville .....

Téléphone .....

Avons obtenu des réponses aux questions que nous avons posées et sommes d'accord pour apporter notre contribution à la thèse de Marine Le Mené et aux travaux de son équipe de recherche dans les limites des conditions acceptées ci-dessous.

Acceptons que notre enfant prénomné(e) .....  
ainsi que (autres enfants pouvant apparaître dans l'enregistrement) .....  
et nous-mêmes, soient filmés.

Le tournage est prévu pour une période de 1 an à compter du ....., à raison d'une séance par mois (ou autre fréquence), au domicile de Mr et Mme .....

Acceptons et autorisons expressément Marine Le Mené et son équipe de recherche à utiliser les images et les sons qui représentent (prénom(s)) de l'enfant ..... et renonçons, en conséquence, à demander tout dédommagement ou indemnité du fait de ces utilisations.

Acceptons que le prénom réel de notre enfant soit mentionné dans l'utilisation du corpus de notre enfant.

Souhaitons que seule l'initiale du prénom de notre enfant soit mentionnée dans l'utilisation du corpus de notre enfant.

Autorisons Marine Le Mené et son équipe de recherche à :

utiliser ces données sous leur forme transcrite et anonymée à des fins de recherches scientifiques (thèse, articles scientifiques, exposés à des congrès, séminaires) :  *oui* // *non*

reproduire des séquences audio pour étayer les hypothèses qui seront défendues dans leurs recherches :  *oui* // *non*

FORMULAIRE D'AUTORISATION PARENTALE

à reproduire des images ou des séquences vidéo pour étayer les hypothèses qui seront défendues dans leurs recherches :  *oui* // *non*

Acceptons la diffusion des données recueillies à un public de chercheurs en acquisition du langage en France et à l'étranger dans des bases de données informatiques pour la recherche en acquisition du langage (CHILDES par exemple)  des transcriptions seulement  des transcriptions et des données audio  des transcriptions et des données vidéo.

Nous réservons le droit de demander l'effacement d'un enregistrement ou d'une partie d'un enregistrement si celui-ci ne nous convient pas.

Nous prenons acte que :

Aucune information confidentielle sur l'identité complète de l'enfant ou sur les membres de sa famille ne pourra figurer dans la mutualisation des données recueillies, ni dans les travaux réalisés à partir des enregistrements. Seuls le prénom ou les initiales de l'enfant ainsi que son âge figureront dans les en-têtes de fichiers. Les enregistrements présentés à des conférences ou des cours seront bipés lors de la mention d'une adresse ou d'un numéro de téléphone identifiables.

A ....., le .....

Signatures des deux parents ou responsable légal, précédées de la mention manuscrite « Lu et  
approuvé »

# CONVENTIONS DE TRANSCRIPTION

*Extraits du guide élaboré au sein du projet Léonard (responsable : A. Morgenstern)*

## 1. Les en-têtes

Les lignes d'en-têtes commencent par @.

Les en-têtes suivantes sont obligatoires et toujours dans l'ordre suivant :

@Begin

@Languages: [tabulation] fr pour français, eng pour anglais, etc.

@Participants: [tab.] CHI prénom Target\_Child, MOT Mother, FAT Father, BRO prénom Brother

*Exemples de participants possibles :*

*Child Target, Mother, Father, Brother, Sister, Sibling, Grandmother, Grandfather, Aunt, Uncle, Cousin, Playmate, Visitor, Student, Teacher, Babysitter, Caretaker, Housekeeper, Investigator, Examiner, Observer, Clinician, Therapist Informant, Camera Operator, Doctor, Nurse, Patient, Client, Subject, Unidentified (UNI), Teenager, ...*

@ID: [tab.] (reprend des informations sur chaque participant (une ligne par participant))

@Birth of CHI: [tab.]

@Date: [tab.]

@Time duration: [tab.]

@Coder: [tab.]

@Situation: [tab.]

@End (dernière ligne)

En cours de transcription, l'indication suivante doit être ajoutée pour tout changement de situation/ nouvel épisode:

@G: [tab.] description de la situation

●\* **FORMAT A RESPECTER** ●\*:

Pas d'espace entre @ et en – tête ni entre en – tête et : / tab après :

## 2. Lignes principales

La ligne principale est symbolisée par \*. On y code les énoncés en orthographe (sauf rares exceptions).

●\* **FORMAT A RESPECTER** ●\* : **REGLES DE BASE DE TRANSCRIPTION**

- Un seul énoncé par ligne. Si la transcription ne peut contenir sur une ligne, commencer la ligne suivante avec une [TAB] (éventuellement automatique).

- Chaque ligne principale commence par \*NOM: [TAB]

- Le nom (ou rôle) du participant est un code de 3 caractères majuscules (CHI, MOT, BRO, etc.)

<ul style="list-style-type: none"> <li>- Pas d'espace entre * et nom du participant</li> <li>- Pas d'espace entre nom et :</li> <li>- TAB après les :</li> <li>- Ex : *CHI : moi veux le nounours !</li> </ul> <ul style="list-style-type: none"> <li>- Un énoncé se termine toujours par un marqueur de fin d'énoncé (voir tableau ci-dessous).</li> <li>- L'espace avant le marqueur de fin d'énoncé est conseillé.</li> <li>- Espace obligatoire après le marqueur de début d'énoncé.</li> <li>- Espace obligatoire entre marqueur de fin d'énoncé et balise.</li> <li>- La balise figure toujours sur la ligne principale.</li> <li>- Les mots sont codés orthographiquement sur la ligne principale.</li> <li>- Aucune ponctuation au sein de l'énoncé (sauf marqueur de fin d'énoncé) (ne jamais utiliser de , virgule ou ; point virgule à l'intérieur ou en fin d'énoncé). Les crochets [] et chevrons &lt; &gt; sont utilisés pour du métacodage (cf. ci-après).</li> <li>- Majuscules seulement aux noms propres.</li> </ul>
--

### Marqueurs de fin d'énoncé

Marqueur de fin d'énoncé	Signification	Exemple
.	énoncé affirmatif	*MOT: non ça c'est une pomme .
?	énoncé interrogatif	*MOT: tu veux une banane ?
!	énoncé exclamatif	*MOT: oh t(u) es toute belle !
+...	<i>Énoncé en suspend</i> (a) énoncé inachevé non repris ensuite par le locuteur.  (b) ou interruption dans la conversation, le même locuteur reprend ensuite la parole et il termine son énoncé. Dans ce cas, combiné avec +, (cf ci-dessous)	(a) *CHI: ça c'est +... *MOT: tu le vois là ? *CHI: je peux le prendre ? (b) *CHI: et ça aussi +... *MOT: d'accord. *CHI: +, c'est des enfants .
+..?	<i>Question en suspend</i>	*MOT: et ça c' est un +..?
+//.	<i>Interruption par le locuteur :</i> le locuteur s'interrompt brusquement et commence immédiatement à parler d'autre chose. Il n'y a pas de pause dans la conversation.	*MOT: ah oui je vois c'est le +//. *MOT: tu te souviens d'Anna ?

+//?	<i>Le locuteur interrompt son interrogation.</i>	*CHI: je peux prendre le +//? *CHI: regarde j'ai mon chapeau .
+/. .	<i>Interruption par quelqu'un d'autre</i>	*CHI: maman j' ai pris +/. *MOT: fais attention tu vas renverser .
+/?	<i>Le locuteur est interrompu par quelqu'un d'autre dans son interrogation.</i>	*CHI: je peux prendre le +/? *MOT: non ce n'est pas un jouet .

### Marqueurs de début d'énoncé

Marqueur de début d'énoncé	Signification	Exemple
+,	énoncé complétant un énoncé précédent produit par le même locuteur.	*CHI: et ça aussi +... *MOT: d'accord . *CHI: +, c'est des enfants .
++	énoncé complétant un énoncé précédent produit par un autre interlocuteur.	*MOT: et ça c' est un +..? *CHI: ++ mouton .
+^	<i>le locuteur embraye tout de suite sans que l'autre n'ait le temps de prendre la parole (le schéma conversationnel n'est pas respecté).</i>	*CHI: comment ça s' appelle ça ? *CHI: +^ à l' école on a fait du bricolage .

### Codages particuliers en ligne principale

#### Mots composés

On les écrit normalement mais sans tirets et sans majuscule (par exemple père Noël). Les noms propres composés seront codés avec un \_ (par exemple : Marie\_Ange).

#### Mot(s) ou énoncés inintelligibles

- yy : Production d'un mot que l'on ne peut pas écrire orthographiquement sur la ligne principale mais qui sera transcrit sur la ligne %pho (cf. ci-après). Quand il y a plusieurs "mots" (identifiés du point de vue sémantique) indécodables, on note autant de yy que de mots.

## CONVENTIONS DE TRANSCRIPTION

*Exemples :*

\*CHI: le yy est tombé .

%pho: lə mila ε tōbe

\*CHI: papa parti yy yy .

%pho: papa pati tame kada

\*CHI: yy xx papa .

%pho: bogo \$ papa

\*CHI: yy papa .

%pho: aga papa

- **yyy** : énoncé entier (ou un long bout d'énoncé) impossible à transcrire en orthographe.

*Exemples :*

\*CHI: yyy .

%pho: agapobalejemu

- **xx** : mot inintelligible, non transcriptible ni orthographiquement ni phonologiquement . xx sera transcrit par \$ dans la ligne %pho

*Exemples :*

\*CHI : xx papa .

%pho: \$ papa

(voir explication du codage en ligne %pho plus loin).

- **xxx** : production inintelligible, sur tout l'énoncé (pas de ligne %pho dans ce cas là ).

*Exemple :*

\*CHI: xxx .

### ☛\* FORMAT A RESPECTER ☛\*:

- xx / xxx / yy / yyy sont toujours écrits en lettres minuscules

- Comme pour tout énoncé, ne pas oublier le marqueur de fin d'énoncé après yyy ou xxx.

## Elisions

Les élisions réalisées à l'oral et les parties de mots non prononcées en cas d'ébauches phonologiques sont marquées entre parenthèses.

*Exemples:*

\*CHI: t(u) as vu le p(e)tit chat ?

\*MOT: ça c'est un m(outon) +..?

\*CHI : ++ mouton !

Formes particulières ne figurant pas dans le dictionnaire

Pour coder des formes particulières, le mot est codé orthographiquement sur la ligne principale directement suivi du symbole (sans espace):

- @c : mot inventé par l'enfant (ex : le tictic@c pour les ciseaux)
- @f : mot inventé par la famille (ex : le papoum@f= nom du doudou)
- @i: interjection (ex : ouh@i)
- @o: onomatopées (ex : coincoin@o)
- @l : lettre (ex : elle dit le r@l maintenant)

Les interjections et onomatopées figurant dans le dictionnaire ne sont pas nécessairement suivies de @i et @o

### Les fillers

Forme phonétique suivie de @fs

*Exemple :*  
si l'enfant dit [ə gato] → \*CHI: ə@fs gateau

●\* FORMAT A RESPECTER ●\*:  
- le mot est directement suivi de @ et de son extension, sans espace (la@fs)

### Répétition et reprise d'un ou plusieurs mot(s) au sein d'un énoncé

[/] pour répétition, reprise sans correction de ce qui est dit avant

*Exemple :*  
\*MOT: <le lapin> [/] le lapin c'est le blanc .

[//] pour reprise avec correction (correction syntaxique)

*Exemple :*  
\*MOT: j'ai mangé un [//] le lapin .

[///] pour reprise avec reformulation (auto – correction sémantique)

*Exemple :*  
\*MOT: lapin [///] le lièvre court !

Si la répétition porte sur plusieurs mots, mettre ceux-ci entre < > (les chevrons ne sont pas nécessaires si la répétition ne porte que sur un mot, par défaut le mot juste avant ce symbole.)

On ne code pas de répétition d'énoncé complet. On ne code pas la répétition quand elle porte sur un énoncé avec une structure syntaxique correcte.

●\* FORMAT A RESPECTER ●\*:  
- [/] écrit sans espace mais espace avant et après  
- pas d'espace entre les < > et le groupe de mots qui y figure (par exemple : <le chocolat>)

### Evénements para-linguistiques

L'indication figure entre crochets (voir format à respecter)

ex : [=! rit], [=! sourit], [=! crie], [=! pleure], [=! chuchote], [=! chante], [=! tousse], [=! petits bruits], [=! bruits], [= ! lit], [= ! discours rapporté], ...

Cette indication porte sur le mot juste avant. Si elle porte sur plusieurs mot ou l'énoncé en entier, mettre le groupe de mots entre < >.

L'indication peut également figurer seule, précédée de 0 (absence d'énoncé).

*Exemples :*

\*CHI : non [/] non [/] non c'est à moi [=! crie]!

Dans cet exemple : « moi » est dit en criant.

\*MOT: <tout ça> [=! rit] ben dis donc ?

Dans cet exemple : « tout ça » est dit en riant

\*MOT: <je vais te manger> [=! rit].

Dans cet exemple : l'énoncé en entier est dit en riant.

\*CHI: 0 [=! rit].

Dans cet exemple : il n'y a pas d'énoncé.

#### 🔴\* FORMAT A RESPECTER 🔴\*:

- espace avant [
- [=! sans espace
- espace entre ! et mot
- pas d'espace entre mot et ]
- pas d'espace entre ] et marqueur de fin d'énoncé
- ne pas oublier le marqueur de fin d'énoncé même si on code par 0

### Pause à l'intérieur d'un énoncé

Marqueur sommaire : (.)

### Chevauchements

Marqueur sommaire : +< utilisé au début de l'énoncé qui recouvre le précédent.

Le marqueur sommaire sera à privilégier en cas de chevauchements multiples ou en cas de chevauchement avec un énoncé codé en 0 ou comprenant un événement para-linguistique.

*Exemples :*

\*CHI: <c'est> [/] ce sont les miens attends .

\*BRO: +< j' ai tout pris les animaux .

\*MOT: t' es une coquine !

\*CHI: +< 0 [= ! rit].

\*MOT: tu ouvres la porte .

\*CHI: +< 0 .



%act: CHI accourt vers MOT

Codage plus élaboré :

[>] indique un recouvrement avec l'énoncé suivant. Ce symbole figure directement après le mot ou groupes de mots (entre < >) sur lequel porte le chevauchement, prononcé par le premier interlocuteur.

[<] indique un recouvrement avec l'énoncé précédent. Ce symbole figure directement après le mot ou groupes de mots (entre < >) sur lequel porte le chevauchement, prononcé par le deuxième interlocuteur.

Si le chevauchement porte sur plusieurs mots, mettre ceux-ci entre < > (les chevrons ne sont pas nécessaires si le chevauchement ne porte que sur un mot, par défaut le mot juste avant ce symbole.)

Il n'est pas possible de faire porter le chevauchement sur une partie de mot (nécessairement sur le mot en entier).

*Exemples :*

\*CHI: je [>] +/.

\*MOT: <tu vas> [<] lui chercher une grande feuille Arthur .

\*BRO: <moi je> [/] moi j(e) ferai pas une peinture j(e) ferai juste un dessin <pour Zoum> [>].

\*MOT: oui [<].

\*CHI: <un petit coup de l'eau> [>].

\*MOT : <tu veux encore un biscuit> [<]?

●\* FORMAT A RESPECTER ●\*:

- Espace après +< (comme après tout marqueur de début d'énoncé)
- Espace entre [>] ou [<] et la ponctuation de fin d'énoncé. (l'espace est en fait optionnel avant . ? et ! mais obligatoire avant +... +..? +/. +/? +//. +//?)
- Espace entre [>] ou [<] et le mot suivant.
- Pas d'espace entre les < > et le mot (ou groupe de mots) qui y figure
- L'espace est optionnel entre > et [>] ou [<]

### 3. Lignes dépendantes

Les lignes dépendantes comprennent des informations complémentaires, elles commencent par %

●\* FORMAT A RESPECTER ●\*:

- On y code des indications complémentaires correspondant à la ligne principale située juste au-dessus.
- On peut avoir plusieurs lignes dépendantes pour une même ligne principale, mais une seule

ligne de chaque type (par exemple une seule ligne %act dépendante d'une ligne principale donnée).

- Chaque ligne dépendante commence par % suivi de lettres minuscules, ex : %pho: [TAB]

Pas d'espace entre % et nom de la ligne

Pas d'espace entre nom et :

TAB après les :

- Si l'information ne peut contenir sur une ligne, commencer la ligne suivante avec une [TAB] (éventuellement automatique).

Si vous codez plusieurs lignes dépendantes pour une ligne principale, respectez l'ordre de présentation suivant :

### Coder la phonétique

#### Ligne %pho

= Transcription phonétique en API

On découpe la transcription sur la ligne phono en « mots » identiques à ceux de la ligne principale. Il faut exactement le même nombre d'éléments sur les deux lignes. On code les xx par \$ sur la ligne %pho. (Rappelons que les yy, eux, seront transcrits phonétiquement sur la ligne %pho).

*Exemples :*

\*CHI: ah les xx .

%pho: a lɛ \$

\*CHI: ah les yy .

%pho: a lɛ putu

Sur la ligne principale, on ne note pas les liaisons "naturelles" mais bien sur la ligne %pho.

*Exemple :*

\*CHI: c'est mon ami .

%pho: sɛ mɔ̃ nami

On ne code pas l'allongement (:) dans la ligne principale mais bien dans la ligne %pho

### Coder les actions des interlocuteurs

#### Ligne %act

= décrire les actions / activités des participants concernés

La ligne dépendante %act se rapportera à la ligne principale juste au-dessus. Ainsi, pour coder une action sans énoncé, il faudra mettre en ligne principale « 0 ».

*Exemple :*

\*CHI: 0.

%act: CHI court

De même, si un interlocuteur fait une action en même temps ou juste après un énoncé produit par un autre, il faudra lui mettre une ligne principale « 0 ».

*Exemple :*

\*MOT: on peut regarder ensemble si tu veux .

\*CHI: 0.

%act: CHI va chercher un livre

Le regard se code en ligne % act

*Exemple :*

\*CHI: 0.

%act: CHI regarde OBS

### Ligne %sit

= décrire la situation ou des actions non impliquées dans la conversation (plus général que %act) (les lignes %sit peuvent concerner quiconque et pas nécessairement l'interlocuteur de la ligne principale juste avant, il ne sera pas nécessaire d'ajouter une ligne principale « 0 » pour ce codage)

*Exemples :*

\*CHI: attends je regarde ce truc .

%sit: MOT vient de rentrer dans la pièce

\*MOT: on peut le faire ensemble si tu veux .

%sit: CHI est en train de ranger ses gommettes une à une dans une boîte

### **Commentaires divers**

#### %com

= « voix du transcripateur », pour faire un commentaire métalinguistique. On peut notamment y coder des informations sur la voix, par exemple voix chuchotée, mais également pour fournir des explications sur ce qui est dit.

*Exemple :*

\*MOT: tiens tu la mets .

%com:la désigne la veste

\*CHI: oui c'est celui de Babouche .

%com:Babouche est un personnage du dessin animé Dora

On utilise également cette ligne pour faire mention de problèmes techniques.

## CONVENTIONS DE TRANSCRIPTION

### %int

= exprimer un doute sur ce qui est transcrit en ligne principale. La ligne %int permet de proposer une transcription alternative, avec une indication du degré de certitude. La ligne %int figurera après la ligne %pho

%int: mot /1/ → on est presque sûr

%int: mot /2/ → le doute est plus grand, interprétation plausible en contexte et en fonction de la réaction des interlocuteurs

%int: mot /3/ → on a essayé de deviner mais c'est vraiment un gros doute

*Exemples :*

\*CHI : il est pas là le joli cheval

%pho : il ε pa la lə joli səva

%int: joli/1/-cheval/1/

\*CHI: coco

%pho: koko

%int: coco=caméra/2/

\*CHI: fais ça

%pho: fe sa

%int: fe=fais/2/-fait/2/

●\* FORMAT A RESPECTER ●\*:

- pas d'espace sur cette ligne

# DÉCOUPAGE DES TOURS DE PAROLE EN ÉNONCÉS

*Critères retenus dans le cadre du projet DIAREF (responsable : A. Salazar Orvig)*

## 1. Un énoncé est délimité par :

- Les tours de parole : dans la mesure où nous cherchons à rendre compte des énoncés pour chaque locuteur on considère que, par convention, dès lors que l'on change de tour de parole, on change d'énoncé ;
- Des silences, éventuellement par des pauses. Nous ne nous donnons pas pour le moment de critères de longueur, vu la diversité d'opinions dans la littérature. En général la pause est associée à un critère mélodique.
- Par des critères mélodiques : courbe descendante, par exemple.
- Par des critères syntaxiques :
  - Si on adopte un critère de type "micro-syntaxique" : font partie d'un même énoncé des segments qui ont entre eux des relations de rection.

*Exemple :*

une proposition principale commune et deux propositions subordonnées coordonnées  
=> 1 seul énoncé.

"Nous voudrions que vous le regardiez et que vous le racontiez ensemble".

- Coordination : deux énoncés différents ; le connecteur de coordination avec l'énoncé suivant, sauf cas de solidarité sémantique entre les segments : *voir point suivant*
- Juxtaposition : deux énoncés différents

*Exemple :*

"Dis moi tu t'es habillée vite moi j'aurais mis une heure à mettre tout ça"

*Remarque :* Ces critères de découpage aboutissent à considérer des structures comme "il y a un garçon il court" comme deux énoncés distincts. Nous trouverons le moyen d'indiquer par ailleurs, pour les études pertinentes, les liens macro-syntaxiques qui existent entre ces deux énoncés.

- Les ajouts (dans une relation de rection) sont considérés comme intégrés à l'énoncé.

*Exemples :*

Damien : I(ls) vont s'arrêter ou pas / pour les aider ?

Madeleine : Interdit de passer là entre / pour l'instant

- Discours représenté avec verbe introducteur et sans marqueur > deux énoncés.

*Exemple :*

t'as fait le ménage dans ta grotte (énoncé 1) dit Papa Ours (énoncé 2).

- Le cas des dislocations :

- Dislocation du sujet = on ne découpe pas
- Que faire dans des cas comme Dam 8 (et une charrette / i(l) demande d'y aller)
- L'absence de pause et de chute de la courbe mélodique seraient les critères les plus clairs pour considérer que l'on a affaire à une dislocation.

## 2. Quelques cas :

- Les préfixes comme *regarde*, les infixes et suffixes comme *je crois* sont considérés comme intégrant l'énoncé.
- Les interjections, phatèmes et autres *voilà*, *regarde* en préfixe ou préambule sont intégrés dans l'énoncé s'ils sont produits sans pause.
- Des *questions phatiques* telles que *tu vois ? tu te souviens ? tu crois ? tu comprends ? tu vois ce que j(e) veux dire?*, etc... sont intégrés à l'énoncé si la pause est minime. ; les tags (*hein ? n'est-ce pas ? ...*) sont considérés comme partie de l'énoncé.

*Remarque* : on peut envisager des questions phatiques et autres tags avec une pause claire, remarquable, et dans ce cas, il faudra les considérer comme des questions à part. Les critères de catégorisation des pauses seront à affiner.

*Exemple* :

Mer 90 - euh un moule ?

Eva 96 - non . // hein ?

Mer 91 - euh : / un vélo ?

- *Oui*, *non* et *si*, avant un segment. On considère que *oui*, *non* et *si*, font partie du même énoncé que le segment qui les suit :
  - si le segment apparaît comme une paraphrase stricte de la phrase ;
  - si le segment module le oui/non : "*par exemple, oui*".

En revanche, si le segment apporte un élément informatif nouveau, alors on considère qu'il y a deux énoncés.

*Exemples* :

Mère : c'était quand la dernière fois que tu l'as mis tu te souviens ?

Mad : Ah oui c'était pour l'anniversaire de Blandine.

DAM 22 - après i(l) demande au train :::

MER 21 - au conducteur.=((en pointant l'image))

DAM 23 - au conducteur d'y aller à l'intérieur ((tourne une page)) .

MER 22 - ah non non j(e) pense =((en revenant en arrière)) qu'il a dû lui demander s'arrêter pour pouvoir descendre

Pour considérer la séquence "non + segment", comme un seul énoncé il aurait fallu que la mère dise quelque chose comme "*non non il ne demande pas d'y aller à l'intérieur*".

- Les incises autres que les incises phatiques, comme *regarde*, *tu vois* ...etc, correspondent à nouvel énoncé. il faudrait séparer cette incise (Énoncé B) du reste du premier énoncé (Énoncé A) sur le modèle : Énoncé A [Énoncé B] Énoncé A

*Exemples :*

Puis c'est pour ça moi je trouve que [d'un côté j'adhère aussi un petit peu à ton avis] que quelque part ce qui est le plus important c'est d'informer les gens.

Les chats ils en avaient beaucoup avaient chacun leur panier.

Si "ils" n'est pas co-référentiel à chats, alors on a bien une incise du type A[B]A. On a donc deux énoncés. Si "ils" est co-référentiel à chats, alors on est dans un cas de dislocation, pas d'incise, un seul énoncé (deux prédicats avec le même sujet)

### 3. Traitement des hésitations et ruptures de construction

- Les pauses pleines (*emb...*) ne sont pas synonymes de limite d'énoncé.
- Dans le cas d'hésitation et de re-formulation ponctuelle, on ne compte qu'un énoncé :

*Exemples :*

il est il était où Patrick ?

(Damien) Mer2 : non y a pas de : → ((regarde plusieurs pages)) y a pas de texte.

Par contre, s'il y a un changement de projet énonciatif, on compte deux énoncés comme dans "il fait une ... il mange des biscuits".





# TABLE DES MATIÈRES

RÉSUMÉ.....	5
ABSTRACT .....	6
REMERCIEMENTS .....	7
SOMMAIRE.....	9
INTRODUCTION .....	11
<b><u>PARTIE I - APPROCHES THEORIQUES</u></b> .....	<b>17</b>
<b>CHAPITRE I - DES APPROCHES VARIEES SUR LE DEVELOPPEMENT DU LANGAGE DE L'ENFANT</b> .....	<b>19</b>
<b>1. L'ACQUISITION DU LANGAGE : UNE FACULTE INNEE ?</b> .....	<b>22</b>
1.1 Créativité de l'usage du langage et récursivité.....	22
1.2 Opposition compétence-performance.....	24
1.3 La notion de 'pauvreté du stimulus'.....	25
1.4 Les trois composantes de la grammaire générative .....	25
1.5 Du 'Modèle Standard' au 'Programme Minimaliste' : évolution de la perspective générativiste sur l'acquisition du langage.....	26
1.6 Les théories du « bootstrapping ».....	30
1.6.1 <i>Le bootstrapping sémantique</i> .....	31
1.6.2 <i>Le bootstrapping syntaxique</i> .....	31
1.6.3 <i>Le bootstrapping prosodique et phonologique</i> .....	32
<b>2. L'ACQUISITION DU LANGAGE : UNE CONSTRUCTION PROGRESSIVE</b> .....	<b>33</b>
2.1 Autonomie de l'enfant dans le processus d'acquisition langagière.....	33
2.2 Apprendre à parler : un processus social.....	38
2.2.1 <i>Enjeux communicationnels de l'acquisition du langage</i> .....	39
2.2.2 <i>Un processus soutenu par l'interlocuteur</i> .....	42
2.2.3 <i>Des usages langagiers dépendants de l'expérience communicative</i> .....	49
<b>3. DES PERSPECTIVES DIVERGENTES SUR L'ACQUISITION DES CATEGORIES SYNTAXIQUES</b> .....	<b>53</b>
3.1 Perspective innéiste.....	54
3.2 Perspective basée sur l'usage .....	56
<b>4. PERSPECTIVE ADOPTEE POUR APPROCHER LES DONNEES</b> .....	<b>58</b>
4.1 Approche théorique retenue pour l'étude du développement de la catégorie 'déterminant'.....	58
4.2 Une étude linguistiquement et culturellement située.....	60
4.3 Une place faite à la variation individuelle.....	61
<b>CHAPITRE II</b> .....	<b>65</b>
<b>1. GENERALITES SUR LE DETERMINANT CHEZ L'ADULTE</b> .....	<b>68</b>
1.1 La complexité de la cible à atteindre .....	68
1.2 Caractéristiques formelles du déterminant chez l'adulte.....	69
1.3 Caractéristiques sémantico-référentielles du syntagme nominal chez l'adulte .....	71
1.3.1 <i>Traits sémantico-référentiels des syntagmes définis</i> .....	72
1.3.2 <i>Traits sémantico-référentiels des syntagmes indéfinis</i> .....	73
1.3.3 <i>Le modèle adulte de combinaisons forme-fonction</i> .....	77
<b>2. UN PARADIGME EN CONSTRUCTION : L'ETUDE DES FILLERS</b> .....	<b>80</b>
2.1 Remarques préliminaires sur les fillers.....	80
2.2 Un phénomène situé dans le temps du développement .....	82
2.2.1 <i>Transition entre énoncés à un terme et énoncés à deux termes</i> .....	82
2.2.2 <i>Masse lexicale critique et développement des formes pré-nominales</i> .....	83
2.2.3 <i>Concomitance entre fillers, omissions et formes adultes</i> .....	84
2.3 La construction progressive d'une catégorie de transition .....	85

## TABLE DES MATIERES

2.3.1	Du pré-morphologique au (quasi-)morphologique .....	86
2.3.2	Continuité dans le processus de 'grammatisation' des fillers .....	87
2.4	Un phénomène soumis à la variation individuelle .....	89
2.4.1	Des styles individuels dans la proportion des fillers produits .....	89
2.4.2	Un rapport individuel aux contraintes en jeu dans la production des fillers .....	91
2.5	Un phénomène soumis à la variation inter-langues .....	92
2.5.1	Des différences dans le moment d'émergence des fillers et des déterminants .....	92
2.5.2	Influence des propriétés de la langue cible sur la proportion des fillers produits .....	93
<b>3.</b>	<b>LE TRAITEMENT DES FILLERS ET DES DETERMINANTS DANS LA LITTERATURE, REFLET DE LA DIVERSITE DES APPROCHES METHODOLOGIQUES ET THEORIQUES .....</b>	<b>94</b>
3.1	Perspective phonologique sur le développement des déterminants .....	94
3.1.1	La place de la phonologie dans le débat sur le statut des fillers .....	94
3.1.2	Structure prosodique de la langue en cours d'acquisition .....	98
3.2	Influence du contexte lexical et syntaxique .....	101
3.2.1	Propriétés lexico-sémantiques du nom .....	101
3.2.2	Contrastes dans l'usage des formes en position pré-nominale .....	103
3.2.3	Rôle du cadre lexico-syntaxique .....	106
3.3	Production des formes pré-nominales et contextes pragmatico-discursifs .....	108
3.3.1	Influence du contexte discursif sur la production des proto-morphèmes .....	108
3.3.2	Une sensibilité précoce aux fonctions pragmatico-discursives des déterminants ? .....	111
<b>4.</b>	<b>ENJEUX D'UNE ETUDE SUR L'ACQUISITION DU PARADIGME DES DETERMINANTS .....</b>	<b>124</b>
4.1	Insuffisance des approches monodimensionnelles .....	125
4.2	Nécessité de la prise en compte des formes de transition .....	126
4.3	Des résultats parfois contradictoires .....	127
4.4	Fillers et fonctions pragmatico-discursives : une interface à explorer .....	128
	<b><u>PARTIE II - QUESTIONS DE RECHERCHE ET METHODOLOGIE .....</u></b>	<b><u>131</u></b>
	<b>CHAPITRE III .....</b>	<b>133</b>
<b>1.</b>	<b>DESCRIPTION DES FORMES .....</b>	<b>136</b>
<b>2.</b>	<b>CONSTRAINTES DISTRIBUTIONNELLES .....</b>	<b>137</b>
2.1	Structure syllabique du nom .....	137
2.2	Sémantisme du nom .....	137
2.3	Contextes lexicaux et syntaxiques .....	138
<b>3.</b>	<b>INFLUENCES DES FACTEURS FONCTIONNELS .....</b>	<b>138</b>
3.1	Référence et discours .....	139
3.2	Rôle des reprises .....	139
<b>4.</b>	<b>MISE EN PERSPECTIVE DES FACTEURS DISTRIBUTIONNELS ET FONCTIONNELS .....</b>	<b>139</b>
	<b>CHAPITRE IV .....</b>	<b>141</b>
<b>1.</b>	<b>QUESTIONNEMENTS PREALABLES A LA CONSTITUTION DU CORPUS .....</b>	<b>143</b>
1.1	Le type de corpus à recueillir .....	143
1.2	La population à étudier .....	145
1.3	Le choix d'un recueil 'naturel' .....	146
<b>2.</b>	<b>LE RECUEIL DE CORPUS A L'HEURE DU PARTAGE DE DONNEES .....</b>	<b>146</b>
2.1	Le partage de données .....	146
2.2	Le recueil d'un corpus spécifique à la thèse .....	147
<b>3.</b>	<b>PRESENTATION DES ENFANTS .....</b>	<b>149</b>
3.1	Le corpus de Lubin .....	149
3.2	Le corpus de Naël .....	150
3.3	Le corpus d'Antonin .....	151
3.4	Le corpus d'Adrien .....	152

<b>4. LES SITUATIONS D'ENREGISTREMENT .....</b>	<b>154</b>
4.1 Les participants lors des enregistrements.....	154
4.2 Les types d'activités .....	155
<b>5. LA PLACE DE L'OBSERVATEUR .....</b>	<b>155</b>
5.1 Le statut de l'observateur dans les familles .....	156
5.2 Une observation directe mais périphérique .....	158
5.3 La vidéo, reflet infidèle de la réalité.....	159
5.4 Le dépassement du paradoxe de l'observateur.....	160
<b>6. LA TRANSCRIPTION DU CORPUS.....</b>	<b>163</b>
6.1 La sélection des données à analyser .....	163
6.2 L'outil utilisé pour la transcription .....	165
6.3 Les conventions de transcription .....	165
6.4 La transcription du corpus d'Adrien.....	169
6.5 Les questions soulevées par la transcription.....	170
6.5.1 <i>La transcription : un premier pas dans l'analyse.....</i>	<i>170</i>
6.5.2 <i>Questions posées par le choix du format de transcription .....</i>	<i>173</i>
6.5.3 <i>La transcription des fillers.....</i>	<i>174</i>
<b>7. L'ANALYSE DU CORPUS .....</b>	<b>178</b>
7.1 Identification et caractérisation des formes analysées.....	178
7.2 Analyse phono-prosodique des formes pré-nominales.....	179
7.3 Analyse du statut sémantique des noms .....	179
7.4 Analyse de la dépendance au lexème et aux structures syntaxiques.....	180
7.5 Caractérisation pragmatico-discursive des référents.....	181
7.6 Analyse des phénomènes de reprise.....	181
7.7 L'outil utilisé pour l'analyse .....	182
 <b><u>PARTIE III - ANALYSES ET RESULTATS.....</u></b>	 <b><u>185</u></b>
<b>CHAPITRE V.....</b>	<b>187</b>
<b>1. IDENTIFICATION DES POSITIONS DITES 'PRE-NOMINALES' ET CATEGORISATION DES FORMES ....</b>	<b>189</b>
1.1 Identification des positions pré-nominales .....	189
1.2 Catégorisation des formes pré-nominales .....	192
<b>2. DEVELOPPEMENT DES FORMES PRE-NOMINALES CHEZ LES QUATRE ENFANTS DU CORPUS.....</b>	<b>196</b>
2.1 Dans le corpus de Lubin.....	196
2.2 Dans le corpus de Naël .....	198
2.3 Dans le corpus d'Antonin.....	202
2.4 Dans le corpus d'Adrien .....	205
<b>3. BILAN COMPARATIF SUR L'EVOLUTION DE LA POSITION PRE-NOMINALE DANS LE CORPUS.....</b>	<b>207</b>
3.1 Des variations de rythme et de style .....	207
3.2 Une mise en place progressive mais non linéaire des proto-formes et des formes adultes .....	209
3.3 Fillers et déterminants : des catégories (transitionnelle et adulte) en évolution ? .....	211
 <b>CHAPITRE VI .....</b>	 <b>223</b>
<b>1. INFLUENCE DE LA STRUCTURE SYLLABIQUE.....</b>	<b>226</b>
1.1 Décompte du nombre de syllabes des noms et des syntagmes phonologiques.....	227
1.2 Longueur syllabique des syntagmes phonologiques .....	228
1.3 Distribution des formes devant les noms mono- et bisyllabiques .....	231
1.4 Distribution des formes devant les noms de plus de deux syllabes .....	234
1.5 Une préférence pour le pied binaire ?.....	237
<b>2. INFLUENCE DU STATUT SEMANTIQUE DU NOM ASSOCIE A LA FORME PRE-NOMINALE .....</b>	<b>239</b>
2.1 Caractérisation sémantique des noms du corpus.....	240
2.2 Emploi des classes sémantiques de noms dans le corpus.....	241

2.3 Distribution des formes devant les noms animés, inanimés et de procès .....	244
2.4 Distribution des formes devant les noms concrets et abstraits .....	252
2.5 Des combinaisons préférentielles entre formes pré-nominales et statuts sémantiques des noms ? .....	255
<b>3. IMPACT DU CONTEXTE LEXICAL ET SYNTAXIQUE .....</b>	<b>259</b>
3.1 Mesure de l'overlap et caractérisation des structures syntaxiques .....	260
3.1.1 <i>Mesure de l'overlap</i> .....	260
3.1.2 <i>Caractérisation des structures syntaxiques</i> .....	261
3.2 Alternance des formes devant les noms .....	263
3.2.1 <i>Distribution des noms à occurrences uniques ou multiples</i> .....	263
3.2.2 <i>Alternance entre omission, filler et déterminant</i> .....	264
3.2.3 <i>Alternance entre différents types de déterminants</i> .....	272
3.2.4 <i>Alternance des formes entre les séances</i> .....	277
3.3 Cadres lexico-syntaxiques .....	279
3.3.1 <i>Cadres lexico-syntaxiques les plus fréquemment produits</i> .....	279
3.3.2 <i>Combinaisons entre cadres lexico-syntaxiques et formes pré-nominales</i> .....	284
3.4 La construction de la catégorie 'déterminant' : des usages figés aux usages généralisés ? .....	294
<b>CHAPITRE VII.....</b>	<b>299</b>
<b>1. INFLUENCE DES CARACTERISTIQUES PRAGMATICO-DISCURSIVES DES REFERENTS .....</b>	<b>301</b>
1.1 Caractérisation pragmatico-discursive des référents.....	302
1.1.1 <i>Caractère référentiel du syntagme nominal</i> .....	302
1.1.2 <i>Type de référence</i> .....	304
1.1.3 <i>Statut attentionnel et discursif du référent</i> .....	305
1.1.4 <i>Familiarité du référent</i> .....	306
1.2 Gestion des contrastes pragmatico-discursifs chez l'enfant.....	308
1.2.1 <i>L'opposition entre syntagmes référentiels et non référentiels</i> .....	308
1.2.2 <i>Caractère particulier ou indéterminé du référent</i> .....	316
1.2.3 <i>Distinction entre référent nouveau et donné dans le discours</i> .....	324
1.2.4 <i>Impact du degré de connaissance partagée entre les interlocuteurs</i> .....	335
1.3 Une sensibilité de l'enfant aux facteurs pragmatico-discursifs ?.....	338
<b>2. LE ROLE DES REPRISSES DU NOM DANS L'ACQUISITION DES FORMES .....</b>	<b>339</b>
2.1 Description des reprises de l'enfant .....	343
2.1.1 <i>Identification des reprises</i> .....	343
2.1.2 <i>Source et distance de la reprise</i> .....	343
2.1.3 <i>Caractérisation de la forme pré-nominale de la mention précédente</i> .....	344
2.2 Inventaire des contextes de production des noms et des types de reprises.....	344
2.3 Emploi des formes pré-nominales devant les noms repris .....	351
2.3.1 <i>Comparaison de la distribution des formes pré-nominales devant les noms introduits dans le discours et repris</i> ...	351
2.3.2 <i>Comparaison de la distribution des formes pré-nominales dans les contextes de reprises immédiates et à distance</i>	357
2.4 Une influence des reprises sur les formes pré-nominales produites par les enfants ? .....	362
<b>CHAPITRE VIII .....</b>	<b>369</b>
<b>1. TRAITEMENT STATISTIQUE DES DONNEES .....</b>	<b>373</b>
1.1 Intérêts du modèle mixte pour notre étude .....	374
1.2 Elaboration du modèle.....	374
1.3 Arbres de partition.....	376
<b>2. PREVALENCE - CONSTANTE OU PERIODIQUE - D'UN FACTEUR OU COMPLEMENTARITE DES FACTEURS ? .....</b>	<b>376</b>
2.1 Omissions.....	377
2.2 Fillers.....	380
2.3 Déterminants .....	383
2.4 Synthèse des résultats .....	386

<b>CONCLUSION</b> .....	<b>389</b>
<b>1. SYNTHÈSE GÉNÉRALE</b> .....	<b>389</b>
1.1 Description des formes.....	390
1.2 Influence des facteurs distributionnels.....	390
1.3 Influence des facteurs fonctionnels.....	391
1.4 L'acquisition du déterminant : un phénomène au cœur de plusieurs dimensions.....	392
<b>2. UN TABLEAU À COMPLÉTER</b> .....	<b>392</b>
2.1 La catégorie transitionnelle des fillers : que reste-t-il à explorer ?.....	393
2.2 Appréhender le développement des déterminants en lien avec le développement global des expressions référentielles.....	394
2.3 Influence de l'activité et des types de discours.....	395
2.4 Comparaisons inter-langues .....	396
2.5 Extension du corpus analysé.....	397
<b>BIBLIOGRAPHIE</b> .....	<b>399</b>
<b><u>PARTIE IV - ANNEXES</u></b> .....	<b><u>415</u></b>
<b>FORMULAIRE D'AUTORISATION PARENTALE</b> .....	<b>417</b>
<b>CONVENTIONS DE TRANSCRIPTION</b> .....	<b>419</b>
<b>DÉCOUPAGE DES TOURS DE PAROLE EN ÉNONCÉS</b> .....	<b>429</b>
<b>TABLE DES MATIÈRES</b> .....	<b>433</b>

## **L'acquisition d'un paradigme - Éclairage multidimensionnel sur la mise en place des déterminants chez quatre enfants entre 1;6 et 3;5**

De par leur présence à l'interface entre plusieurs domaines linguistiques (phonologie, morpho-syntaxe, sémantique ou pragmatique), les déterminants ont depuis longtemps suscité l'intérêt des chercheurs en acquisition. On trouve ainsi de nombreux travaux portant notamment sur les contraintes phono-prosodiques, lexico-syntaxiques, sémantiques ou pragmatiques à l'œuvre dans le processus d'acquisition de ces formes, mais peu d'entre eux ont cherché à proposer une analyse combinée de ces facteurs, en tenant compte à la fois des morphèmes adultes et des formes de transition (omissions et fillers). Nous avons donc souhaité, dans le cadre de ce travail, décrire l'émergence des déterminants et de leurs proto-formes en essayant de saisir, dans une perspective multidimensionnelle intégrant à la fois des facteurs distributionnels et fonctionnels, quels facteurs influencent le plus ce processus, et à quelles périodes.

Cette thèse repose sur l'analyse d'un corpus longitudinal de quatre enfants, âgés de 1;6 pour le plus jeune lors de la première séance sélectionnée à 3;5 pour le plus âgé lors de la dernière. Nous avons repéré au sein de ce corpus tous les noms produits par les enfants ainsi que les formes les précédant et chaque syntagme a fait l'objet d'une analyse phono-prosodique, lexico-syntaxique, sémantique, pragmatique et discursive.

Les résultats montrent d'une part que tous les facteurs distributionnels peuvent - à des degrés différents - influencer la production des formes pré-nominales, et d'autre part, que les usages que les enfants font des formes témoignent d'une sensibilité précoce à certains contrastes pragmatico-discursifs. Toutefois, aucun des facteurs pris séparément n'a permis de rendre compte de l'intégralité des productions des enfants. L'analyse conjointe de ces facteurs distributionnels et fonctionnels a confirmé les tendances dégagées au préalable et a mis en évidence l'influence d'un ensemble de facteurs (et en particulier, phono-prosodique, syntaxique et pragmatique) plus que d'un facteur unique ou prédominant tout au long du processus d'acquisition des déterminants.

*Mots clés : acquisition du langage, acquisition des déterminants, fillers, phono-prosodie, cadres lexico-syntaxiques, pragmatique, discours*

---

## **Acquiring a paradigm - Multidimensional perspective on the setting up of determiners in a corpus of four children from 1;6 to 3;5**

Given their presence at the crossroads of different linguistic fields (phonology, morphosyntax, semantics or pragmatics), determiners have for a long time aroused the interest of language acquisition researchers. We thus find several studies dealing with the phono-prosodic, lexical and syntactic, semantic or pragmatic constraints involved in the acquisition of determiners. But only few of them have provided a combined analysis of all these factors, considering both adult morphemes and transitional forms (omissions and filler syllables). Our study aims to describe the emergence of filler syllables and determiners and understand - in a multidimensional perspective examining simultaneously distributional and functional factors - which factors may affect the acquisition process, and over which periods of time.

Our study is based on a corpus of four children, aged from 1;6 for the youngest in the first selected session to 3;5 for the eldest in the last session. We first identified all the noun phrases produced in the data and each occurrence was then analysed according to phono-prosodic, lexical, syntactic, semantic, pragmatic and discursive factors.

Our results show that each distributional factor may influence, at different degrees, the production of pre-nominal forms and that children's use of forms shows early sensitivity to pragmatic factors. However, none of the factors taken separately have accounted for all of the children's productions. An additional combined analysis of both distributional and functional factors has supported the initial trends and highlighted the influence of a series of factors (and in particular, phono-prosodic, syntactic and pragmatic), more than a single and predominant one throughout the acquisition process of determiners.

*Keywords : language acquisition, acquisition of determiners, filler syllables, prosodic phonology, lexical and syntactic frames, pragmatics, discourse*

UNIVERSITE SORBONNE NOUVELLE - PARIS 3

ECOLE DOCTORALE 268

MAISON DE LA RECHERCHE - 4, RUE DES IRLANDAIS, 75005 PARIS