

L'ÉMERGENCE DU SYNTAGME NOMINAL EN POLONAIS L1 : CONTRAINTES LINGUISTIQUES ET FACTEURS DISCURSIFS

Ewa Lenart et Isabelle Maillochon

Université Paris 8 & UMR 7023 Structures formelles du langage
et Université Le Havre-Normandie & UMR 7023 Structures formelles du langage

Résumé : L'enfant construit graduellement la grammaire au contact avec la langue de l'entourage (input) en s'appuyant sur ses capacités cognitives d'apprentissage. Cette gradualité de l'acquisition est notamment montrée à travers la grammaticalisation du syntagme nominal (SN) en français (par exemple, Bassano *et al.* 2005, 2008, Salazar Orvig *et al.* 2013, Veneziano 2003, 2019). L'enfant francophone doit comprendre la contrainte d'emploi du déterminant obligatoire du nom et l'intégrer aux autres contraintes morphosyntaxiques du français. Dans une perspective comparative inter-langue, nous nous interrogeons sur le parcours acquisitionnel d'un enfant polonophone pour qui les difficultés acquisitionnelles se situent davantage au niveau discursif, le polonais ne possédant pas la catégorie grammaticale de l'article. De fait, en polonais, le nom est le plus souvent nu et l'ensemble des déterminants facultatifs, indéfinis, démonstratifs ou possessifs, est régi par des contraintes discursives.

La présente étude est basée sur un corpus longitudinal de productions orales d'un enfant polonophone de 24 à 35 mois. Elle examine l'émergence du SN en polonais L1, interroge l'existence de morphèmes grammaticaux comparables aux *fillers* (ou proto-morphèmes) et met en regard le rôle de la structure informationnelle avec la grammaticalisation du SN au sens large. Elle contribue à déterminer la part de la structure de la langue dans la complexification du langage chez l'enfant.

1. Grammaticalisation du SN en L1

Les approches fonctionnalistes et constructivistes de l'acquisition du langage soulignent le rôle prépondérant de facteurs cognitifs et environnementaux dans le développement du langage (Bates & MacWhinney 1989, Karmiloff-Smith 1992, Veneziano 2003, Bassano *et al.* 2008, Tomasello 2000, 2003). En effet, l'enfant construit graduellement la grammaire dans l'interaction avec l'entourage (qui lui fournit l'*input*), tout en s'appuyant sur ses capacités cognitives d'apprentissage, telles que la mémorisation et la catégorisation des éléments de l'*input*. Ces approches insistent sur l'importance de prendre en compte différents facteurs pour déterminer l'acquisition d'une compétence (Pine et Lievens 1997, Veneziano et Sinclair 2000, Veneziano 2019).

Bassano (2010, pour une synthèse) a montré la gradualité de l'acquisition à l'exemple de la grammaticalisation du SN en français en analysant les productions spontanées d'enfants francophones de 20, 30 et 39 mois. Elle a mis en avant des contraintes lexico-sémantiques et grammaticales (notamment l'obligation ou non de l'emploi du déterminant), les facteurs discursifs (la structure informationnelle de l'énoncé, le contexte intra- et extra discursif), ainsi que le rôle de l'*input*.

Parallèlement, l'acquisition de la grammaire dépend d'autres capacités langagières comme la phonologie, la prosodie et le lexique (pour la dépendance entre grammaire et lexique : *the critical mass hypothesis*, Marchman et Bates 1994, Bates et Goodman 1999). Bassano

(2010) prône en faveur d'une continuité inter-domaines, entre les compétences lexicales, grammaticales et phonologiques/prosodiques. En effet, ces composants de la capacité langagière sont interdépendants. Elle accorde également un intérêt particulier à la variabilité, inter-linguistique et inter-/intra-individuelle qui permettent d'apporter un éclairage sur les moments particuliers du développement langagier (changements, transitions, moments critiques, cf. Bassano et Van Geert 2007).

En français, considéré comme une langue particulièrement restrictive, le déterminant doit être antéposé au nom. Il porte dans la plupart des cas des marques du genre, du nombre et du statut défini/indéfini de la référence. Le déterminant peut être omis avec des noms propres ou dans des constructions verbales/prépositionnelles (*avoir faim, sans beurre*). Il s'en suit que l'enfant francophone doit comprendre la contrainte d'emploi du déterminant obligatoire du nom et l'intégrer aux autres contraintes morpho-syntaxiques du français (cf. Veneziano 2003, Salazar Orvig *et al.* 2013). Les travaux de Bassano *et al.* (2008), par exemple, montrent une explosion de l'utilisation des déterminants est observée à partir de deux ans environ. Cette explosion est préparée par des phénomènes précurseurs ou concomitants, à travers la présence de *fillers* (pré- ou proto-déterminants), qui s'insèrent dans la structure prosodique du français, avant l'âge de deux ans. Veneziano (2019 : 12) donne des exemples de *fillers* suivants : /ə/, dans /əsʊ/ pour « sou », dans /əʃd/ pour « bouchon », dans /ə'pik/ pour « pique », ou /a/ dans /adam/ pour « dame », ou encore /o/ dans /oli/ pour « lit ». Voir aussi Veneziano (2003), Yamaguchi *et al.* (2021). A l'âge de 30 mois, les déterminants se diversifient : articles défini et indéfini, masculin singulier, féminin singulier, pluriel ; possessif, démonstratif. Les omissions sont peu fréquentes et des contrastes (*overlaps*) variés : *un gâteau/le gâteau, une feuille/la feuille, les couleurs/des couleurs*. A 39 mois, la catégorie grammaticale des déterminants nominaux est opérante. Ainsi, en français, la capacité à produire des morphèmes pré-nominaux se fait progressivement et marque l'entrée dans la grammaticalisation des noms¹.

Différentes études interrogent également l'influence des facteurs prosodiques et rythmiques sur le processus de grammaticalisation des déterminants. Celles portant sur des langues à accent, avec un patron trochaïque, comme l'anglais et l'allemand, montrent chez les enfants de moins de deux ans une tendance à omettre les syllabes non accentuées, comme les déterminants, devant les syllabes accentuées (Demuth 1996, Gerken 1996, Demuth & Tremblay 2008). En revanche, la structure rythmique du français repose sur la syllabe et le patron iambique (et non pas sur l'accent). Il s'en suit que les enfants francophones devraient produire plus souvent un *filler* ou un déterminant devant des noms monosyllabiques que devant des noms à plusieurs syllabes où la contrainte prosodique est déjà réalisée. Cela est démontré chez les enfants de 20 mois qui emploient des noms monosyllabiques plus fréquemment que plurisyllabiques. Cet écart est moindre dans les enfants plus âgés.

L'hypothèse de l'influence des facteurs lexico-sémantiques sur la grammaticalisation des déterminants se vérifie chez les enfants de 30 et 39 mois. Le trait sémantique animé/inanimé qui constitue une des propriétés centrales du nom (Comrie 1989) sous-tendrait une hiérarchisation des noms communs : les animés susceptibles d'être plus souvent sujets et agents seraient non marqués, les inanimés associés plus souvent à la fonction d'objet auraient davantage besoin d'une détermination nominale pour les

¹ Comme nous le verrons pour le polonais (p. 9), la chronologie et la temporalité de l'acquisition établies par Bassano donnent un contour acquisitionnel à partir d'un échantillon d'apprenants réduit. Voir Veneziano (2019) pour une synthèse sur la grammaticalisation acquisitionnelle en L1.

individualiser. En effet, ces enfants préfèrent employer les noms inanimés avec un déterminant et les animés sans déterminant.

Ainsi, l'enfant acquiert graduellement les déterminants nominaux. Entre 20 et 30 mois, on observe une diversification progressive des formes de déterminants et de leurs contextes d'emploi argumentant en faveur des modèles gradualistes des approches constructivistes (Pine et Lieven 1997, Tomasello 2000, 2003). Parallèlement, les influences prosodiques et lexicales de la langue se manifestent dans l'acquisition des déterminants nominaux : les premières dominent chez les plus jeunes, les secondes jouent un rôle important ensuite (Bassano 2008, Bassano *et al.* 2008). Les travaux de Bassano et les collègues ont impulsé un autre questionnement, à savoir celui de l'impact de la structure informationnelle sur l'émergence des déterminants en français (*cf.* Bassano *et al.* 2013, 2014).

Dans une perspective interlangue, d'autres travaux sur l'acquisition de la détermination nominale interrogent les mêmes phénomènes développementaux. Ils montrent que les déterminants sont produits plus tôt dans les langues romanes que dans les langues germaniques où l'omission des déterminants est plus fréquente et dure plus longtemps que dans les premières (Lleó 2001, allemand et espagnol ; Van der Velde 2004, néerlandais et français, par exemple).

Ces comparaisons inter-langues permettent de dégager des facteurs spécifiques aux langues qui influencent le parcours acquisitionnel de l'enfant. Ces facteurs s'ajoutent à l'effet de l'âge, lié aux caractéristiques cognitives de l'enfant en cours de développement, et à l'effet de l'*input* linguistique.

2. Facteurs discursifs et pragmatiques et émergence des déterminants nominaux

Nous savons que les enfants sont tôt sensibles au statut informationnel des référents dans le discours (Greenfield et Smith 1976 ; Baker et Greenfield 1988). Chez les très jeunes enfants, la réalisation linguistique des arguments dans le discours interactionnel dépend de leur accessibilité cognitive : une forme linguistique plus informative est réservée aux référents peu accessibles, et *vice versa* pour les référents accessibles (*cf.* Allen, Skarabela & Hughes 2008, pour une revue). Après une étude pionnière de Brown (1973), l'acquisition de la distinction défini/indéfini, importante pour étudier l'accessibilité des référents, a été abordée plus récemment, par exemple, par De Cat (2004), Serratrice (2000), Kupisch (2006), Rozendaal & Baker (2008). Cependant, ces études ne s'intéressent pas au rôle des facteurs discursifs dans l'émergence des déterminants, ni, concrètement, aux relations entre la configuration informationnelle du référent et la forme du SN qui l'encode (avec ou sans déterminant).

Cette interface a été interrogée par Bassano *et al.* (2013, 2014) qui ont examiné l'influence de la *configuration informationnelle* sur la production des déterminants en français. Trois dimensions interdépendantes en font partie : la structure informationnelle de l'énoncé, le statut informationnel du référent dans le discours et le statut informationnel du référent dans le contexte extra-linguistique. La première dimension correspond à l'opposition *topique/focus* (*commentaire*), définie, d'après le modèle de la *quaestio* (Klein et von Stutterheim 1991), comme la distinction entre ce qui est présupposé par la *quaestio* (question implicite ou explicite) à laquelle l'énoncé répond, et l'information spécifiée. Les deux autres dimensions ont trait à l'accessibilité du référent dans le contexte discursif et extra-discursif (entité nouvelle ou ancienne/donnée par la situation de communication (ou les connaissances du monde) ou déjà introduite dans le discours). Cette grille

tridimensionnelle est appliquée au (même) corpus de données longitudino-transversales de six enfants de 20, 30 et 39 mois, trois garçons et trois filles, chacun enregistré en situation d'interaction avec sa mère pendant des séances de 30 minutes intégralement transcrites.

Partant du fait que les jeunes enfants en situation interactionnelle réservent des formes plus informatives à des référents peu accessibles et *vice versa*, des formes moins informatives dans des contextes où le référent est supposé accessible (cf. *supra*), l'apparition des déterminants devrait être favorisée pour les référents en fonction de focus (qui possède une plus grande valeur informationnelle), nouveaux dans le discours ou non donnés par le contexte (les moins accessibles). En effet, les résultats montrent une forte prédominance de l'un des pôles de chaque dimension sur la production des SN. Ils sont plus fréquents en fonction de *focus* que de *topique*, pour les référents donnés dans le discours que pour les référents nouveaux, et pour les référents donnés par la situation de communication. Les enfants se limitent principalement à la situation de communication et des objets qui s'y trouvent et n'emploient que peu de prédicats. Ces effets sont d'autant plus marqués que les enfants sont plus jeunes.

L'analyse de cette dimension tripartite associée à l'émergence des déterminants à l'âge de 20 mois confirme l'hypothèse que les jeunes enfants produisent un *filler* ou un déterminant proportionnellement plus fréquemment avec des SN en *focus* et qui réfèrent à des entités nouvelles dans le contexte discursif. En revanche, le statut de l'information par rapport au contexte situationnel ne paraît pas une dimension pertinente, étant donné que les jeunes enfants sont très dépendants de la situation de communication et mentionnent très peu d'entités hors contexte. Ces résultats apportent des éléments de discussion sur le rôle des facteurs discursifs dans l'émergence de la grammaticalité dans le langage de l'enfant. Nous avons observé un effet significatif de la configuration informationnelle (structure *topique/focus*) sur l'acquisition du contraste défini/indéfini. A 20 mois, plus de la moitié des morphèmes pré-nominaux sont des *fillers* et, parmi les articles, l'article indéfini est plus fréquent que l'article défini qui, proportionnellement, est plus fréquent en topique qu'en focus. En effet, l'article indéfini n'est jamais utilisé en topique et constitue la majeure partie des déterminants accompagnant des SN en focus.

Pour résumer, les facteurs discursifs influencent très tôt la production du SN et l'émergence des déterminants. L'impact majeur de la configuration informationnelle *topique/focus* est visible à travers l'emploi privilégié de *fillers* ou de déterminants en focus (principalement l'article indéfini).

Dans une perspective comparative inter-langue, une question fondamentale se pose de savoir quel est le processus de grammaticalisation du SN dans des langues sans détermination nominale obligatoire. Quelle est alors l'importance des facteurs discursifs et pragmatiques pour les mêmes phénomènes observés dans des langues à articles, comme le français ?

Afin de répondre à ces questionnements, disposant d'enregistrements longitudinaux de productions orales d'un enfant polonophone entre 21 et 35 mois, nous avons étudié l'émergence du SN en polonais, langue sans articles et sans déterminants nominaux obligatoires.

3. L'acquisition du polonais L1

Depuis les années 1980, les études inter-langues sont de plus en plus nombreuses et les chercheurs comparent l'acquisition d'une variété de L1s peu étudiées, à côté des recherches sur l'anglais qui restent majoritaires. Slobin reste le pionnier des études inter-langues et sa présentation des « principes opérationnels » (1973, 1982) montre les stratégies de compréhension, de production et de construction d'un système de règles mises en place par les enfants lorsqu'ils acquièrent leur L1. Slobin écrit ensuite un ouvrage dans lequel il rassemble des études monographiques sur des langues typologiquement variées (Slobin 1985). Smoczyńska y présente une synthèse de travaux sur le polonais et analyse des données provenant du corpus de Kraków (dix enfants, Smoczyńska 1985) et de Poznań (quatre enfants, Weist *et al.* 1984) (cf. corpus polonais de CHILDES, MacWhinney et Snow 1985). Les travaux de Weist sur le polonais qui s'appuient sur une approche fonctionnelle de l'acquisition, *Role and Reference Grammar* (Van Valin 1991, 2005) montrent l'interface entre la syntaxe et la sémantique, en lien avec des facteurs pragmatiques et cognitifs pendant la construction de la grammaire par l'enfant.

Le polonais étant une langue flexionnelle, les difficultés de l'enfant polonophone acquérant sa L1 se situent ailleurs que celles de l'enfant francophone. En effet, en polonais, la morphologie flexionnelle joue un rôle important pour exprimer les différentes fonctions syntaxiques. Les morphèmes nominaux expriment le cas, le genre et le nombre, les morphèmes verbaux la personne et le nombre. Smoczyńska (1985 : 601) souligne que « *the individual suffixes can only be understood within the context of a declensional pattern* ». Autrement dit, une forme fléchie du mot (nom ou verbe) doit être considérée en opposition avec toutes les formes du même paradigme plutôt que par rapport à sa forme de base (le cas nominatif singulier pour le nom, l'infinitif pour le verbe), comme c'est le cas pour l'anglais ou le français. Par conséquent, « *morphological development should be conceived of as the acquisition of the ability to REPLACE grammatical morphemes according to the rules of language, rather than the ability to ADD them to basic forms when required* » (*idem* : 597).

Les études de Weist (par exemple, 1990, 2010), montrent que les enfants polonophones âgés entre 1;7 et 1;11 ans², distinguent les noms des verbes et les verbes transitifs des intransitifs. Ils distinguent le perfectif de l'imperfectif et les traits sémantiques statique *vs* dynamique et télique *vs* atélique. Ces connaissances sont à la base de l'émergence de la morphologie flexionnelle (*finite morphology*). A l'âge de deux ans, les verbes sont fléchis et portent la marque du temps, de l'aspect et du mode. Les enfants polonophones doivent faire attention à la fin des mots, puisque les suffixes permettent la découverte des accords grammaticaux (fonctions grammaticales). De plus, le polonais étant une langue *pro-drop* (pronom sujet non obligatoire), le suffixe du verbe peut être le seul indice pour identifier la personne et le nombre du sujet, qui constitue « *privileged syntactic argument* » (PSA),

2 La temporalité et la chronologie de l'acquisition données dans cette section sont à relativiser, compte tenu du petit nombre d'enfants étudiés et de l'absence de résultats ayant pris compte les différences interindividuelles. Les choix méthodologiques peuvent aussi induire des résultats différents, complémentaires. Par exemple, l'acquisition du datif en polonais est attestée à l'âge de 2;5 par Smoczyńska (1985) (données naturelles), contrairement à Dąbrowska et Szczerbiński (2006) (données expérimentales) où le datif neutre n'est pas encore maîtrisé à l'âge de 4;8. Les résultats dépendent également de la précision de l'analyse, ce qui rend difficile leur généralisation. Par conséquent, le parcours acquisitionnel présenté dans cette section a pour but de mettre en exergue les difficultés auxquelles les enfants polonophones sont confrontés, difficultés liées aux spécificités du polonais.

assigné au cas nominatif. Lorsque le sujet est présent, les enfants contrastent le nominatif (fonction sémantique agent) avec l'accusatif (fonction sémantique patient), qui est acquis entre 1;6 et 2;0 (cf. Smoczyńska 1985). Ainsi, l'« agent » se voit assigné le rôle thématique (cf. *Case Grammar* de Bowerman 1973). Pour plus de détails sur l'acquisition de la morphologie nominale voir aussi Krajewski (2008), Krajewski *et al.* 2012).

En bref, la morphologie verbale permet de récupérer la structure argumentale de la phrase. La *deixis* personnelle commence à être contrôlée à l'âge de deux ans et quelques mois : les enfants utilisent des morphèmes permettant d'identifier la personne, le nombre et parfois le genre du sujet dans un contexte conversationnel.

Quant à la structure interne du syntagme nominal, aucune contrainte grammaticale n'est imposée au niveau de la présence du déterminant. Le nom est le plus souvent nu, le polonais ne possède pas la catégorie grammaticale de l'article. Il peut toutefois être accompagné d'un ensemble de déterminants facultatifs, indéfinis, démonstratifs ou possessifs. L'emploi de ces déterminants relève davantage des compétences discursives des locuteurs. Il en résulte que la grammaticalisation des noms en polonais doit être regardée différemment, par rapport au même phénomène observé en français et dans d'autres langues qui ont bénéficié de nombreuses études, biaisant quelque part l'importance de certains phénomènes linguistiques, notamment l'importance des articles.

Compte tenu de cette différence fondamentale entre le français et le polonais dans la réalisation du SN et en nous inspirant des travaux de Bassano *et al.* (2005, 2008), nous voulons répondre aux questions suivantes :

- Quel est le processus de grammaticalisation du SN en polonais ?
- Quel est le rôle de facteurs discursifs dans la grammaticalisation du SN en polonais ?
- Peut-on observer des similitudes dans l'émergence du SN en français et en polonais ?

Nous émettons l'hypothèse que la grammaticalisation du SN en polonais possède un contour spécifique, étant donné une morphologie flexionnelle complexe du polonais. Il s'ensuivrait une autre temporalité de l'acquisition que, par exemple, en français, une singularité des formes premières (au lieu de *fillers*), une spécialisation des formes des noms selon le système casuel du polonais, ainsi qu'une correspondance spécifique entre la forme du nom et la structure informationnelle. Autrement dit, nous nous attendons à une fréquence élevée de SN au nominatif et en topique, suivis de SN à l'accusatif en focus, tous les deux cas remplissant les fonctions sémantiques fondamentales d'agent et de patient.

4. Méthodologie

Recueil des données et corpus³

L'étude repose sur trois enregistrements transcrits d'un corpus de production spontanée d'une jeune enfant polonaise. La méthode de recueil est celle dite de la « production naturelle » : l'enfant a été enregistrée en vidéo à son domicile, au cours de sessions régulièrement espacées (deux par mois de 21 à 35 mois), durant lesquelles elle était en

³ Nous remercions toutes les personnes ayant contribué à la réalisation de cette recherche : Ola et sa mère, l'enquêtrice, Urszula Paprocka-Piotrowska, Marzena Watorek, Dominique Bassano. La constitution du corpus a pu être réalisée grâce au soutien financier de la Commission de la Recherche de l'Université de Paris 8 et de l'UMR 7023 Structures formelles du langage.

situation de vie quotidienne et d'interaction avec son entourage. Cette méthode permet une perspective longitudinale depuis les débuts de la production de syntagmes nominaux. Pour plus de précisions sur la production naturelle, voir Bassano 2005, Kail & Bassano 2000/03, Boulton *et al.* 2013, Krajewski *et al.* 2012, Goldenberg *et al.* 2022.

Le corpus de Ola a été enregistré durant les années 2009-2010. L'enfant, monolingue de langue polonaise, ayant une sœur cadette, est une fille vivant à Lublin en Pologne, de la classe moyenne. Les séances sont non structurées et durent environ une heure. L'enfant interagit avec sa mère ou avec sa tante, qui filme et est l'observatrice, exceptionnellement avec d'autres membres de la famille.

Transcription

Les transcriptions du corpus ont été informatisées en conformité avec les standards internationaux du format CHAT de CHILDES (MacWhinney 2000). La transcription fournit pour chaque session un compte rendu fidèle des productions de l'enfant et de ses interlocuteurs, avec les indications et commentaires requis sur les situations, gestes ou mimiques qui permettent l'interprétation des énoncés. Les formes des noms spécifiques à l'enfant ont été transcrites en API, par exemple, /nem/ pour *krem* (*crème*), /lal/ pour *szal* (*écharpe*), /pam/ pour *pan* (*monsieur*). De même que pour les études du français de Bassano et ses collègues (par exemple, Bassano & Maillachon 1994), le découpage des productions en énoncés remplit une double condition : correspondre à une unité prosodique et comporter au moins un élément identifiable du polonais, excluant ainsi babillages et vocalisations.

A titre d'exemple, nous donnons ci-dessous la transcription d'un énoncé de Ola à 24 mois :

*MOT: Nawet więcej, bo na wsi jak żeśmy byli, to mówię Ci, że jak chciała siusiu, to „mama Ola siusiu cie”
 %*mer: *Même plus, car quand nous étions à la campagne, Ola disait „maman, Ola veut faire pipi” au moment où elle voulait aller aux toilettes*

*OLA: Mama siusiu cie [= chcę]

%*olf: *Maman, je veux faire pipi*

%com : Ola powtarza słowa matki. Wymowa słowa „chcę” zostaje zmiękczone

%com: *Ola répète ce que sa mère vient de dire. La prononciation du mot « chcę » (en fr. veux) fait l'objet du processus de l'adoucissement.*

Sélection des données et codage des énoncés

Pour les besoins de la présente recherche, un codage spécifique, présenté ci-après, a été réalisé sur les énoncés produits par l'enfant OLA dans trois sessions. Le tableau 1 fournit un état récapitulatif des sessions analysées :

Tableau 1. Caractéristiques des données du corpus de Ola (sessions de l'étude)

Session	Date d'enregistrement	Durée d'enregistrement	Age de OLA	Nombre d'énoncés produits par OLA
24 mois	15/07/2009	56 mn	2;00.08	238
30 mois	07/01/2010	57 mn	2;06:00	207
35 mois	07/06/2010	50 mn	2;11:00	281

Ces sessions ont été initialement choisies pour optimiser les conditions de la recherche. Elles permettent de montrer le contour de la grammaticalisation du SN en polonais et de la comparer au français, sachant qu'il s'agit ici d'une première étude de ce riche corpus. Le choix du point de départ à 24 mois est associé au début de la complexification des énoncés de Ola. Avant cet âge, l'enfant produit des énoncés sans verbe et sans forme nominale, par exemple, *a to ?* (*et ça ?*, (seulement 3 occurrences de verbes à 23 mois (*ma* (= *masz*), (*tu as*)) et des (rares) SN au nominatif ou vocatif (*mama, Ola, /lali/* (= *szalik, écharpe*)). Ayant choisi une perspective comparative avec la grammaticalisation du SN en français, c'est aussi l'âge où, en français, est observée une explosion de l'utilisation des déterminants devant le nom. A notre connaissance, aucune étude sur l'acquisition du SN en polonais n'avait croisé différents niveaux d'analyse, à savoir la structure interne du SN et l'organisation informationnelle de l'énoncé (topique/focus). A ce double niveau d'analyse sera ultérieurement ajoutée l'analyse du mouvement référentiel dans le domaine des entités (introduction, maintien et réintroduction des référents).

Nous avons retenu une suite de 120 énoncés par session, soit 360 énoncés, selon le principe de base employé dans les études longitudinales de productions naturelles en français de Bassano et ses collègues (cf. Bassano 2000, Bassano *et al.* 1998), dans lesquels tous les SN (formes nominales et pronominales) sont soumis à un codage à quatre dimensions.

Nous avons codé la position du SN par rapport au verbe (1) : préverbale ou postverbale ; énoncés avec verbe seul ; énoncés sans verbe. En effet, l'ordre des mots relativement libre en polonais permet d'exprimer des fonctions pragmatiques/discursives du SN.

En ce qui concerne la structure du SN (2), nous avons codé les formes nominales (nom animé / nom inanimé/ nompropre) et les formes pronominales (pronom personnel / démonstratif / indéfini/ interrogatif), ainsi que le marquage casuel des formes nominales (3) : NOMinatif / GENitif / DATif / ACCusatif / INSTRumental / LOCatif / VOCatif. Dans des énoncés sans verbe, nous avons regardé le contexte de l'échange verbal pour interpréter le marquage casuel, compte tenu du fait que certains cas sont formellement identiques, par exemple, le nominatif et l'accusatif des noms inanimés. La dimension discursive a été analysée du point de vue de la structure informationnelle de l'énoncé (topique/focus) (4). Comme pour le marquage casuel, le statut informationnel du SN a été décidé en tenant compte du contexte. Dans l'exemple 1 (OLA 30 mois), le SN lode [= brode] (*barbe*) a été interprété comme un SN à l'accusatif et focus, en tenant compte de l'énoncé 92 qui donne le contexte de l'emploi :

(1) [92b⁴]*OLA: o tutaj ja odlalam jednego wlonsia [= o tutaj ja oderwałam jednego wąsa].
(*Oh ici, j'ai coupé une moustache*)

[93]*OLA: o to (*Oh ça!*)

[94]*OLA: lode [= brode] (*barbe*)

5. Analyse des résultats

Dans le corpus analysé, nous observons tout d'abord un saut qualitatif à 30 mois en ce qui concerne la longueur moyenne des énoncés (LME), en nombre de mots. La LME passe de

⁴ Les chiffres entre [] indiquent le numéro de l'énoncé dans le corpus ; b indique ici le deuxième SN analysé (wlonsia = moustache), le premier étant *ja* (*je*).

1,491 à 24 mois à 2,458 à 30 mois et ne change pas à 35 mois. La longueur maximale relevée est de 5 mots à 25 mois, 9 mots à 30 mois (1 occ.) et 7 mots à 35 mois (2 occ.), ce qui peut indiquer une fluctuation entre 30 et 35 mois.

Au niveau syntaxique, la complexification des productions est visible dans l'augmentation progressive des énoncés avec SN entre 24 mois et 35 mois (respectivement, 68%, 80% et 89%), par exemple, « *Mama siusiu cie* » [= *chcę*] (= *Maman, je veux faire pipi*) ou « *Nemo jeszcze mam i Klifolda* [= *Kliforda*] (= *J'ai encore Némó et Clifford*) et, par conséquent, la diminution des énoncés sans SN, par exemple à 24 mois « *Mniam.* » (= *miam*), « *Nie !* » (= *non*). Lorsque le SN est produit par l'enfant, il apparaît le plus souvent dans des énoncés sans verbe (0vb) et cela aux trois âges étudiés. Lorsque l'énoncé comporte un verbe et un SN, la position préverbale (prevb) est la plus fréquente, avec une forte domination à 30 mois, par exemple :

(2) [6a]*OLA: *amoci totuje* [= *pomocy potrzebuje*]. (= aide (j)'ai besoin « J'ai besoin d'aide. »).

Enfin, l'énoncé peut comporter un verbe seul, sans SN (Vb), par exemple :

(3) [28]*OLA: *ta, mozies* [= *tak, mozesz*]. (= « Oui, (tu) peux. »)

La distribution des SN dans l'énoncé est illustrée dans la Figure 1.

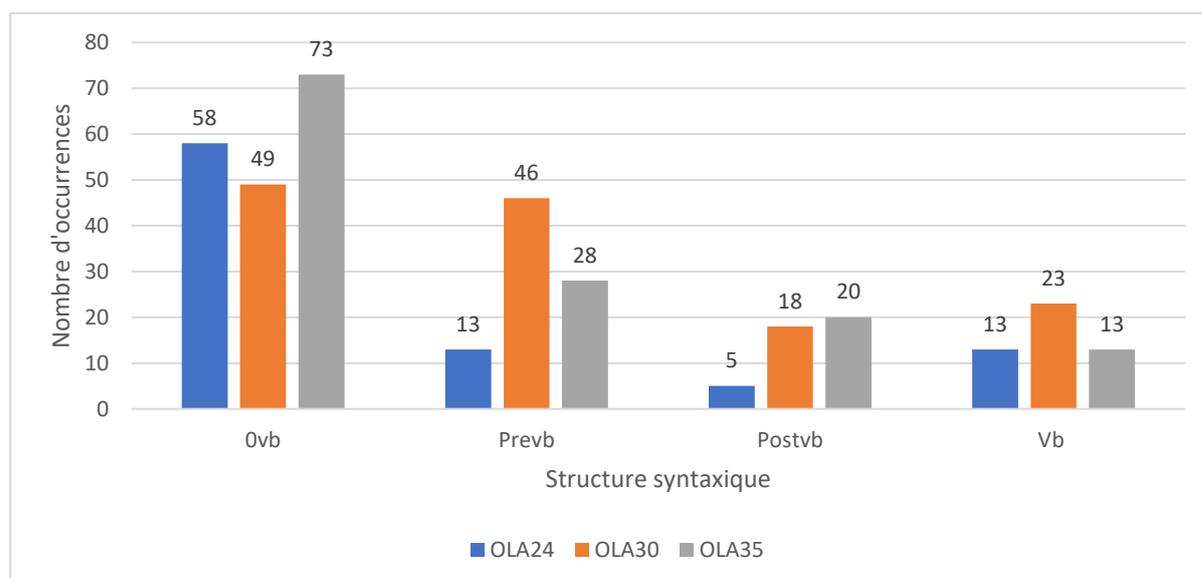


Figure 1. Structure syntaxique des énoncés : ordre des mots (ensemble du corpus)

En ce qui concerne la structure (interne) du SN, les SN sont le plus souvent des noms à 24 et 30 mois, et des pronoms à 35 mois (voir Figure 2). Par exemple, à 35 mois, l'enfant emploie des noms (ex. 4) mais également très souvent le pronom démonstratif impersonnel *to* (*ça*) (ex. 5)

(4) [28]*OLA: *a ja lubię motylki*. (= et moi (j)'aime (les) papillons)

(5) [37]*OLA: *a to?* » (= et ça ?)

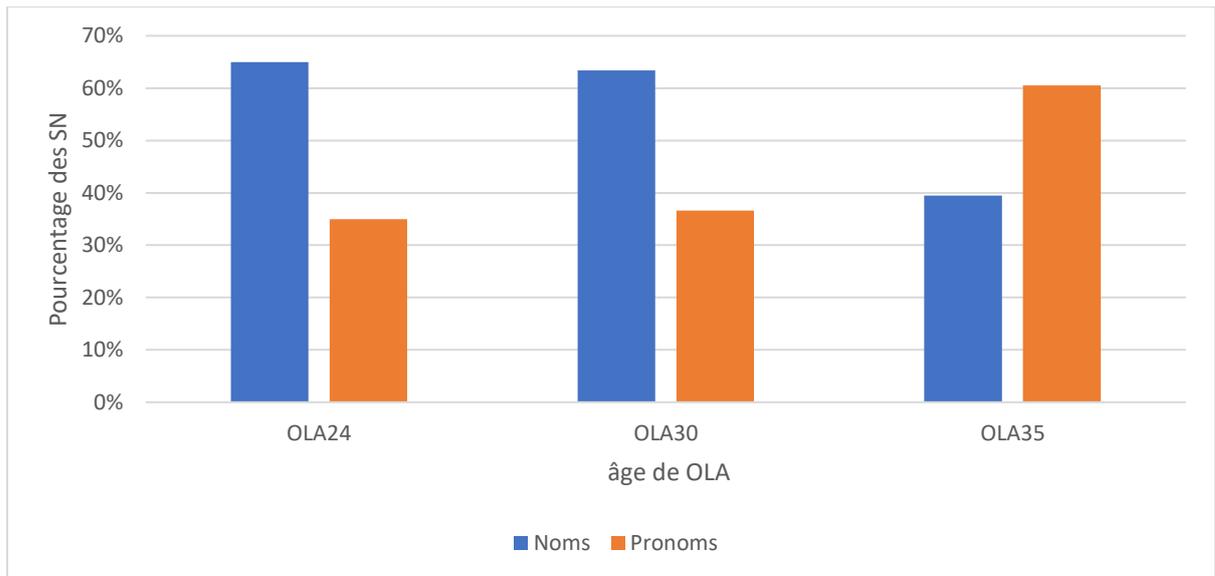


Figure 2. Répartition des formes nominales et pronominales (ensemble du corpus)

Les SN produits par Ola sont soit des noms – animés, inanimés ou propres, soit des pronoms – démonstratifs, indéfinis ou interrogatifs (Figure 3). A 24 mois, elle produit toutes les catégories, la plus fréquente étant celle des pronoms démonstratifs impersonnels (*a to = et ça ?*), suivie de celle des noms animés (*Ju dzie ? [= wujek gdzie] = oncle où ?*). Le nombre de sous-catégories de pronom diminue à 30 mois et les noms inanimés dominent. A 35 mois, toutes les catégories de nom sont encore présentes mais les noms propres sont les moindres. En revanche, la diversité des pronoms est réduite au profit des démonstratifs impersonnels (ex. 4).

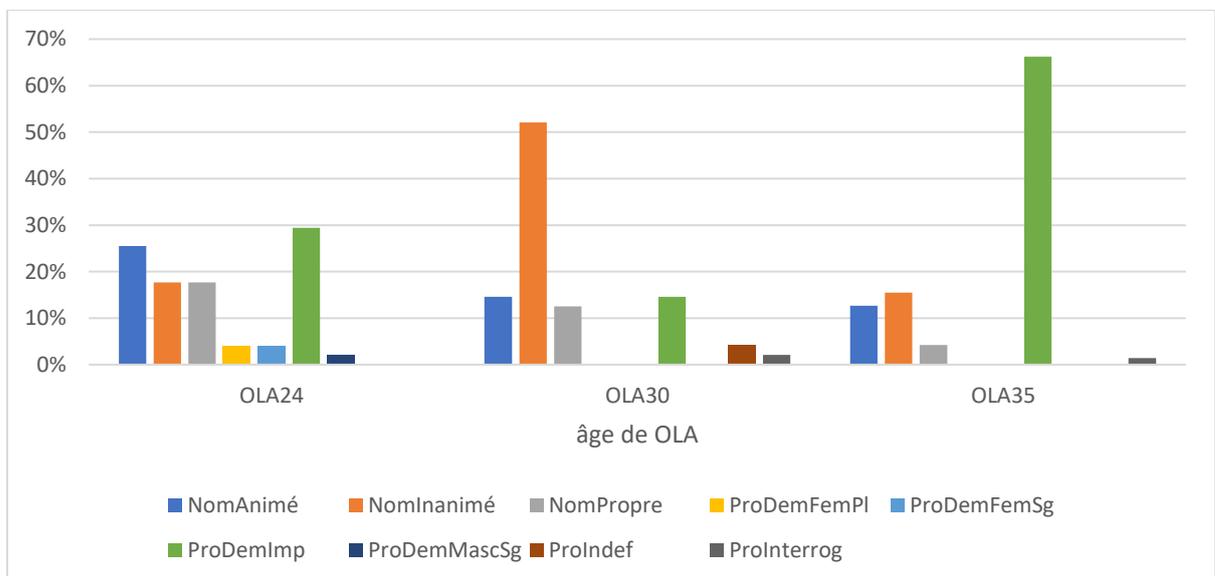


Figure 3. Structure interne du SN

Les SN sont le plus souvent des topiques à 24 et 35 mois, plus de 70 % dans les deux cas (Figure 4). A 30 mois, ils sont aussi souvent topiques que focus (environ 50%). Dans

l'exemple (5), les trois SN sont à l'accusatif et font partie du focus. En revanche, dans (6), le nom *pan* (=monsieur) au nominatif est topique :

(6) [100a]*OLA: ziuś, nogi i lence ma [= brzuch, nogi i ręce ma]. (= « ventre, jambes et mains (il) a »)

(7) [12]*OLA: a no bo pan jeśt, o tutaj i tat [= a no bo pan jest, o tutaj i tak]. (= et parce que monsieur est ici et oui »).

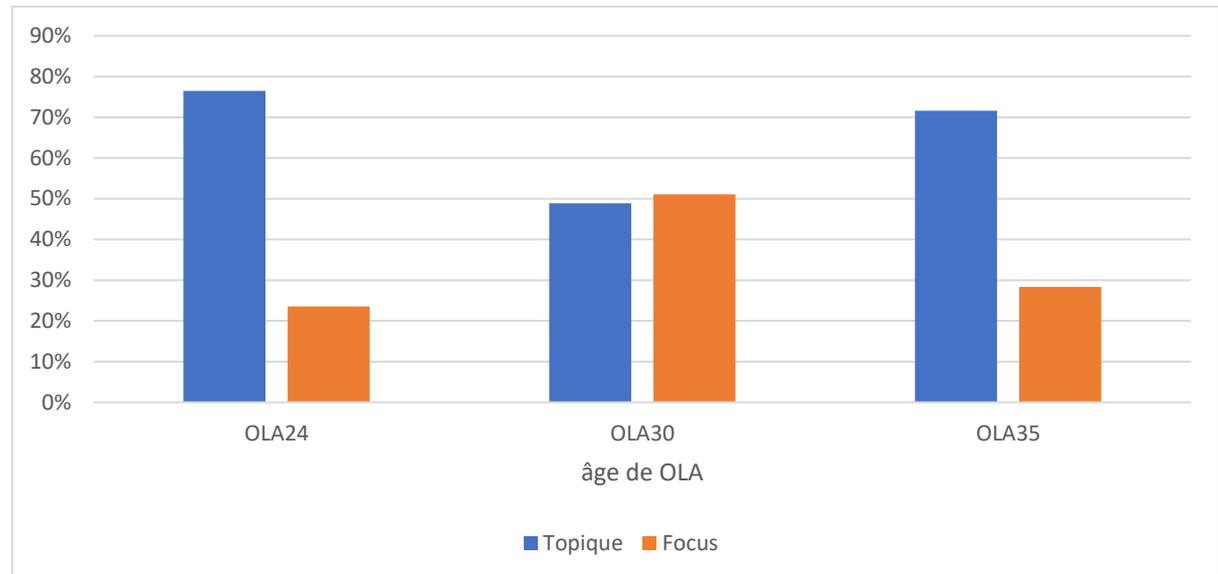


Figure 4. Statut informationnel des SN

Pour ce qui est de la morphologie casuelle, les noms produits par Ola sont en grande majorité au nominatif (*ręce* = *mains*) dès 24 mois et plus généralement aux trois âges étudiés. L'accusatif est présent également aux trois âges, mais dans une moindre mesure. La diversité des cas est la plus nette à 30 mois (Figure 5).

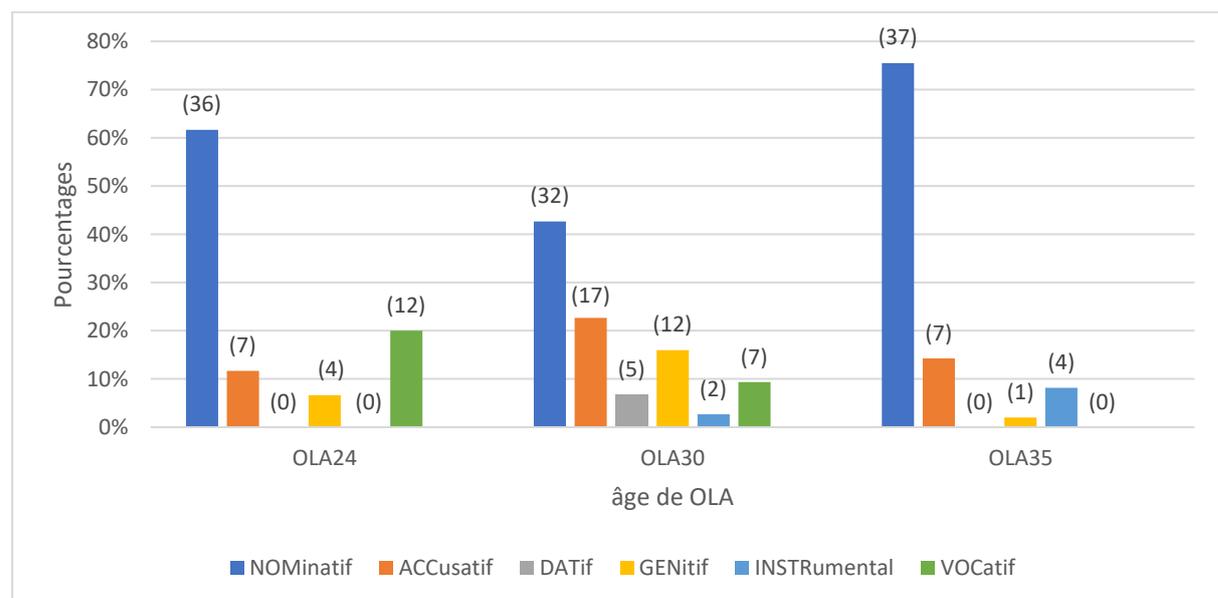


Figure 5. Structure des noms (en pourcentage des noms avec marquage casuel)

En ce qui concerne la corrélation *statut informationnel et marquage casuel* (Tableau 2), elle fait apparaître une forte présence de SN au nominatif en topique à 24 mois et une tendance inverse à partir de 30 mois. L'âge de 30 mois constitue également un âge charnière auquel l'enfant diversifie le marquage casuel qui est aussi lié à davantage d'interactions verbales avec l'interlocutrice, contrairement à un nombre important de dénominations simples (étiquettes) à 24 mois. Ces différentes formes de SN font plus souvent partie de l'information focale.

Tableau 2. Statut informationnel et marquage casuel (en nombre d'occurrences et en pourcentage des noms comportant un marquage casuel pour chaque catégorie de statut)

	Topique		Focus	
	occ.	%	occ.	%
24 mois				
NOMinatif	24	67%	12	33%
ACCusatif	4	57%	3	43%
GENitif	0	0%	4	100%
VOCatif	12	100%	0	0%
N=59 SN à 24 mois	40	67,8%	19	32,2%
30 mois				
NOMinatif	4	12,5%	28	87,5%
ACCusatif	0	0%	17	100%
DATif	4	80%	1	20%
GENitif	0	0%	12	100%
INSTRumental	0	0%	2	100%
VOCatif	7	100%	0	0%
N=75 SN à 30 mois	15	20%	60	80%
35 mois				
NOMinatif	8	22%	29	78%
ACCusatif	1	14%	6	86%
GENitif	1	100%	0	0%
INSTRumental	1	25%	3	75%
N=49 SN à 35 mois	11	22,4%	38	77,6%

6. Synthèse et conclusion

Compte tenu des caractéristiques typologiques du polonais, une langue ayant une morphologie flexionnelle complexe, la grammaticalisation du SN possède un contour spécifique, différent du français, auquel nous avons fait référence dans cet article. La référence l'acquisition du français est en quelque sorte une toile de fond qui permet de voir et mettre en exergue un autre parcours acquisitionnel déterminé par les propriétés typologiques du polonais, langue sans articles. Nous avons aussi voulu montrer le biais qui existe dans la littérature si l'on considère comme point de comparaison ce qui se passe dans l'acquisition d'une langue à articles. Or, les langues dites « à articles » ont été beaucoup étudiées mais non pas celles sans la catégorie grammaticale de l'article, qui sont majoritaires, mais qui n'ont pas bénéficié d'autant d'études. Nous n'avons pas observé de

formes singulières de SN, comparables aux *fillers*, ni de déterminants (deux occurrences seulement) mais le marquage casuel des noms correct, dès l'âge de 24 mois, avec une fréquence élevée du nominatif. L'âge de 30 mois semble constituer un âge charnière où l'enfant emploie un large éventail de formes casuelles des noms (sauf le locatif), tout comme en français où une diversification progressive des formes de déterminants et de leurs contextes d'emploi a été observée entre 20 et 30 mois. C'est à cet âge aussi que des énoncés produits par Ola se complexifient (LME, structure syntaxique, structure interne du SN) et les SN sont à quasi égalité des topiques et des focus (environ 50%). Avant cet âge, à 24 mois, l'enfant produit beaucoup d'étiquettes (0vb), qui sont des topiques. Après, à 35 mois, elle produit beaucoup de pronoms démonstratifs impersonnels (ProDemImp) qui sont aussi des topiques. On observe donc une période de fluctuation entre 30 et 35 mois dans la grammaticalisation du SN chez Ola.

La nouveauté de cette étude consiste à avoir comparé la complexification morphosyntaxique avec des facteurs discursifs. Contrairement à nos attentes, la majorité des noms, quel que soit leur cas et donc leur fonction syntaxique, font partie de l'information focale, et ce à partir de 30 mois où les échanges avec les interlocutrices sont fréquents. Pouvons-nous établir ici un parallèle entre la diversité des formes casuelles en focus en polonais et l'emploi privilégié de *fillers* et de déterminants en focus en français ?

Ces résultats ne peuvent évidemment être généralisés mais nous espérons qu'ils contribuent à rendre plus visible la spécificité du parcours acquisitionnel du polonais L1. Nous n'avons pu développer l'importance de l'ordre des mots en polonais. Si l'on prend en compte l'ordre des mots, les SN préverbaux, quel que soit leur cas, possèdent davantage le statut informationnel de topique et ceux postverbaux, quel que soit leur cas, sont davantage des focus. Cela reflète une des caractéristiques du polonais où l'ordre des mots est relativement libre et régi par des contraintes discursives.

La corrélation *statut informationnel – structure morphosyntaxique* dépend également d'autres facteurs comme les traits sémantiques du nom (nom animé/inanimé, genre), sa forme (nom ou pronom), l'accentuation, le type de phrase. Il s'ensuit que nous projetons d'affiner nos analyses, tout en élargissant le corpus à toutes les sessions filmées entre 21 et 35 mois. Nous souhaitons aussi nous pencher sur une autre dimension discursive, à savoir le mouvement référentiel et l'influence du contexte sur la production des SN. Cela nous permettra de croiser différents résultats pour voir les corrélations entre la structure interne du SN (forme nominale ou pronominale/avec ou sans déterminant) – sa fonction syntaxique (marquage casuel) – son statut informationnel (topique/focus) et sa fonction dans le discours (entité nouvelle, maintenue ou réintroduite).

Pour résumer, cette étude de l'acquisition du SN par une enfant polonaise nous a permis d'observer différentes dimensions linguistiques et discursives et leur graduelle mise en place dans l'acquisition du polonais L1. Néanmoins, nous sommes conscientes que seule l'analyse des mêmes dimensions chez d'autres enfants polonophones aux mêmes âges permettrait de généraliser ces résultats et questionner l'influence des différences inter- et intra-individuelles sur l'acquisition du polonais L1. Un projet de longue haleine...

Références

- Allen, S., Skarabela, B. & Hughes, M. (2008). Using corpora to examine discourse effects in syntax. In H. Behrens (ed.). *Corpora in Language Acquisition Research: History, methods, perspectives*. Amsterdam, John Benjamins, 99-137.

- Baker, N. D. & Greenfield, P. M. (1988). The development of new and old information in young children's early language. *Language Sciences* 10(1), 3-34.
- Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between the lexicon and grammar. *Journal of Child Language* 27(3), 521-559.
- Bassano, D. (2005). Développement du lexique et émergence de la grammaire. Hommage à Elisabeth Bates. *Le Langage et l'Homme* XXXX, 2, 7-22.
- Bassano, D. (2008). Acquisition du langage et grammaticalisation : le développement des noms et des verbes en français. In F. Labrell & G. Chasseigne (eds.), *Aspects du développement conceptuel et langagier*. Paris, Edition Publibook Université, 17-50.
- Bassano, D. (2010). L'acquisition du déterminant nominal en français : une construction progressive et interactive de la grammaire. *CogniTextes* 5 [En ligne], <http://journals.openedition.org/cognitextes/315>
- Bassano, D. & Maillloch, I. (1994) Early grammatical and prosodic marking of utterance modality in French: a longitudinal case study. *Journal of Child Language* 21, 649-675.
- Bassano, D., Maillloch, I. & Eme, E. (1998). Developmental changes and variability in early lexicon: A study of French children's naturalistic productions. *Journal of Child Language*, 25, 493-531.
- Bassano, D. & Van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Developmental Science* 10(5), 588-612.
- Bassano, D., Maillloch, I. & 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, 403-438.
- Bassano, D., Lenart, E., Trévisiol, P. & Maillloch, I. (2014). *Discourse factors and the emergence of grammar: Example from the acquisition of determiners in French*. Communication orale à la 8th Annual International Conference Languages & Linguistics, 6-9 July 2015, Athens, Greece
- Bassano, D., Lenart, E., Maillloch, I. & Trévisiol, P. (2013). *L'acquisition des déterminants nominaux en français L1 : quel est l'impact de la configuration informationnelle ?* Communication orale au XXVII^e Congrès international de Linguistique et Philologie Romanes, Nancy, 15-20 juillet. (Pré-Actes : p. 338)
- Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language*. Lawrence Erlbaum Associates Publishers, 29-79.
- Bates, E. & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (eds). *Language acquisition: The state of the art*. Cambridge. Cambridge University Press, 173-2018.
- Bates, E. & B. MacWhinney (1989). « Functionalism and the competition model ». In B. MacWhinney & E. Bates (éds), *The Crosslinguistic Study of Sentence Processing*. New York, Cambridge University Press, 3-76.
- Brown, R. (1973). *A first language: the early stages*. London, George Allen and Unwin.
- Bowerman, M. (1973). *Early Syntactic Development*. Cambridge, Cambridge University Press.
- Boulton, A., Canut, E., Guerin, E., Parisse, Ch. & Tyne, H. (2013). Corpus et appropriation de L1 et L2. *Lynx* (en ligne) 68-69, 9-32.
- Comrie, B. (1989). *Language universals and linguistic typology*. 2nd Edition. Chicago, University of Chicago Press.
- Dąbrowska, E. & Szczerbiński, M. (2006). Polish children's productivity with case marking: The role of regularity, type frequency, and phonological diversity. *Journal of Child Language* 33, 559-597.
- De Cat, C. (2004). A fresh look at how young children encode new referents. *International Review of Applied Linguistics* 42, 11-127.
- Demuth, K. (1996). The Prosodic Structure of Early Words. In J. Morgan and K. Demuth (eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates, 171-184.
- Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, 35(1), 99-127.
- Eisenbeiss, S. (2010). Production methods in language acquisition research. In E. Blom, & S., Unsworth (eds). *Experimental Methods in Language Acquisition Research*. Amsterdam, John Benjamins, 11-34.
- Gerken, L. (1996). Prosodic structure in young children's language production. *Language* 72(4), 683-712.
- Greenfield, P. & Smith, J. H. (1976). *The structure of communication in early language development*. New York, Academic Press.
- Goldenberg, E. R., Repetti, R. L., & Sandhofer, C. M. (2022). Contextual variation in language input to children: A naturalistic approach. *Developmental Psychology*, 58(6), 1051-1065.

- Kail & Bassano (2000/03). Méthodes d'investigation et perspectives heuristiques. In M. Kail & M. Fayol (eds). *L'acquisition du langage* Vol. 1 : *Le langage en émergence*. Presses Universitaires de France, 29-59.
- Karmiloff-Smith (1992). *Beyond modularity. A developmental perspective on cognitive science*. Cambridge MA, MIT Press.
- Klein, W. & (von) Stutterheim, C (1991). Text structure and referential movement. *Sprache und Pragmatik* 22, Lund, 1-32.
- Krajewski, G. (2008). *A constructivist investigation into the development of Polish noun inflections in children between two- and three-and-a half years of age*. Ph.D. Thesis, University of Manchester.
- Krajewski, G., Lieven, E. & Theakston, A. (2012). Productivity of a Polish child's inflectional noun morphology: A naturalistic study. *Morphology* 22(1), 9-34.
- Kupisch, T. (2006). The emergence of article forms and functions in the language acquisition of a German-Italian bilingual child. In C. Lleó (ed.). *Interfaces in Multilingualism. Acquisition and representation*. Amsterdam, John Benjamins, 139-177.
- Lleó, C. (2001). The interface of phonology and syntax: The emergence of the article in the early acquisition of Spanish and German. In J. Weissenborn, et B. Höhle (eds) *Approaches to bootstrapping*. Vol. II, Amsterdam, Philadelphia, John Benjamins, 23-44.
- MacWhinney, B. & Snow, C. (1985). The Child Language Data System (CHILDES). *Journal of Child Language* 12, 271-296.
- MacWhinney, B. (2000). *The CHILDES project: tools for analysing talk*. Hillsdale, NJ: Erlbaum.
- Marchman, V. & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language* 21, 339-366.
- Pine, J. M., & Lieven, E. (1997). Slot and frames patterns and the development of the determiner category. *Applied Psycholinguistics* 18(2), 123-138.
- Rozendaal, M. I. & 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, 773-807.
- 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, 88-112.
- Serratrice, L. (2000). The emergence of functional categories in bilingual language acquisition. Unpublished doctoral dissertation, University of Edinburgh, UK.
- Slobin, D. (1973). Cognitive prerequisites for the development of grammar. In A. Ferguson & D. Slobin (eds). *Studies of child language development*. New York, Holt, Rinehart & Winston, 175-208.
- Slobin, D. (1982). Universal and particular in the acquisition of language. In E. Wanner & L. R. Gleitman (eds). *Language acquisition: The state of the art*. Cambridge, Cambridge University Press, 128-173.
- Slobin, D. (1985). *The cross-linguistic study of language acquisition*. Hillsdale, N.J. Lawrence Erlbaum.
- Smoczyńska, M. (1985). The Acquisition of Polish. In D. Slobin (ed.). *The cross-linguistic study of language acquisition*. Hillsdale, N.J. Lawrence Erlbaum, 595-686.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3), 209-253.
- Tomasello, M. (2003). *Constructing a language: A Usage-Based Theory of Language Acquisition*. Cambridge, Massachusetts: Harvard University Press.
- Van der Velde, M. (2004). L'acquisition des articles définis en L1 : une étude comparative entre le français et le néerlandais. *Acquisition et Interaction en Langue Étrangère* 21, 9-46.
- Van Valin, R. D. (1991). Functionalist Linguistic Theory and Language Acquisition. *First Language* 11, 7-40.
- Van Valin, R. D. (2005). *Exploring the syntax- semantics interface*. Cambridge, Cambridge University Press.
- Veneziano, E., & Sinclair, H. (2000). The changing status of 'filler syllables' on the way to grammatical morphemes. *Journal of Child Language* 27(3), 461-500.
- Veneziano, E. (2003). The Emergence of Noun and Verb Categories in the Acquisition of French. *Psychology of Language and Communication* 7(1), 23-36.
- Veneziano, E. (2019). Les débuts de la grammaire chez le jeune enfant : Combinaison des mots et morphologie grammaticale. Dans S. Kern (Ed.), *Le développement du langage : théorie, clinique, pratique*. Paris, De Boeck Supérieur, 143-174.
- Weist, R., Wysocka, H., Witkowska-Stadnik, K., Buczkowska, E. & Konieczna, E. (1984). The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language* 11, 347-374.
- Weist, R. (1990). Neutralization and the Concept of Subject in Child Polish. *Linguistics* 11, 1332-1348.
- Weist, R. (2010). Properties of Acquisition: Child Polish Perspective. In B. Bogus, (ed.). *Studies in the Psychology of Language and Communication*. Warszawa, Matrix, 55-72.

Yamaguchi, N., Salazar Ovig, A., Le Mené, M., Caët ; S. & Rilland, A. (2021). Filler syllables as precursors of referring expressions. In A. Salazar Orvig, G. de Weck, R. Hassan & A. Rilland (eds). *The Acquisition of Referring Expressions: a dialogical approach*. Amsterdam, John Benjamins, 42-80.