

# La composition du lexique productif dans les trois premières années de vie : revue systématique

D. LABERTONNIÈRE, K. SKORUPPA

## RÉSUMÉ : La composition du lexique productif dans les trois premières années de vie : revue systématique

Le lexique productif précoce semble contenir plus de noms que d'autres classes de mots. Cependant, des questions demeurent autour du caractère universel de ce biais et de ses causes, et certaines classes de mots (adjectifs, mots sociaux...) sont peu étudiées. Ainsi, nous présentons une revue systématique de la littérature sur la répartition des classes de mots dans le lexique productif d'enfants de 1 à 3 ans et les facteurs d'influence possibles (type d'activité pendant la production, mode de recueil).

**Mots clés :** Biais pour les noms – Vocabulaire – Productif – Lexique précoce – Classe de mots.

## SUMMARY: The composition of the productive lexicon in the first three years of life: a systematic review

The early productive lexicon seems to contain more nouns than other word classes. However, questions remain with regards to the universality of this bias and to its causes, and other word classes (adjectives, social words...) are less frequently studied. We thus present a systematic review of the literature on the frequency of different word classes in the productive lexicon of children aged 1 to 3 and discuss the role of some factors (type of activity during production, data collection method).

**Key words:** Noun-bias – Vocabulary – Productive – Early lexicon – Part of speech.

## RESUMEN: La composición del léxico productivo en los tres primeros años de vida: revisión sistemática

El léxico productivo temprano parece contener más sustantivos que otras clases de palabras. Sin embargo, sigue habiendo dudas sobre la universalidad de este sesgo y sus causas, y otras clases de palabras (adjetivos, palabras sociales...) están poco estudiadas. Así, presentamos una revisión sistemática de la literatura sobre la proporción de las clases de palabras en el léxico productivo de los niños de 1 a 3 años y los posibles factores de influencia (tipo de actividad durante la producción, método de recogida de datos).

**Palabras clave:** Sesgo sustantivo – Vocabulario – Productivo – Léxico infantil – Clase de palabras.

Institut des Sciences Logopédiques,  
Université de Neuchâtel, Suisse.  
Adresse de correspondance :  
Institut des Sciences Logopédiques,  
Université de Neuchâtel,  
rue de la Pierre-à-Mazel, 7,  
Neuchâtel, Suisse.  
dahliane.labertoniere@unine.ch

.....  
Conflicts d'intérêts : les auteures déclarent n'avoir aucun conflit d'intérêts.

Pour citer cet article : Labertonnière, D., & Skoruppa, K. (2022). La composition du lexique productif dans les trois premières années de vie : revue systématique. A.N.A.E., 181, 000-000.

## Introduction

Les premiers mots qu'un enfant produit constituent un pas important dans l'acquisition du langage, et leur apparition tardive constitue un facteur de risque pour un trouble du langage (voir Desmarais *et al.*, 2008 pour une revue). Ces dernières décennies, de nombreux outils ont été développés afin de documenter les trajectoires de croissance lexicale chez les très jeunes enfants francophones, sous forme de questionnaires parentaux (IFDC, Kern, 2018 ; DLPP ; Bassano *et al.*, 2020) mais aussi de tests (par exemple, EVALO-BB, Coquet *et al.*, 2010). De plus, l'augmentation du lexique productif est une cible importante pour de nombreux programmes de soutien au langage précoce (voir Kern & Fekete, 2019 pour une revue internationale). Au-delà du niveau quantitatif, la composition du premier lexique retient également l'attention de la communauté scientifique.

Un biais pour les noms, c'est-à-dire une prépondérance de noms par rapport aux verbes (biais faible) ou par rapport aux autres catégories grammaticales (biais fort) dans le lexique productif des enfants, a été mis en évidence par de nombreuses chercheuses et chercheurs (Nelson, 1973 ; Gentner, 1982 par exemple). Une des questions fondamentales qui se sont posées suite à cette observation a été de savoir si le biais pour les noms était universel ou bien propre à certaines langues seulement, dont les caractéristiques typologiques favoriseraient l'apprentissage des noms (langues qui seraient donc « *noun-friendly* »).

En effet, une des hypothèses proposées pour expliquer l'existence de ce biais nominal défend l'aspect universel du biais en arguant les caractéristiques perceptives universelles des noms, comme le fait qu'ils désignent très souvent des objets, qui seraient plus faciles à conceptualiser que les actions (souvent désignées par des verbes). D'autres hypothèses semblent dépendantes de la langue. Certaines langues présenteraient une saillance perceptive (due à la position du mot dans la phrase) mettant en avant les noms (Nicoladis, 2001) alors que d'autres seraient plus « *verb-friendly* » (Choi & Gopnik, 1995), accordant par exemple une place de choix aux verbes dans l'*input* (langues SOV<sup>1</sup>, *pro-drop*<sup>2</sup>).

Face à ces deux types d'hypothèses, le premier objectif de cette revue est une synthèse des travaux récents relatifs à la composition nom

vs verbe du premier lexique. De plus, nous essaierons d'étendre notre analyse au-delà de ces deux classes de mots très largement étudiées en incluant des travaux sur des classes plus variées.

Concrètement, nous examinerons les données concernant le développement du lexique d'enfants d'environ 1 à 3 ans dans de nombreuses langues, avec un axe particulier autour des noms et des verbes, mais également autour d'autres catégories grammaticales et sémantiques, à savoir les adjectifs et les mots sociaux. Nous nous intéresserons de plus aux facteurs extralinguistiques (externes et internes à l'enfant) pouvant avoir une influence sur ce développement. Enfin, le lien avec d'autres domaines langagiers sera évoqué, particulièrement la morphosyntaxe et la sémantique, susceptibles de nous donner des pistes d'interprétation quant aux causes du biais pour les noms. Nous finirons par une discussion autour des limites de notre travail, et nous évoquerons les implications de nos résultats pour l'évaluation et l'intervention au niveau du lexique productif précoce.

## Méthodologie

« Une revue systématique essaie d'identifier, évaluer et synthétiser toutes les preuves empiriques remplissant des critères d'éligibilité préétablis afin de répondre à une question de recherche spécifique »<sup>3</sup>. Pour mener à bien cette revue, nous avons donc suivi des étapes prédéfinies dans un ordre précis. La première étape de recherche exhaustive de littérature, conduite en novembre 2020, s'est faite dans les quatre bases de données suivantes : *PsychInfo*, *Scopus*, *Pubmed* et *ASHA* par les deux autrices et six étudiantes du Master de Logopédie à l'Université de Neuchâtel. Ces bases de données ont été choisies sur des critères d'accessibilité et de pertinence pour le domaine de recherche et parce qu'elles présentaient des fonctionnalités de recherche avancées. Le processus de sélection a été mené en parallèle, de sorte que chaque élément a été jugé soit par les deux autrices, soit par la première autrice et deux étudiantes. En cas de désaccord inter-juge, des discussions ont eu lieu afin d'arriver à un consensus.

Afin de mener une recherche suffisamment spécifique, nous avons défini des mots clés précis utilisés dans toutes les bases de données. Nous avons ainsi restreint notre recherche à la littérature s'intéressant à la *classe des mots* (ou

1 ● Avec un ordre des mots de type sujet-objet-verbe, comme le japonais par exemple.

2 ● Omission possible du sujet, exprimé par la flexion verbale.

3 ● Traduction libre d'après <https://www.cochranelibrary.com/about/about-cochrane-reviews>

partie du discours ou nom ou verbe ou adjectif), au lexique (ou sémantique ou vocabulaire) en production chez les bambins ou enfants<sup>4</sup>.

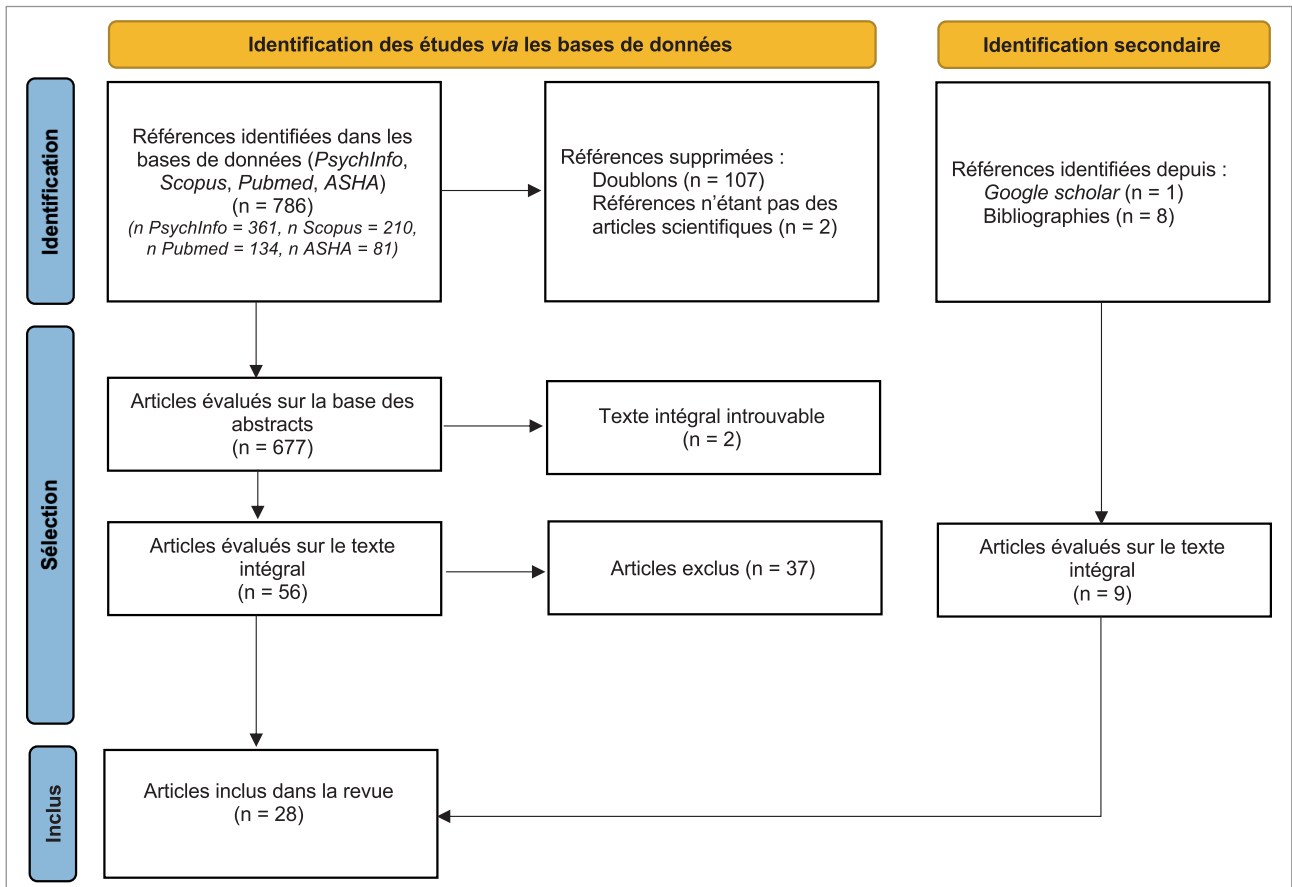
Pour un meilleur ciblage, nous avons également ajouté (pour les bases de données le permettant) des critères d'exclusion. Nous avons ainsi exclu les articles se concentrant sur une population trop jeune ou âgée (nourrissons et enfants d'âge scolaire) ou présentant des troubles développementaux et les articles étudiant le lexique uniquement en compréhension. Nous nous sommes concentrées sur les articles publiés après 1990 en anglais, français, espagnol, italien et japonais. Enfin les livres et les publications non revues par des pairs ont aussi été exclues.

Le diagramme de flux (figure 1) synthétise le nombre d'articles présents à chaque étape du processus. Pour la première recherche, 786 articles ont été recueillis. Après suppression des doublons et des publications n'étant pas des articles scientifiques, 677 articles ont été retenus. Pour chaque étape, les articles et leurs informations ont été reportées dans un tableau Excel.

4 • Mots clés originaux avec les opérateurs : (Infant\* OR Child\* OR toddler\*) AND (semantic\* OR vocabular\* OR lexic\*) AND (production OR productive) AND ("part of speech" OR "word class" OR noun\* OR verb\* OR adjectiv\*).

Une première sélection sur la base des titres et des résumés, en appliquant les critères d'exclusion, nous a permis de réduire ce nombre à 56 publications. Pour cette étape, nous avons notamment exclu les études ne traitant pas de la population d'intérêt ou n'effectuant pas de comparaison entre les catégories grammaticales ainsi que les études non expérimentales. Une recherche a été répliquée dans ces mêmes bases de données en juin 2022 par les deux autrices et nous a permis d'inclure un article récent en sus, ainsi que huit articles pertinents provenant de sources supplémentaires (Google Scholar et bibliographies des articles déjà retenus). Après ajout de neuf études supplémentaires, nous obtenons 65 articles. Finalement, une évaluation critique de tous ces articles a été effectuée, selon des critères de qualité basés sur les propositions faites par PRISMA (Moher et al., 2009) ainsi que les recommandations INESS (Martin et al., 2013). Un article a été jugé de *bonne qualité*, *qualité moyenne*, ou de *mauvaise qualité* en fonction de ces critères ; seuls les articles de bonne et de moyenne qualité ont pu passer à l'étape de sélection finale. En particulier, nous avons porté attention à la répliquabilité méthodologique, l'absence de biais expérimental ainsi qu'à la présence d'analyses statistiques, ces trois critères étant éliminatoires si non remplis.

Figure 1. Diagramme de flux PRISMA illustrant les étapes de la recherche de littérature et les résultats quantitatifs.



À la fin du processus, nous avons donc obtenu 28 articles (19 articles issus de la recherche dans les bases de données et neuf publications supplémentaires issues de la recherche en juin 2022) qui seront analysés en détail ci-après.

## Résultats

Un bref descriptif des articles sélectionnés est disponible dans le *tableau 1* (voir *annexe*), récapitulant leur population, leur méthodologie et leurs résultats. Une majorité, 22/28, traite de langues indo-européennes (10 pour l'anglais, 8 pour l'italien, 5 pour le français, 4 pour l'espagnol, 2 pour le néerlandais, 1 pour l'allemand, 1 pour le polonais). Les autres langues analysées appartiennent à des familles variées. Nous avons des langues sino-tibétaines (mandarin, 6 articles, cantonais, 1 article), le coréen (3 articles), et le japonais (1 article), mais aussi des langues plus rarement étudiées comme l'hébreu (2 articles), le turc (1 article), le wichi (1 article) et deux langues bantoues (kiswahili et kigiriama, 1 article). La majorité des participant·e·s étudié·e·s sont monolingues, mais deux articles s'intéressent à une population bilingue (cf. section « *Le cas du bilinguisme* »).

Au niveau méthodologique, nous avons choisi de distinguer les études utilisant des questionnaires parentaux de type MCDI (*MacArthur Communicative Development Inventory*, Fenson et al., 2000) et celles étudiant des pro-

ductions spontanées de l'enfant, bien que des variations parfois importantes existent au sein de ces deux grandes catégories. Le MCDI est un outil qui mesure le développement du vocabulaire des enfants selon l'âge sur la base d'un inventaire d'*items* à cocher par les parents si leur enfant les produit. Les *items* sont regroupés dans des catégories sémantiques. En calculant le total d'*items* cochés, on a ainsi accès à une bonne approximation du niveau langagier de l'enfant. Ainsi, quatorze articles font usage de questionnaires, neuf d'enregistrements et cinq utilisent les deux méthodes conjointement.

Enfin, la tranche d'âge combinée de toutes les publications se situe entre 8 mois et 3 ans et 11 mois (cf. *figure 2*).

### Variations translinguistiques dans l'acquisition des catégories grammaticales

On observe que l'ordre d'acquisition des catégories grammaticales et leur proportion dans le lexique précoce varient selon la langue. Nous nous intéresserons en premier lieu à la représentation des noms et des verbes dans le lexique des enfants apprenant les langues présentées ci-dessus, et en particulier au phénomène de biais nominal. En second lieu, nous examinerons les données concernant les autres catégories grammaticales.

Figure 2. Frise développementale du biais pour les noms dans les langues indo-européennes.

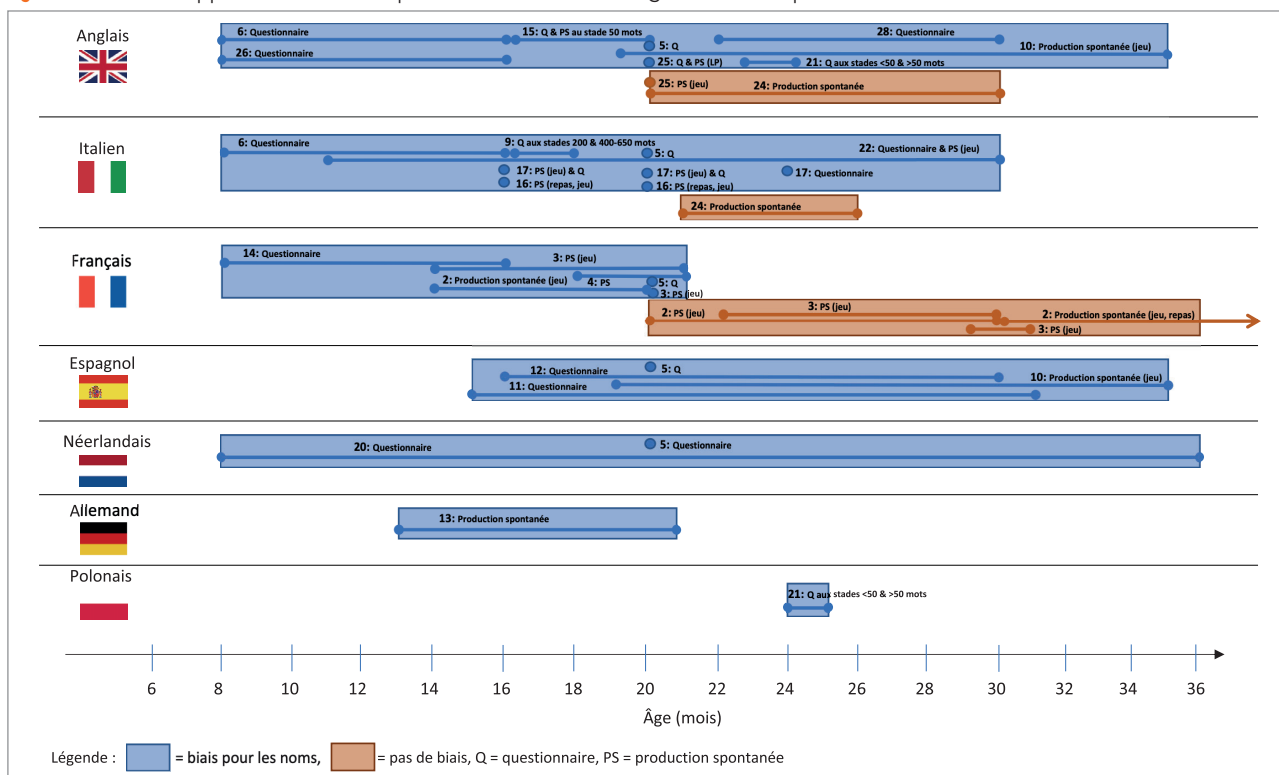
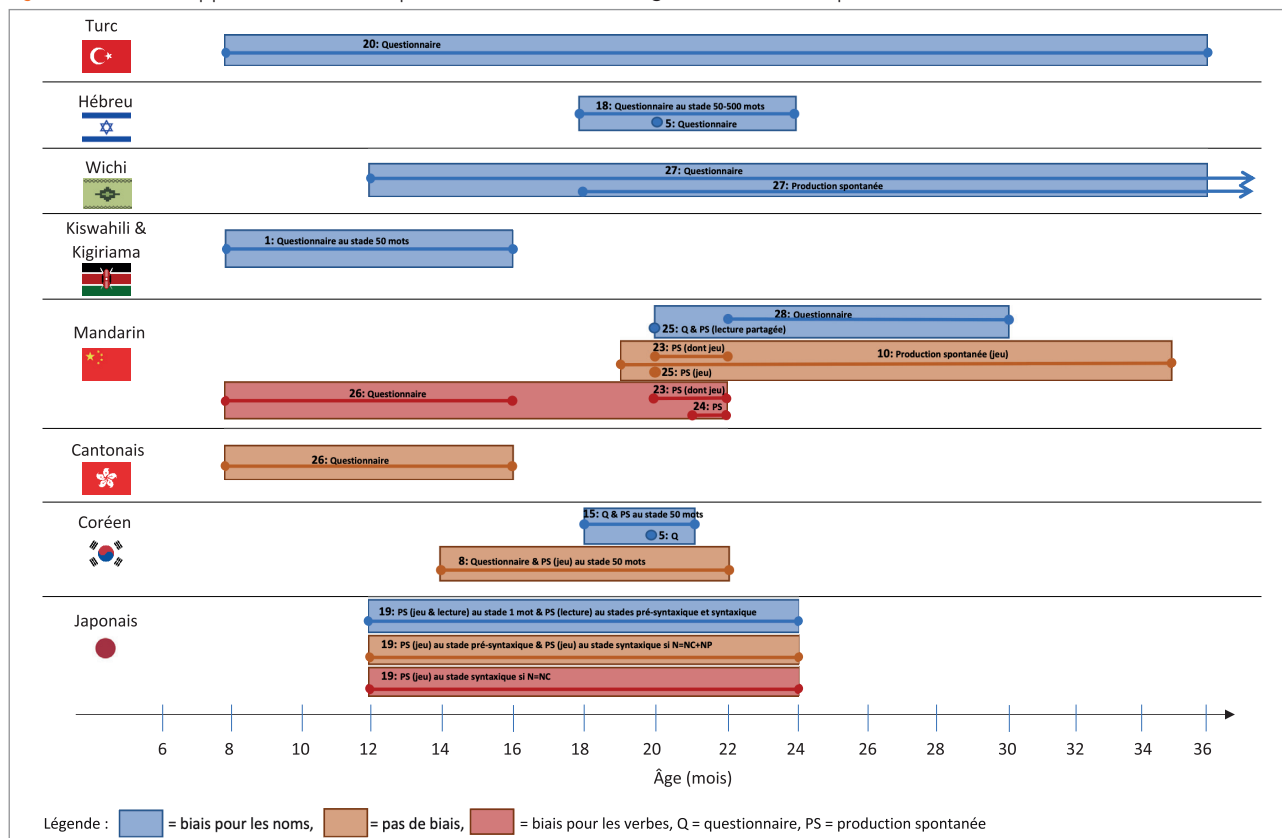


Figure 3. Frise développementale du biais pour les noms dans les langues non indo-européennes.



### Noms et verbes

Comme on peut le voir dans la frise présentée en figure 2, la grande majorité des articles analysés (24/28) mettent en évidence une forme de biais pour les noms à un moment du développement et dans diverses langues.

Certaines de ces études (6) présentent des résultats nuancés avec la présence d'un biais pour les noms uniquement à certaines périodes du développement, ou dans certaines situations. Ces résultats seront détaillés dans une seconde partie, avec ceux des trois articles qui ne trouvent pas de biais nominal ou bien à l'inverse un biais verbal. Enfin, nous finirons par examiner les articles effectuant une comparaison directe entre plusieurs langues.

### Biais pour les noms

Pour toutes les langues indo-européennes étudiées, la majorité des articles (21/22) font état d'un biais pour les noms au cours de la période de temps étudiée (8 mois à 3 ans). Il s'agit d'articles sur l'anglais (8/10), l'italien (6/7), le français (5/5), l'espagnol (4/4), le néerlandais (2/2), et des articles uniques sur l'allemand et le polonais. Entre la première et la deuxième année de vie, on retrouve un biais faible à fort, les données étant recueillies au moyen de questionnaires parentaux (Caselli *et al.*, 1995 ; D'odorico & Fasolo, 2007 ; Jackson-Maldonado *et al.*, 1993 ; Junyent *et al.*, 2020 ; Kauschke &

Hofmeister, 2002 ; Kern, 2007 ; Özcan *et al.*, 2016 ; Tardif *et al.*, 2008).

En français par exemple, entre 0;8 et 1;4 ans, le biais pour les noms est fort à tous les stades de l'acquisition du vocabulaire et s'exacerbe même après le stade 50 mots produits (Kern, 2007). L'article de Caselli *et al.* (1995) montre que la force du biais pour les noms peut dépendre de la définition que l'on donne à ceux-ci. En effet, pour des enfants anglophones et italo-phones, le biais nominal est faible avec une définition restreinte des noms mais devient fort si on élargit la classe des noms communs en incluant entre autres les noms propres et les onomatopées désignant des animaux ou objets. Chez des enfants italo-phones également (entre 1;4 - 1;6 ans), D'odorico et Fasolo (2007) rapportent un biais pour les noms à deux stades développementaux assez avancés (200 mots et 400 - 650 mots produits). Entre ces deux stades, on peut noter que le pourcentage de noms par rapport aux autres classes stagne, alors que celui des verbes augmente. Ainsi, de nombreuses études sur le français, l'anglais, l'allemand et l'italien rapportent qu'entre 1;3 et 2;0 ans, les verbes commencent en effet à se développer mais sont toujours en infériorité proportionnelle par rapport aux noms (Bassano *et al.*, 1998, 2005 ; Bassano, 2000 ; Bornstein *et al.*, 2004 ; Kauschke & Hofmeister, 2002 ; Kim *et al.*, 2000 ; Longobardi *et al.*, 2015, 2017 ;

Rescorla et al., 2017 ; Salerni et al., 2007 ; Tardif et al., 1999). Dans cette période de développement, les questionnaires parentaux sont toujours beaucoup utilisés mais on trouve également des articles se basant sur l'enregistrement de productions spontanées. Il peut s'agir de situations de lecture partagée (Tardif et al., 1999 pour l'anglais), de jeux (Bassano et al., 1998 ; Bassano, 2000 ; Longobardi et al., 2017 ; Salerni et al., 2007), nouveaux ou familiers (Longobardi et al., 2015), de repas (Longobardi et al., 2015).

Entre la deuxième et la troisième année, plusieurs articles rapportent toujours une dominance des noms sur les verbes (Dhillon, 2010 ; Jackson-Maldonado et al., 1993 ; Junyent et al., 2020 ; Özcan et al., 2016 ; Rescorla et al., 2017 ; Salerni et al., 2007 ; Xuan & Dollaghan, 2013), par exemple un biais fort entre 1;7 et 2;11 ans en anglais dans une situation de jeu (Dhillon, 2010) et entre 2;0 et 2;1 ans en polonais (Rescorla et al., 2017), quel que soit le stade lexical atteint (stock supérieur à 50 mots ou inférieur à 50 mots).

Pour les langues d'Asie de l'Est, les résultats sont plus nuancés. Deux études sur six trouvent un biais pour les noms en mandarin (Tardif et al., 1999 ; Xuan & Dollaghan, 2013) et deux sur trois en coréen (Bornstein et al., 2004 ; Kim et al., 2000). Le seul article s'intéressant aux enfants japonais, Ogura et al. (2006), détecte un biais pour les noms à 1;0 - 2;0 ans dans certaines situations uniquement.

En hébreu, turc, wichi, kiswahili et kigirima, toutes les études montrent un biais pour les noms ; la trajectoire développementale langagière paraît donc semblable à celle des enfants apprenant les langues indo-européennes. Pour l'hébreu, un consensus en faveur d'un biais fort pour les noms est obtenu à partir d'un recueil via des questionnaires parentaux dans les deux articles suivants : Bornstein et al. (2004) pour des enfants de 1;8 an et Maital et al. (2000) pour des enfants de 1;6 à 2;0 ans (langues lexical en production de 50 à 500 mots). En turc, Özcan et al. (2016) arrivent à la même conclusion qu'en néerlandais pour leur population bilingue de 0;8 à 3;0 ans : les noms sont plus présents que les verbes. En wichi, le biais pour les noms est également présent via le recueil par questionnaires ainsi qu'en production spontanée, comme le montrent Taverna et Waxman (2020) pour des enfants de 1;0 à 3;11 ans, mais celui-ci diminue en intensité avec l'âge. Enfin en kiswahili et kigirima, (langues bantoues), Alcock (2017) détecte un biais pour les noms au stade des 50 mots pour des enfants de 0;8 à 1;4 an.

### Absence de biais

Certains articles ne mettent pas en évidence de biais pour les noms dans le lexique productif. La plupart traitent de langues asiatiques, mais quelques langues indo-européennes montrent également une absence de biais dans certaines situations.

En mandarin, les études se sont penchées sur la production d'enfants de 1;8 à 2;11 ans, et il est particulièrement intéressant de noter que la situation dans laquelle le biais nominal est absent est une situation de jeu (semi-dirigée ou spontanée).

L'étude de Dhillon et collègues (2010) s'intéresse à trois groupes d'enfants répartis par tranches d'âge et révèle une tendance (non significative) chez les enfants les plus jeunes à produire plus de noms que de verbes, alors que pour le groupe le plus âgé il y aurait plus de verbes que de noms. Quant à l'article de Tardif (1996), l'auteur ne trouve pas de biais en faveur des noms ou des verbes quand elle prend une définition large des noms incluant les noms propres. La seconde langue sino-tibétaine de notre corpus, le cantonais, n'est malheureusement étudiée que dans un seul article (Tardif et al., 2008), et là non plus, aucun biais n'est trouvé, que ce soit pour les noms (définition stricte avec seulement des noms communs) ou les verbes. Pour le coréen et le japonais, il n'y a pas de biais décrit entre 1 et 2 ans en situation de jeu (Ogura et al., 2006) ni via l'analyse des questionnaires parentaux (Choi & Gopnik, 1995). Pour Ogura et al. (2006), cette absence de biais se retrouve chez des enfants au stade pré-syntaxique quelle que soit la définition des noms, mais seulement si on utilise une définition élargie (noms communs et noms propres) chez des enfants au stade syntaxique.

Dans les langues indo-européennes où les articles ne trouvent pas de biais pour les noms, la méthodologie de collecte des données est toujours l'enregistrement de productions naturelles. La population étudiée se trouve également dans la limite d'âge supérieure de notre revue, allant de 1;8 à 3;4 ans. Les langues concernées sont l'anglais (Tardif et al., 1997, 1999), l'italien (Tardif et al., 1997), le français (Bassano, 2000 ; Bassano et al., 1998, 2005) et l'allemand (Kauschke & Hofmeister, 2002). Pour prendre l'exemple du français, dans l'étude de cas de Bassano et al. (1998), les noms et les prédicats<sup>5</sup> sont en proportions à peu près égales dans le lexique à partir de 1;10 jusqu'à 2;6 ans en situation de jeu. Ces résultats sont confirmés dans la seconde étude du même

<sup>5</sup> Les prédicats sont utilisés pour décrire le rôle des verbes et des adjectifs (ce qu'on dit du sujet) ; dans les articles utilisant cette classe, il s'agit donc d'un regroupement des classes Verbes et Adjectifs.

article pour le groupe d'enfants le plus âgé (2;5 - 2;7 ans), pour l'enfant entre 1;8 et 2;6 ans de Bassano (2000), et pour les enfants entre 2;6 et 3;4 ans en situations naturelles de jeu et de repas dans l'article de Bassano et collègues (2005). En allemand, Kauschke et Hofmeister (2002) observent une tendance à l'équilibrage du lexique similaire à celui du français dans la troisième année.

#### Biais pour les verbes

Seuls quelques articles mettent en évidence un biais en faveur des verbes par rapport aux noms. Ceux-ci sont uniquement focalisés sur des langues aux caractéristiques dites « *verb-friendly* » d'Asie de l'Est. En mandarin, il s'agit de trois articles sur six. Au début du développement, entre 0;8 et 1;4 an, Tardif *et al.* (2008) rapportent que les verbes prédominent sur les noms communs dans des questionnaires parentaux. En utilisant une définition stricte des noms, les études de Tardif et collègues (Tardif, 1996 ; Tardif *et al.*, 1997) relèvent un biais verbal chez des enfants de respectivement 1;8 à 1;10 an et 1;9 à 1;10 an dans des situations interactives. En japonais, seuls les enfants au stade syntaxique<sup>6</sup> produisent plus de verbes dans une situation de jeu, en utilisant une définition stricte des noms (Ogura *et al.*, 2006).

#### Différences interlangues

Une partie des articles analysés ont étudié deux populations linguistiques ou plus, autorisant une comparaison directe entre elles. De plus, certains articles ont fait une comparaison entre leur population d'intérêt et celle d'un autre article usant d'une méthodologie similaire et s'intéressant à des enfants de même niveau développemental. Kim *et al.* (2000) et Choi et Gopnik (1995) trouvent par exemple plus de verbes dans le lexique productif coréen que dans celui de l'anglais à tous les stades de vocabulaire examinés. Inversement, une proportion plus élevée de noms est retrouvée dans le lexique des enfants anglophones que dans celui des enfants en voie d'acquisition du coréen. L'article de Kim *et al.* (2000) met aussi en évidence une explosion nominale chez tous les enfants anglophones, mais seulement chez la moitié des enfants coréens, et celle-ci se manifeste plus tardivement que pour les anglophones. Quatre articles effectuent quant à eux des comparaisons de langues indo-européennes avec le mandarin, et trouvent un avantage des noms pour ces premières. Ainsi, les études rapportent qu'il y a plus de noms dans le lexique des enfants anglophones de 1;8 à 2;11 ans (Dhillon, 2010 ; Tardif *et al.*, 1999 ; Xuan & Dollaghan, 2013), dans celui des italo-phones de 1;9 à 2;2 ans (Tardif *et al.*, 1997) et

des hispanophones de 1;7 à 2;11 ans (Dhillon, 2010).

En comparant son étude avec celle de Caselli *et al.* (1995), Alcock (2017) estime que la proportion de verbes dans le lexique des apprenant·e·s des langues bantoues kiswahili et kigiriama (qui ont des caractéristiques « *verb-friendly* ») augmenterait de façon plus rapide que chez les anglophones ou les italo-phones. L'autrice se base sur le fait que les enfants kenyans étudiés au stade lexical supérieur à 50 mots produisent plus de verbes que les enfants américains ou italiens au même stade de vocabulaire. Maital et collègues (2000), en faisant une comparaison de leur population hébréophone avec la population anglophone d'une autre étude (Fenson *et al.*, 1994), trouvent que le biais pour les noms est légèrement plus marqué chez les hébréophones, et ce malgré certaines propriétés « *verb-friendly* » de l'hébreu (saillance perceptuelle et richesse morphologique des verbes).

Enfin, dans des comparaisons directes entre langues indo-européennes, Rescorla et collègues (2017) trouvent une proportion égale de noms et de verbes en anglais et en polonais. Dhillon (2010) rapporte une proportion supérieure de noms en anglais par rapport à l'espagnol et Caselli et collègues (1999) ne rapportent pas de biais verbal en italien par rapport à l'anglais, malgré les caractéristiques « *verb-friendly* » de l'italien (*pro-drop*). À l'inverse, les autrices trouvent plus de verbes mais aussi de noms en anglais par rapport à l'italien. Ces résultats sont cependant opposés à ceux de Caselli *et al.* (1995) qui trouvent une proportion plus importante de noms (catégorie large) en italien qu'en anglais (et pas de différence pour les verbes).

Ainsi, les apprenant·e·s des langues indo-européennes et d'autres langues typologiquement éloignées (comme l'hébreu, le wichi...) semblent avoir un biais pour les noms plus ou moins marqué en fonction de la situation de collecte des données, les productions spontanées contenant plus de verbes que les questionnaires. En revanche, pour les langues d'Asie de l'Est, le biais est présent dans certaines situations, mais moins évident. On y retrouve parfois aussi un biais en faveur des verbes.

#### Autres catégories

##### Adjectifs

Parmi les catégories grammaticales ayant intéressé les auteur·rice·s des articles sélectionnés, celle des adjectifs est traitée par sept études (Bassano *et al.*, 1998, 2005 ; Jackson-Maldonado

6 ● Qui produisent donc des phrases et pas simplement des mots isolés.

et al., 1993 ; Junyent et al., 2020 ; Kern, 2007 ; Maital et al., 2000 ; Tardif et al., 1999). Trois articles (Bassano et al., 2005 ; Bornstein et al., 2004 ; Salerni et al., 2007) rapportent des différences significatives entre la proportion de verbes et la proportion d'adjectif présents : pour tous, la proportion de verbes dépasse celle des adjectifs. Il s'agit plus précisément dans ces études d'enfants francophones entre 1;6 et 3;4 ans ou âgés de 1;8 an aux stades entre 100 et 500 mots (Bassano et al., 2005 ; Bornstein et al., 2004), d'enfants italo-phones au stade 100 - 200 mots et 500 mots<sup>7</sup> (Bornstein et al., 2004 ; Salerni et al., 2007) et d'enfants coréanophones, anglophones, hispanophones, néerlandophones et hébreophones de 1;8 an au stade 100 - 500 mots (Bornstein et al., 2004<sup>8</sup>). Pour tous, il s'agit d'échantillons de langage spontané et l'on peut remarquer qu'avec une autre méthodologie (questionnaires parentaux), l'équipe de Salerni (2007) ne retrouve pas cette différence entre verbes et adjectifs au stade des 200 mots chez les italo-phones. Enfin, les adjectifs semblent très peu présents dans le lexique d'enfants italo-phones et anglo-phones de 0;8 - 1;4 an (Caselli et al., 1995). Malgré cela, la classe des adjectifs augmente toujours avec la taille du vocabulaire.

#### Mots fonctionnels

La classe des mots fonctionnels<sup>9</sup> est traitée par huit articles. Ceux-ci semblent rares au début du développement en français, anglais, italien et espagnol (Bassano, 1998 ; Bassano et al., 2005 ; Bornstein et al., 2004 ; Caselli et al., 1995 ; D'odorico & Fasolo, 2007 ; Junyent et al., 2020). Ils sont tout de même en proportion supérieure aux verbes, derrière les noms, en kiswahili et kigiriama entre 0;8 et 1;4 an (Alcock, 2017). Plus précisément, ils restent inférieurs aux noms et prédicats en espagnol (du moins jusqu'à 2;6 ans cf. Junyent et al., 2020). En français, leur proportion augmente conjointement avec celle des prédicats pour arriver à un équilibre des classes étudiées (noms, prédicats, mots fonctionnels et para-lexicaux<sup>10</sup>) puis ils deviennent la classe dominante après 2 ans et jusqu'à 3 ans environ (Bassano et al., 1998). Les autrices discutent de ce phénomène d'« explosion grammaticale » qui serait plus important en français qu'en anglais et serait dû à une richesse spécifique du français dans ce domaine. Quant à l'italien, Salerni et al. (2007) proposent un *pattern* différent et complexe : dans les questionnaires parentaux, les mots

fonctionnels sont soit en proportion égale aux adjectifs et verbes (stade 200 mots produits), soit en proportion égale aux seuls adjectifs (stade 500 mots, les verbes ayant pris une place plus importante). Cependant, en production spontanée, on a plus de types de mots fonctionnels que d'adjectifs (stade 500 mots) ou de verbes (stade 200 mots).

#### Mots sociaux

Dans l'analyse du premier lexique infantin, les noms communs ne sont pas toujours la classe dominante. Les catégories des noms de personnes et des onomatopées mais aussi les routines<sup>11</sup> comprennent une proportion plus élevée d'items cochés dans les questionnaires parentaux (Tardif et al., 2008 ; Caselli et al., 1995). C'est le cas jusqu'à 50 mots produits (Kern, 2007) où les trois catégories sémantiques les plus représentées sont les noms de personnes, les jeux et routines et les onomatopées. Certains articles mettent également en avant une probabilité de produire un mot de type « nom de personne » très haute pour les trois langues étudiées au stade 7 - 10 mots et les noms de personnes sont toujours très présents vers la fin de la deuxième année (Tardif, 1996 ; Tardif et al., 2008). À la fin de la première année, les classes les plus importantes dans le lexique semblent être les noms (incluant les onomatopées) et les routines (Alcock, 2017).

En plus des noms de personnes et des onomatopées, d'autres mots plus difficiles à catégoriser semblent prendre une place non négligeable dans le premier lexique. La catégorie des « jeux et routines » est très présente, avec une proportion égale à celle des noms pour les enfants anglo-phones (Tardif et al., 2008). On retrouve également la catégorie regroupant les « para-lexicaux », contenant les mots vides, les routines, les interjections et les particules affirmative et négative (« oui » et « non ») (Bassano, 1998 ; Bassano et al., 2005). Cette catégorie est majoritaire au début du développement (vers 1 an) en français puis sa proportion dans le lexique diminue drastiquement.

Ces diverses catégories sémantiques pourraient être qualifiées de mots sociaux, c'est-à-dire des mots produits dans des contextes particuliers d'interactions sociales, ou servant à désigner des êtres sociaux (par exemple « maman », « coucou » ...) et analysée en tant que telle (Caselli et al., 1999). Cette catégorie comprend spécifiquement les onomatopées, noms de personnes et jeux et routines et elle représente une plus grande partie du lexique d'enfants italo-phones par rapport à des anglo-

7 ● Productions spontanées.

8 ● Les enfants coréens présentant ce profil de production dès le stade 50 mots.

9 ● Définition variable : sont inclus généralement les pronoms, mots interrogatifs, prépositions, articles, déterminants et conjonctions au moins.

10 ● Mots vides, interjections, particules affirmative et négative.

11 ● Catégorie du MCDI, par exemple « faire les courses ».

phones. Les résultats de Kauschke & Hofmeister (2002) montrent que cette catégorie est prépondérante dans le premier lexique germanophone (entre 1,1 et 1,3 an) lorsqu'on y ajoute les assertions (« oui ») et interjections (« oh ! »).

Les mots sociaux semblent donc universellement occuper une place importante dans le tout premier lexique, alors que les adjectifs se développent plus tardivement. Quant aux mots fonctionnels, leur acquisition serait peut-être plus dépendante de la langue parlée.

### Facteurs extralinguistiques

En plus de la langue en cours d'acquisition, d'autres facteurs externes ou internes à l'enfant peuvent influencer sur le développement des catégories grammaticales au cours de l'acquisition. Nous rapportons ici de nombreux facteurs d'influence de type interactionnels et sociaux (*input*, pragmatique, situation), et nous essaierons également de traiter le sujet de façon qualitative.

#### Facteurs externes

##### • L'*input*

En premier lieu, l'un des facteurs les plus étudiés dans les articles sélectionnés est l'influence de l'*input* parental (7 études). Intimement liée aux caractéristiques linguistiques de la langue donnée, cette influence est difficilement analysable séparément de celle des caractéristiques typologiques de la langue en cours d'acquisition. Dès lors, la fréquence des *items* lexicaux, la saillance perceptive, la simplicité ou la saillance morphologique semblent jouer un rôle dans l'acquisition différentielle des classes de mots.

Dans Ogura (2006), les parents japonais produisent plus de verbes que de noms lors d'une activité de jeu. Les verbes sont effectivement dans des positions saillantes (en fin de phrase) plus souvent que les noms en japonais. Cependant, le groupe d'enfants au stade d'acquisition le moins avancé montre tout de même un biais pour les noms. C'est seulement à partir des stades pré-syntaxique et syntaxique que l'*input* semble influencer sur la fréquence d'utilisation des noms et verbes par les enfants, le biais pour les noms s'atténuant, voire devenant inexistant. Dans Longobardi et al. (2015), l'*input* des parents italiens montre une prédominance des verbes sur les noms, ces derniers se retrouvant pourtant plus souvent en position saillante en fin d'énoncé. L'effet de saillance semble donc positif pour l'acquisition des noms mais négatif pour la production de verbes. Par contre, la position initiale des verbes a un effet positif sur leur acquisition, ce qui pourrait être un effet de saillance retrouvé

également en position initiale. Dans les études de Kim et al. (2000) et de Choi et Gopnik (1995), l'*input* des parents anglophones semble présenter un biais nominal alors que celui des parents de langue coréenne semble présenter un biais verbal. Toujours dans Kim et al. (2000), la même tendance est observée au niveau de la saillance (position finale mais aussi flexion prédictible des verbes en coréen). Cependant, un biais pour les noms est observé dans les deux groupes d'enfants.

Dans Dhillon (2010), les productions parentales ne permettent pas de conclure à un biais pour les noms ou les verbes en anglais, espagnol et mandarin ; cependant, le pourcentage de noms produits en anglais est supérieur à celui des hispanophones qui est lui-même supérieur à celui du mandarin (ce qui se retrouve chez les enfants en production).

Dans Tardif et al. (1997), les verbes sont plus saillants dans l'*input* des parents mandarins, alors que ce sont les noms les plus saillants dans celui des anglophones, et les italophones se trouvent entre les deux. En wichi, d'après Taverna et Waxman (2020), les verbes dominent dans l'*input* parental alors qu'un biais pour les noms est présent chez les enfants ; cependant, celui-ci s'atténue au cours du développement.

Certains paramètres interactionnels peuvent en effet influencer sur la production de tel ou tel type de classe de mots. Tardif (1996) nous indique ainsi qu'en mandarin, l'enfant produit plus de noms si le parent parle d'objets ; l'inverse est également vérifié avec les routines sociales et la production de verbes.

##### • Les facteurs socio-culturels et pragmatiques

Les influences de l'aspect socio-culturel et de l'aspect pragmatique sont également étudiées dans plusieurs articles.

Par exemple, les deux langues bantoues, le kiswahili et le kigiriyama, sont parlées en majorité dans des zones rurales du Kenya, où les parents, occupés à d'autres tâches, s'engagent rarement dans des dénominations actives d'objets auprès de leurs enfants (Alcock, 2017). Aussi, dans le contexte culturel où sont parlées ces langues, les attentes sociales incluent un degré élevé d'obéissance de la part des enfants, ce qui influe sur l'usage du langage en contexte (pragmatique). L'*input* des enfants contient ainsi de nombreux ordres, ce qui pourrait avoir une influence positive sur leur compréhension des verbes et éventuellement influencer sur leur production. On pourrait donc penser que les enfants aient des facilités au niveau des verbes et des compétences moindres au niveau des noms par rapport à d'autres langues. Pourtant,

dans les premiers mots produits dans ces deux langues, on retrouve un biais pour les noms.

Tardif et collègues (2008) relèvent quant à elles des différences culturelles entre la Chine et les États-Unis d'Amérique, rappelant que la culture chinoise a des attentes fortes au niveau des formes d'adressage aux personnes par rapport à la culture américaine. Les enfants chinois auraient donc besoin plus tôt d'un grand stock de noms. Les résultats de cette étude montrent que les enfants chinois apprenant le cantonais ou le mandarin produisent en effet plus de noms de personnes que les enfants américains anglophones. Les autrices proposent également que les différences observées au niveau de la classe des noms communs seraient au moins en partie liées à des différences culturelles, citant par exemple l'importance du jeu de dénomination pour les anglophones.

#### • La situation de collecte des données

Comme indiqué au début de cette section, cinq articles ont utilisé conjointement des questionnaires et des enregistrements de production spontanée et ont donc eu l'opportunité de comparer ces deux méthodes. Malheureusement, seuls deux articles nous offrent une vraie comparaison.

Dans l'étude de Salerni et collègues (2007), les noms sont majoritaires par rapport aux autres catégories grammaticales dans les deux situations. Cependant, des différences dans les proportions de noms et de verbes existent : le biais pour les noms est plus marqué dans le questionnaire et il y a également une plus grande proportion d'adjectifs avec cette méthode. La production spontanée montre quant à elle une plus grande proportion de verbes et de mots fonctionnels. Il est intéressant de noter que pour les adjectifs et les verbes, ces constats sont vrais seulement pour le premier stade de développement étudié (200 mots produits), la différence devenant non significative pour une taille supérieure de vocabulaire (500 mots).

Dans l'étude de Tardif et al. (1999), la proportion de noms obtenus via les questionnaires parentaux semble plus haute que celle obtenue grâce aux mesures observationnelles, mais la différence n'est pas statistiquement significative. Des analyses complémentaires sur les occurrences révèlent que les noms cochés par les parents apparaissent plus souvent en production spontanée que les verbes (en anglais et en mandarin) et les verbes produits spontanément par les enfants sont eux moins cochés que les noms (en anglais).

#### • Le contexte de production

Dans Ogura et al. (2006), deux types d'activités sont utilisés pour éliciter une production

spontanée de la part de l'enfant : une lecture partagée et un jeu. Les résultats comparés indiquent que les enfants produisent plus de noms dans la situation de lecture partagée (et c'est d'ailleurs aussi le cas pour les mères). On peut également noter que l'effet du contexte de production sur la proportion des catégories grammaticales surpasse, dans cette étude, l'effet du stade développemental.

La méthodologie de Tardif et al. (1999) est similaire avec en plus une différenciation entre deux types de jeux (avec des jouets motorisés ou non). La lecture partagée est ici aussi plus favorable à la proportion de noms<sup>12</sup>.

Dans l'étude de Salerni et al. (2007), on retrouve le même type de résultats avec une proportion plus élevée de verbes dans les productions spontanées que dans les questionnaires parentaux (le type d'activité utilisée est un jeu spontané en dyade).

L'avantage pour les noms pendant la lecture pourrait être lié à la plus grande « imageabilité » (« *imageability* », facilité de conceptualisation) des objets par rapport aux actions ou au fait que le parent se mettrait en situation de tutrice en dénommant plus d'objets que d'habitude (Ogura et al., 2006).

En résumé, les facteurs méthodologiques que sont la situation de collecte des données et le contexte de production ont donc une influence certaine sur l'obtention d'un biais, tandis que l'influence de l'*input* et des facteurs socio-culturels et pragmatiques semble une piste encore à approfondir.

#### Facteurs internes

##### • Les facteurs cognitifs

Finalement, des aspects cognitifs relatifs à des heuristiques d'apprentissage langagier peuvent entrer en jeu dans la composition du lexique précoce. Certains enfants se focaliseraient sur les fonctions référentielles du langage, donc les objets et les noms, et d'autres seraient plus portés vers la partie sociale et instrumentale du langage, donc les routines et un lexique varié (style référentiel vs expressif, Bassano, 2005). Plus tard, la diversité disparaîtrait avec un seul style dominant : le style grammatical (avec beaucoup de mots fonctionnels), lié aux particularités structurelles du français. Cela va dans le sens de l'hypothèse selon laquelle la variabilité dans le lexique productif est plus importante au début du développement.

<sup>12</sup> • Dans cette étude, les parents produisent plus de noms dans la situation de lecture, et produisent également plus de verbes dans les situations de jeu (les enfants dans ce contexte produisent environ la même proportion de noms et de verbes).

Enfin, comme évoqué dans Longobardi et al. (2015), des facteurs neurodéveloppementaux pourraient être liés au biais pour les noms. Des régions cérébrales différentes semblent être activées lors de la production de noms ou de verbes (Shapiro et al., 2006) et pourraient avoir une trajectoire développementale différenciée.

### Le cas du bilinguisme

Seuls deux articles sélectionnés étudient une population plurilingue. Il s'agit d'enfants bilingues turc-néerlandais (Özcan et al., 2016) et bilingues anglais-mandarin (Xuan & Dollaghan, 2013). Dans les deux cas, bien que les deux langues parlées par l'enfant soient relativement éloignées typologiquement (c'est surtout le cas de la deuxième étude), un biais pour les noms est observé dans les deux langues. Les données sont obtenues via des questionnaires parentaux et le biais est présent quelle que soit la méthode de catégorisation des noms utilisée. Dans les deux cas également, ce biais est plus marqué dans une des deux langues (néerlandais pour la première et anglais pour la seconde), ce qui semble lié aux propriétés des langues (cf. section « *Variations translinguistiques dans l'acquisition des catégories grammaticales* »).

L'étude de Xuan et Dollaghan est donc particulièrement intéressante car elle compare directement chez les mêmes sujets deux langues présentant des proportions différentes de biais catégoriels lexicaux.

On peut regretter l'absence d'autres études comparant une population bilingue à deux populations monolingues apprenant chacune des langues de la population bilingue. De telles comparaisons seraient utiles pour déterminer l'éventuelle existence d'une influence interlangue chez les bilingues.

### Lien entre le développement des catégories grammaticales et les autres domaines langagiers

Choi et Gopnik (1995) se sont intéressées aux liens entre l'acquisition des catégories grammaticales et la morphologie. Dans leur étude, les enfants coréens semblent acquérir les verbes et les noms en tant que classes morphologiques distinctes dès 1,3 an. En effet, certain·e·s de leurs participant·e·s produisent déjà des flexions verbales correctes et d'autres utilisent plusieurs morphèmes négatifs pour un même verbe. Concernant les noms, ces enfants ne sont pas encore au stade de morphologie productive (il n'y a pas encore de marquage casuel ou le pluriel est non utilisé). Un seul suf-

fixe est utilisé, et c'est un possessif spécifique aux êtres humains ; la nature sociale du mot semble donc jouer un rôle. Les enfants anglophones quant à eux ne montrent pas une telle tendance productive au niveau de la flexion verbale, certainement parce qu'elles ou ils n'utilisent pas ou très peu de verbes à cet âge.

Au niveau sémantique, en plus des mots sociaux, la dichotomie objet/action représente une autre facette du contraste noms vs verbes. En effet, la cause conceptuelle majeure du biais pour les noms est le fait que les objets (désignés majoritairement par des noms) sont plus faciles à conceptualiser que les actions (souvent désignées par des verbes).

En français, les noms animés sont une sous-classe de noms dominante, et dans la classe des prédicats, ce sont les verbes d'action qui prédominent (Bassano et al., 2005).

Pour Kern (2007), au niveau sémantique, les mots d'action sont peu présents au stade 1 - 50 mots mais augmentent petit à petit au cours du développement. Pour des enfants apprenant le mandarin, Tardif (1996), en concordance avec ses résultats sur les noms et verbes, trouve une proportion plus élevée de mots d'actions que de mots d'objets (définition stricte) ou une proportion égale (noms de personnes comptés comme objets). Par ailleurs, Choi et Gopnik (1995) analysent les 10 premiers mots dans le lexique coréen au niveau du sens : il s'agit principalement de noms désignant des objets concrets et des personnes, et de verbes rentrant en majorité dans les catégories des actions et activités. En comparant avec l'anglais, les autrices trouvent que les enfants coréens utilisent plus de mots d'action et moins de mots d'objets, des résultats similaires à une comparaison noms vs verbes. Ainsi, l'acquisition morphosyntaxique semble être corrélée à celle des classes de mots tandis que la sémantique semble directement influencer leur acquisition.

### Discussion générale

Au vu de l'ensemble des travaux considérés, il semblerait que l'on ne puisse pas parler de biais pour les noms au sens strict dans toutes les langues, car les noms ne sont pas toujours la classe dominante dans le lexique des enfants, que ce soit au début ou à la fin de la période de temps étudiée.

En effet, les catégories dominantes dans le tout premier lexique semblent être celles des onomatopées et des mots sociaux (Caselli et al., 1995 ; Kauschke & Hofmeister, 2002 ; Kern, 2007 ; Tardif et al., 2008), des catégories diffi-

ciles à classer dans les classes grammaticales « noms » et « verbes »<sup>13</sup>. De ce fait, on peut noter que la présence ou non de cette catégorie de mots dans les analyses du premier lexique influe sur la mesure du biais nominal. Quant à la fin de la période développementale étudiée (3 ans environ), il semble que les classes de mots arrivent à un certain équilibre (Bassano, 2000), les noms perdant petit à petit leur dominance. Entre autres, les adjectifs et les mots fonctionnels, qui sont en général peu présents dans le premier lexique (on en trouve moins que de verbes) se développent au fil du temps. Quelques résultats vont même dans le sens d'une absence de biais pour les noms (ou même d'un biais pour les verbes) pour certaines langues « *verb-friendly* » avec certaines méthodologies et à certaines périodes du développement.

Cependant, quand on s'intéresse uniquement aux classes des noms et des verbes, la première semble dominer sur la seconde au début du développement dans presque toutes les langues étudiées. Le mandarin et le cantonais se démarquent sur ce point (Tardif et al., 2008), mais il faut nuancer en considérant que les noms de personnes ne sont pas comptés par les autrices dans la classe des noms, alors que les enfants locuteur-riche-s de ces deux langues semblent justement en produire beaucoup plus que les anglophones.

Il est intéressant de noter que dans les études interlangues qui ont comparé directement le coréen ou le mandarin (dont le biais catégoriel lexical est plus contrasté) à des langues à biais plus systématique, les auteur-riche-s rapportent une proportion plus faible de noms et inversement une proportion plus importante de verbes dans ces deux langues (Kim et al., 2000). Ces études corroborent donc le statut orienté vers les verbes de ces langues non indo-européennes.

Si on peut émettre l'hypothèse selon laquelle le biais pour les noms aurait des causes universelles, celles-ci semblent jouer un rôle surtout au début de l'apprentissage, dans toutes les langues. Ces causes pourraient s'ancrer dans des caractéristiques sémantiques des noms/verbes, liés aux objets/actions comme discuté dans la section « *Lien entre le développement des catégories grammaticales et les autres domaines langagiers* ».

Un autre argument en faveur de l'universalité du biais pour les noms est le fait que les enfants

plurilingues semblent suivre une trajectoire de développement des différentes classes de mots semblable à celle de leurs pairs monolingues.

Ces causes universelles seraient contrebalancées par certains facteurs internes ou externes à l'enfant. Outre les caractéristiques structurales de la langue en cours d'acquisition (ordre canonique des mots...) et les caractéristiques culturelles de la langue en contexte, la situation de production influe sur l'*input*. On remarque notamment dans plusieurs études que dans les cas de jeux, le biais pour les noms est moins marqué (ou remplacé par un biais pour les verbes) par rapport aux situations de lecture partagée (Ogura et al., 2006).

Ainsi, la combinaison de facteurs divers, concernant leur langue, le contexte dans lequel l'*input* est donné, etc., conduirait à un biais plus ou moins marqué chez les enfants. Au fil du développement, ce biais s'atténuerait, les catégories grammaticales se développant différemment pour arriver aux proportions qu'elles occupent dans le lexique adulte de la langue donnée.

Quant aux moyens de mesurer ce biais, plusieurs études vont dans le sens d'une surestimation du biais pour les noms par les questionnaires parentaux (cf. Tardif et al., 1999). Si ces questionnaires sont très utiles, leur structure même pourrait présenter un biais pour les noms.

En effet, la majorité des questionnaires utilisés dans les articles sont la version originale anglaise ou des adaptations du MCDI (Fenson et al., 1993) et la majorité des *items* présents dans ces questionnaires sont des noms. Dans la version française<sup>14</sup>, on compte précisément 419 noms<sup>15</sup> et 125 verbes<sup>16</sup> sur un total de 691 mots (soit 60 % et 18 %). De plus, les parents auraient tendance à « oublier » de cocher les verbes que leur enfant produit, et cela même si l'on en rajoute dans la liste (Tardif et al., 1999). Une autre difficulté liée aux questionnaires parentaux est qu'il en existe peu qui soient adaptés à une population bilingue (voir O'Toole, 2022 pour un exemple récent).

En outre, les mesures observationnelles ont aussi leurs limites. Les enregistrements de productions spontanées ont par exemple un biais en faveur des mots que l'enfant préfère utiliser plutôt que ceux qu'elle ou il est capable

<sup>13</sup> ● Bien que certains articles comme Caselli et al. (1995) forment une super-catégorie des noms pour certaines analyses et y comptent une partie des onomatopées.

<sup>14</sup> ● Inventaires Français du Développement Communicatif (IFDC, Kern, 2007).

<sup>15</sup> ● Catégories 1 à 13 de l'IFDC 16 - 30 mois.

<sup>16</sup> ● Catégorie 15 (actions) et 18 (auxiliaires) de l'IFDC 16 - 30 mois + 5 expressions verbales de type « avoir faim ».

de produire (Caselli *et al.*, 1995), et en faveur de ceux utilisés dans des situations spécifiques. Malgré une certaine comparabilité statistique des mesures (Tardif *et al.*, 1999), si l'on veut étudier le début du développement langagier de jeunes enfants, il semble pertinent de combiner les deux méthodes pour maximiser la précision des données obtenues.

### Limites

Les comparaisons directes entre les articles étudiés dans cette revue sont rendues délicates par la diversité des méthodologies utilisées, non seulement au niveau de la collecte et l'analyse des données mais aussi et surtout dans la manière de catégoriser les mots produits par les enfants.

Tout d'abord, comme on l'a vu plus haut (« *Variations translinguistiques dans l'acquisition des catégories grammaticales* »), la proportion des différentes catégories grammaticales varie en fonction de l'âge mais aussi en fonction du stade développemental des enfants. Ainsi, des enfants de 2 ans pourront avoir un écart de niveau de vocabulaire important mais être classés dans le même groupe dans certaines études. Il est donc ardu d'effectuer des comparaisons directes entre les études se focalisant uniquement sur l'âge et celles ayant choisi de distinguer les enfants selon leur niveau de vocabulaire.

Au niveau des méthodes d'analyse des données, pour comparer les proportions des différentes classes de mots dans le lexique, certains articles utilisent par exemple des pourcentages parmi les 100 premiers (ou 50 premiers) mots les plus reportés aux âges étudiés (Rescorla *et al.*, 2017) ou une proportion en termes d'*items* cochés dans une catégorie sur les *items* possibles dans la catégorie (Bornstein *et al.*, 2004), alors qu'une majorité s'intéresse aux pourcentages de mots présents dans une catégorie sur le total de mots produits par l'enfant (Kern, 2007) ou bien au ratio  $N/N + V$ <sup>17</sup> (cf. Tardif *et al.*, 1999). Ainsi, la méthode de Bornstein et collègues (2004) aura tendance à contrebalancer l'effet de biais nominal intrinsèque au questionnaire. D'ailleurs Bornstein et collègues (2004) et Kern (2007) diffèrent justement sur le français au stade 50 mots, où les premier-ère-s ne trouvent pas de biais pour les noms alors que la seconde en rapporte.

Enfin, les méthodologies de catégorisation utilisées dans les articles analysés sont assez éclectiques. En effet, les premiers mots produits par les enfants sont difficiles à classer

dans les catégories grammaticales du lexique adulte. Chaque auteur-riche se retrouve donc à devoir décider, parfois un peu arbitrairement, de la catégorie dans laquelle les placer. Bornstein et collègues (2004) incluent par exemple la catégorie du MCDI « jeux et routines » dans celle des verbes pour leur analyse alors que Kern (2007) non. Pour les noms, c'est cette fois Kern qui inclut les noms de personnes et les « endroits » et pas Bornstein et ses collègues. Comme discuté dans Tardif (1996), le fait que la catégorie des noms de personnes soit parfois incluse dans celle des noms, et parfois non, est problématique : d'une part parce que la fonction sémantique des noms communs et des noms de personnes est différente (les premiers se référant à des classes, les seconds à des individus), d'autre part parce qu'on a vu que les noms de personnes, dans la catégorie sémantique des mots sociaux, sont très présents dans le premier lexique et peut donc peser considérablement sur les résultats.

### Conclusion et perspectives

Dans cette revue systématique de la littérature sur le lexique productif des enfants de 1 à 3 ans, nous avons pu examiner un total de 28 articles récents traitant de la question dans 16 langues différentes. Nous avons pu mettre en évidence que les noms sont une classe sur-représentée dans le premier lexique, un phénomène que l'on nomme « biais pour les noms ». Les verbes ont également une place majeure, même si moindre en général. Dans certains contextes et dans certaines langues, ils peuvent l'emporter sur les noms. Au début du développement lexical, on retrouve assez peu d'adjectifs et de mots fonctionnels, mais une autre catégorie sémantique a attiré notre attention, celle des mots sociaux, qui semble majoritaire au tout début de l'acquisition du langage.

Les causes du biais en faveur des noms semblent en partie universelles. Elles seraient liées à des paramètres conceptuels des catégories sémantiques d'objet et d'action. Ces causes universelles semblent pourtant atténuées par divers paramètres : les caractéristiques « *verb-friendly* » de certaines langues (position saillante du verbe dans la parole, par exemple) ainsi que la situation de production.

Les résultats de notre revue systématique démontrent ainsi que tous les enfants tendent à présenter un biais nominal au début de l'acquisition, indépendamment de la langue, mais que les facteurs translinguistiques entrent en jeu progressivement et conduisent à un développement lexical différent selon la langue acquise. De plus, nous avons également

<sup>17</sup> ● Noms/(noms + verbes).

démontré que le biais nominal est plus marqué quand on évalue le lexique de l'enfant grâce à un questionnaire parental.

Pris ensemble, ces résultats indiquent un certain nombre d'implications pour les professionnels-le-s gravitant autour des jeunes enfants. Premièrement, lors d'évaluations lexicales mêmes précoces, il est important de prendre en compte le lexique dans son ensemble, et pas seulement les noms, afin de vérifier si la diversification lexicale typique dans la deuxième et troisième année de vie suit son cours.

De plus, la mise en évidence des facteurs translinguistiques semble tout à fait pertinente à prendre en compte, notamment dans la prise en soin orthophonique d'enfants en situation multilingue, si possible dans toutes les langues de l'enfant.

Dans les programmes de stimulation langagière, il faudra également veiller à ajouter aux pratiques de lecture partagée, biaisée pour les noms, des activités de jeu dynamique impliquant des routines qui favorisent la production de mots sociaux, ainsi que des mouvements et des actions diverses impliquant la production de verbes, classe au rôle central pour le développement morphosyntaxique (Choi & Gopnik, 1995).

## RÉFÉRENCES

- Alcock, K. J. (2017). Production is only half the story-First words in two East African languages. *Frontiers in psychology*, 8, 1898.
- Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of child language*, 27(3), 521-559.
- Bassano, D., Eme, P. E., & Champaud, C. (2005). A naturalistic study of early lexical development: General processes and inter-individual variations in French children. *First Language*, 25(1), 67-101.
- Bassano, D., Labrell, F., & Bonnet, P. (2020). Le Développement du langage de production en français (DLPF) entre 18 et 42 mois : une synthèse. *Enfance*, 2(2), 151-174.
- Bassano, D., Maillolochon, I., & Eme, E. (1998). Developmental changes and variability in the early lexicon: a study of French children's naturalistic productions. *Journal of Child language*, 25(3), 493-531.
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., Pécheux, M., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child development*, 75(4), 1115-1139.
- Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10(2), 159-199.
- Caselli, C., Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of child language*, 26(1), 69-111.
- Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of child language*, 22(3), 497-529.
- Coquet, F., Ferrand, P., & Roustit, J. (2010). *Evalo BB*. Isbergues: Ortho Éditions.
- Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International journal of language & communication disorders*, 43(4), 361-389.
- Dhillon, R. (2010). Examining the Noun Bias: A structural approach. *University of Pennsylvania Working Papers in Linguistics*, 16(1), 7.
- D'odorico, L., & Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language*, 34(4), 891-907.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), 1-185.
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *MacArthur Communicative Inventories: User's guide and technical manual*. San Diego.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Language*, 2, 301-334.
- Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of child language*, 20(3), 523-549.
- Junyent, A., Blume, M., Fernández Flecha, M., & Tijero Neyra, T. (2020). El vocabulario productivo y su relación con la gramática en niños hablantes de castellano peruano entre los 16 y los 30 meses. *Interdisciplinaria*, 37(2), 143-158.
- Kauschke, C., & Hofmeister, C. (2002). Early lexical development in German: A study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of child language*, 29(4), 735-757.
- Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, 27(3), 227-250.
- Kern, S., & Fekete, G. (2019). De l'évaluation à l'intervention. Le développement du langage chez le jeune enfant. In *Théorie, clinique, pratique* (pp. 233-267). Bruxelles, De Boeck.
- Kern, S., & Gayraud, F. (2018). *IFDC : Inventaire français du développement communicatif*. Les éditions de la Cigale.
- Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English and Korean-speaking children: Language-general and language-specific patterns. *Journal of child language*, 27(2), 225-254.
- Longobardi, E., Rossi-Arnaud, C., Spataro, P., Putnick, D. L., & Bornstein, M. H. (2015). Children's acquisition of nouns and verbs in Italian: Contrasting the roles of frequency and positional salience in maternal language. *Journal of child language*, 42(1), 95-121.
- Longobardi, E., Spataro, P., Putnick, D. L., & Bornstein, M. H. (2017). Do early noun and verb production predict later verb and noun production? Theoretical implications. *Journal of child language*, 44(2), 480-495.
- Maital, S. L., Dromi, E., Sagi, A., & Bornstein, M. H. (2000). The Hebrew Communicative Development Inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language*, 27(1), 43-67.
- Martin, V., Renaud, J., & Dagenais, P. (2013). *Les normes de production des revues systématiques : Guide méthodologique*. National Institute of Excellence in Health and Social Services.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group (2009). Reprint-preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Physical therapy*, 89(9), 873-880.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38(1-2, Serial No 149), 136
- Nicoladis, E. (2001). Finding first words in the input. *Trends in bilingual acquisition*, 1, 131.
- Ogura, T., Dale, P. S., Yamashita, Y., Murase, T., & Mahieu, A. (2006). The use of nouns and verbs by Japanese children and their caregivers in book-reading and toy-playing contexts. *Journal of Child Language*, 33(1), 1-29.
- O'Toole, C. (2022). Developing an assessment of vocabulary proficiency for Irish-English bilingual children: The Irish crosslinguistic lexical tasks. *Journal of Multilingual and Multicultural Development*, 1-16.
- Özcan, F. H., Altinkamiş, F., & Gillis, S. (2016). Early lexical composition of Turkish-Dutch bilinguals: Nouns before verbs or verbs before nouns. *Poznan Studies in Contemporary Linguistics*, 52(4), 583-604.
- Rescorla, L., Constants, H., Bialecka-Pikul, M., Stepień-Nycz, M., & Ochal, A. (2017). Polish vocabulary development in 2-year-olds: Comparisons with english using the language development survey. *Journal of Speech, Language, and Hearing Research*, 60(4), 1029-1035.
- Salerni, N., Assanelli, A., D'odorico, L., & Rossi, G. (2007). Qualitative aspects of productive vocabulary at the 200-and 500-word stages: A comparison between spontaneous speech and parental report data. *First Language*, 27(1), 75-87.

Shapiro, K. A., Moo, L. R., & Caramazza, A. (2006). Cortical signatures of noun and verb production. *Proceedings of the National Academy of Sciences*, 103(5), 1644-1649.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental psychology*, 32(3), 492.

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, 44(4), 929.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child development*, 70(3), 620-635.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, 24(3), 535-565.

Taverna, A. S., & Waxman, S. R. (2020). Early lexical acquisition in the Wichi language. *Journal of Child Language*, 47(5), 1052-1072.

Xuan, L., & Dollaghan, C. (2013). Language-specific noun bias: Evidence from bilingual children. *Journal of child language*, 40(5), 1057-1075.

## ANNEXE

Tableau 1. Caractéristiques méthodologiques et résultats des articles de la revue.

| N° | Article                 | Nb de participants | Âge     | Profil linguistique               | Méthodologie         | Influences translinguistiques                    |  | Autres facteurs                       | Plurilinguisme & autres domaines                   |
|----|-------------------------|--------------------|---------|-----------------------------------|----------------------|--|--|---------------------------------------|--|
|    |                         |                    |         |                                   |                      | Noms et Verbes                                   | Autres   |                                       |  |
| 1  | Alcock (2017)           | n = 208            | 0;8-1;4 | monolingues kiswahili & kigiriana | questionnaire        | N > V  |  | résultats différents en compréhension |  |
| 2  | Bassano (2000)          | n = 1              | 1;2-2;6 | monolingue français               | production spontanée | N > V puis N = V au cours du développement       |  |                                       |  |
| 3  | Bassano et al. (1998)   | n = 1              | 1;2-2;6 | monolingues français              | production spontanée | N > Pred puis N = Pred (2 ans)                   | Mots grammaticaux > N & Pred après 2 ans   |                                       |  |
|    |                         | n = 24             | 1;8-2;6 | monolingues français              | production spontanée | N > Pred puis N = Pred (2 ans)                   |  |                                       |  |
| 4  | Bassano et al. (2005)   | n = 60             | 1;6-3;4 | monolingues français              | production spontanée | N > Pred puis N = Pred au cours du développement | V > Adj ; Mots gram. > N & Pred après 2 ans  |                                       | N réfèrent maj. à des noms animés, V à des actions |
| 5  | Bornstein et al. (2004) | n = 47             | 1;8     | monolingues espagnol              | questionnaire        | N > V  | N = V > Adj puis N > V > Adj au cours du développement   |                                       |  |
|    |                         | n = 30             | 1;8     | monolingues français              | questionnaire        | N > V  |  |                                       |  |
|    |                         | n = 28             | 1;8     | monolingues néerlandais           | questionnaire        | N > V  |  |                                       |  |
|    |                         | n = 37             | 1;8     | monolingues hébreu                | questionnaire        | N > V  |  |                                       |  |
|    |                         | n = 34             | 1;8     | monolingues italien               | questionnaire        | N > V  |  |                                       |  |
|    |                         | n = 42             | 1;8     | monolingues coréen                | questionnaire        | N > V  |  |                                       |  |
|    |                         | n = 51             | 1;8     | monolingues anglais               | questionnaire        | N > V  |  |                                       |  |
| 6  | Caselli et al. (1995)   | n = 659            | 0,8-1,4 | monolingues anglais               | questionnaire        | N > V  | % N en italien > % N en anglais (pour une certaine catégorisation large des N seulement) ; pas de différences pour les V entre les 2 langues |                                       |  |
|    |                         | n = 195            | 0,8-1,4 | monolingues italien               | questionnaire        | N > V  |  |                                       |  |
| 7  | Caselli et al. (1999)   | n = 1001           | 1;6-2;6 | monolingues anglais               | questionnaire        |  | % N & V en anglais > % N et V en italien   |                                       |  |
|    |                         | n = 386            | 1;6-2;6 | monolingues italien               | questionnaire        |  | % mots sociaux en italien > % mots sociaux en anglais  |                                       |  |

| N° | Article                                | Nb de participants | Âge      | Profil linguistique        | Méthodologie                         | Influences translinguistiques   |  | Autres facteurs                        | Plurilinguisme & autres domaines  |
|----|--|--------------------|----------|----------------------------|--------------------------------------|---|--|--|---|
|    |  |                    |          |                            |                                      | Noms et Verbes  | Autres   |  |   |
| 8  | Choi & Gopnik (1995)                   | n = 9              | 1;2-1;10 | monolingues coréen         | questionnaire & production spontanée | N = V   | % V en coréen > % V en anglais ; % N en coréen < % V en anglais (comparaison avec une autre étude)                             | production parentale                   | morphologie productive dès 1;3 chez les enfants coréens, mais pas chez les enfants anglais. V référent maj. à des actions, N à des objets |
| 9  | D'odorico & Fasolo (2007)              | n = 24             | 1;4-1;6  | monolingues italien        | questionnaire                        | N > V   |  |  |   |
| 10 | Dhillon (2010)                         | n = 30             | 1;7-2;11 | monolingues anglais        | production spontanée                 | N > V   | % N en anglais > % N en espagnol > % N en mandarin   | production parentale                   |   |
|    |  | n = 30             | 1;7-2;11 | monolingues mandarin       | production spontanée                 | N = V (tendance n.s. N > V au début du développement)   |  | production parentale                   |   |
|    |  | n = 10             | 1;7-2;11 | monolingues espagnol       | production spontanée                 | N > V   |  | production parentale                   |   |
| 11 | Jackson-Maldonado <i>et al.</i> (1993) | n = 68             | 1;3-2;7  | monolingues espagnol       | questionnaire                        | N > Pred  |  |  |   |
| 12 | Junyent <i>et al.</i> (2021)           | n = 104            | 1;4-2;6  | monolingues espagnol       | questionnaire                        | N > Pred  | Pred > Mots fonct.   |  |   |
| 13 | Kauschke & Hofmeister (2002)           | n = 32             | 1;1-3;0  | monolingues allemand       | production spontanée                 | N > V puis N = V et V > N au cours du développement mais non testé explicitement                        | Mots sociaux > N au début du développement (N apparaissent avant V mais jamais la catégorie la + imp. à cause de mots sociaux) |  |   |
| 14 | Kern (2007)                            | n = 548            | 0;8-1;4  | monolingues français       | questionnaire                        | N > Pred  |  |  | peu de mots d'action au stade < 50 mots   |
| 15 | Kim <i>et al.</i> (2000)               | n = 8              | 1;4-1;8  | monolingues anglais        | questionnaire & production spontanée | N > V   | % V en coréen > % V en anglais   | production parentale                   |   |
|    |  | n = 8              | 1;6-1;9  | monolingues coréen         | questionnaire & production spontanée | N > V   |  | production parentale                   |   |
| 16 | Longobardi <i>et al.</i> (2015)        | n = 26             | 1;3-1;8  | monolingues italien        | production spontanée                 | N > V   |  | production parentale                   |   |
| 17 | Longobardi <i>et al.</i> (2017)        | n = 30             | 1;2-2;0  | monolingues italien        | questionnaire & production spontanée | N > V   |  |  |   |
| 18 | Maital <i>et al.</i> (2000)            | n = 253            | 1;6-2;0  | monolingues hébreu         | questionnaire                        | N > Pred  | biais pour les N légèrement > en hébreu qu'en anglais ; % V en hébreu = % V en anglais (comparaison avec une autre étude)      |  |   |
| 19 | Ogura <i>et al.</i> (2006)             | n = 31             | 1;0-2;0  | monolingues japonais       | production spontanée                 | N > V (contexte lecture) et N > V puis N = V puis V > N suivant le stade développemental (contexte jeu) |  | type d'activité + production parentale |   |
| 20 | Özcan <i>et al.</i> (2016)             | n = 55             | 0;8-3;0  | bilingues turc-néerlandais | questionnaire                        | N > V   |  |  | résultats valables dans les 2 langues des enfants, mais différence est plus marquée en néerlandais  |
| 21 | Rescorla <i>et al.</i> (2017)          | n = 199            | 2;0-2;1  | monolingues polonais       | questionnaire                        | N > V   | pas de différence de % de N et V entre les 2 langues   |  |   |
|    |  | n = 422            | 1;11-2;0 | monolingues anglais        | questionnaire                        | N > V   |  |  |   |

| N° | Article                      | Nb de participants | Âge      | Profil linguistique        | Méthodologie                         | Influences translinguistiques                        |  | Autres facteurs  | Plurilinguisme & autres domaines |
|----|------------------------------|--------------------|----------|----------------------------|--------------------------------------|--|--|--|----------------------------------|
|    |                              |                    |          |                            |                                      | Noms et Verbes                                       | Autres   |  |                                  |
| 22 | Salerni <i>et al.</i> (2007) | n = 30             | 0;11-2;6 | monolingues italien        | questionnaire & production spontanée | N > V (expliquer les abréviations en bas de tableau) | V > Adj et Mots fonct. au stade 500 mots             | situation de collecte des données (questionnaire vs. production spontanée)   |                                  |
| 23 | Tardif (1996)                | n = 10             | 1;8-1;10 | monolingues mandarin       | production spontanée                 | V > N (V = N + NP)                                   |  |  | actions > ou = à objets          |
| 24 | Tardif <i>et al.</i> (1997)  | n = 6              | 1;8-2;6  | monolingues anglais        | production spontanée                 | N = V (tendance n.s. N > V)                          | % N en mandarin < % N en italien                     | production parentale   |                                  |
|    |                              | n = 6              | 1;9-2;2  | monolingues italien        | production spontanée                 | N = V (tendance n.s. N > V)                          |  | production parentale   |                                  |
|    |                              | n = 10             | 1;9-1;10 | monolingues mandarin       | production spontanée                 | V > N  |  | production parentale   |                                  |
| 25 | Tardif <i>et al.</i> (1999)  | n = 24             | 1;8      | monolingues anglais        | questionnaire & production spontanée | N > V (et N > Pred) et N = V                         | ratio N/(N + V) plus élevé en anglais qu'en mandarin | situation de collecte des données (questionnaire vs. production spontanée) + type d'activité + biais de notification |                                  |
|    |                              | n = 24             | 1;8      | monolingues mandarin       | questionnaire & production spontanée | N > V et N = V                                       |  |  |                                  |
| 26 | Tardif <i>et al.</i> (2008)  | n = 265            | 0,8-1,4  | monolingues anglais        | questionnaire                        | N > V  |  |  |                                  |
|    |                              | n = 336            | 0,8-1,4  | monolingues mandarin       | questionnaire                        | V > N  |  |  |                                  |
|    |                              | n = 369            | 0,8-1,4  | monolingues cantonnais     | questionnaire                        | N = V  |  |  |                                  |
| 27 | Taverna & Waxman (2020)      | n = 22             | 1;0-3;11 | monolingues wichi          | questionnaire                        | N > V  |  | production parentale   |                                  |
|    |                              | n = 3              | 1;4-3;11 | monolingues wichi          | production spontanée                 | le ratio N/(N + V) diminue au cours du développement |  | production parentale   |                                  |
| 28 | Xuan & Dollaghan (2013)      | n = 50             | 1;10-2;6 | bilingues anglais-mandarin | questionnaire                        | N > V  | % N en anglais > % N en mandarin                     |  | bilingues équilibrés             |

notes sur les abréviations : N = noms, V = verbes, Adj = adjectifs, Pred = prédicats, Mots gram. = mots grammaticaux, Mots fonct. = mots fonctionnels, NP = noms propres.