DOI : 10.35662/unine-thesis-3156

Université de Neuchâtel

Faculté des Lettres et Sciences Humaines
Institut des Sciences Logopédiques

Factors in the acquisition of new meaning in human and nonhuman great apes

Thèse présentée à la Faculté des Lettres et Sciences Humaines
pour l'obtention du titre de docteure ès Sciences Humaines et Sociales, discipline Logopédie

par

Dahliane Labertoniere

Sous la direction de :

Prof. Katrin SKORUPPA, Université de Neuchâtel, Suisse
Prof. Klaus ZUBERBÜHLER, Université de Neuchâtel, Suisse

Rapporteur·es externes :

Emmanuel CHEMLA, Directeur de recherche, CNRS, France
Sophie KERN, Directrice de recherche, CNRS, France

Soutenue le 10 février 2025

unine
Université de Neuchâtel

Faculté des Lettres et Sciences Humaines
Espace Tilo-Frey 1
CH - 2000 Neuchâtel
Tél. +41 32 718 1700
secretariat.lettres@unine.ch

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de Mme Katrin Skoruppa, co-directrice de thèse, professeure, Université de Neuchâtel ; M. Klaus Zuberbühler, co-directeur de thèse, professeur, Université de Neuchâtel ; Mme Sophie Kern, directrice de recherches CNRS, Laboratoire Dynamique du Langage, Lyon ; M. Emmanuel Chemla, directeur de recherches CNRS, Laboratoire de Sciences Cognitives et Psycholinguistique, Paris autorise l'impression de la thèse présentée par Mme Dahliane Labertoniere en laissant à l'autrice la responsabilité des opinions énoncées.

Neuchâtel, le 10 février 2025

Le doyen
Loris Petris

p. O.

Acknowledgements

First of all, I want to thank my advisors, Katrin Skoruppa and Klaus Zuberbühler, for the many discussions that we had concerning the different experiments presented in the thesis and for their dynamic guidance over the years. I am particularly grateful that you agreed to collaborate on such an interdisciplinary project.

I am also very grateful to the members of my jury, Sophie Kern for her insightful comments and Emmanuel Chemla, who has observed this project's evolution from my mid thesis in 2020 and always had pertinent advice, thought-provoking remarks and encouraging feedback.

I would like to thank Sarah Brocard and Vanessa Wilson for their insights into animal research, support and the interactions that we had in and outside the zoo. This thesis was also made possible by contributions from my student assistants: Katia Damache and Marie Dumoulin by their help in recruiting participants, Mathilde Wenger for the same task as well as testing, Hadrien Laforest and Matylda Florez by acting in the videos stimuli, and Claudia Ferrario by filming those videos. Thank you as well to all my Master students from the 'TP recherche en Logopédie': Rim Benkirane, Thays Cavalcanti, Fiona De Virgilio, Mathilde Iseli and Martha Vacher who helped during the design and testing of the first adult pilot experiment; Julia Berger, Mélina Boyer, Loïse Frutiger, Marine Jonat and Belinda Keller who helped conduct the systematic review; Justine Blanc, Oriane Catillaz, Lucile Charmillot, Oriane Rieder, Jeanne Volery, Léane Zbinden and Florence Zulauff who ran and analysed the second adult pilot; Elena Blauel and Justine Dornier, who recruited and tested many participants for Experiment 2 and Agathe Agostini, Olivia Bähler Laure Ramstein and Céline Rousselot who modified the existing stimuli; lastly Nadia Besselma, Camille Lefebvre, Elisa Medici, Amélie Steiner, Eleonora Vaccaro and Téa Waecker for the coding and preliminary analyses of the corpus study.

I also thank my current and former colleagues from PAM7, the Institute of biology and in particular Julie Bodard, Géraldine Jean-Charles and Nuzhat Tabassum and I also extend my thanks to Suzanne Lesage.

Last but not least, all my thanks go to Gabrielle Aguila-Multner for her continuous support over the years and her thorough proofreading.

Abstract

This doctoral dissertation aims to shed light on the influence of different factors on the acquisition of new meaning in closely related species of great apes: humans, gorillas and orangutans. First, the ability to rapidly learn to associate meaning to form will itself be investigated and compared in adult humans and a captive population of nonhuman great apes, with the aim to gain insight into the evolution of this capacity and thus the evolution of language. Further, it will inquire into the role of visual context (stable visual context vs. variable visual context) and the role of semantics (acquisition of words for actions vs. words for objects) on rapid word learning in young infants during their second year of life. Finally, it will explore the temporal evolution of different classes of words (in particular nouns and verbs) in young children's productive lexicon, and take a look at the different factors (like semantic class, social information and situational context) influencing word production by presenting the results of a systematic review of the literature. The same questions will be examined in an ecological corpus study conducted on children aged from one to three years. A discussion of children's first speech acts and their relevance to nonhuman animal communication will also be presented. The results of the four research articles presented in this dissertation fail to demonstrate fast mapping in nonhuman apes, but propose possible changes in method for further comparative studies. Moreover, they indicate that a changing visual context could be beneficial for rapid word learning in infants. Finally, they confirm that the noun bias is quasi-universal across all human languages. They show that social words are an important part of children's early lexicon and confirm that social information plays a key role in language acquisition.

keywords: language evolution, comparative cognition, language acquisition, lexicon, fast mapping, eye-tracking

Résumé

Cette thèse de doctorat vise à mettre en lumière l'influence de différents facteurs sur l'acquisition de nouveaux sens chez des espèces de grands singes étroitement apparentées : l'être humain, le gorille et l'orang-outan. Tout d'abord la capacité d'apprendre rapidement à associer un sens à une forme sera elle-même étudiée et comparée chez des humain-es adultes et dans une population captive de grands singes non humains, dans le but de mieux comprendre l'évolution de cette capacité et, par conséquent, l'évolution du langage. En outre, elle examinera le rôle du contexte visuel (contexte visuel stable vs. contexte visuel variable) et le rôle de la sémantique (acquisition de mots désignant des actions vs. des objets) sur l'apprentissage rapide de mots chez les jeunes enfants au cours de leur deuxième année de vie. Enfin, elle étudiera l'évolution temporelle de différentes classes de mots (en particulier les noms et les verbes) dans le lexique productif des jeunes enfants, et examinera les différents facteurs (tels que la classe sémantique, l'information sociale et le contexte situationnel) qui influencent la production de mots en présentant les résultats d'une revue systématique de la littérature. Les mêmes questions seront examinées dans le cadre d'une étude de corpus écologique menée sur des enfants âgés de un à trois ans. Une discussion sur les premiers actes de langage des enfants et leur pertinence pour la communication animale sera également présentée. Les résultats des quatre articles de recherche présentés dans cette thèse ne suffisent pas à démontrer l'existence du fast mapping chez les primates non humains, mais proposent d'éventuels changements de méthode pour des études comparatives ultérieures. De plus, ils indiquent qu'un contexte visuel changeant pourrait être bénéfique pour l'apprentissage rapide des mots chez les très jeunes enfants. Enfin, ils confirment que le biais pour les noms est quasi-universel dans toutes les langues humaines. Ils montrent que les mots sociaux constituent une part importante du lexique précoce des enfants et confirment que l'information sociale joue un rôle clé dans l'acquisition du langage.

mots clefs: évolution du langage, cognition comparée, acquisition du langage, lexique, fast mapping, eye-tracking

Contents

Introduction

The goal of this dissertation is to explore the factors influencing the early stages of the acquisition of new meaning in human and nonhuman great apes. On the one hand, in an evolutionary perspective, I want to investigate whether the process of rapid acquisition of meaning (fast mapping), crucial to human language acquisition, evolved before language did, and is thus present in our closest cousins, nonhuman great apes. To this end, the first chapter will present a comparative eye-tracking study looking into this ability in adult humans and nonhuman apes. On the other hand, I explore more specifically the role of visual context and of semantics on the rapid acquisition of new meaning in children. Is a stable or changing visual background better for learning? Are words for objects easier to learn than words for actions? This will be the subject of the second chapter, where I present two eye-tracking fast mapping studies performed on 14- to 19-month-old children. Finally, I examine the productive lexicon of young children: what kinds of words are acquired first? how does the lexicon evolve? which other factors have an effect on children's productions? The third chapter will explore these questions in a systematic review of children's productive lexicon and a naturalistic corpus study on early productions. I will also discuss the acquisition of speech acts in children and how certain early acquired categories of speech acts can be compared with communication in nonhuman apes.

This introduction will present the theoretical literature necessary to understand the subject matter of this dissertation, namely the comparative approach, children's lexical acquisition, including the mechanisms involved (focusing on fast mapping) and the factors influencing it (focusing on word classes, social information and visual context), and finally the eye-tracking method. A detailed plan of the dissertation will also be laid out.

A note on human language and its evolution

Human language is a complex recursive system that can be described and analysed across many different domains like phonetics, phonology, syntax, morphology, semantics and pragmatics. It is a powerful social tool that allows us to express any concept that forms in our mind to other members of our species. None of the other animal species use such a complex system of communication – not even nonhuman great apes, which raises the question of how it has evolved.

Humans and other great apes are very closely related phylogenetically, forming the Hominidae family. By extension, they are also closely related genetically. Chimpanzees (*Pan troglodytes*) in particular (along with bonobos, *P. paniscus*) are, from an evolutionary standpoint, our closest living relatives (Mikkelsen et al., 2005). Studying a cognitive ability (its presence and its properties) in nonhuman apes can not only be interesting in and of itself (Schlenker et al., 2022), but through comparison with its human counterpart, it can shed light on how this ability evolved from our common ancestor to modern humans (Berthet et al., 2023).

A few components of human language have been investigated in nonhuman animal communication, such as call referentiality (semantics), syntax and pragmatics. Even though nonhu-

man primates have a small vocal repertoire compared with humans and songbirds for example (Cheney & Seyfarth, 2018), and cannot expand it (Hobaiter et al., 2022), they use their calls in flexible ways. Chimpanzees use specific calls in specific contexts (e.g. hunting, travelling), suggesting that such calls might be functionnally referential (Crockford & Boesch, 2003), while Diana monkeys (*Cercopithecus diana*) and Campbell's monkeys (*C. campbelli*) use specific calls when encountering different kinds of predators (Zuberbühler, 2000, 2001). Campbell's monkeys also use their limited vocal repertoire to form combinations in a form of proto-syntax (Ouattara et al., 2009). Call compositionality has also been found in other species of primates, although the syntax involved is quite simple and can be modeled with very simple finite state grammars (Schlenker et al., 2016). Moreover, call production in chimpanzees can be influenced by specific audiences, in a form of pragmatics (Townsend et al., 2008), and nonhuman primates seem to produce specific alarm calls based on an informativity principle (Schlenker et al., 2016). Therefore, nonhuman primate communication systems do share certain properties with human language, providing evidence of the evolution of those properties before humans diverged from our common ancestor. It is then relevant to delve further into other early precursors of human language to see if we can find them in nonhuman apes as well.

Fast mapping is one such precursor, as it is a learning ability that is essential during human language acquisition. It is defined in the literature as the ability of an individual to form hypotheses about the meaning of a new sound after as few as one exposure to it (Carey, 1978). It is possible that fast mapping is simply a type of learning, albeit one requiring only minimal exposure (over a very brief period of time) to the novel associations, with the hypothesis (or hypotheses) formed directly after these exposures. Some amount of retention of this initial hypothesis exists, however, a further step of consolidation (that is separate from fast mapping) is needed to encode it in long term memory. I extrapolate that a rapid association of any sign (not just a sound) to a referent can qualify as fast mapping. However, it is likely that only certain types of signs are considered relevant (to have meaning) by a given species. Furthermore, it is probable that individuals of the same species form the same guesses about the meaning of the novel sign (Chemla, personal communication).

Chimpanzees and bonobos, like the well know Kanzi (Savage-Rumbaugh et al., 1985), as well as grey parrots (*Psittacus erithacus*) like Alex (Pepperberg, 1990), have been able to learn several associations between signs and referents (noniconic graphic symbols and, for example, food objects in the case of Kanzi, and words and objects in the case of Alex), however, both these cases involved extensive training (nonexplicit in the case of Kanzi, as he acquired his 'lexicon' of symbols by watching his mother learn). To my knowledge, however, nothing is available from nonhuman great apes on the matter of fast mapping, even if this phenomenon has been studied in domestic animals (Fugazza et al., 2021; Kaminski et al., 2004; Takagi et al., 2024). By using a comparative approach, we can gain insight into the evolutionary history of communication at large. If fast mapping can be demonstrated in great apes, then this will have far-reaching consequences for our understanding of the evolution of language, by indicating that this ability most likely[1] appeared before humans and thus before language.

In humans, children as well as adults are particularly good at this kind of task. In the following section, I will go into detail into the acquisition of meaning – the acquisition of the lexicon

[1] If we exclude convergent evolution.

– in children, exploring what we know about the fast mapping mechanism and looking at the role of certain factors on the composition and acquisition of this lexicon.

Acquisition of new meaning in human children

During the acquisition of their mother tongue(s), children are confronted with several difficult tasks: they not only need to learn an immense vocabulary (the lexicon), but before that also to learn to parse speech into meaningful units (phonological development), learn how those units can be combined (morphosyntactic development), learn to produce them and in which contexts to use them (pragmatical development). Here, we will focus on the acquisition of the lexicon, or the acquisition of new meanings. While children begin to produce their first words around the start of their second year of life (Hoff, 2013), the rate of acquisition of novel words is slow at first. The rapid vocabulary growth known as the *naming explosion*, which can be observed in some children, only begins around 18 months or around the 50-word mark[2] (Carey, 1978; Hoff, 2013). During this period, children acquire new words at a much faster rate than before, and can acquire a new word after minimal exposure to it. It is debated whether this increase in the number of new acquisitions per day is abrupt (hence the term 'explosion'), or gradual (Bloom, 2002; Ganger & Brent, 2004; Hoff, 2013). How can a child acquire around nine new words per day (or even just nine per week, Bloom, 2002) during this period? Even for children who acquire words at a slow rate, how do they manage it? What mechanisms are at play?

Fast mapping

The first step in the acquisition of new meanings in children is most often called fast mapping, the mapping of a form (e.g. a sound) onto a concept. This default mechanism, first described by Carey (1978), enables infants to rapidly infer the meaning of a novel word, and is essential for them to form a lexicon from scratch. What is considered 'fast' or 'rapid' can vary across studies, with children being exposed to the novel associations from one to up to 40 times before being tested on them (Kan & Windsor, 2010). Since fast mapping consists in forming a first hypothesis about a new label's meaning, further steps are needed to consolidate this hypothesis and encode it in the mental lexicon. First, children have to retain the novel association in their long term memory (see for example Carey, 1978; Vlach & Sandhofer, 2012) and then they have to learn to extend the novel label to other examplars of the referent: this is called generalisation (Behrend et al., 2001).

The process of fast mapping has been investigated in several studies since being first described. For example, in Behrend et al. (2001)'s interactive study, two-year-old children correctly retrieved the target object among distractors after only three exposures to the novel label-object association. In Heibeck and Markman (1987)'s study, two to four-year-olds could rapidly learn new words pertaining to diverse semantic domains (colour, shape and texture), showing that fast mapping is not limited to novel nouns. As an object manipulation task experiment, their study was once again interactive. In more recent years, the rapid mapping of a novel label to an object has been thoroughly investigated using implicit methods as well. Such methods, like eye-tracking with a preferential looking paradigm, allow the testing of younger children. This has contributed to proving the existence of fast mapping in children as young as 12-months, a period

[2] Children reach vocabulary milestones at different ages.

at the very beginning of language acquisition. For example, in Smith and Yu (2008), 12- and 14-month-old infants could rapidly infer the meaning of novel labels through cross-situational statistical learning.

Finally, even though fast mapping is an ability that is crucial during language development for young children, it does not fade away during the child's development. Indeed, adults also have the ability to fast map and do not differ in this from children in word learning studies (Bloom, 2002) when tested either with direct (Markson & Bloom, 1997; Vlach & Sandhofer, 2012) or indirect (Halberda, 2006; Trueswell et al., 2013) measuring methods.

What is more, several learning strategies can be employed during word learning along with fast mapping to better guess the meaning of a novel word. One of those strategies is mutual exclusivity (Markman, 1994; Markman & Wachtel, 1988). If an already known concept is presented to the child at the same time as a novel one and they have to guess whether a novel word applies to one of them, they will use mutual exclusivity to infer that this novel word should be paired with the novel concept: since the known concept already has a label, the novel label should apply to the novel object (operating under the hypothesis that an object should only be designated by one label). However, it is not always possible to use mutual exclusivity during word learning, and mutual exclusivity is not necessary for successful fast mapping. It is possible for example to be presented with an image of a bunny in a field and a novel word in a completely new language (e.g. 'blicket') and infer that it refers to the bunny. At play here is another strategy or bias of human learning known as the whole object bias (Markman, 1994): it has been shown that infants prefer to associate a novel label to a whole object rather than its parts (Hollich et al., 2007). Supported by fast mapping as well as additional constraints (i.e. cues) on word meaning (Trueswell et al., 2013), like mutual exclusivity, whole object bias and syntactic knowledge, children will rapidly develop a vocabulary of thousands of words in just a few years.

Many factors can have an influence on the learning (rapid or not) of novel words. For example advanced phonemes in the word (MacLeod et al., 2011), segmental information (Havy et al., 2011), knowledge of syntax (Gleitman et al., 2005), use of child-directed speech (prosody, see Matychuk, 2005), word frequency in the input (Lieven, 2010), but also internal factors like memory capacity and non-linguistic external factors like parental socio-economical status. In the following parts, I will focus on psycholinguistic child-external factors and detail the role of the grammatical category of a new word, the role of the meaning itself (semantics) of a new word and the role of social information on learning. These factors help explain why the lexicon is first comprised of many social terms and nouns (that are often objects). I will also review the role of visual context on rapid word learning.

The role of grammatical categories on learning: verbs and nouns

During the acquisition of their mother tongue's lexicon, children of various ages have been shown to have a 'bias' towards nouns over verbs in production and comprehension. In a cross-linguistic study (Bornstein et al., 2004), 20 month-old children of various native languages were found to produce more nouns than verbs. Moreover, in proportion, names for objects constitute a larger part of children's early vocabularies compared to adults' and older children's (Bloom, 2002). In experimental settings, novel nouns are easier to learn for children than novel verbs (Arunachalam & Waxman, 2011). Finally, the meaning of some high frequency English verbs

(e.g. ‘pour’, ‘fill’) is not fully understood until middle childhood, while nouns are less problematic (Bloom, 2002).

It has been proposed (Bloom, 2002) that the bias towards nouns could be explained by the different morphosyntactic properties of verbs and nouns in certain languages (called ‘noun-friendly’ by certain authors[3]). Regarding syntax, in English, for example, finding the verb in the VP in the following sentence

(1) The otters [chase [the butterfly.]$_{NP}$]$_{VP}$

requires identifying multiple content words (the NP *the butterfly* is contained inside the VP) whereas in the NP *the otters* there is only one content word and thus finding the noun is easier.

For comparison, here is the same sentence in Japanese (a ‘verb-friendly’ language):

(2) Kawauso-tachi wa [[chouchou o]$_{NP}$ oikake-teiru]$_{VP}$
otter-PL TOP butterfly OBJ chase-PROG
‘The otters are chasing the butterfly.’

The verb here is easier to locate than in the corresponding English sentence, being in final position.

In some languages, verbs are also morphologically more complex and thus vary more from instance to instance, making them harder to extract from speech than nouns. Take for example the following sentences in French:

(3) Le chat est dans le jardin.
lə ʃa ɛ dɑ̃ lə ʒaʁdɛ̃
the cat be.PRS.3SG in the garden
‘The cat is in the garden.’

(4) Les chats sont dans le jardin.
lɛ ʃa sɔ̃ dɑ̃ lə ʒaʁdɛ̃
the cat be.PRS.3PL in the garden
‘The cats are in the garden.’

(5) Les ratons-laveurs finissent les croquettes pendant que le chat finissent
le ʁatɔ̃lavœʁ finis lɛ kʁokɛt pɑ̃dɑ̃ kə lə ʃa fini
the racoon finish.PRS.3PL the kibble while that the cat finish.PRS.SG
l’ eau.
l o
the water
‘The racoons finish the kibble while the cat finishes the water.’

Phonetically, in both (3) and (4), the noun ‘cat’ does not change, even though it is plural in (4). The common verb *être* ‘be’ however, is highly variable. Note that this is still the case in English, where the plural mark in (4) would be heard, but the noun *cat* still has the same stem, while *be* takes a novel form like in French. In a less extreme example (5) involving a more regular verb like *finir* ‘finish’, the verb is still more morphologically complex than the noun.

[3] See for example Imai et al. (2008).

Verbs could also be harder to learn because they are not heard as often as nouns. Indeed, in 'noun-friendly languages', caretakers' productions sometimes exhibit a bias towards nouns rather than verbs (Choi & Gopnik, 1995; Kim et al., 2000), though function words are also highly frequent in the input, and are still not learnt early (Gentner, 1982).

However, certain studies on Mandarin and Korean – 'verb-friendly' languages that put more emphasis or saliency on the verb than on the noun, e.g. presenting more verbs in utterance-final position (see Bornstein et al., 2004) and where parental input does not present a noun bias (Choi & Gopnik, 1995; Kim et al., 2000) – have still found a bias towards nouns (if somewhat attenuated) for young native learners in both production (Bornstein et al., 2004) and during learning (Imai et al., 2008). This suggests that morphosyntactic complexity and input frequency are not the sole factors at play in the difficulty to learn verbs.

The role of semantics on learning: actions and objects

Attempts to explain the noun bias in infancy have also focused on the conceptual difference between verbs and nouns (Gentner, 1982). Nouns have in general higher imageability than verbs (Ma et al., 2009), meaning that it is easier to picture them mentally. Indeed, in a semantic perspective, verbs tend to denote actions or events (activities or changes in state), labelling relations between things, while most nouns tend to refer to objects or people, labelling things themselves.[4] Concepts for objects could be argued to be easier to grasp than concepts for actions. Actions are also harder to observe. In a study on adults, Gillette et al. (1999) found that linguistic cues like syntactic frame were essential in order to correctly guess the meaning of a verb compared to cross-situational observation only. Young children's first verbs usually refer to actions or events that encode physical motion, and not the mental status of an agent (Snedeker & Gleitman, 2004).

Alternatively, the bias might be attentional: humans might be drawn more towards objects, or believe that other people would draw attention to objects and not actions most of the time (proposed in Bloom, 2002).

Experimentally, children learning 'verb-friendly' and 'noun-friendly' languages are better at learning new nouns (referring to objects) than new verbs (referring to actions) when the task requires generalisation (Imai et al., 2005, 2008). This seems to be true cross-linguistically, as Waxman et al. (2013) reviews that children acquiring many different languages can easily map novel nouns to novel objects, but have much more difficulties in mapping novel verbs to novel actions, with infants' performance varying with the language they acquire and the learning method. One relevant study found that English infants were better at mapping a novel label to an object rather than an action; however, Mandarin learning infants were better at mapping novel labels to actions (no generalisation was required – Chan et al., 2011). This suggests a complex interplay between semantics, syntax and morphology in the acquisition of novel words.

[4] In fact, the majority of the most common nouns used by caregivers in the CHILDES corpora refer to concrete objects (Sandhofer et al., 2000).

The role of social information on learning

Aside from the grammatical category or semantic class of a word, social information also plays an important part in word learning, from interaction with others (joint attention) to the nature of the word itself (social words).

Joint attention is thought to be critical for the learning of a new language. An example of this phenomenon is when both the child and an 'expert' adult (e.g. the parent) are focused on a particular object. Joint attention can involve, among other things, shared gaze and pointing. Tomasello and Farrar (1986) found that children aged 17 months would learn new words better in a context of joint attention with an experimenter than without joint attention. Moreover, in naturalistic situations between mother and child (aged between 15 and 21 months), they observed a higher number of utterances and a longer conversation time in episodes of joint attention compared to other situations. Live social interaction also seems to be important for learning in young infants, as Kuhl et al. (2003) showed that 9-month-olds were able to learn novel consonant contrasts with a live teacher present, but not via a video recording of the same teacher.[5] It remains to be seen which kind of social cues were essential here, but a probable hypothesis would be that infants were aided by the contingent responses of the teacher in the live condition.

However, most learning situations could be argued to happen without joint attention (Tomasello & Barton, 1994). A number of researchers have engineered learning situations in which various types of social contexts were absent to find out if children would still be able to learn new words.[6] In Arunachalam (2013), it is argued that as long as the toddlers are exposed to syntactic context (in this case for verb learning), the presence of social context like eye gaze might be less important.

Regardless, social information can also be embedded in the meaning of words themselves. For example, interjections, terms for social (or play) routines and greetings (e.g. *hey, thank you, bye bye, go away, 'ainsi font font'*...), onomatopeias and affirmative and negative particles (e.g. *yes*, *no*) are strongly connected with interactions and only make sense to utter in the presence of another person. Finally, person nouns refer to social beings. All such terms, that I propose to call 'social words', seem to form a major part of children's first lexicon (Kauschke & Hofmeister, 2002; Kern, 2007; Tardif et al., 2008).

The role of visual context on learning

Finally, the visual context in which a word is produced, or on which the novel referent is presented to the child, seems to have an influence on word learning.

To start with, when infants first begin to produce words (at about 10 to 14 months), they seem to produce them only in specific contexts. Specifically, if the child's first occurence of the word was produced in a certain situation (for example in the bath or watching from a window – Barrett, 1986; Bloom, 2002, respectively), during the following days or weeks, they will only say it again in exactly the same context. This suggests that infants could have trouble with generalising the meaning of novel words to different situations. In general, the situation in

[5] Interestingly, grey parrots trained by humans to label objects also failed to learn via video recordings and only succeeded with live modeling, see Pepperberg et al. (1999).

[6] For precise examples, see Arunachalam (2013).

which the child first produces a novel word is the one where their caregiver uses that word most often (Harris et al., 1988). However, the apparent attribution of a restricted meaning to the word by the child happens in spite of the caregivers' usage, as caregivers use words in many different situations or visual contexts (Hoff, 2013).

Additionally, experimental studies have depicted a complex picture regarding the role of visual context on rapid word learning by children. Indeed, the stability of the visual context during the learning of novel words can either help or hinder learning. In Twomey et al. (2018), the colour of the background on which a novel object was presented either varied or remained stable (the novel object was also accompanied with two known objects). Two-year-old children were better at making the correct associations between target and label in the stable condition, but in a test of retention five minutes later, were better in the varied condition. In the same vein, the two-and-a-half- to three-year-old children of Vlach and Sandhofer (2011)'s study performed better in the stable condition (the colour of the cloth on which the object was presented did not change during learning) than in the varied condition, but three- to four-year-olds displayed the opposite pattern. In contrast, when the stability of visual context concerned the objects presented around the target object (e.g. in a stable context, with an object on the left side and an object on the right side never changing across all learning instances), Axelsson and Horst (2014) found that three-year-olds learnt better in a stable context condition. Lastly, in a study comparing three conditions, the first with a stable background (coloured cloth) during learning, the second with a varied one and the third one with interleaved varied and stable contexts, two-year-old children could only learn novel associations in the last condition (Goldenberg & Sandhofer, 2013).

Finally, the context of production also plays a role on the type of words likely to be produced by a child. Reading contexts seem to favour nouns, while play activities seem to generate more verbs from children (Ogura et al., 2006; Salerni et al., 2007; Tardif et al., 1999). This might be due to the fact that caretakers are more likely to name multiple objects during shared-reading (Ogura et al., 2006), while play activities are more conducive to the presence of actions and therefore to their naming.

A note on the eye-tracking method

As previously seen, eye-tracking, and in particular the preferential looking paradigm, is a well tested method to study novel word learning in young children (see Halberda, 2006; Piot et al., 2024; Trueswell et al., 2013; Twomey et al., 2018). It can also be used with adults, with or without explicit instructions (see for example Trueswell et al., 2013).

Eye-tracking has already been used successfully in apes: see Kano and Call (2014) with orangutans, bonobos and chimpanzees, Myowa-Yamakoshi et al. (2012) with chimpanzees, Kano et al. (2012) with gorillas and orangutans, and Wilson et al. (2024) with chimpanzees, gorillas and orangutans. Although a few differences exist between the gaze movement of the different species longer saccade latency in humans compared to other apes, see Kano and Call, 2014, it is still possible to make a comparison between different species.

Therefore, eye-tracking seems to be an appropriate method to use in a comparative study

between human and nonhuman great apes.

Plan

In conclusion, we can note that while many factors play a role in early word learning, multiple questions remain unanswered in the domain of early first language acquisition. The syntactic and semantic natures of a word seem to play a part in children's ability to learn it, but it remains to be proven that this extends to all human languages. The role of social information on word learning should be explored further, as is the case with the role of visual and situational contexts, in particular on very young children at the beginning of language development. Finally, although fast mapping abilities are essential for human language acquisition, there is a lack of empirical evidence, especially for nonhuman apes. By directly comparing great apes with humans with an eye-tracking paradigm, I aim to gain a deeper understanding of the evolution of fast mapping, which is key to human language.

In chapter 1, I will take an evolutionary perspective and compare human adults to non-human apes (gorillas and orangutans) in a fast mapping eye-tracking task, which will be similar to the one presented in chapter 2. This task, subject of a article submitted in *Animal Cognition*, in addition to inquiring into the presence of fast mapping in nonhuman apes, also investigates the role of the stability of visual context on associative learning in both populations.

In chapter 2, the effect of visual context stability on the fast mapping of novel associations will be explored in an experimental manner in a research article. This article was submitted to the *Journal of Child Language*. It presents an eye-tracking experiment comparing two conditions, a stable visual context condition and a changing context condition, with naturalistic stimuli in 14- to 20-month-old children. After having discussed the possible effect of the ecological backgrounds on the results by presenting the results of a second experiment, I will present a similar third experiment comparing the rapid learning of actions versus the rapid learning of objects. Lastly, I will address the rapid learning of actions and objects in adults as well, by presenting the results of two forced-choice tasks.

Finally, in chapter 3, I will present a systematic review of the productive lexicon of children aged one to three, in the beginning of language development (Labertoniere & Skoruppa, 2022). This article, which was published in the journal *A.N.A.E.*, inquires into the noun bias, but also takes a look at the place of other grammatical categories (like adjectives) and of social words in the early lexicon of learners of various languages. It also details the role of certain situational contexts and the role of different methods of data collection on children's productions. After the article, I will present a naturalistic corpus study looking to summarise the effects of grammatical category, semantic category and context of production in French learning one- to three-years-olds' productions. This study uses data from the CHILDES database as a mean to categorise productions in an ecological setting (in naturalistic situations, most often at home). To conclude, I will then discuss the distribution of speech acts in children's first productions, while making a link with nonhuman ape communication.

1. Evolutionary perspective

1.1. Research article

Status: Submitted to *Animal Cognition* on January 6th 2025.

Fast mapping in hominids

Dahliane Labertoniere[1, 2], Vanessa Wilson[1, 3, 4, 5], Carla Pascual-Guàrdia[1], Katrin Skoruppa[2], & Klaus Zuberbühler[1, 6]

[1] Department of Comparative Cognition, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland
[2] Institute of Speech and Language Therapy, University of Neuchâtel, Neuchâtel, Switzerland
[3] Department of Comparative Language Science, University of Zurich, Zurich, Switzerland
[4] Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, Zurich, Switzerland
[5] School of Psychology and Social Work, University of Hull, United Kingdom
[6] School of Psychology & Neuroscience, University of St Andrews, St Andrews, Scotland

Address for correspondence

dlab101@proton.me

ORCID

0009-0008-6757-5948

Conflict of Interest

The authors have no conflicts of interest to declare.

Ethics statement

The Cantonal Veterinary Office of Basel Stadt and the Animal Welfare Officer at Basel Zoo authorised the ape research (approval number 3059). The human research was approved by the Ethics Committee of the University of Neuchâtel (CER-UNINE, approval number 133/2024).

Acknowledgements

We thank Adrian Baumeyer, Fabia Wyss, Stephan Lopez, Gaby Rindlisbacher, Rene Buob, Roland Kleger, Markus Beutle, Nicole Fischer and the technician team at Basel Zoo for all their support and assistance. We thank Géraldine Jean-Charles for data cleaning and script writing, Sarah Brocard for support and suggestions during data collection, Letizia Volpin for running the lab visits while the human experiment was ongoing and Simone Marty for help with the audio stimuli.

Funding

This research was supported by funding from the Institute of Speech and Language Therapy (K.S.), the University of Neuchâtel (K.S., K.Z.) and the National Centre for Competence in Research 'Evolving Language' (SNSF agreement 295 number 51NF40_180888, K.Z.) and Top-Up Grant (grant number N603-18-01, V.A.D.W., K.Z.).

Author contributions

Conceptualisation: D.L., K.S., K.Z. ; **Methodology**: D.L., V.A.D.W., K.S., K.Z. ; **Formal analysis**: D.L. ; **Investigation**: D.L., V.A.D.W., C.P-G. ; **Resources**: D.L. ; **Writing - Original Draft**: D.L. ; **Writing - Review & Editing**: V.A.D.W., C.P-G., K.S., K.Z. ; **Visualisation**: D.L. ; **Data curation**: D.L. ; **Supervision**: K.S., K.Z. ; **Project administration**: D.L. ; **Funding Acquisition**: V.A.D.W., K.S., K.Z.

Abstract

Fast mapping is essential when children acquire language, but whether the required cognition is uniquely human or shared with animals is debated. Although documented in dogs and cats, both species have a history of domestication of social cognition, so that it remains unclear whether fast mapping is naturally present in animals. Here, we used an eye-tracking paradigm to test three species of hominids – gorillas, orangutans and humans – in their ability to rapidly learn to associate novel sounds with objects in their everyday noisy environment. The task was difficult for all participants, but while adult humans showed evidence of fast mapping, we could not detect any sign of learning in the other hominids. These species differences could have trivial causes, such as problems with attention or motivation, but it is also possible that fast mapping requires a preexisting lexicon before becoming an effective learning mechanism, or that it has simply evolved after the shared ancestor of all great apes.

Keywords

Fast mapping, word learning, language evolution, comparative cognition, language acquisition, meaning, mental representation

1.1.1. Introduction

Fast mapping is a learning ability regarded as essential during language acquisition in human children. It is defined as the ability to form hypotheses about the meaning of a new form (a spoken or gestural unit), often after only one or a few exposures (Carey, 1978) and is often seen as the first step in the word learning process. Fast mapping is based on the association of a form – typically a sound – to a referent, typically an object in the child's surroundings. Fast mapping is probably essential in the acquisition of lexical meaning and could explain how children efficiently learn their immense vocabulary of tens of thousands of words in just a few years (Bloom, 2002).

In empirical studies of fast mapping, children will typically be exposed to a few instances of a new word in conjunction with a referent and subsequently tested to see if they correctly learnt the association (Behrend et al., 2001; Heibeck & Markman, 1987; Vlach & Sandhofer, 2011). In some cases, children will be retested a few weeks later to check for signs of retention (Carey, 1978; Vlach & Sandhofer, 2012), a memory consolidation mechanism that is, along with generalisation, involved in building the lexicon (Vlach & Sandhofer, 2012) after the first step of fast mapping. Although fast mapping is important during language acquisition in children, human adults continue to deploy this ability, as demonstrated by various word learning studies (Bloom, 2002; Halberda, 2006; Markson & Bloom, 1997; Vlach & Sandhofer, 2012).

There is an ongoing debate whether fast mapping specific to language acquisition or whether

the mechanism is cognitively more general and also relevant in other domains (Behrend et al., 2001; Markson & Bloom, 1997). In the latter case, this might suggest that fast mapping evolved before modern humans and should therefore be present in some other animal species, which face similar challenges when establishing various sound-event relations during ontogeny (León et al., 2022). Relevant here is evidence for fast mapping in dogs (*Canis familiaris*). In one study (Kaminski et al., 2004), a border collie was able to form novel word-object associations, based on an exclusion learning paradigm. In the experiment, the dog encountered an unfamiliar object within an array of familiar items after a 'fetch' command by his owner uttered in conjunction with an unfamiliar word. The dog then preferentially retrieved the novel object and, moreover, showed evidence for retention when retested a few weeks later. In a more recent study (Fugazza et al., 2021), two dogs (a border collie and a Yorkshire terrier) were tested on their ability to choose between two newly learnt object labels. First the dogs were exposed by their owners to two new toys associated with novel labels, then the owners asked the dogs to bring back one of the novel toys, labeling it. Both individuals performed above chance in this task, specifying the conditions under which fast mapping operates in dogs. Specifically, the dogs learnt the new labels while playing with their owners, who explicitly labeled the objects, and both dogs were already very familiar with object-name learning. Importantly, 20 other family dogs that had no object-naming experience failed the task. There is also evidence for fast mapping in domestic cats (*Felis catus*). Here, in a habituation-dishabituation manual gaze-tracking experiment (Takagi et al., 2024), cats were first exposed to a novel word-object association during the habituation phase and then responded more strongly to the incorrectly labeled object than to the correctly labeled object in the dishabituation phase,[7] indicating that they had associated the stimuli with their correct labels.

These studies suggest that fast mapping is not limited to humans and might also exist in other species. However, the caveat with this interpretation is that cats and dogs have undergone substantial artificial selection over thousands of years, generally to facilitate interspecies communication and social interactions with their human owners. Whether fast mapping is naturally present in animals is still unknown, as we are not aware of any comparable research with phylogenetically relevant groups of animals, especially great apes.

To address this, we used an eye-tracking paradigm (a method that was successfully used in both humans and nonhuman apes, see Kano & Call, 2014; Kano et al., 2012; Wilson et al., 2024) that allowed us to directly compare human adults and nonhuman great apes in a simple fast mapping task. We used a protocol similar to one used in infant studies (see Labertoniere et. al, submitted; Schafer & Plunkett, 1998), requiring no explicit learning. To improve the ecological validity of the study, we did not follow the standard protocol, which is to present objects in front of simple, featureless backgrounds to enhance their salience (e.g. Schafer & Plunkett, 1998). Instead, objects were presented in how they are usually encountered in real life, that is, embedded in a complex, visually-structured natural background that sometimes changes between each encounter. This is important because of evidence suggesting that the nature and variability of the background can impact learning (Axelsson & Horst, 2014; Twomey et al., 2018; Vlach & Sandhofer, 2011, and Labertoniere et. al, submitted), and that input variability facilitates the generalisation of novel concepts (Bourgoyne & Alt, 2017).

[7] For example, they looked longer at a red object paired with the label which had been previously associated with a blue object.

Here, we explored fast mapping in three species of hominids, human adults, gorillas (*Gorilla gorilla*) and orangutans (*Pongo abelii*) and how learning was influenced by the variability in the background in which the referents were presented. If fast mapping evolved before language, we expected to find evidence in all three species. If variation in background mattered, we expected it to make learning less difficult (Bourgoyne & Alt, 2017).

1.1.2. Method

Participants

Great apes

N=4 gorillas (3 F, mean age = 8.2 years, range = 2.8-15.9 years) and N=1 orangutan (M, age = 19.9 years) served as participants and were tested in their indoor enclosures at Basel Zoo, Switzerland. All contributed to the experiment voluntarily, could leave at any time and were never separated from the rest of the group or food deprived. Participation was rewarded with sugar-free syrup, in amounts approved by the zoo's veterinary team. Since participants were not separated from their peers, social distractions were common, simulating the conditions under which learning takes place in nature.

Humans

N=5 French-speaking bachelor students (mean age = 22.4 years, range = 19-26 years, gender/sex was not recorded) at the University of Neuchâtel, Switzerland, served as participants following informed consent. To provide a learning environment similar to the one experienced by the apes at the zoo, we collected the data during an ongoing lab visit during which some of the participants' peers walked by and talked, which likely generated comparable social distraction.

Stimuli

To investigate whether participants could learn novel sound-object associations, we presented them with four different yoked pairs of objects associated with four different yoked pairs of sounds. Objects were pictures of exotic fruits, presented on four different naturalistic backgrounds (rocks, leaves, soil, grass) as it might happen in real-life. We chose exotic fruits as objects to keep apes interested and to minimise the risk that humans already had a name for them. See all pairs of objects in Supplementary Table 1.2, and see Supplementary Figure 1.1 for one pair of objects on all possible backgrounds. As sound stimuli, we generated pseudo-words that complied with French and Swiss-German phonotactic rules (e.g. /if/, /ovo/),[8] produced by a bilingual speaker of both languages. See Supplementary Table 1.1 for a list of all paired pseudo-words.

Procedure and apparatus

For the apes, gaze was recorded at a distance of approximately 60 cm with a Tobii Spectrum eye-tracker (300 Hz sampling rate), positioned under a monitor (23.8", 1920 x 1080 pixels), which displayed the images using Tobii Pro Lab. The pseudo-words were played by a loudspeaker (Fostex 6301B 10W amplifier) hidden behind the screen. The experiment was coded in

[8] Nonhuman apes were used to hearing Swiss-German from the zoo keepers and were thus familiar with such sounds and human adults were to be tested in a French-speaking part of Switzerland.

Tobii Pro Lab.

For N=2 gorillas, the apparatus was positioned in a small enclosure separated from the main enclosure by a sliding door (open during testing) and a plexiglass window, with the experimenter present with access to controls, see Supplementary Figure 1.2. For N=2 gorillas and the orangutan, the apparatus was part of a mobile setup, consisting of a height-adjustable, purpose-built plexiglass box, which could be placed in front of the enclosure, see Supplementary Figure 1.3. In this setup, a drinking nozzle attached to the vertical mesh ensured that eyes remained positioned between the mesh to avoid interference with gaze detection.

For the humans, gaze was recorded in the same way with a Tobii Pro X3-120 eye-tracker (120 Hz sampling rate) positioned under a monitor (Dell, 24.0", 1920 x 1080 pixels). Here, the pseudo-words were presented via two stereo loudspeakers (Fostex PM0.5n, 70W amplifier) hidden under the table supporting the monitor. This setup was placed inside a sound-attenuated booth with the door open to allow for distractions. The script running the experiment was coded in Matlab version 2019b (The Math Works, Inc., 2019) using the Psychtoolbox (Brainard & Vision, 1997; Kleiner et al., 2007; Pelli, 1997) and the Titta toolbox (Niehorster et al., 2020). The experimenter had access to the controls from outside the booth. To match the ape condition as closely as possible, human participants were told that they could pause or abort the experiment at any time (none did) but otherwise they were not given any instructions. Also important, in contrast to similar experiments (e.g. Markson & Bloom, 1997), we decided not to use any carrier sentences or attention getters (e.g., "hey" or "This is a xxx"), as they would be meaningless to nonhuman apes and would only give humans an unfair advantage.

For the human participants, we carried out a full 5-point calibration with pulsating dots before starting the experiment. For the ape participants, a session began with a green screen displayed for 2 s and a 4-point validation of a previously recorded calibration. This calibration was recorded prior to the testing sessions following an established protocol in Wilson et al., 2024: we installed a drinking nozzle so that the participant's head ended up in a central position in front of the screen, and once a participant was properly positioned, we carried out a 2-point calibration.

Experimental design

The entire experiment consisted of four sessions (two sessions with invariable backgrounds and two sessions with variable backgrounds). Each session, which corresponded to one specific pair of object-pseudo-word, consisted of six repeated blocks. One block lasted at least 23 s, as follows.

Invariable background (Pairs 2 and 3)

Each block started with a first 6,000 ms *learning phase* during which one exotic fruit was presented with the same background three times in a row, each time for 2,000 ms accompanied by the same pseudo-word, 500 ms after image onset. This was then followed by a *second learning phase* during which another exotic fruit was presented in the same way, again accompanied by its own pseudo-word. The two paired fruits were chosen to be maximally different in size, colour and shape. We also made sure that the two paired pseudo-words differed in the number of syllables and were as phonetically distant as possible from each other.

The first and second learning phases were separated by a gaze-contingent (manual for the nonhuman apes, automatic for the humans) visual attention getter (AG).

Provided the participant paid attention to the screen during a second attention getter, this was then followed by a *first test phase*, during which the two fruits appeared side-by-side for 2,500 ms without any sound. Subsequently, an attention getter was presented between the two fruits for 500 ms to re-center the participant's gaze. Immediately after this, the two fruits were shown for another 2,500 ms, but this time with one of the two pseudo-words broadcast 500 ms after stimulus onset. This was then followed by the *second test phase*, identical to the first one, during which the *other* pseudo-word was broadcast.

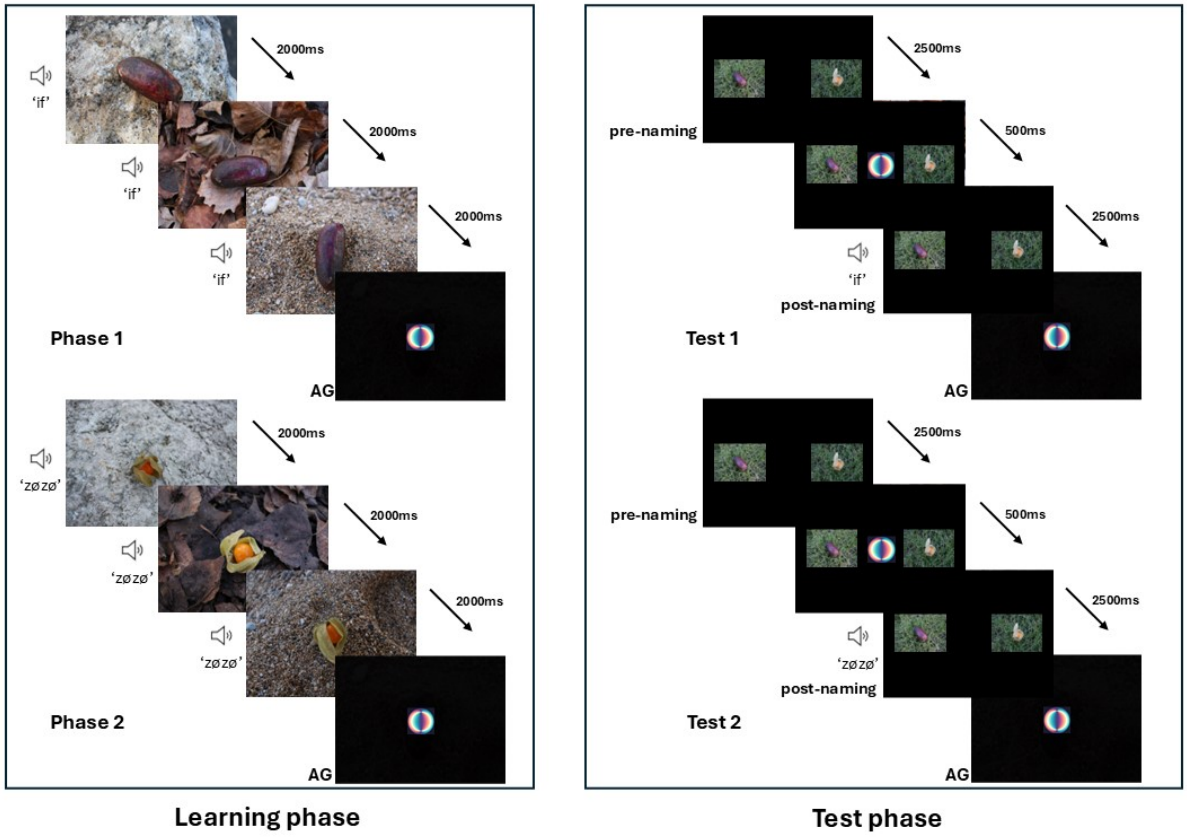


Figure 1.1.: Schematic representation of the experimental design for one block in the variable background condition. The same block is repeated five times with a change in the order of the stimuli to form one session (here corresponding to Pair 1). For the invariable condition, each fruit would be presented on the same background and at the same angle during Learning and Test.

Once a block (i.e., two learning and two test phases) was completed, it could be repeated five times, up to block six (depending on the participant's motivation).[9] A participant who had completed all six blocks in a session had received N=18 expositions to the two novel fruit-word associations during the learning phases and had been tested N=6 times on each association during the test phases.

[9] The order of presentation of the stimuli was changed between blocks.

Variable background (Pairs 1 and 4)

To test the impact of the visual environment on learning, the same design was used, but this time the fruit was presented on four visually different backgrounds and photographed from different angles for each of its appearances during the learning and test phases (see Supplementary Figure 1.1). Figure 1.1 recapitulates the procedure schematically for the variable background condition.

1.1.3. Data analysis

For ape participants, data files supplied by Tobii Pro Lab were transformed in Matlab for subsequent analyses in R Studio (running R version 4.2.2). Our files provided gaze hits in our two areas of interest (AOIs),[10] the left and right images in the test, every 3 ms. For human participants, data analysis was carried out similarly, except that raw data were supplied directly in Matlab format. We used a Matlab script to transform those raw files into files providing gaze hits in the two AOIs every 8 ms. For both species, we computed the proportion of target looks for the pre- and post-naming phases of each trial.[11]

Human participants completed all four sessions of the experiment on the same day, with short pauses between sessions (for a total duration of approximately 15 min). The apes often struggled to remain focused, but still completed up to two sessions per day, depending on motivation.[12] We excluded test trials where the participant did not pay any attention to the screen (2) or where the sound was not playing due to technical issues (3). In both cases, the experiment was stopped and the block run again at a later time.

If fast mapping took place, we predicted relatively increased looking times towards the fruit whose pseudo-word was played during the (first or second) test phase. Specifically, we predicted looking time to the named fruit to increase between the pre- and the post-naming phases in the test phases, expressed in terms of mean proportion of target looks (PTL).

1.1.4. Results

Great apes

We managed to record gaze data in N=130 of 216 test trials (60.2%) in both the pre- and post-naming condition.[13] Since the data were not normally distributed (Shapiro-Wilk test: $W = 0.798$, $p < 0.001$) but had a zero-one inflated beta distribution, we applied a slight shrinkage to the proportional data (zeros were transformed to 0.0001 and ones to 0.9999). We built a generalised linear mixed model for beta distributions with a logit link using the glmmTMB package (Brooks et al., 2017). PTL was the dependent variable, with 'background' as a fixed effect, 'naming' (pre- vs. post-) as a nested fixed effect within 'background' and 'participant' as a random effect. The factor 'background' was sum-contrasted for the model. Adding block number did not improve the model, so it was not included in the final model, which was as

[10] AOI regions were defined as the area around an image – a 350*620 px rectangle.
[11] Data files and codes can be provided in the supplementary materials.
[12] One subject had to come back two days later to finish one of her block.
[13] For the exact number of trials per session provided by each participant, see Supplementary Table 1.3. Some participants did not go up to five repetitions of a block.

follows:

PTL ~ Background / Naming + (1 | Participant)

We found no significant increase in looking time between pre- and post-naming phases, neither for invariable nor with variable backgrounds (see Table 1.1). A graphical overview of the data (with PTL averaged by subject) showing comparisons between pre- and post-naming phases of both background conditions can be found in Figure 1.2 (a, b).

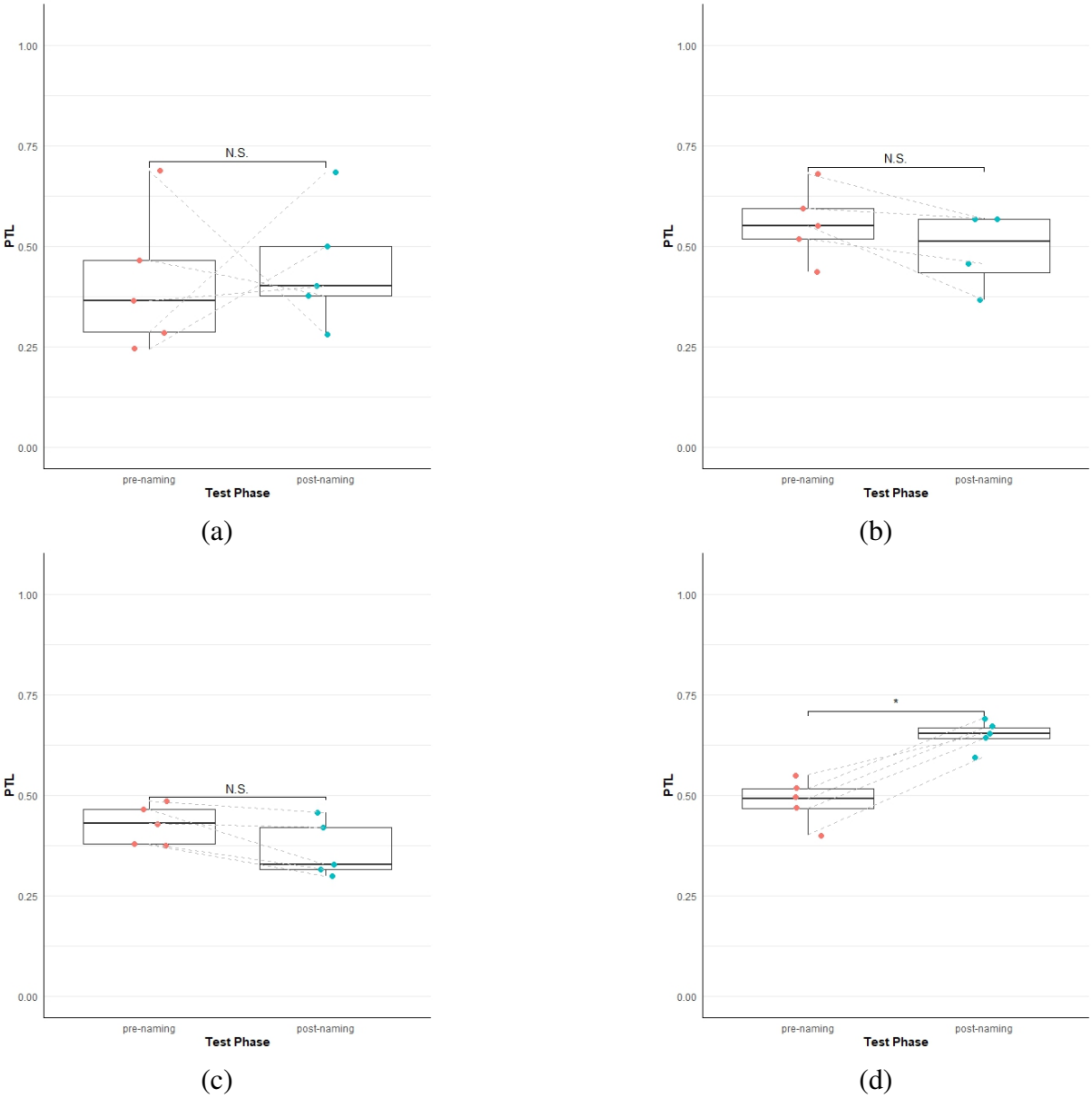


Figure 1.2.: Mean gaze proportions in pre- and post-naming phases in the context-change condition (a, c) and invariant condition (b, d) for apes (a, b) and humans (c, d).

Humans

Participants provided a total of N=240 test trials (N=24 per background condition per participant). As with the apes, data were not normally distributed (Shapiro-Wilk test: W = 0.877, p <

0.001) but followed a zero-one inflated beta distribution. We carried out operations analogous to the ape data and set up a model using PTL was the dependent variable, 'background' as a fixed effect, 'naming' (pre- vs. post-.) as a nested fixed effect within 'background', and 'participant' as random effect. Adding block number did not improve the model and was not considered any further.

We found a significant main effect of naming (β = 0.577, SE = 0.180, p = 0.001), although in the invariable background condition only (see Table 1.1). Figure 1.2 (c, d) shows comparisons between pre- and post-naming PTL averaged by subject for both background conditions.

Table 1.1.: Results of the models looking into the presence of a naming-effect in both conditions in human and nonhuman great-apes

Predictor	Estimate	SE	*p*
Nonhuman apes			
(Intercept)	0.0101	0.131	0.939
Background	-0.150	0.132	0.255
Background_Variable:Naming	0.038	0.305	0.901
Background_Invariable:Naming	-0.321	0.330	0.330
Humans			
(Intercept)	-0.089	0.092	0.330
Background	-0.075	0.092	0.415
Background_Variable:Naming	-0.292	0.179	0.103
Background_Invariable:Naming	0.577	0.180	**0.001**

Abbreviation: SE, standard error.

1.1.5. Discussion

How a signal obtains its meaning is of cross-disciplinary interest, particularly in the areas of language acquisition and evolutionary studies of cognition. A key mechanism is fast mapping (Carey & Bartlett, 1978), a mental process by which a representation is formed following exposure to a referent in conjunction with a signal. Fast mapping likely plays an important role in word learning and provides an explanation for the astonishing speed at which children acquire their vocabularies. Although animal communication also involves learning (Vernes et al., 2021), the problem is somewhat simplified by the fact that most communication systems consist of limited signal repertoires, usually produced as part of biologically fixed functions (Oller et al., 2013). Nevertheless, when hearing an unfamiliar signal produced by a more experienced conspecific, the problem for an infant ape is identical to that of a human child, that is, to identify which aspect of reality the signal refers to and to retain this relation for the future.

The current literature is equivocal as to the nature of fast mapping. Either it is simply a special case of cognition based on one-trial learning, for which there are countless examples in the animal literature, ranging from conditioned taste aversion (Garcia et al., 1955) to rapid associations of alarm calls to arbitrary referents (Curio et al., 1978; León et al., 2022). Alternatively, fast mapping may be a mechanism specifically evolved for word learning (Behrend et al., 2001) or a mere byproduct of it. Although analogues of fast mapping have been demonstrated in dogs

and cats, both species have been under strong artificial selection to facilitate interspecies communication with humans - the domestication of social cognition (Hare et al., 2002). At the same time, we are not aware of systematic research on fast mapping in natural animal communication, including humans' closest relatives, the great apes.

We addressed this by directly comparing human and nonhuman great apes in a fast mapping paradigm based on eye-tracking. We chose this method to facilitate direct cross-species comparison using one type of data. We also made an effort to test participants under equivalent conditions and with ecologically relevant stimuli, i.e., an unfamiliar food item encountered on its natural substrate of varying visual complexity. To this end, participants experienced three exposure trials during which an unfamiliar fruit was paired with its pseudo-word, a simple (di- or tri-)syllabic sound sequence. We subsequently tested each participant on whether they had been able to retain the association, by presenting pairs of fruits whilst playing the name of one, predicting that participants' gazes should be guided by the previously established sound-referent relation.

We found that the task was difficult for the participants of all three species. Adult humans, however, showed some evidence of fast mapping of novel sounds to novel objects. Interestingly, however, this was only the case if the fruit was presented on an invariable background, with no evidence for learning if the background changed across the N=6 learning trials. This finding went against our initial prediction, i.e., that encountering new objects in variable ways enhanced generalisation and concept formation. Nevertheless, the fact that we found a learning effect, even though we deliberately made the task difficult with a distracting social environment and lack of explicit instructions, shows that, in principle, fast mapping is possible with our paradigm. In contrast, we found no evidence for fast mapping the association of an acoustically distinct utterance and a specific fruit in great apes. Bearing in mind that it is difficult to draw firm conclusions from negative evidence, we can think of the following explanations for this apparent human-ape divide.

First, although the task appeared to be cumbersome for all species, it appeared to be even more difficult for the apes to remain focused, which may have limited their ability to learn novel associations. The sample sizes were exceedingly small for both human and nonhuman participants (the human sample size was chosen to match the ape one), suggesting that a larger participant pool might have washed out eventual outliers due to low motivation. It is also possible that apes would have needed more training instances to elicit learning (like some children sometime receive, see Kan & Windsor, 2010), although this would have likely only worsened their already prevalent motivation problem (Wilson et al., 2023). The number of learning exposures did not play a role in the adults' ability to fast map, suggesting that the apes' underperformance had other roots.

Second, and related to the above, we noted that the stimulus set was generally not particularly effective in eliciting learning. Although human participants could learn new associations if the background did not change, this was not the case if the background changed. Interestingly, this appears to be different in 14-20-month-old children tested in a similar paradigm (Labertoniere et. al, submitted). Here, learning was only observed when the background changed, which is arguably much closer to what children experience in real-life word learning situations. The adult

participants probably guessed that the goal of the experiment was to test for learning of novel names for objects, which might have incentivised them to look towards the object indicated by the pseudo-word during the tests. Here, the invariant background condition might have been more effective in revealing the goal of the experiment, suggesting that participants performed better because less attention was needed. Interestingly, for adults learning new languages, explicit teaching strategies appear to be more effective (DeKeyser, 2003), suggesting that word learning strategies change during ontogeny from childhood to adulthood.

Third, it is possible that fast mapping did take place in the apes but that our way of assessing learning was simply not fine-grained enough. It is important to note that, even when the apes were focused on the task, their total looking time to the screen was short,[14] which could have masked a potential learning effect in the pre- vs post-attention calculations. In a typical trial, apes quickly scanned the stimuli but then preferred to look elsewhere, a pattern also found in other great ape studies (Kano & Tomonaga, 2011; Wilson et al., 2024). One way to address this might be to assess participants with forced choices on a touch screen, to make the task more interactive and rewarding (see Brocard et al., 2024; Wilson et al., 2023), even though getting the apes to make explicit choices poses its own challenges.

Fourth, while the learning situation modelled in our experiment is likely common in real life, and certainly ecologically relevant (a referent is spotted by an individual, who utters speech and then shares attention with a bystander, allowing the bystander to form an association between sound and referent, and remembers it), it is not clear whether such situations also occur in wild animals. One documented domain in which this has been observed is in the context of anti-predator behaviour, when young individuals rapidly learn to apply one of their species-specific alarm calls to a new referent (Deshpande et al., 2022; León et al., 2022), suggesting that events involving some sort of danger might have been more effective in triggering fast mapping. Regarding our design, we are not aware of any systematic research investigating how primates or other animals learn to use vocal signals to refer to food.

Fifth, it is possible that the nature of the stimuli, in our case spoken pseudo-words, promoted learning in humans, but not in great apes. Relevant here is that captive orangutans respond better to their keepers' gestures compared to words (Dezecache et al., 2019), suggesting that human speech sounds may not be optimal for comparative fast mapping research. In another relevant experiment (Déaux et al., 2021), chimpanzees were faster to learn in a touchscreen-based visual discrimination task if the images were paired with conspecific food calls as opposed to other sounds or with silence, suggesting that some sound-referent associations appear to be more natural than others, with conspecific, referentially correct signals most likely being more efficient than other sounds. Regarding the objects, future studies might seek to present the novel sound-object associations during brief video clips, which tend to increase attention in apes compared to still images (e.g., Wilson et al., 2024), and perhaps to switch to animate objects or to at least present the fruits with some sort of natural motion.

Sixth, fast mapping may require crucial ontogenetic experience to scaffold on. For instance, the fact that it has been observed in domesticated animals suggests that the ability has emerged

[14] For the apes: a mean of 965 ms in pre-naming and 548 ms in post-naming ; for the humans: a mean of 1814 ms in pre-naming and 2178 ms in post-naming.

from individuals having had prolonged contact with humans, which might have also led to a small lexicon of words (see Takagi et al., 2024). Whether this is the main explanation for the species difference is doubtful, mainly because fast mapping also seems to be involved in domains other than language learning (Behrend et al., 2001; Markson & Bloom, 1997).

Finally, the language acquisition literature is full of evidence highlighting the important role of social interactions during learning. A particularly effective situation is when two participants share their attention towards a referent whilst the name is produced (Tomasello & Farrar, 1986). Specifically, children can learn novel consonant contrasts when tutored by a person, but not from a video recording of the same person Kuhl et al., 2003. Human adults also seem to benefit from learning in a social context (Jeong et al., 2010) and in all likelihood similar processes play a role in great apes (Howard et al., 2017). In our design, none of this was the case since the name was broadcast from a speaker, with no real-life reference provider present. This may in fact have made learning harder for all participants, and explain that fast mapping did not seem to happen in the apes.[15] Crucially, both studies on dogs (Fugazza et al., 2021; Kaminski et al., 2004) in which fast mapping was reported were conducted in socially more adequate situations. Not only were the dogs already familiar with the process of new label learning from early puppyhood, the experiment was embedded in a deeply social setting, with dogs and owners playing and interacting, including joint attentional episodes. In the cat study (Takagi et al., 2024), there was also a social component, as the sound stimuli were spoken by their owners. Indeed, cats failed to fast map in the same experiment when sound stimuli were electronic sounds.

1.1.6. Conclusion

We demonstrated that our very basic design was sufficient to trigger fast mapping in hominids, but that various factors might have made it more difficult for great apes, including species differences in motivation, attention, the effectiveness of the stimuli and the importance of social interaction. Until these hypotheses have been addressed by future research the tentative conclusion is that spontaneous fast mapping of novel names for objects in the environment, and the concept formations that go along with this process, is a uniquely human process not naturally present in animal communication. The fact that some domesticated species have provided positive evidence, however, suggests that fast mapping does not essentially require a primate brain. It also suggests that fast mapping does not evolve along phylogenetic branches, but emerges as a consequence of special social cognition in which learning takes place during directed social interactions rather than the default way of observing signal-referent relations in the surroundings. Future research is likely to produce progress on elucidating these factors, and to provide more of a definite answer about the status of what appears to be a human universal.

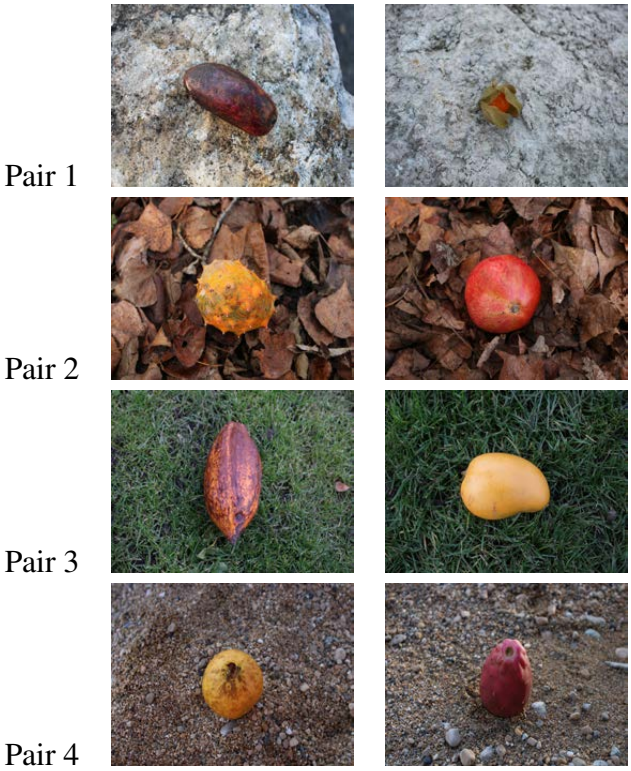
[15] However, many fast mapping studies with human infants show positive results using this kind of protocol.

Supplementary information

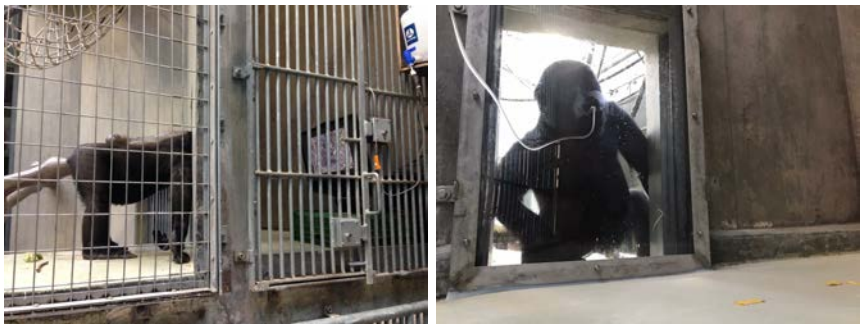
Supplementary Table 1.1.: Pairs of pseudo-words used in the experiment

Pair 1	/if/	/zøzø/
Pair 2	/fa/	/ømø/
Pair 3	/yʃ/	/vuvu/
Pair 4	/za/	/ovo/

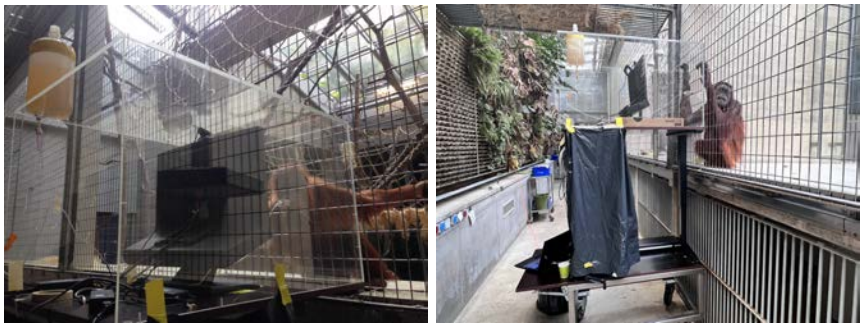
Supplementary Table 1.2.: The four pairs of objects on the four different backgrounds



Supplementary Figure 1.1.: One pair of objects with different orientations on the four different backgrounds



Supplementary Figure 1.2.: Enclosure setup



Supplementary Figure 1.3.: Mobile setup

Supplementary data for the nonhuman great apes:

Since we lost a number of trials in which the apes either did not pay attention during the test phase or the eye-tracker could not register their gaze, we decided to collect more data for the N=4 gorillas, from one week to three months after they had completed a session (therefore they repeated some sessions). When running the model on these data combined with the previous test

set (for a total of 218 test trials), we still did not find any significant increase in the proportion of target looks between pre- and post-naming.

Supplementary Table 1.3.: Number of test trials each ape provided data for per session

	Session 1	Session 2	Session 3	Session 4
Subject 1	8	7	6	6
Subject 2	7	6	2	1
Subject 3	12	10	11	2
Subject 4	5	11	9	10
Subject 5	6	5	6	0

References

Axelsson, E. L., & Horst, J. S. (2014). Contextual repetition facilitates word learning via fast mapping. *Acta Psychologica*, *152*, 95–99. https://doi.org/10.1016/j.actpsy.2014.08.002

Behrend, D. A., Scofield, J., & Kleinknecht, E. E. (2001). Beyond fast mapping: Young children's extensions of novel words and novel facts. *Developmental Psychology*, *37*(5), 698. https://doi.org/10.1037/0012-1649.37.5.698

Bloom, P. (2002). *How children learn the meanings of words*. MIT press.

Bourgoyne, A., & Alt, M. (2017). The effect of visual variability on the learning of academic concepts. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1568–1576. https://doi.org/10.1044/2017_JSLHR-L-16-0271

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433–436.

Brocard, S., Wilson, V. A., Berton, C., Zuberbühler, K., & Bickel, B. (2024). A universal preference for animate agents in hominids. *iScience*, *27*(6). https://doi.org/10.1016/j.isci.2024.109996

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). Glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9*(2), 378–400. https://doi.org/10.3929/ethz-b-000240890

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). MIT press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Curio, E., Ernst, U., & Vieth, W. (1978). Cultural transmission of enemy recognition: One function of mobbing. *Science*, *202*(4370), 899–901. https://doi.org/10.1126/science.202.4370.899

Déaux, E. C., Baumeyer, A., & Zuberbühler, K. (2021). Food calls enhance visual discrimination learning in chimpanzees (pan troglodytes verus). *Journal of Comparative Psychology*, *135*(3), 420. https://doi.org/10.1037/com0000286

DeKeyser, R. (2003). Implicit and explicit learning. *The handbook of second language acquisition*, 312–348. https://doi.org/10.1002/9780470756492

Deshpande, A., Van Boekholt, B., & Zuberbuhler, K. (2022). Preliminary evidence for one-trial social learning of vervet monkey alarm calling. *Royal Society Open Science*, *9*(8), 210560. https://doi.org/10.1098/rsos.210560

Dezecache, G., Bourgeois, A., Bazin, C., Schlenker, P., Chemla, E., & Maille, A. (2019). Orangutans' comprehension of zoo keepers' communicative signals. *Animals*, *9*(6), 300. https://doi.org/10.3390/ani9060300

Fugazza, C., Andics, A., Magyari, L., Dror, S., Zempléni, A., & Miklósi, Á. (2021). Rapid learning of object names in dogs. *Scientific reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-81699-2

Garcia, J., Kimeldorf, D. J., & Koelling, R. A. (1955). Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science*, *122*(3160), 157–158. https://doi.org/10.1126/science.122.3160.157

Halberda, J. (2006). Is this a dax which i see before me? use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive psychology*, *53*(4), 310–344. https://doi.org/10.1016/j.cogpsych.2006.04.003

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, *298*(5598), 1634–1636. https://doi.org/10.1126/science.1072702

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child development*, 1021–1034. https://doi.org/10.2307/1130543

Howard, L. H., Wagner, K. E., Woodward, A. L., Ross, S. R., & Hopper, L. M. (2017). Social models enhance apes' memory for novel events. *Scientific reports*, *7*(1), 1–7. https://doi.org/10.1038/srep40926

Jeong, H., Sugiura, M., Sassa, Y., Wakusawa, K., Horie, K., Sato, S., & Kawashima, R. (2010). Learning second language vocabulary: Neural dissociation of situation-based learning and text-based learning. *Neuroimage*, *50*(2), 802–809. https://doi.org/10.1016/j.neuroimage.2009.12.038

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping". *Science*, *304*(5677), 1682–1683. https://doi.org/10.1126/science.1097859

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kano, F., & Call, J. (2014). Cross-species variation in gaze following and conspecific preference among great apes, human infants and adults. *Animal Behaviour*, *91*, 137–150. https://doi.org/10.1016/j.anbehav.2014.03.011

Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas (Gorilla gorilla), orangutans (Pongo abelii), and humans (Homo sapiens): Unique eye-viewing patterns in humans among hominids. *Journal of comparative psychology*, *126*(4), 388. https://doi.org/10.1037/a0029615

Kano, F., & Tomonaga, M. (2011). Species difference in the timing of gaze movement between chimpanzees and humans. *Animal cognition*, *14*, 879–892. https://doi.org/10.1007/s10071-011-0422-5

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

León, J., Thiriau, C., Bodin, C., Crockford, C., & Zuberbühler, K. (2022). Acquisition of predator knowledge from alarm calls via one-trial social learning in monkeys. *Iscience*, *25*(9). https://doi.org/10.1016/j.isci.2022.104853

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*(6619), 813–815. https://doi.org/10.1038/385813a0

Niehorster, D. C., Andersson, R., & Nyström, M. (2020). Titta: A toolbox for creating psychtoolbox and psychopy experiments with tobii eye trackers. *Behavior research methods*, *52*, 1970–1979. https://doi.org/10.3758/s13428-020-01358-8

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6318–6323. https://doi.org/10.1073/pnas.1300337110

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897x00366

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child development*, *69*(2), 309–320. https://doi.org/10.1111/j.1467-8624.1998.tb06190.x

Takagi, S., Koyasu, H., Nagasawa, M., & Kikusui, T. (2024). Rapid formation of picture-word association in cats. *Scientific Reports*, *14*(1), 23091. https://doi.org/10.1038/s41598-024-74006-2

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development*, 1454–1463. https://doi.org/10.2307/1130423

Twomey, K. E., Ma, L., & Westermann, G. (2018). All the right noises: Background variability helps early word learning. *Cognitive Science*, *42*, 413–438. https://doi.org/10.1111/cogs.12539

Vernes, S. C., Janik, V. M., Fitch, W. T., & Slater, P. J. (2021). Vocal learning in animals and humans. https://doi.org/10.1098/rstb.2020.0234

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, *108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in psychology*, *3*, 46. https://doi.org/10.3389/fpsyg.2012.00046

Wilson, V. A., Bethell, E. J., & Nawroth, C. (2023). The use of gaze to study cognition: Limitations, solutions, and applications to animal welfare. *Frontiers in Psychology*, *14*, 1147278. https://doi.org/10.3389/fpsyg.2023.1147278

Wilson, V. A., Sauppe, S., Brocard, S., Ringen, E., Daum, M. M., Wermelinger, S., Gu, N., Andrews, C., Isasi-Isasmendi, A., Bickel, B., et al. (2024). Humans and great apes visually track event roles in similar ways. *PLoS Biology*, *22*(11). https://doi.org/10.1371/journal.pbio.3002857

2. The effect of visual context stability and semantics on the rapid acquisition of new mearnings

2.1. Research article

Status: Submitted on December 11th 2024 to the *Journal of Child Language*

Influence of visual context stability on word learning in 14- and 19-month-old children

Dahliane Labertoniere[a] Géraldine Jean-Charles[a] & Katrin Skoruppa[a]
[a] Institute of Speech and Language Therapy, University of Neuchâtel, Neuchâtel, Switzerland

Adress for correspondence

Correspondence to Dahliane Labertoniere: dlab101@proton.me

Conflict of Interest

The authors have no conflicts of interest to declare.

Ethics statement

The research was evaluated as unobjectionable by the local ethics committee on research involving humans (CER-VD, Commission cantonale d'éthique de la recherche sur l'être humain du canton de Vaud, Req-2020-00065).

Data Availability Statement

Data will be made available on request.

Acknowledgements

We thank all of our participants and their families for taking the time to participate in this study. Special thanks go to Simone Marty for the recordings and to Erik Ringen for the statistical advice. Many thanks to the students from the 'TP recherche en logopédie' for their help with the pilot experiment and to Katia Damache and Mathilde Wenger for their help with the recruitment. This research was made possible by funding from the Institute of Speech and Language Therapy and the University of Neuchâtel.

Abstract

When hearing a label for an object, toddlers have to encode the visual context surrounding it. Our study investigates the role of the variability of this context during fast mapping in young children. Specifically, we compare word learning in French-learning fourteen- and nineteen-month-olds (N = 41) using visually distinct and identical object pictures in a fast mapping eye-tracking paradigm. The results show a learning effect only in the visually distinct condition. This suggests that toddlers benefit from a variability in visual context during word learning in this crucial developmental period of early lexical acquisition.

Keywords

Fast mapping, Word learning, Toddlers, Eye-tracking, Context, Language development

2.1.1. Introduction

Despite the inherent difficulties in acquiring a language, children develop a surprisingly large vocabulary in just a few years (an estimated average of 40,000 words at ten years for English-speaking children, cf. Bloom, 2002). This is why the existence of a default mechanism that permits children to rapidly guess the meaning of a novel word has been proposed (Carey, 1978) and investigated (Behrend et al., 2001; Heibeck & Markman, 1987). This mechanism, which constitutes the first step in the acquisition of new meanings (before retention and generalisation), is most often called fast mapping, i.e. the mapping of a form (e.g. a sound) onto a concept.

Anecdotal evidence suggests that infants may at first have difficulties generalising the meaning of a new word to different situations. The first words that children use seem to occur only in the particular context in which the word was first produced, at least for the first occurences of that word: Bloom (2002) gives an example of a child only saying the word 'car' when watching cars from the window, while Barrett (1986)'s child begins using the word 'duck' in contexts other than the one where the word was first produced (bathing) only two weeks after first production. These words – for which the children seem to have inferred a particularly narrow meaning – are used by the caregivers in a variety of contexts (Hoff, 2013). Therefore, the fact that children's first words are context-bound is not entirely linked to the input they receive (the most frequent context in which the caregiver uses the word is generally the one where the child will say it as well, cf. Harris et al., 1988), but seems to be a characteristic of early word learning. Like adults (Dautriche & Chemla, 2014), infants seem to encode the situation in which they hear a word, but that context is at first too strongly associated with the word for it to be used in another situation.

The role of visual context on word learning has also been explored in experimental settings. In Axelsson and Horst (2014)'s experiment, three-year-old children were presented with names for novel objects in two different conditions. In the first condition, the children would see the novel object three times, always accompanied by the same two other (known) objects. In the second condition, the children would be exposed to the novel object three times as well, but accompanied by different pairs of objects each time. Children were better at learning the name of the target object in condition 1 than in condition 2. In Vlach and Sandhofer (2011)'s study, the experimenters used a similar word-learning paradigm, but the two conditions differed in that the novel object was either presented on the same background (a colored cloth) or on different backgrounds (different colored cloths) during learning. Again, stable context during learning seemed to help young children (two-and-a-half- to three-year-olds), although older children (three- to four-year-olds) performed better when context changed. However, these results contrast with those of Goldenberg and Sandhofer (2013), whose two-year-old subjects were unable to learn new words either in a context change condition (objects on different coloured cloths) or in an unchanging context condition. Only when the two contexts were interleaved (a few presentations of the novel object on the same background, a few on novel ones) did the children perform above chance.

The above studies present us with contradictory results, with some suggesting that invariance of visual context helps word-learning in young children, while varying contexts could help slightly older children, and the last one suggesting that a combination of variance and invariance could be the best learning aid. Furthermore, data concerning younger children is lacking,

especially since it has been shown that changing context could help retention in three-month-olds (Rovee-Collier & Dufault, 1991). The methods used in those studies, while designed to maximise children's performance, diverge considerably from a day-to-day learning situation, where background information might be more complex, for example. While this is typical in fast mapping studies (e.g. eye-tracking experiments presenting objects on a grey background as in Schafer & Plunkett, 1998), we feel that using a more naturalistic approach to better reflect on children's naturalistic abilities is important.

In the present study, we seek to remedy these biases and find out if fast mapping is still possible when learning the meaning of a new word in a more ecological setting, using pictures of real objects presented on a naturalistic background, to approach a real-life word-learning situation. Our experiment was designed to assess the influence of visual context in word learning in very young children (thus filling the age gap in the literature), from fourteen month-old, in the beginnings of production, to nineteen month-old, when children are just entering the vocabulary spurt phase (Bloom, 2002). We use eye-tracking to explore whether these children need as few visual changes as possible when learning a new word, or if on the contrary their ability to learn is enhanced when context changes during learning. We also examine the fast mapping process in further detail by comparing performance after three (a minimal number for word learning, see Kan & Windsor, 2010) and six exposures to the novel sound-object associations. We expect our infants to perform better in the unchanging context condition, as they are closer in age to Vlach and Sandhofer (2011)'s younger children, and that our oldest group might outperform the youngest one in the changing context condition. Since increasing the number of exposures to a novel association should reduce the difficulty in learning and retaining that association, we also expect that our subjects will show stronger learning effects after six exposures than after three.

2.1.2. Method

Participants

We tested 41 French-learning infants (with at least 50% exposure to French), a sample size that we determined suitable by comparing it to the mean number of participants in the eye-tracking datasets on young children's word recognition present in the peekbank database (excluding two big datasets with over 250 participants, the mean number of participants was 44 as of 10/12/24, Zettersten et al., 2023). Their ages ranged from 1;1.27 (27 subjects tested around 1;2.0, 13 F) to 1;8.16 (14 subjects tested around 1;7.0, 3 F), with a mean age of 1;4.3. We asked that participants present no developmental delay, hearing or visual impairments or other strong disabilities, and checked that they were born full-term. Data collection occurred in Switzerland where data on ethnicity is not usually collected. Nine additional infants were tested but excluded from the analyses because they did not complete at least three blocks of the experiment (7) or because of data loss due to a computer crash (1). The participants were recruited via advertisement on social networks and university mailing lists. A consent form was filled by their caretakers before the beginning of the study.

Stimuli

We used a total of 8 different yoked pairs of pictures and 8 different yoked pairs of (pseudo)words.

Pictures were photos of objects (exotic fruits) on four different naturalistic backgrounds (sand, leaves etc.). Exotic fruits were chosen so as to be unfamiliar to the children and interesting enough. For each pair, fruits were chosen to be maximally different in size, colour and shape. An example of a pair of objects with the four different backgrounds can be seen in Figure 2.1.

The sound stimuli were maximally different CV/CVCV (pseudo)words following French phonotactic rules, produced by a native speaker (a different token of the same sound was used for each of the four instances of labelling in a trial). For each pair, each word was different in number of syllables and as phonetically distant as possible from the other. See Table 2.1 for a list of all paired (pseudo)words.

Table 2.1.: Pairs of (pseudo)words used in the experiment

/if/	/zøzø/
/fa/	/ømø/
/yʃ/	/vuvu/
/za/	/ovo/
/un/	/ʃoʃo/
/fy/	/eve/
/us/	/lele/
/mi/	/asa/

The order of presentation of the stimuli and their sound/image associations was pseudo-randomised in four groups.

Procedure and apparatus

Our subjects were tested in their caretakers' presence at our Babylab in Neuchâtel. Infants sat on their caretaker's lap in a sound-attenuated, dimly lit room. In front of them, at a distance of approximately 60cm, a monitor (24", 1920 x 1080 pixels) displayed the experiment. The audio stimuli were presented via two loudspeakers (Fostex PM0.5n, 70W amplifier) hidden under the table suporting the monitor. The experimenter monitored the progress of the study from outside the booth via video feedback. The experiment could be paused or terminated in case of discomfort from the child. Under the monitor, a Tobii eye-tracker (Tobii Pro X3-120, 120Hz sampling rate) recorded the eye movements of the participants. To ensure that the gaze measured was that of the infants and not their caretakers', the latter were asked to close their eyes or look at a fixation cross in a corner of the booth during the experiment. This, combined with worn headphones emitting a constant stream of background conversation, reduced the amount of unconscious influence that the caretaker could have on the child.

Once the participant was in place, a manually validated 5-point calibration with pulsating dots was carried out. During the experiment, we presented infants with eight blocks of two novel object-(pseudo)word associations in the following two conditions:

1. **Context-change condition**: in both the learning phase and the test phase, the target object is presented at a different angle and on a different background (see Figure 1.1) – four blocks

2. **Invariant condition**: no background or angle change – four blocks

The order of the blocks was pseudo-randomised between the four groups so that each subject got no more than two blocks back to back in the same condition. One block consisted of a learning phase, a test phase, another learning phase and a final test phase. During the first learning phase (the top part of Figure 2.2), each object was presented three times with its label (500ms after image onset, without a carrier sentence). During the following test phase (the bottom part of Figure 2.2), the two novel objects appeared side-by-side with one (pseudo)word in two trials. Learning and test phase were then repeated with a change in the order of the stimuli so that infants got up to six exposures to the novel associations per block (there were two test trials for each test phase, and four test trials total for each pair of objects). In a test trial, two phases of interest were defined: the pre-naming phase (duration: 2500ms), before labelling, and the post-naming phase (2500ms as well) after onset of the label. Between the two phases, an abstract geometrical figure appeared in the middle of the screen for 500ms to recentre participants' attention. Figure 2.2 recapitulates this procedure schematically for half of a block of the context-change condition.

The experiment, if completed in one go, lasted approximately ten minutes.

The script running the experiment was coded on Matlab version 2019b (The Math Works, Inc., 2019) using the Psychtoolbox (Brainard & Vision, 1997; Kleiner et al., 2007; Pelli, 1997) and the Titta toolbox (Niehorster et al., 2020) for implementing the stimulus presentation.

2.1.3. Results

A Matlab script was used to tranform raw data into files providing gaze hits in our two areas of interest, the left and right images in the test, every 8ms. Short tracking losses (no hit < 88ms) and blinks (no hit between 88ms and 304ms) were considered continuous gaze, whereas longer absences (no hit > 304ms) were considered an absence of gaze (Hollander & Huette, 2022). Trials with looking times deemed too low to be informative were filtered: we excluded trials where looking time towards the screen in the post-naming phase was less than 800ms or where looking time towards both objects in post-naming was less than 600ms. We then controlled for any already established baseline preference by excluding trials where children looked more than 90% towards one or the other object during pre-naming (as was done in Piot et al., 2024). In the end, a total of 214 trials were excluded from the analyses (18.5% of the initial 1154 trials). For pre- and post-naming phases, we computed the mean proportion of target looking (PTL) for each remaining trial. Figure 2.3 shows the mean PTL during a whole test trial (first part: pre-naming, second part: attention getter, third and fourth part: post-naming) for each condition. The third part of the graph represents the duration of 367ms which is necessary for infants and toddlers to program eye movements after hearing the target sound (Swingley & Aslin, 2000).

Data was analysed in R Studio (running R version 4.2.2). They were not normally distributed (Shapiro-Wilk test: $W = 0.9373$, $p < 0.001$) and followed a zero-one inflated beta distribu-

tion. After slight shrinkage (zeros were transformed to 0.0001 and ones to 0.9999) of the data, a generalised linear mixed model for beta distributions with a logit link was built using the glmmTMB package (Brooks et al., 2017). The model had PTL as the dependant variable, the variable subject as random effect, condition as a fixed effect and naming (pre- vs. post-.) as a nested fixed effect within condition. The factor condition was sum-contrasted. The equation for the model was

PTL ~ Condition / Naming + (1 | Subject)

The results are presented in Table 2.2 and indicate a significant main effect of naming (β = 0.180, SE = 0.089, p = 0.043) in the context-change condition only. Adding age group (14- vs.

Table 2.2.: Results of the model looking into the presence of a naming-effect in the two conditions

Predictor	Estimate	SE	*p*
(Intercept)	0.007	0.045	0.878
Condition	0.063	0.045	0.163
Condition1:Naming	0.180	0.089	**0.043**
Condition2:Naming	-0.058	0.090	0.523

Abbreviation: SE, standard error; Condition 1 = context-change, Condition 2 = invariant.

19-month-olds) and number of exposures (3 vs. 6) as fixed effects in interaction with naming did not improve the model.[16] We performed *post-hoc* Wilcoxon signed-rank tests for paired data with a Holm-Bonferroni correction for multiple testing (to avoid type I errors) on pre- and post-naming PTL, averaged by subject. We found a significant (p = 0.007) increase in PTL in post-naming corresponding to a naming effect in the context-change condition but not in the invariant condition (p = 0.418). A graphical overview of the data (with PTL averaged by subject) showing comparisons between pre- and post-naming phases of both conditions can be found in Figure 2.4.

2.1.4. Discussion

The present findings confirm that fast mapping is possible in an experimental setting with complex stimuli approaching real-life variability. However, unlike traditional word-learning paradigms (e.g. Schafer & Plunkett, 1998), our study seems to have been particularly difficult for children, as they showed a significant learning effect only in one condition. Like the older group in Vlach and Sandhofer (2011), our infants benefited from the change in context during learning, and could generalise to another context at the time of testing. However, their participants were older than ours (three- to four-year-olds vs. 14-19-month-olds) and the group closest in age to ours (two-and-a-half- to three-year-olds) actually performed better in a stable context. Maybe the type of variability and the methods could explain these diverging results. Our study used an indirect method (eye-tracking), while theirs used a direct one (forced choice task). As direct methods are more difficult for children, maybe different mechanisms are involved in indirect tasks. Interestingly, in young adults, variability in the visual input during learning facilitates

[16] Exploratory analyses indicated that the only factor that could improve the model was adding test number as a nested fixed effect within condition in interaction with naming, with a larger effect observed in the second test trials, probably due to mutual exclusivity.

the generalisation and acquisition of novel concepts (Bourgoyne & Alt, 2017). Alternatively, it could be that during the period of the naming explosion, context variance helps when learning new words; children would then benefit from invariance after their second birthday then once again from variance after three (but this goes against the proposal that context effects are linked to memory development, see Vlach & Sandhofer, 2011).

In any case, our toddlers' being able to learn the new words after three exposures heightens the robustness of these results, as this is a relatively low number (Kan & Windsor, 2010). An explanation for the positive effect of context change on learning could be that our context-change condition, unlike the invariant condition, resembles word learning tasks with object manipulations (even though the toddler cannot touch the novel object, they can see it under multiple angles) and thus is more akin to a real word learning situation. It is closer to what a child could experience during a self-produced action, and could activate the same neural sensory-motor pathways. This more embodied attention might be beneficial (Yu & Smith, 2012). A related explanation could be that this condition is better at maintaining toddlers' attention throughout the task. However, we found no difference between the conditions when comparing total looking time towards the screen during either the learning or the test phase.

The fact that our participants could not fast map sounds to objects in the invariant condition is somewhat surprising. However, several factors could explain this. First, our study was particularly difficult as the words were presented without a carrier sentence or interjections to help get the child's attention. The toddlers, already familiar with the syntactic structure of their native language, could thus not extract syntactic cues to infer that the target sound was a noun for example. Second, our stimuli, even though they were chosen to be maximally different from one another, all pertained to the same semantic category (fruits). Toddlers are capable of encoding the visual similarity of novel words (and thus a facet of the semantic relationship between them, see Wojcik & Saffran, 2013) and it is possible that learning new names for 16 semantically related objects in such a short time might have posed a difficulty for our participants. Even when objects belong to different categories, learning and retaining multiple associations is a difficult task (Axelsson & Horst, 2014). Our infants had to learn and retain new labels for two novel objects before beeing tested on them. Moreover, Goldenberg and Sandhofer (2013)'s children were also unable to learn new words in an invariant condition. Their study showed that presenting children with objects both in a changing and unchanging context was the best learning aid. A follow-up study could investigate if this remains true for our age range as well, using our paradigm. Finally, a problem present in many fast mapping eye-tracking studies is that they are blind to the interactive component of language learning. Indeed, joint attention with a caretaker has a positive influence on the number of words learnt by infants (Tomasello & Farrar, 1986). Even though our infants were seated with their caretaker in the booth, they could not interact with them (this was discouraged before the start of the experiment, so as not to bias results), and the caretaker did not share the child's attention (being blind to the images and sounds being presented). Despite this, many of our participants tried to get their caretaker's attention during testing, and failed to establish lasting contact. This in turn, increased the difficulty of the task.

Nevertheless, our study confirms that visual context plays a crucial role during the fast mapping process of early word learning. We suggest that further studies delve deeper into the subject, particularly with toddlers between one and three years of age, and with varied methods,

to verify that variance during learning is beneficial to children. Particular attention should be paid to the type of context change used during learning. In our study, we chose to maximise the variability by changing object orientation and background. It would be interesting to see if only changing the orientation of the object (thus mimicking object manipulation word-learning paradigms) while keeping the same background would be enough of a change to elicit the facilitating effect on word learning. Our study shows that using naturalistic methods is important in acquisition research. While our method only approached real-life situations of learning, the use of other methods like head-mounted paradigms to study the role of visual context might prove fruitful.

References

Axelsson, E. L., & Horst, J. S. (2014). Contextual repetition facilitates word learning via fast mapping. *Acta Psychologica*, *152*, 95–99. https://doi.org/10.1016/j.actpsy.2014.08.002

Barrett, M. D. (1986). Early semantic representations and early word-usage. In *The development of word meaning: Progress in cognitive development research* (pp. 39–67). Springer. https://doi.org/10.1007/978-1-4612-4844-6_2

Behrend, D. A., Scofield, J., & Kleinknecht, E. E. (2001). Beyond fast mapping: Young children's extensions of novel words and novel facts. *Developmental Psychology*, *37*(5), 698. https://doi.org/10.1037/0012-1649.37.5.698

Bloom, P. (2002). *How children learn the meanings of words*. MIT press.

Bourgoyne, A., & Alt, M. (2017). The effect of visual variability on the learning of academic concepts. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1568–1576. https://doi.org/10.1044/2017_JSLHR-L-16-0271

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433–436.

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). Glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9*(2), 378–400. https://doi.org/10.3929/ethz-b-000240890

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). MIT press.

Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 892. https://doi.org/10.1037/a0035657

Goldenberg, E. R., & Sandhofer, C. M. (2013). Same, varied, or both? contextual support aids young children in generalizing category labels. *Journal of Experimental Child Psychology*, *115*(1), 150–162. https://doi.org/10.1016/j.jecp.2012.11.011

Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language*, *15*(1), 77–94.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child development*, 1021–1034. https://doi.org/10.2307/1130543

Hoff, E. (2013). *Language development*. Cengage Learning.

Hollander, J., & Huette, S. (2022). Extracting blinks from continuous eye-tracking data in a mind wandering paradigm. *Consciousness and Cognition*, *100*, 103–303. https://doi.org/10.1016/j.concog.2022.103303

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Niehorster, D. C., Andersson, R., & Nyström, M. (2020). Titta: A toolbox for creating psychtoolbox and psychopy experiments with tobii eye trackers. *Behavior research methods*, *52*, 1970–1979. https://doi.org/10.3758/s13428-020-01358-8

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897x00366

Piot, L., Chen, H., Picaud, A., Dos Santos, M., Granjon, L., Luo, Z., To, A. W. H., Lai, R. Y., Cheung, H., & Nazzi, T. (2024). Tonal interference in word learning? a comparison of cantonese and french. *Journal of Experimental Child Psychology*, *242*, 105883. https://doi.org/10.1016/j.jecp.2024.105883

Rovee-Collier, C., & Dufault, D. (1991). Multiple contexts and memory retrieval at three months. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *24*(1), 39–49. https://doi.org/10.1002/dev.420240104

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child development*, *69*(2), 309–320. https://doi.org/10.1111/j.1467-8624.1998.tb06190.x

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166. https://doi.org/10.1016/S0010-0277(00)00081-0

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development*, 1454–1463. https://doi.org/10.2307/1130423

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, *108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011

Wojcik, E. H., & Saffran, J. R. (2013). The ontogeny of lexical networks: Toddlers encode the relationships among referents when learning novel words. *Psychological science*, *24*(10), 1898–1905. https://doi.org/10.1177/0956797613478198

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Zettersten, M., Yurovsky, D., Xu, T. L., Uner, S., Tsui, A. S. M., Schneider, R. M., Saleh, A. N., Meylan, S. C., Marchman, V. A., Mankewitz, J., et al. (2023). Peekbank: An open, large-scale repository for developmental eye-tracking data of children's word recognition. *Behavior Research Methods*, *55*(5), 2485–2500. https://doi.org/10.3758/s13428-022-01906-4

Figures

Figure 2.1.: One pair of objects in all possible contexts

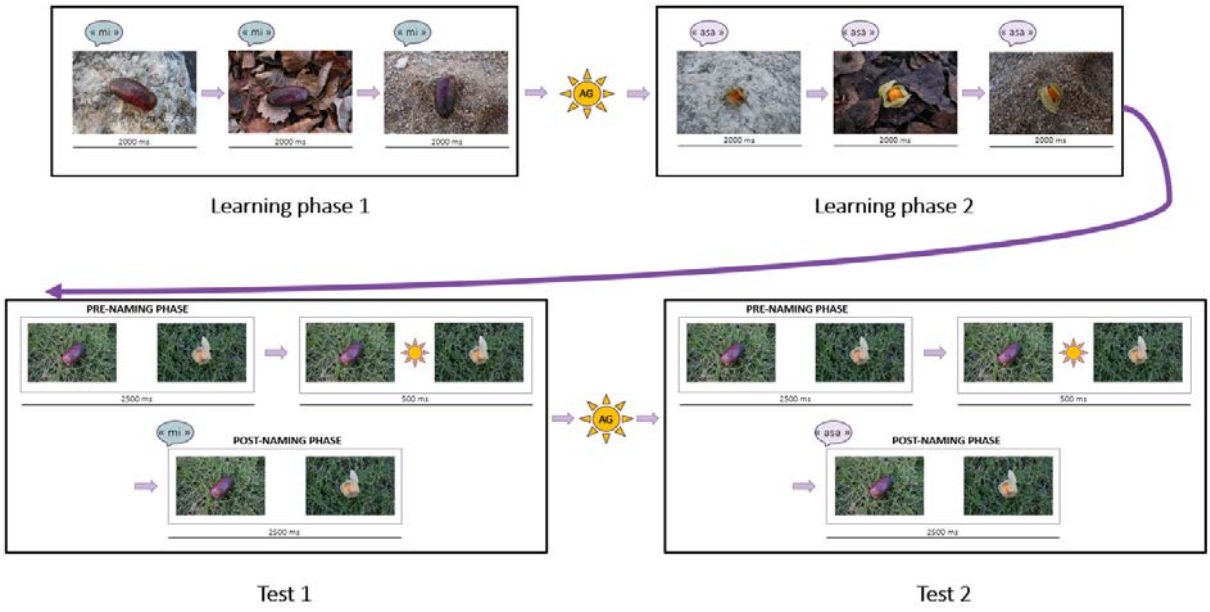
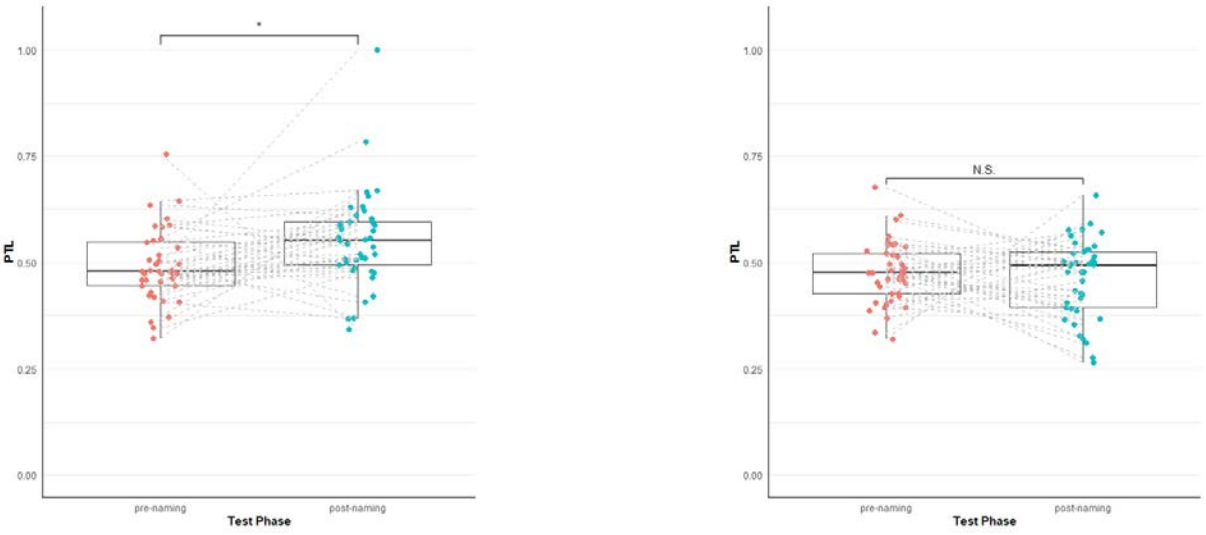


Figure 2.2.: Procedure of a context-change trial for three exposures: one half of a block

Figure 2.3.: Mean proportion of target looks over the whole test trial in both conditions (± SE)



(a) (b)

Figure 2.4.: Gaze proportions in pre- and post-naming phases in the context-change condition (a) and invariant condition (b)

2.2. A note on the ecological backgrounds

2.2.1. Introduction

During the preliminary discussion of the results of the experiment on the effect of context, we hypothesised that children's performance in the context-change condition was better than in the invariant condition because the change in background and in object orientation during learning helped children extract the object from the complex ecological background and thus supported the acquisition of novel associations. To test this hypothesis, we designed an experiment similar in all respects to the previous one (which we will now call Experiment 1), except that we tried to facilitate the extraction of the object in the invariant condition by replacing the complex ecological background with a grey background similar to those used in usual word-learning experiments (see for example Schafer & Plunkett, 1998). We expected that if we could observe a learning effect in this grey-invariant condition (while we could not in the invariant condition of Experiment 1), this would confirm the hypothesis that background change and change in orientation of the object helped children extract the object and associate the pseudoword with it. It would suggest that children failed the fast mapping task in the invariant condition of Experiment 1 because they had difficulties in extracting the object from the complex background.

2.2.2. Method

Participants

Seventeen 16 to 23 month-old French-learning infants (age range: 1;4.1–1;11.13, mean age 1;7.12, 10 F, with at least 50% exposure to French) participated in this experiment at our Babylab in Neuchâtel. We were aiming to test at least 22 infants in total, but did not find enough participants in time. The infants we recruited did not present any developmental delay or hearing or visual impairments and were born full-term. The caretakers filled out a consent form before the beginning of the study.

Stimuli and Procedure

The procedure, apparatus and stimuli were exactly the same as in Experiment 1 except that for the invariant condition, we replaced the ecological background with a simple, grey background. An example of a pair of objects in this condition can be found in Figure 2.5.

Figure 2.5.: One pair of objects on a grey background in the invariant condition

2.2.3. Results

As previously mentioned, we could not recruit enough participants to reach our desired sample size, but we will still report here the results of the data collected. The data were pre-processed as in Experiment 1. They were not normally distributed (Shapiro-Wilk test: $W = 0.921$, $p < 0.001$) and followed a zero-one inflated beta distribution. The results from a generalised linear mixed model similar to the one used in Experiment 1 are presented in Table 2.3. We did not find any significant effect of naming in either condition. To better visualise the data, we averaged PTL data in pre- and post-naming phases by subjects. The averaged data for both conditions can be found in Figure 2.6.

Table 2.3.: Results of the model looking into the presence of a naming-effect in the context-change and grey-invariant conditions

Predictor	Estimate	SE	*p*
(Intercept)	0.042	0.084	0.615
Condition	-0.007	0.074	0.919
Condition_Change:Naming	-0.039	0.148	0.790
Condition_Invariant:Naming	-0.017	0.146	0.908

Abbreviations: SE, standard error; Condition_Change = context-change condition, Condition_Invariant = invariant condition.

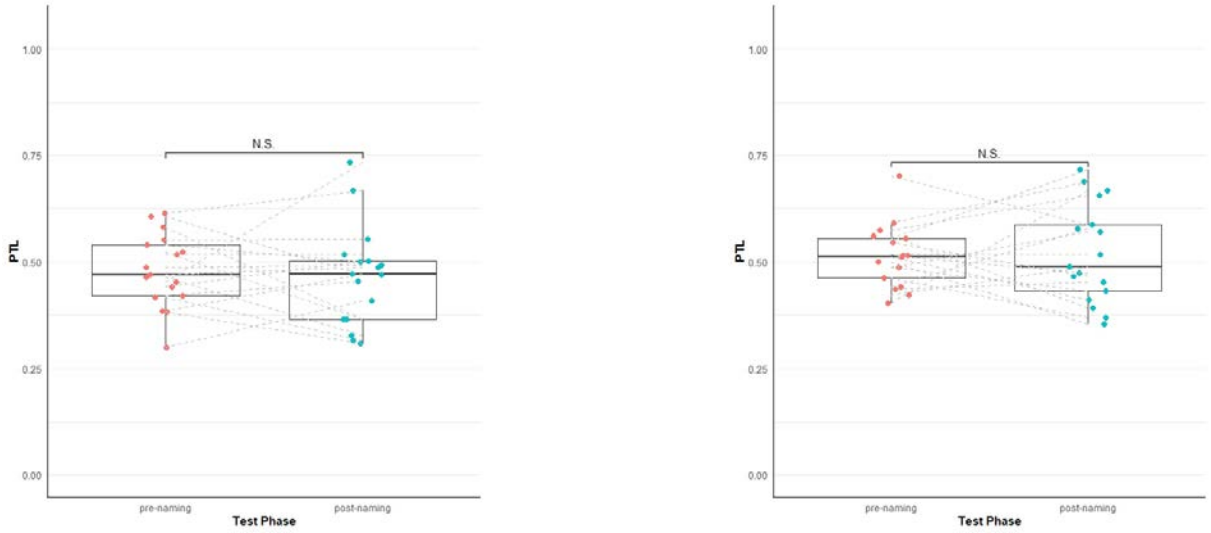


Figure 2.6.: Gaze proportions in pre- and post-naming phases in the context-change condition (a) and in the grey-invariant condition (b)

2.2.4. Discussion

As we could not test our desired sample, we will be prudent in interpreting the results. We did not replicate the results of Experiment 1 with a smaller sample: we could not find a learning effect in the context-change condition. This could be because of the rather small effect size found for the effect of context change in Experiment 1 (odds-ratio = 1.2), preventing us from

observing it with a smaller sample size. Unfortunately, we could not observe a learning effect either in the invariant condition. This indicates that even if an effect could be observed with a larger sample, the effect size should be small. These results suggest that the invariant condition with grey backgrounds was not particularly beneficial for children in that if children had been particularly supported in their learning of new associations by the grey background, we probably could have observed a learning effect even with a small sample. It is thus more likely that the positive effect on the learning of new associations observed for the context-change condition in Experiment 1 is due to other factors mentioned in the discussion of this experiment.

2.3. Semantics experiment: actions vs. objects

2.3.1. Introduction

This study, which was designed in parallel to Experiment 1, and which we will call Experiment 3, inquired into the role of semantics on learning. In particular, its purpose was to compare rapid word learning (fast mapping) of novel object words to the learning of words for novel actions. According to the natural partitions hypothesis (Gentner, 1982), verbs are less easy to learn than nouns because nouns correspond to a category of concepts such as objects and beings that are simpler or more basic than the category of concepts corresponding to verbs, such as activity, change of state or causal relations. In a cross-linguistic perspective, a language has less degrees of freedom in lexicalising objects than actions. For example, depending on the language, verbs can encode the manner of motion or not, the direction of motion or not etc. For an object, the same perceptual characteristics of an object will be encoded in all languages. In a study comparing 3- and 5-year-old children (Imai et al., 2005, 2008), the youngest group failed to fast map novel verbs to events, but succeeded in mapping novel nouns to novel objects; the older group, while succeeding in both conditions, had much more trouble with verbs. Their pseudowords were embedded in specific linguistic contexts which clearly identified them as verbs or as nouns. For example, a verb could be in progressive form and be placed between a subject and a complement (e.g. *the girl is blicking something'*). Adults are also better at guessing what word has been produced in a naturalistic situation when the sentence has been beeped out if the word refers to a concrete object (Medina et al., 2011).

To test whether the previous hypothesis is correct and whether at least part of what explains the tendency of nouns to outnumber verbs in children's productive lexicon (Gentner, 1982; Labertoniere & Skoruppa, 2022; Nelson, 1973) is linked to the perceptual differences in the concepts that they refer to, we want to compare the fast mapping of novel words onto novel actions and novel objects. For the same reasons as in Experiment 1, we wanted our eye-tracking task to approach a day-to-day learning situation as much as possible, so we used videos filmed in a natural environment as visual stimuli. Moreover, social information is important for learning, and joint attention in particular is thought to be critical when learning new words (Tomasello & Farrar, 1986). While it is difficult to create true joint attention in an eye tracking task, we tried to approximate it as much as possible by having the people in our object videos look at the target object. We also chose to use socials actions as our action stimuli. In an exploratory approach, we designed two kinds of actions: specifically 'transitive' actions (a person performing an action on another person) and 'continuous' or 'intransitive' actions (two people performing a joint action) to see if there could be a difference in the ease of associating new words to the two kinds of actions.

2.3.2. Method

Participants

Thirty-six French-learning infants aged between 1;1.27–2;1.0 (mean age 1;5.13, 19 F) participated in this experiment at our Babylab in Neuchâtel (one additional infant was tested but excluded from the analyses because they did not complete at least 3 blocks of the experiment). As in Experiment 1 and Experiment 2, the infants had at least 50% exposure to French and did not present any developmental delay or hearing or visual impairments and were born full-term.

Their caretakers filled out a consent form before the beginning of the study.

Stimuli

We used 8 different yoked pairs of videos and 8 different yoked pairs of pseudowords. The videos were either videos of two actors in a realistic scene (in a garden) sitting and looking at an object, or videos of two actors in the same scene performing an action. The actions chosen for this experiment follow the requirements listed below:

- They are all 'continuous' actions meaning that the action is performed in a continuous manner by the actors.
- They have no designated word to describe the action associated to it, in French at least.
- Half of the actions are 'transitive' (one of the actors is performing the action on the other), half are 'intransitive' (both actors are performing the action together).
- Half of the actions are performed in a sitting position, half are performed standing up.

As for the food objects, they were ('exotic') fruits that were chosen so as to not be familiar to Swiss children aged 20 months or below as well as to Basel zoo great apes, so that no word form was already associated with the fruit as a concept. All videos had the same duration of about 9 seconds. Example frames of one 'object' video (Figure 2.7) and two 'action' videos (one 'transitive': one agent, one patient – Figure 2.8, one 'intransitive': two agents – Figure 2.9) can be seen below.

Figure 2.7.: Still of an 'object' video

Figure 2.8.: Still of a trans. 'action' video

Figure 2.9.: Still of an intrans. 'action' video

Sound stimuli were maximally different CV/CVCV pseudo-words following French and Swiss-German phonotactic rules, produced by a native speaker of both languages. See Table 2.4 for a list of all paired sound stimuli used in this experiment.

Procedure and apparatus

The testing procedure and the experimental apparatus were similar to that of Experiment 1 and Experiment 2 with the following modifications. During the experiment, infants were presented with eight blocks of two novel referent–pseudoword associations in the following conditions:

1. **Object condition**: two actors, sitting on the ground, looking at a novel object (see Figure 2.7) – four blocks
2. **Action condition** : two actors performing a social action (see Figure 2.8 and Figure 2.9) – four blocks

During the learning phase, the infants watched a video in either condition and heard a pseudoword three times (the first time 500ms after video onset, the second at 3500ms and the last at 6500ms). This was the first association. They then saw the other (paired) video of the second association to be learnt. During the following test phase, screencaps of each video appeared side-by-side on the screen. The pre- and post-naming phases were similar to those of Experiments 1 and 2. If completed in one go, the experiment lasted a little over ten minutes.

Table 2.4.: Pairs of sound stimuli for Experiment 3

/af/	/mymy/
/ze/	/yʃy/
/yl/	/føfø/
/ʃaj/	/øzø/
/ol/	/vava/
/maj/	/ifi/
/øs/	/nini/
/ʃø/	/ulu/

Table 2.5.: Results of the model looking into the presence of a naming-effect in the action and object conditions

Predictor	Estimate	SE	*p*
(Intercept)	0.023	0.047	0.624
Condition	0.021	0.047	0.649
Condition_Act:Naming	0.072	0.093	0.435
Condition_Obj:Naming	0.007	0.096	0.944

Abbreviations: SE, standard error; Condition_Act = action condition, Condition_Obj = object condition.

2.3.3. Results

The data were pre-processed as in Experiment 1 and Experiment 2. They were not normally distributed (Shapiro-Wilk test: $W = 0.887$, $p < 0.001$) and followed a zero-one inflated beta distribution. The results from the same generalised linear mixed model as in previous experiments are presented in Table 2.5. We did not find any significant effect of naming in either condition. To better visualise the data, we averaged PTL data in pre- and post-naming phases by subjects. The averaged data for both conditions can be found in Figure 2.10.

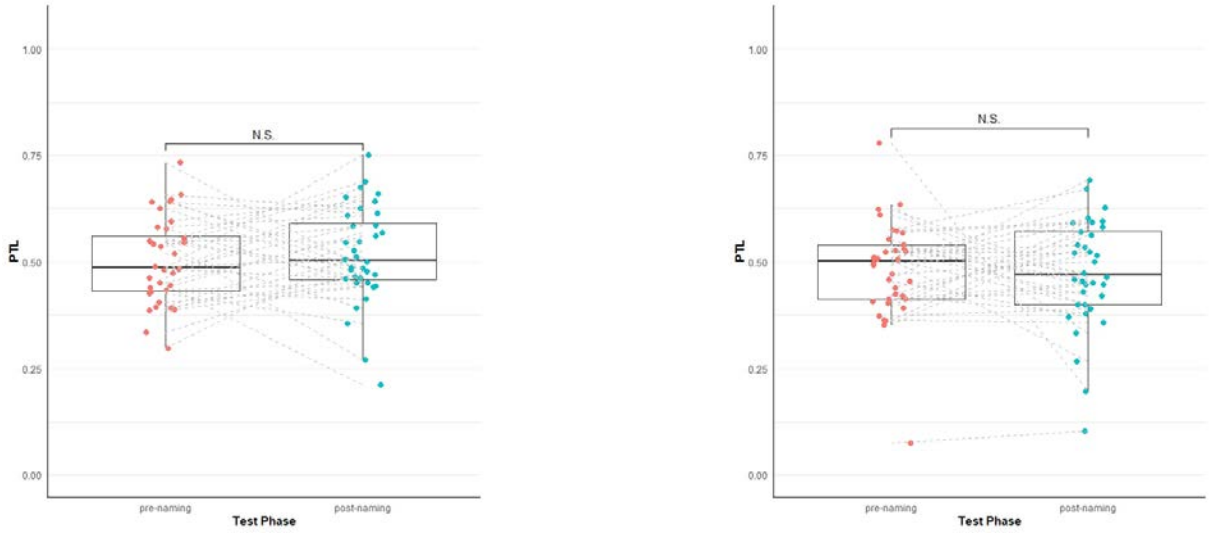


Figure 2.10.: Gaze proportions in pre- and post-naming phases in the object condition (a) and in the action condition (b)

2.3.4. Discussion

As we did not manage to observe a learning effect either in the object condition or the action condition, we can conclude that our experiment was very difficult for children. In addition to the same difficulties as Experiments 1 and 2 (complex background, pseudowords presented without a carrier sentence, lack of attention-getting interjections, lack of interaction and necessity to learn and retain two novel associations before being tested on them), several factors could explain these results.

In the object condition, due to the choice of having two actors present in the scene with the object, to better control for the similarity between the two conditions, the object occupied a very small portion of the screen. Even though a scene like the ones presented in our videos is closer to what a child would experience in real life, when a word is said by an adult and the child has to decide what its referent is, it is genuinely a complicated learning situation. A modeling study using data from head-mounted cameras on 18-month-olds suggests that an uncluttered visual scene was needed for infants to learn new words (Yu & Smith, 2012). Moreover, when they have to guess the meaning of a word uttered in a naturalistic scene between a caretaker and a child, and the linguistic context is blocked (the sentence and target word is replaced by a beep), adults perform poorly (Medina et al., 2011). We tried to mitigate this problem by having both actors look at the object for the duration of the videos, to simulate joint attention, but it seems this was not enough for our participants to extract the correct referent from the scene. Alternatively, it is possible that they were able to correctly infer that the pseudoword referred to the object in the middle of the screen, but that once the objects were presented side by side during testing, they were too small for the infants to correctly identify which object was which.

In the action condition, as discussed in the introduction, words for actions should be more difficult to learn than words for objects. Here, we believe that our actions were more salient than our objects, because they occupied more of the screen. Because of this, we believe that the failure of infants to fast map in this condition was not due to a limitation of the design like in

the object condition, but more to the nature of actions themselves.

In Waxman et al. (2009)'s study, 2-year-old infants successfully mapped novel nouns to objects and novel verbs to events, but as in Imai et al. (2008), the novel words were embedded in sentences that made their grammatical category salient. The participants exhibited a different behaviour based on the grammatical category of the novel word. When hearing novel nouns, they focused their gaze on objects, while when hearing verbs, they focused on actions. In another study, Oshima-Takane et al. (2011) found that novel verbs (pseudowords embedded in a sentence such that they were marked as verbs) appear to direct attention specifically to actions in 20-month-olds. This suggests that syntax might be necessary to the novel learning of verbs, and that is why our infants, who only had isolated words and no carrying sentence, failed in this condition.

Unfortunately, we cannot answer our primary research question which was to verify that words for objects were easier to learn than words for actions. We cannot say anything about the role of the transitive or intransitive nature of actions in learning either. While the choice of not using carrier sentences was motivated by the wish to test this paradigm with nonhuman apes, we would suggest future experiments maximise the potential performance of children by using as much linguistic and non-linguistic information as possible. We could also imagine zooming in on both actions and objects by having actors sitting at a table performing actions with less amplitude of movement, only using their hands in the action condition, and the object placed on the table between them. The framing would be very close to the actors and would not go lower than the table.

2.4. A note on human aduts

During the design of the previous experiments, we decided to run two pilot studies on human adults to verify that fast mapping tasks using the kinds of stimuli that we wanted to employ in the infant experiments were (easily) feasible by adults, thereby supporting the use of these stimuli for infants. The results of these studies are presented below.

2.4.1. First pilot

In a pilot experiment conducted in early 2020, 26 native French speakers between the ages of 18 and 54 were tested in France and Switzerland.

The experiment was designed as a forced-choice task (instead of measuring gaze data, we collected manual responses from the participants). Subjects accessed the experiment via the software Microsoft Powerpoint on a personal computer and wore headphones. They were instructed to choose in the test phases the stills that they felt was right. In this experiment, the effect of context change and of semantics were tested at the same time. The stimuli were videos similar to those used in Experiment 3, with two people either looking at an object or performing an action. In the context-change condition, the ecological background varied during learning and test, while in the invariant condition it stayed the same. The invariant condition blocks, which were deemed particularly easy for adults, were always put at the end of the experiment and were intended as a control to verify participants' ability to do the task. The total number of blocks for this experiment was 9 and sounds associated to the videos were pseudowords generated to follow French phonotactics, of either 1, 2 or 3 syllables, maximally different and with a CV structure, produced using child-directed speech.

Analysis of the results of this pilot experiment indicates that subjects excelled at the task (a clear ceiling effect was observable) and performance was similar across all videos.

2.4.2. Second pilot

A second adult pilot experiment was conducted in 2022. Participants were 39 French speakers (20F, age range: 18-59 years). This was again a forced-choice task. This time, the stimuli and procedure were the same as those of Experiment 1. The participants performed at ceiling, with a 98,23% success rate, therefore validating the stimulus material for the infant version. There was no effect of condition (context-change vs. invariant), number of exposures (3 vs. 6) or block number (no fatigue effect) on the scores.

Both of these pilots confirm that adults are very capable at fast mapping when given semi-explicit instructions, and when the measuring method is direct. They can easily learn to associate novel words with novel objects and novel actions, and are not impacted by the background behind the target referent, or the orientation of the object.

The visual stimuli for Experiment 1 and 2, an image of the setup with an infant and their parent, and stills of the videos used in Experiment 3 can be found in Appendix A.

3. Factors influencing the acquisition of new meaning in children

3.1. Research Article

Status: Published

A.N.A.E. ◆ 2022 - 181 - 001-017

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

D. Labertonière, 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 bais 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

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

Pour citer cet article : Labertonière, 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 ; DLPF ; 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[1], *pro-drop*[2]).

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 »[3]. 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 interjuge, 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*[4].

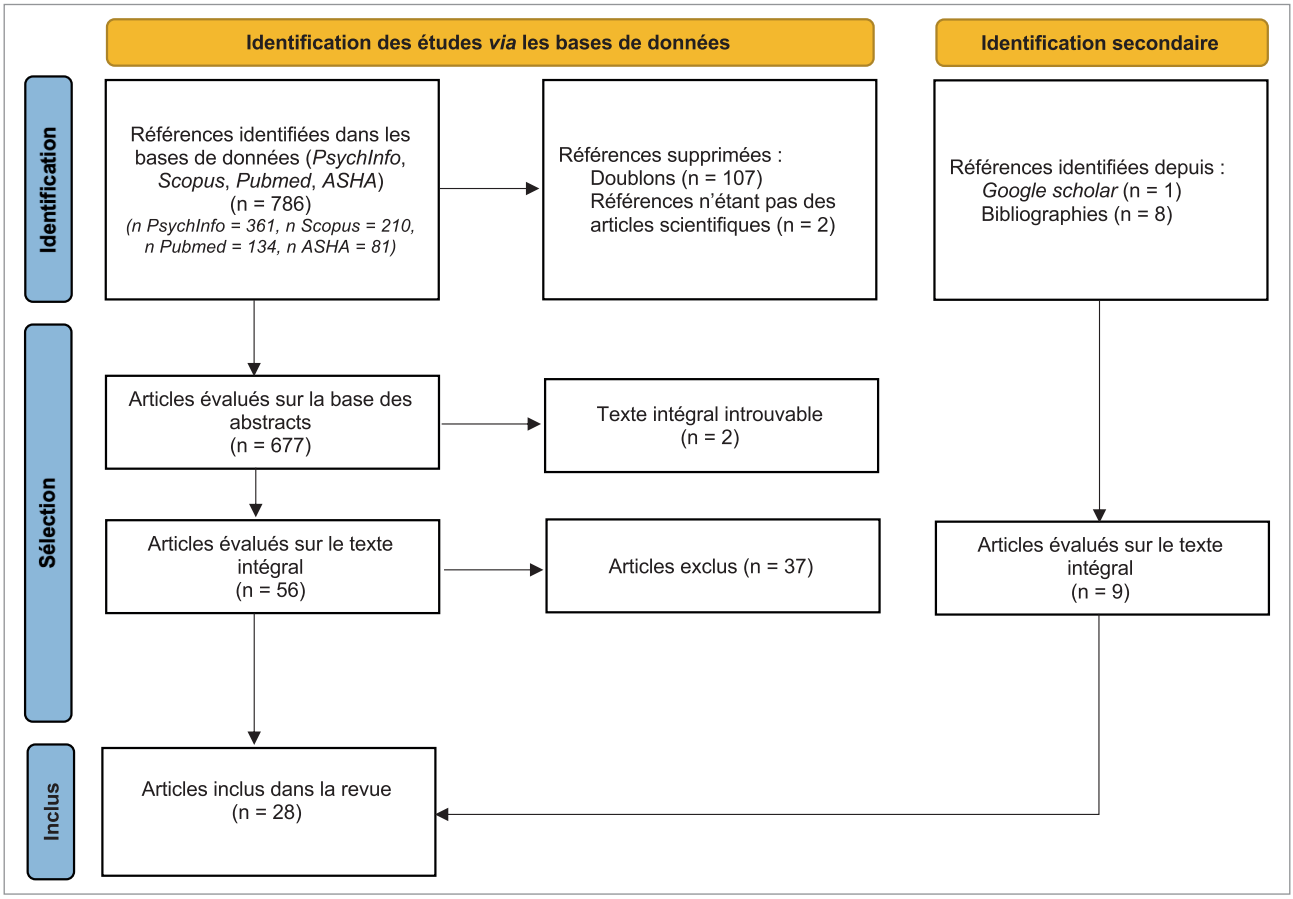
Pour un meilleur ciblage, nous avons également ajouté (pour les bases de données le permettant) des critères d'exclusion. Nous avons ainsi exclu les articles se concentrant sur une population trop jeune ou âgée (nourrissons et enfants d'âge scolaire) ou présentant des troubles développementaux et les 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.

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 réplicabilité méthodologique, l'absence de biais expérimental ainsi qu'à la présence d'analyses statistiques, ces trois critères étant éliminatoires si non remplis.

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*).*

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 productions 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.

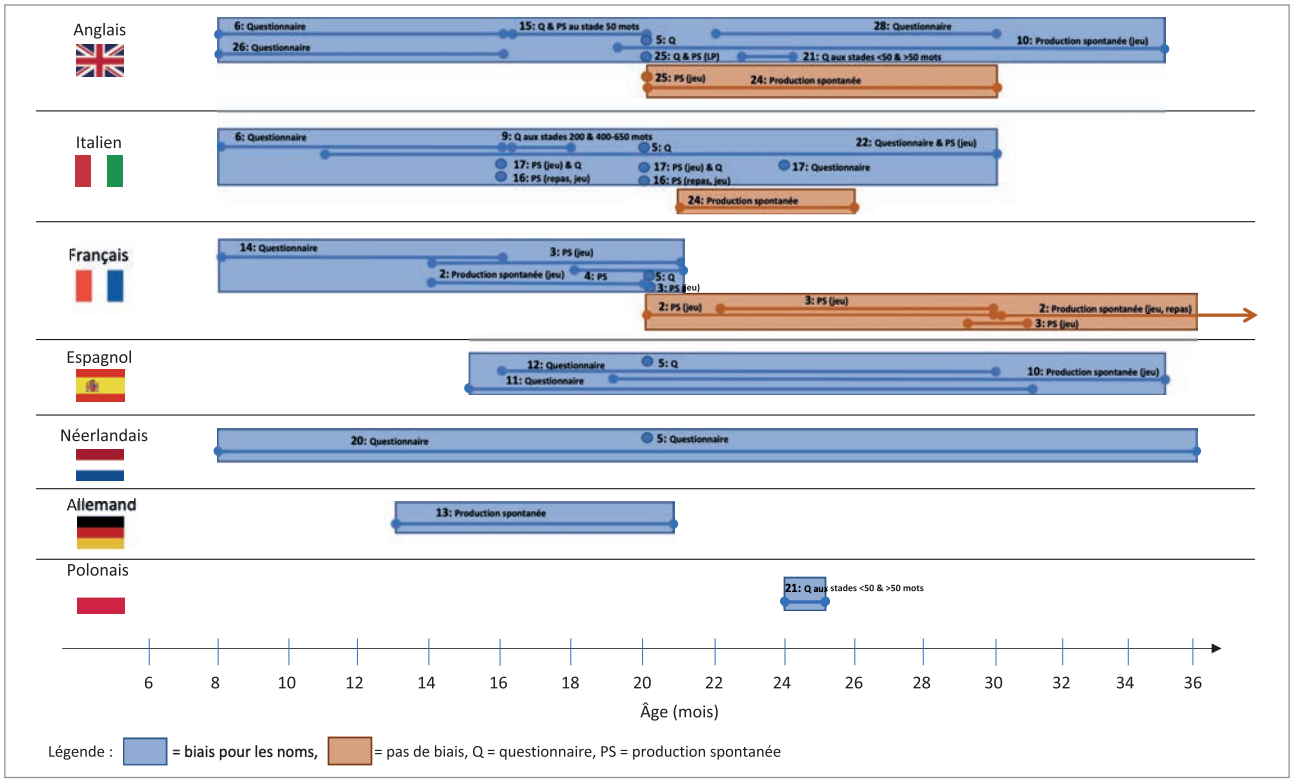
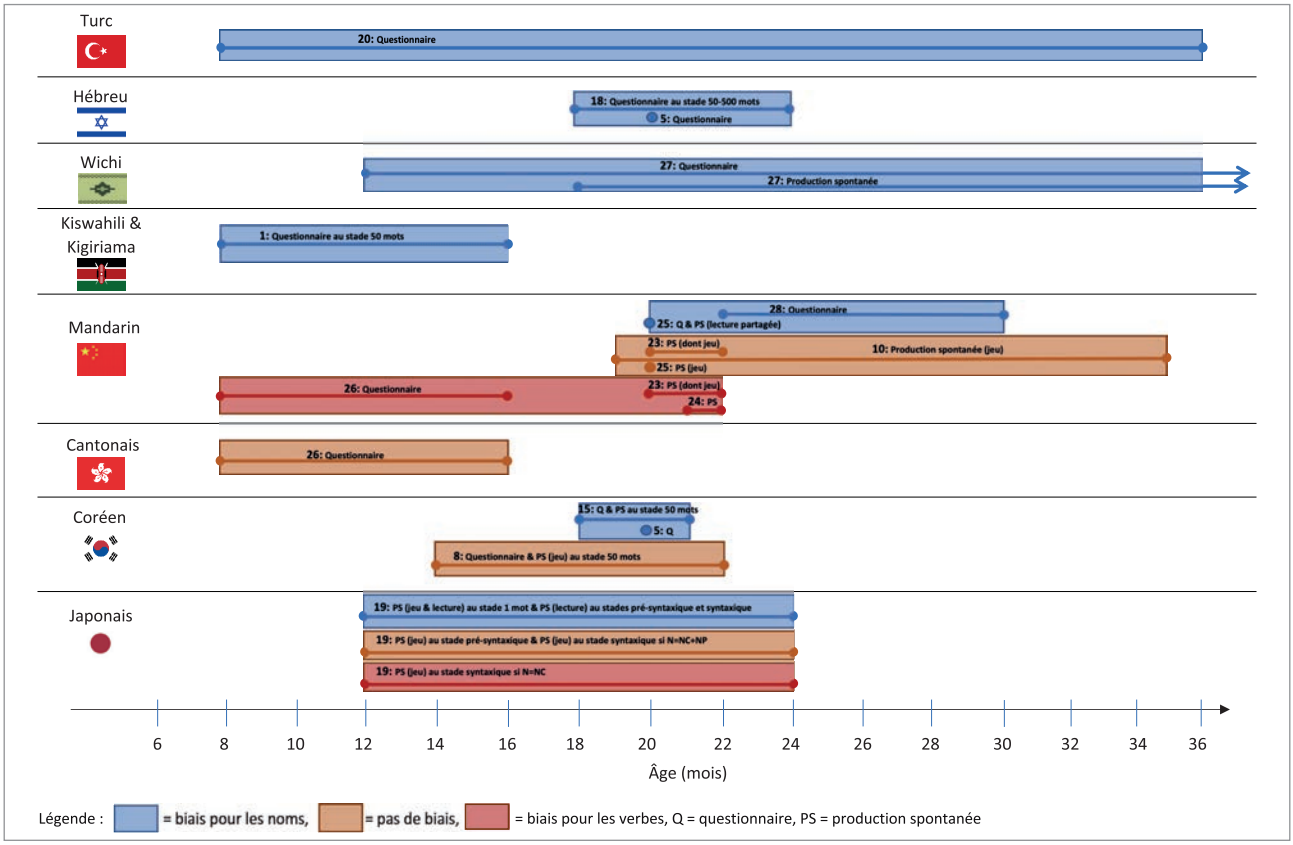


Figure 3. Frise développementale du biais pour les noms dans les langues non indo-européennes.



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 italophones, 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 italophones 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 kigiriama, 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 kigiriama, (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'autrice 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[5] 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

5 ● 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[6] 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 italophones 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 italophones. 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 italophones au stade 100 - 200 mots et 500 mots[7] (Bornstein *et al.*, 2004 ; Salerni *et al.*, 2007) et d'enfants coréanophones, anglophones, hispanophones, néerlandophones et hébréophones de 1;8 an au stade 100 - 500 mots (Bornstein *et al.*, 2004[8]). 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 italophones. Enfin, les adjectifs semblent très peu présents dans le lexique d'enfants italophones et anglophones 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[9] 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[10]) 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 enfantin, les noms communs ne sont pas toujours la classe dominante. Les catégories des noms de personnes et des onomatopées mais aussi les routines[11] 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 anglophones (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 italophones 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 influer 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 influer 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 influer 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 kigiriama, 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 influer 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[12].
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 tuteur·rice 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.

12 • 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 différentiée.

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 suffixe 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 »[13]. 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·rice·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·rice·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 structurelles 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 différentiellement 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[14], on compte précisément 419 noms[15] et 125 verbes[16] 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

13 • 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.

14 • Inventaires Français du Développement Communicatif (IFDC, Kern, 2007).

15 • Catégories 1 à 13 de l'IFDC 16 - 30 mois.

16 • Catégorie 15 (actions) et 18 (auxiliaires) de l'IDFC 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[17] (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·rice 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

17 • 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 professionnel·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., Maillochon, 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., Pêcheux, 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., Białecka-Pikul, M., Stępień-Nycz, M., & Ochał, 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 & kigiriama	questionnaire	N > V		résultats différents en compréhen-sion	
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 référent 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 catégo-risation 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 réfèrent 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 *et al.* (1993)	n = 68	1;3-2;7	monolingues espagnol	questionnaire	N > Pred			
12	Junyent *et al.* (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 *et al.* (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 *et al.* (2015)	n = 26	1;3-1;8	monolingues italien	production spontanée	N > V		production parentale	
17	Longobardi *et al.* (2017)	n = 30	1;2-2;0	monolingues italien	questionnaire & production spontanée	N > V			
18	Maital *et al.* (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 *et al.* (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 *et al.* (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 *et al.* (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 *et al.* (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 *et al.* (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 *et al.* (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 *et al.* (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 cantonais	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.

3.2. Corpus study: Temporal evolution of the composition of the productive lexicon of young children and effect of situational contexts

This study was conducted with the help of six students from the Master's programme *cursus Logopédie* at the University of Neuchâtel. It will be the object of a future publication in a scientific journal, and is now currently in revision, with further analyses being conducted. The authors are Dahliane Labertoniere, Camille Lefebvre, Amélie Steiner and Katrin Skoruppa. Nadia Besselma, Elisa Medici, Eleonora Vaccaro and Téa Waecker participated in the study but will not be full authors.

Abstract

The acquisition of the lexicon of a child's first language is marked by a fundamental period where the first words are produced. During the first three years of life, after a period of slow vocabulary growth until around 50 words acquired or around 18 months of age, an acceleration in the rhythm of acquisition can be observed, with some children showing a sudden increase, the *vocabulary spurt*, and some showing a more gradual one. By observing this period, we can document the nature of the first productions, the influence of activity contexts on them as well as the evolution of the lexicon's composition across time. In this work, we analyse video data of French-learning children aged 1;0 to 2;11 from the CHILDES database to explore 1) the role of activity context on the child's kinds of productions in terms of grammatical categories and semantic categories (like actions, objects and social terms); 2) the evolution of the same kinds of production across time. We report that shared-reading activities have a positive influence on the frequency of nouns in children's productions. Moreover, the frequency of nouns is higher than that of verbs in the whole developmental period observed, even though this effect lessens over time. The pattern looks to be the same for objects and actions, indicating a strong link between grammatical and semantic categories. Finally, we find that social terms are predominant in children's early productions.

3.2.1. Introduction

The first words a child produces appear around the start of their second year of life. For most children, the acquisition of novel words begins slowly: for several months, the rhythm of acquisition is around 10 new words per month. Around the 50 word cap (or at around 1;6) in some (but not all, see Ganger & Brent, 2004) children, this rhythm accelerates sharply in what is called the *vocabulary spurt* or *naming explosion* (Nelson, 1973). This acceleration could be caused by several factors. Bloom (2002) hypothesises that it is highly dependent on the context of production or activity context (see Tardif et al., 1999), that is to say the activity the child is engaged in when the production takes place. McShane (1980) proposes that children begin to understand denomination – the fact that every thing has a name – around this age. Others suggest that the naming explosion is linked to children's cognitive maturation; for example, children aged around 1;6 understand that objects can be grouped into categories (Gopnik & Meltzoff, 1986). Most likely, the development and improvement of several learning mechanisms from the moment the child produces their first words allows for a gradually better acquisition of new words (Hoff, 2013).

The composition of the first lexicon has been thoroughly studied in many different languages, in particular concerning grammatical categories. It has been shown that nouns comprise a major part of children's first lexicon, a phenomenon named the *noun bias*. This phenomenon, while it can be attenuated by certain typological characteristics of certain languages (called *verb-friendly*), seems to be universal (see Labertoniere & Skoruppa, 2022, for a review). The main explanation considers that concreteness or imageability is a major factor in the acquisition of novel words (see the natural partition hypothesis in Gentner, 1982). Nouns usually refer to concrete elements like objects or animate beings and their concepts are easier to grasp than verbs' which usually refer to actions, less stable in time, or mental states, which are not directly observable. Indeed, the more concrete verbs, action verbs, are the most used by children from 1;9 to 2;5 (Bassano, 2000; Bassano et al., 2005), and nouns referring to objects or animals and people are among the first words acquired (Bassano et al., 2005; Choi & Gopnik, 1995). Therefore, semantics certainly play a role in the noun bias in the early lexicon. As the child develops, the noun bias gradually decreases, with the frequency of verbs, then of adjectives, increasing, and function words appearing in the lexicon (Bassano, 2007; Bates et al., 1994; Labertoniere & Skoruppa, 2022). Between 2;6 and 3;0, the frequency of the different word classes in the lexicon stabilises (Bassano, 2000). Moreover, if we take a semantic perspective to look at children's very first productions, we can see that the predominant category is that of social terms, that is to say words produced in specific social contexts (e.g. words for social routines, onomatopeias, interjections, greetings, affirmative and negative particles) or referring to social beings (see Caselli et al., 1995; Kauschke & Hofmeister, 2002; Kern, 2007; Labertoniere & Skoruppa, 2022; Tardif et al., 2008). This highlights the importance of social information and social interactions in language acquisition. During development, the total number of social terms in children's productions remains stable, while other categories of words gradually expand. This means that the frequency of social terms decreases with time, which can be explained by the fact that there is a limited number of social terms, and many more predicates and function words to be acquired (Bassano et al., 2005; Kauschke & Hofmeister, 2002).

Lastly, the situational context or activities the child is engaged in at the time seems to have an influence on their productions. Indeed, activities like shared reading seem to promote the production of nouns by children (Ogura et al., 2006; Tardif et al., 1999), while activities like playing with toys seems to favour verbs (Salerni et al., 2007). This is also the case in adults as well, as caretakers produce more nouns during shared reading (Altınkamış et al., 2014; Ogura et al., 2006), and could influence children's productions in that direction, as in Altınkamış et al. (2014), where caretakers had a tendency to use more object-oriented utterances (compared to action-oriented ones) during shared reading and conversely during playing.

Few studies have been conducted on the role of situational or activity contexts on the productions of children, especially in an ecological setting. Moreover, most studies on the composition of children's early lexicon are based on data obtained via parental questionnaires (Caselli et al., 1995; Choi & Gopnik, 1995; Kern, 2007; Salerni et al., 2007; Tardif et al., 1999, 2008) or spontaneous productions during lab experiments (Choi & Gopnik, 1995; Kauschke & Hofmeister, 2002; Ogura et al., 2006; Salerni et al., 2007; Tardif et al., 1999), and semantic categories like objects, actions and social terms are not often taken into account. While the existing literature already gives us a fascinating picture of children's early productions, we want to expand this knowledge base by analysing naturalistic video data of children in real-life situations interacting

with their caretaker(s). To this end, we will work on two corpora in the CHILDES database: Lyon (Demuth & Tremblay, 2008) and Paris (Morgenstern, 2006). Firstly, we will document the evolution of different categories of words, from grammatical categories (nouns, verbs, adjectives...) to semantic categories (actions, objects and social terms) across an early developmental period: 1 year to 2 years and 11 months. Secondly, we will investigate the effect of situational context on children's productions, once again in terms of grammatical and semantic categories.

3.2.2. Method

Participants

We selected 10 French-learning children (6 F) between the age of 1;0 and 2;11 (mean age = 2;0) –to cover the crucial developmental period of the naming explosion– from the Lyon (Demuth & Tremblay, 2008) and Paris (Morgenstern, 2006) corpora in the CHILDES database (MacWhinney & Snow, 1985). All selected children were followed longitudinally by researchers, which means that we get data points at multiple developmental stages for each child, and incidentally, that they came from high socioeconomic backgrounds. The selection process followed the following rules: French as the first language spoken at home, video data available child productions already transcribed. Out of a total of 402 videos corresponding to these criterion, we selected 101 for the analyses so that we would have an approximately equivalent number of data points across the designated developmental period. Each video consist of an interaction between a child and their caregiver, lasting from around 30 minutes to one hour.

Coding

The 102 videos were distributed between 6 coders using the software CLAN. The coding consisted in adding information about the semantic categories used by the child and the situational context during production. Information concerning the grammatical categories of produced words was already available via the %mor line. For each utterance of the child, we added a dependent line %sem to code actions and objects, a dependent line %soc for social terms and a dependent line %ctx to code for context. In the semantics and social terms lines, we added the code for an action, an object or a social word respectively every time there was one in the utterance; in the context line, we added the code for the specific context in which the utterance occurred. The list of codes in each category (semantics, social terms and context) is available in Table 3.1.

Table 3.1.: List of codes

Dependent line	Code	Signification
%sem	ACT	action
	OBJ	object
%soc	MS	social word
%ctx	LP	shared reading
	JEU	play
	SNO	care

We used strict definitions for each coding category. Objects were defined as concrete elements that are perceptually visible (Abeillé & Godard, 2021, pp. 398–427) and separable from their environment (e.g. *chocolat* 'chocolate', *dessin* 'drawing', but not *mer* 'sea' would be included). They can also refer to living beings such as people (excluding proper nouns) and animals (e.g. *maman* 'mum', *pingouin* 'penguin'). Our definition for Action words was based on Abeillé and Godard (2021, pp. 180–204)'s dynamic verbs, which include activity verbs and verbs expressing that an event is taking place. For example, verbs such as *porter* 'carry' and *regarder* 'look' would be included but state verbs like *savoir* 'know' or *rester* 'stay' would not. Social words were defined as person nouns (e.g. *papa* 'dad'), first names, professional names (e.g. *nounou* 'nanny'), relational names (e.g. *sœur* 'sister'), onomatopoeias, interjections (e.g. *oh*), greetings (e.g. *coucou* 'hi'), social routines (e.g. *merci* 'thanks', *s'il te plait* 'please'), assertions (e.g. *oui* 'yes', *non* 'no') and cuddly toys (*le doudou*). As for the contexts, shared reading consisted in caretaker and child reading a book together, play contexts had the child playing with toys, tinkering or playing an instrument, and care consisted of moments like bathing, dressing and meal times. An example of a coded exchange between mother and child can be found in Figure 3.1.

```
*MOT:   tu t'assieds . •
%mor:   pron|toi-Prs-Acc-S2 pron|toi-Prs-Acc-S2~verb|asser-Fin-Ind-2-Pres .
%gra:   1|3|NSUBJ 2|3|OBJ 3|3|ROOT 4|3|PUNCT
*CHI:   non maman . •
%xpho:  /na~ mama~/
%mor:   adv|non noun|maman&Fem .
%gra:   1|2|ADVMOD 2|2|ROOT 3|2|PUNCT
%soc:   MS MS
%sem:   OBJ
%acl:   RD
%ctx:   JEU
```

Figure 3.1.: Extract showing the coding of a child's utterance from the Lyon corpus

3.2.3. Results

Data were extracted via CLAN and analysed in SPSS. Concerning the grammatical categories, we limited our analysis to nouns, verbs and adjectives. For each video, we calculated the frequency of each grammatical category, the frequency of each semantic category and the frequency of social words in the total number of words produced by the child. We repeated the process for each context, by calculating the frequency of each grammatical category, each semantic class and social terms in the total number of words produced. Videos were separated into one of three age groups, following the work of Bassano et al. (2005), for simplicity in the analyses. Group 1 was children before and at the very beginning of the naming explosion: 1;0-1;7. Group 2 consisted of children at the beginning of the naming explosion: 1;8-2;5. Group 3 consisted of children during the continuing naming explosion: 2;6-2;11. For each age group and for each context, we computed the mean frequency of each class of interest. We used Kruskal-Wallis tests to ascertain if there was a statistically significant difference in the frequency of relevant word classes between age groups or contexts, then, when appropriate, we used Mann-Whitney tests between age groups or contexts.

Temporal development

Grammatical categories: As we can observe in Table 3.2, in all age groups there were more nouns than verbs and more verbs than adjectives being produced. There was a significant difference in the frequency of nouns relative to other grammatical categories between the different age groups ($H = 18.22$, $p < 0.001$) and more specifically, a significant difference between group 2 and group 3 ($U = 138.0$, $p < 0.001$), indicating a significant decrease in the frequency of nouns between the second and third age groups. Concerning verbs, we found a significant difference in their frequency between the groups ($H = 32.89$, $p < 0.001$). More specifically, there was a significant difference between group 1 and group 2 ($U = 287.50$, $p < 0.001$), group 2 and 3 ($U = 280.50$, $p < 0.001$) and group 1 and 3 ($U = 160.0$, $p < 0.001$), indicating a continuous increase in their relative frequency. Finally, there was a significant difference in the frequency of adjectives across age groups ($H = 35.38$, $p < 0.001$), and more specifically, a significant difference between group 1 and 2 ($U = 203.0$, $p < 0.001$) and group 1 and 3 ($U = 124.0$, $p < 0.001$), indicating an increase in their frequency between the youngest age group and the two oldest. These results are summarised in Figure 3.2.

Table 3.2.: Total number of each word category across all age groups

	Age group					
Word class	Group 1		Group 2		Group 3	
	raw number	frequency	raw number	frequency	raw number	frequency
Nouns	790	27,9%	4562	25,4%	5598	16,0%
Verbs	252	8,9%	1850	10,3%	4527	13,0%
Adjectives	34	1,2%	544	3,0%	969	2,8%
Objects	406	14,4%	3062	17,1%	3624	10,4%
Actions	181	6,4%	1301	7,3%	2768	7,9%
Social words	969	34,3%	4036	22,5%	6106	17,5%
Total words	2827		17941		34922	

Group 1 = 1;0-1;7, Group 2 = 1;8-2;5, Group 3 = 2;6-2;11

Semantic classes: Looking at the raw data from Table 3.2, we see that there were more object words produced than action words across all age groups. A significant difference in the frequency of objects relative to other words was found between the different age groups ($H = 8.02$, $p = 0.018$). More specifically, there was a significant increase of their frequency between group 1 and 2 ($U = 435.0$, $p = 0.039$), then a significant decrease between group 2 and 3 ($U = 310.50$, $p < 0.001$). Concerning actions, we found a significant difference in the frequency of actions relative to other words between all age groups ($H = 14.26$, $p < 0.001$). More specifically, there was a significant increase in their frequency during development, between group 1 and 2 ($U = 373.0$, $p = 0.006$), between 2 and 3 ($U = 406.0$, $p = 0.026$) and between 1 and 3 ($U = 292.0$, $p < 0.001$). Figure 3.3 recapitulates these results. Moreover, there were many social words in children's productions: in age group 1 in particular their raw number exceeded that of the noun category in itself (remember that there are nouns in the social word class). More than

one third of children's productions in this group were social words (see Table 3.2). There was a significant difference between their frequency between the three age groups ($H = 15.33$, $p < 0.001$) and more specifically, a significant decrease of their relative frequency: between group 1 and 2 ($U = 361.0$, $p = 0.004$), 2 and 3 ($U = 383.0$, $p = 0.013$), and between 1 and 3 ($U = 290.0$, $p < 0.001$). This is summarised in Figure 3.4.

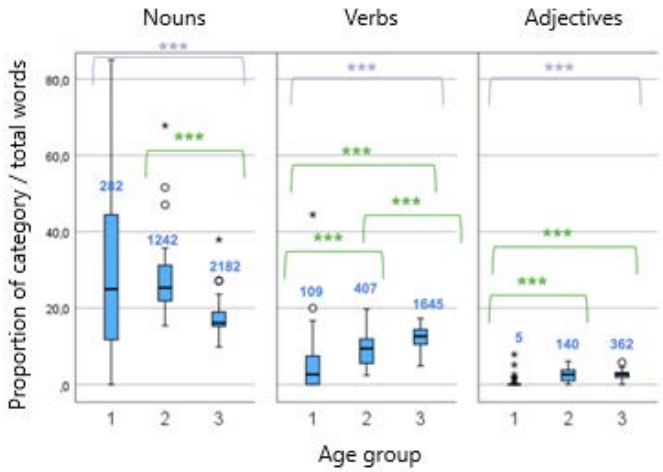


Figure 3.2.: Frequency of nouns, verbs and adjectives for each age group

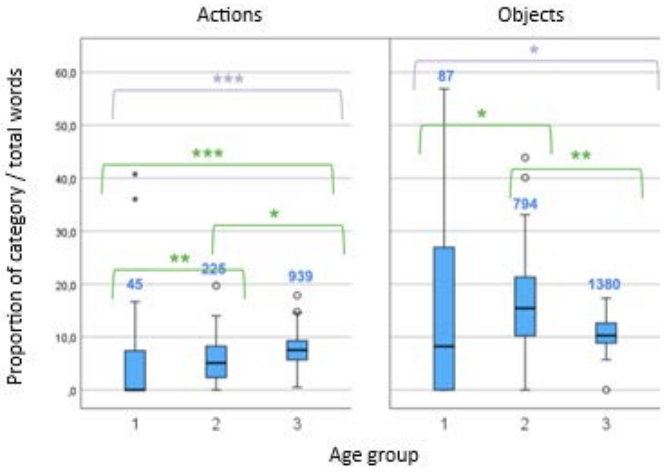


Figure 3.3.: Frequency of actions and objects for each age group

Grammatical categories: As can be seen in Table 3.3, many of our children's productions took place during playing, and we had fewer productions in care contexts. Nevertheless, some tendencies seem to emerge: during play contexts we find a high number of nouns, but also a high number of verbs, while during shared reading and care contexts, nouns more dramatically outnumber verbs. The frequency of adjectives is constant across all contexts. There is a significant difference in the frequency of nouns across the different contexts ($H = 6.35$, $p = 0.042$): more specifically, there are significantly more nouns produced during shared reading compared to play contexts ($U = 2068.0$, $p = 0.006$). The difference between play and care contexts is not significant ($U = 1863.0$, $p = 0.197$). The difference in the frequency of verbs between all

contexts only approaches significance (H = 5.48, p = 0.065). These results are summarised in Figure 3.5.

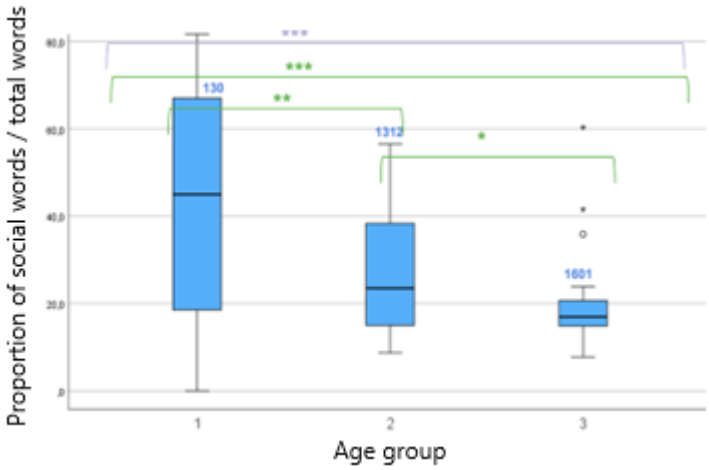


Figure 3.4.: Frequency of social words for each age group

Effect of the context of production

Table 3.3.: Total number of each word class across all contexts

Word class	Context					
	Play		Shared reading		Care	
	raw number	frequency	raw number	frequency	raw number	frequency
Nouns	4484	17,5%	3084	23,1%	642	23,9%
Verbs	3247	12,6%	1258	9,4%	259	9,6%
Adjectives	748	2,9%	384	2,9%	79	2,9%
Objects	2879	11,2%	1986	14,9%	414	15,4%
Actions	2022	7,9%	778	5,8%	176	6,5%
Social words	4695	18,3%	2502	18,8%	729	27,1%
Total words	25668		13340		2688	

Semantic classes: In Table 3.3, a tendency can be observed with objects words outnumbering action words in the shared reading and care contexts. However, more action words seem to be produced during play contexts. When comparing the frequency of objects words across all contexts, we did not find a significant difference (H = 3.22, p = 0.200). This is also true for action words, although the difference only approaches significance (H = 5.16, p = 0.076). Finally, there seem to be more social words in the care contexts compared to the shared reading and play contexts, especially when we compare the frequency of nouns between shared reading and care contexts, which is similar (23,1% vs. 23,9% respectively) while in those same contexts, the frequency of social words differs (18,8% vs. 27,1% respectively). However, the difference

in the frequency of social words in all contexts only approaches significance levels (H = 5.69, p = 0.058).

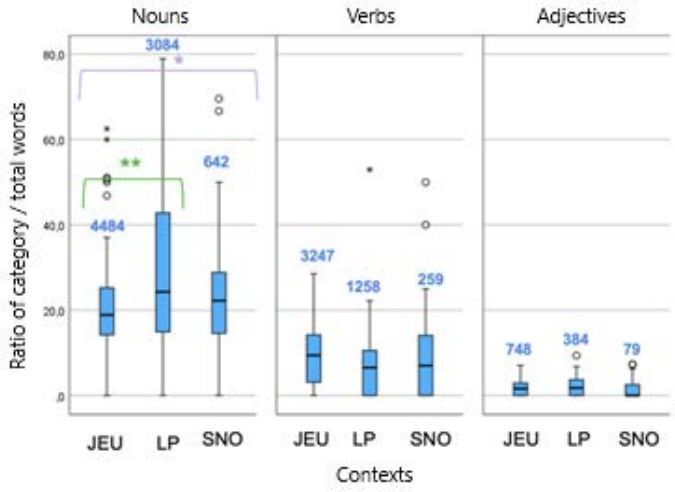


Figure 3.5.: Frequency of nouns, verbs and adjectives for each age group

3.2.4. Discussion

Concerning the evolution of the frequency of grammatical categories across development, we have been able to observe that nouns, while forming a major part of the productive lexicon from 1;0 to 2;5, begin to become less prominent in our third age group from 2;6 to 2;11, as their frequency diminishes. This is in accordance with results from Bassano et al. (2005) and constitutes renewed evidence for the presence of a noun bias in children's early lexicon (see Bassano, 2000; Caselli et al., 1995; Kauschke & Hofmeister, 2002; Labertoniere & Skoruppa, 2022), even in day to day interactions. We can confirm that verb and adjectives are less frequent at the beginning of development, but that their frequency increases (Bassano, 2007; Bassano et al., 2005; Bates et al., 1994) from 1;0 to 2;11. Regarding semantic classes, we noticed that the object class does not follow exactly the same pattern as the noun category. Specifically, there is a relative increase between the first and second age group (which is not the case for nouns). This observation might be linked to the naming explosion, which begins around the period of our second age group (Nelson, 1973). At that time, children begin to rapidly acquire new words, and could in particular acquire many object words compared to other kinds of words. We also observed that the action class begins to develops later than the object class (like in Benedict, 1979), in concordance with results from grammatical categories and in line with the hypothesis that the imageability of a word is key to its early acquisition by children and that actions are less easily conceptualised (Gentner, 1982). Moreover, the number of actions was very close to the number of verbs, especially in the first two age groups, which seems to confirm that action verbs are acquired later than other verbs (like in Bassano, 2000).

Concerning the special category of semantic terms that are social words, we can confirm that their frequency, while high at first, decreases continually across the observed developmental period, in agreement with Labertoniere and Skoruppa (2022). These words, comprising words referring to social beings and words produced under specific conditions of social interactions

constitute a large part of children's first productive lexicon. These situations of social interaction, particularly prominent during the first steps of lexical acquisition (Lytle & Kuhl, 2017), lead to a strong frequency of social words in the first lexicon by offering frequent opportunities to participate in verbal exchanges. The diminution of the frequency of social words during development could be explained by the transition to a more complex communication mode, involving an increase in the frequency of predicates and function words (Bassano et al., 2005; Kauschke & Hofmeister, 2002).

Concerning the role of activity contexts on the frequency of our various word classes in children's productions, we replicated findings from Ogura et al. (2006) and Tardif et al. (1999) confirming that more nouns are produced in situations of shared reading compared to play contexts. This might be explained by the fact that objects are more imageable than actions and are thus more represented in children's picture books (Labertoniere & Skoruppa, 2022), and that caretakers might play more of a tutor role during shared reading, by naming more objects than usual (Ogura et al., 2006). However, we could not statistically confirm that more object words are produced by children in share reading contexts. With respect to verbs and actions, we found that children tend to produce more verbs and action words during play contexts (like mothers in Altınkamış et al., 2014). When they are playing, children are performing actions themselves, which might lead to a higher frequency of action words compared to during shared reading, where they are less active. We did not observe any difference in the frequency of adjectives across activity contexts. While this might be due to their relatively low total number in all our children's productions, this might indicate that activity context has less effect on the production of adjective than on other word classes. Finally, while this was not statistically significant, we found an interesting tendency towards social words being more present in care contexts than in both shared reading and play contexts. While shared reading and playing are both interactive situations calling for joint attention between child and caregiver (thus leading to the production of many social words), the activity contexts that comprised our care category might have been even more conducive to the production of social words. According to Hu et al. (2021), mealtime (which was included in the care contexts) is a good opportunity to develop language and a privileged moment of interaction between caregiver and child.

3.2.5. Conclusion

In this study, we analysed children's early productions during ecological interactions with their caregivers across two levels, namely grammatical categories (nouns, verbs and adjectives) and semantic classes (object, action and social words). These different word classes were analysed in terms of their temporal development from 1;0 to 2;11 and in relation to the activity context in which they were produced. We found more evidence in favour of a noun bias in young children's productive lexicon by confirming that children produce more nouns that other word classes even in ecological settings. We verified that the lexicon gradually diversifies with the frequency of verbs and adjectives increasing over time. In terms of semantic classes, we found that the same pattern holds for object and action words, except that object nouns take a more important place at the beginning of the naming explosion. We noted that social words are highly common in the child's first year and that their frequency decreases thereafter. Concerning the effect of activity contexts on productions, we confirmed that more nouns are produced during shared reading than during play contexts, and found tendencies suggesting that both verbs and action words

are more prevalent during play contexts and that care contexts promote the production of social words compared to play and shared reading. This research confirms the important role of social interactions in children's early language acquisition, by demonstrating the importance of social words in the early productive lexicon. We managed to study a poorly documented domain in the acquisition literature, namely the role of activity contexts on children's productions. In further studies, including an analysis of caregivers' productions could be interesting. If they could show the same pattern as children's, that similarity could be linked to reciprocity in interactions, children's imitative tendencies, adults' interpretative tendencies and joint attention.

A note on additional analyses

During coding, we originally coded for three additional contexts: exploration, when children explore novel things or look out the window (EXPL); discussions and comforting between caretaker and child (DISC); and moments of transition between two activities or of movement between two places (TRANS). However, we decided not to include them in our analyses because the limits between each of these contexts were not well defined. This led us to focus our analysis on very interactive contexts were the caretaker is focused on the child, which might bias our results. Ideally, we would want to add one more category of contexts to the analyses, which would for example include both exploration and transition moments. This way we can explore the child's productive lexicon when there is less joint attention and the activity is less ideal for language development. Moreover, we want to run a generalised linear mixed model on the data, so that we can treat age as a continuous variable and have a better picture of the evolution of the different categories of interest across development. This model might also help us better identify the role of context on the production of different word classes. Finally, we will also code 10% of the video data anew and calculate the inter-rater reliability.

3.3. Speech acts in infants and their relevance to nonhuman communication

In Section 3.2, I presented a corpus study of children's naturalistic productions during their first few years of life. An additional parameter was analysed in this study: speech acts. At the end of this section, I propose to compare the speech acts that are first produced by children to the gestural and vocal communications that occur in nonhuman primates in the wild, in an exploratory and evolutionary perspective. To be more accurate, I will consequently use the term *communicative acts* instead of speech acts in the following section.

Language acquisition involves much more than learning words. The child must also learn to use words and non verbal communication appropriately, i.e. learn pragmatic rules. According to Ninio et al. (1994), linguistic ability is described in terms of speakers' repertoires of communicative intentions and the rules that govern them. Cameron-Faulkner (2014) explains that communicative acts are a way of representing children's intentions in face-to-face interactions with a conversation partner in everyday life. Dore (1974) lists the communicative acts identifiable at the pre-verbal stage (between 1;5 and 1;7), which he calls the primitive communicative acts: labelling, responding, repeating, requesting action, asking, calling, protesting, practising language and greeting. These are based on isolated words or on cues like prosody. Using the DLPF tool (questionnaire sur le Développement du Langage de Production en Français), Bassano et al. (2020) found that the first communicative acts (expression of desires and emotions, request for an object, request for information, simple assertion, event report and explanation) emerge and continue to develop in a distinctive way during the first year of life. The ability to express emotions emerges very early. It is already present at 1;6 and is established at 2;0 in the majority of children. The ability to request objects, obtain information and express comments using simple assertions is also present at 1;6 but is only completely acquired by 2;8. The ability to recount events does not emerge until 1;8 and further develops after 2;8. Finally, Ninio et al. (1994) developed an abridged version of an Inventory of Communicative Acts,[17] which proposes a complete classification of the different speech acts used by children in communicative situations. Based on this inventory, Snow et al. (1996) studied 1;2–2;8 children in interaction with their caregiver. The authors proposed an analytical grid for categorising communicative acts according to production contexts and found that at 14 months, the majority of communicative acts could be classified as utterances without clear function, while those used by most children between 14 and 20 months can be categorised as declarations, verbally accompanying an action, refusals, answers and attracting attention.

In the corpus study, in addition to coding for grammatical categories, semantic classes and activity contexts, we also coded each utterance of the children as a particular communicative act,[18] to observe the effect of development on their acquisition and of context on their production. We used Ninio et al. (1994)'s existing code list, which is reproduced in Appendix B. This list was chosen because of its solid theoretical and methodological foundations as well as its extensive categorisation of communicative acts (67 in total). We chose to group these communicative acts into the following classes, according to their relevance in our study:

[17] Based on the more comprehensive list developed by Ninio and Wheeler (1986).
[18] Via a %acl dependent line.

1. answers
2. questions
3. emotions
4. requests
5. declarations
6. refusals
7. attracting attention
8. verbally accompanying an action
9. sounds/imitations/repetitions

The distribution of these different communicative acts in our data can be found in Table 3.4.

Table 3.4.: Distribution of communicative acts across all age groups

Classes	**Age group**					
	Group 1		Group 2		Group 3	
	raw number	frequency	raw number	frequency	raw number	frequency
Answers	294	4,9%	1477	14,0%	2693	21,1%
Questions	73	1,2%	602	5,7%	923	7,2%
Emotions	572	9,6%	411	3,9%	330	2,6%
Requests	260	4,4%	737	7,0%	829	6,5%
Declarations	909	15,3%	3641	34,6%	5537	43,3%
Refusals	109	1,8%	368	3,5%	438	3,5%
AA	400	6,7%	369	3,5%	443	3,5%
VAA	67	1,1%	100	1,0%	145	1,1%
S/I/Rs	3267	54,9%	2810	26,7%	1443	11,3%
Tot. utterances	5951		10515		12781	

Group 1 = 1;0-1;7, Group 2 = 1;8-2;5, Group 3 = 2;6-2;11, AA = attracting attention, VAA = verbally accompanying an action, S/I/Rs = sounds/imitations/repetitions

While we have not yet statistically analysed the data, some tendencies can be observed. First, we observe, as Snow et al. (1996) did, that the category sounds/imitations/repetitions is dominant in very young children, comprising more than half of their utterances in our corpus. The categories of declarations, emotions and attracting attention are then the most used by infants from 1;0 to 1;7 (matching the results of Bassano et al., 2020; Ninio & Wheeler, 1986). In the

middle age group and oldest age group, we find that sounds/imitations/repetitions notably decrease, while declarations, answers, requests and questions gradually increase, most probably in accordance with the development of a more complex lexicon.

I argue that nonhuman animal communication also expresses communicative acts, and that these could be classified similarly to the ones used by children. In fact, they might be similar to the ones that are most used by young human infants. One of the most studied forms of primate communication is alarm calls (see e.g. Deshpande et al., 2022; Schlenker et al., 2016; Zuberbühler, 2000, 2001). These could be classified as 'attracting attention' communicative acts (in fact, one kind of communicative act in this class is expressly 'warn of danger'). Bonobos' contact calls (Cheney & Seyfarth, 2018; Schamberg et al., 2016) and 'contest hoots' (Genty et al., 2014) as well as chimpanzees' greeting signals (Laporte & Zuberbühler, 2011) might be examples of declarative communicative acts, while chimpanzees' 'play solicitation' gestures (Fröhlich et al., 2016) might be classified as requests. Bonobos' 'peep' calls at least express some amount of emotional valence (Clay et al., 2015), and could be a form of 'emotions' communicative acts, but they also might serve to 'comment' on novel items or environmental events (closer to 'declarations'). Knox et al. (2019) found that orangutans (*Pongo pygmaeus*) use gestures to achieve eight distinct goals, six of which are positive requests and the last two negations (which could be categorised in our 'refusals' class). Captive gorillas also seem to use 'requests' type gestures and gestures for 'attracting attention' (Genty et al., 2009).

As the pragmatic development of children can be assessed by measuring their repertoire of communicative intents (Snow et al., 1996), linking communicative acts in children to communicative acts in nonhuman primates could shed light on the evolution of pragmatics in humans. Researchers in animal communication have intentionally based their categorisations of apes' communication, like gestures (Knox et al., 2019; Tomasello et al., 1985), on young children's communicative acts. While categorising intentional communication in nonhuman primates is challenging, the same can be said for young children's communicative acts: the signaler's intention has to be translated by the coder based on the signaler's observable behaviour and the recipient's behavioural or verbal (in the case of humans) response. I believe that future studies should definitely look into comparing young children's and young and adult nonhuman primates' communicative acts (in the verbal and gestural modalities), and that we should try to unify the methods and terminology, like Berthet et al. (2023) proposes to do in all areas of animal and human communication studies.

References

Abeillé, A., & Godard, D. (2021). *La grande grammaire du français*. Éditions Actes Sud.

Altınkamış, N. F., Kern, S., & Sofu, H. (2014). When context matters more than language: Verb or noun in french and turkish caregiver speech. *First language*, *34*(6), 537–550. https://doi.org/10.1177/0142723714560179

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. https://doi.org/10.1017/S0305000900004396

Bassano, D. (2007). Emergence et développement du langage: Enjeux et apports des nouvelles approches fonctionnalistes. *L'acquisition du langage et ses troubles*, 13–46.

Bassano, D., Eme, P.-E., & Champaud, C. (2005). A naturalistic study of early lexical development: General processes and inter-individual variations in French children. *First Language*, *25*(1), 67–101. https://doi.org/10.1177/0142723705049119

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. https://doi.org/10.3917/enf2.202.0151

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of child language*, *21*(1), 85–123. https://doi.org/10.1017/S0305000900008680

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of child language*, *6*(2), 183–200. https://doi.org/10.1017/S0305000900002245

Berthet, M., Coye, C., Dezecache, G., & Kuhn, J. (2023). Animal linguistics: A primer. *Biological reviews*, *98*(1), 81–98. https://doi.org/10.1111/brv.12897

Bloom, P. (2002). *How children learn the meanings of words*. MIT press.

Cameron-Faulkner, T. (2014). The development of speech acts. *Pragmatic development in first language acquisition*, 37–52.

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. https://doi.org/10.1016/0885-2014(95)90008-X

Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, *115*(9), 1974–1979. https://doi.org/https://doi.org/10.1073/pnas.1717572115

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of child language*, *22*(3), 497–529. https://doi.org/10.1017/S0305000900009934

Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, *3*, e1124. https://doi.org/10.7717/peerj.1124

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of french determiners. *Journal of child language*, *35*(1), 99–127. https://doi.org/10.1017/S0305000907008276

Deshpande, A., Van Boekholt, B., & Zuberbuhler, K. (2022). Preliminary evidence for one-trial social learning of vervet monkey alarm calling. *Royal Society Open Science*, *9*(8), 210560. https://doi.org/10.1098/rsos.210560

Dore, J. (1974). A pragmatic description of early language development. *Journal of psycholinguistic Research*, *3*(4), 343–350. https://doi.org/10.1007/BF01068169

Fröhlich, M., Wittig, R. M., & Pika, S. (2016). Play-solicitation gestures in chimpanzees in the wild: Flexible adjustment to social circumstances and individual matrices. *Royal Society Open Science*, *3*(8), 160278. https://doi.org/10.1098/rsos.160278

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental psychology*, *40*(4), 621. https://doi.org/10.1037/0012-1649.40.4.621

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report; no. 257*.

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (gorilla gorilla): Repertoire, intentionality and possible origins. *Animal cognition*, *12*, 527–546. https://doi.org/10.1007/s10071-009-0213-4

Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi-modal use of a socially directed call in bonobos. *PloS one*, *9*(1), e84738. https://doi.org/10.1371/journal.pone.0084738

Gopnik, A., & Meltzoff, A. N. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development*, 1040–1053. https://doi.org/10.2307/1130378

Hoff, E. (2013). *Language development*. Cengage Learning.

Hu, J., Torr, J., Wei, Y., & Jiang, C. (2021). Mealtime talk as a language learning context: Australian chinese parents' language use in interactions with their preschool-aged children at the dinner table. *Early Child Development and Care*. https://doi.org/10.1080/03004430.2019.1621862

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. https://doi.org/10.1017/S0305000902005330

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, *27*(3), 227–250. https://doi.org/10.1177/0142723706075789

Knox, A., Markx, J., How, E., Azis, A., Hobaiter, C., van Veen, F. J., & Morrogh-Bernard, H. (2019). Gesture use in communication between mothers and offspring in wild orangutans (pongo pygmaeus wurmbii) from the sabangau peat-swamp forest, borneo. *International Journal of Primatology*, *40*, 393–416. https://doi.org/10.1007/s10764-019-00095-w

Labertoniere, D., & Skoruppa, K. (2022). La composition du lexique productif dans les trois premières années de la vie : Revue systématique. *A.N.A.E.*, *181*, 689–705.

Laporte, M. N., & Zuberbühler, K. (2011). The development of a greeting signal in wild chimpanzees. *Developmental Science*, *14*(5), 1220–1234. https://doi.org/10.1111/j.1467-7687.2011.01069.x

Lytle, S. R., & Kuhl, P. K. (2017). Social interaction and language acquisition: Toward a neurobiological view. *The handbook of psycholinguistics*, 615–634. https://doi.org/10.1002/9781118829516.ch27

MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of child language*, *12*(2), 271–295. https://doi.org/10.1017/S0305000900006449

McShane, J. (1980). *Learning to talk*. Cambridge University Press.

Morgenstern, A. (2006). *Un je en construction: Genèse de l'auto-désignation chez le jeune enfant*. Editions OPHRYS.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the society for research in child development*, 1–135.

Ninio, A., Snow, C. E., Pan, B. A., & Rollins, P. R. (1994). Classifying communicative acts in children's interactions. *Journal of communication disorders*, *27*(2), 157–187. https://doi.org/10.1016/0021-9924(94)90039-6

Ninio, A., & Wheeler, P. (1986). A manual for classifying verbal communicative acts in mother-infant interaction. *Transcript analysis*, *3*(1), 1–82.

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.

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. https://doi.org/10.1177/0142723707067545

Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, *122*, 109–116. https://doi.org/10.1016/j.anbehav.2016.10.003

Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinović, D., Arnold, K., Cäsar, C., Keenan, S., et al. (2016). Formal monkey linguistics. *Theoretical Linguistics*, *42*(1-2), 1–90. https://doi.org/10.1515/tl-2016-0001

Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning how to say what one means: A longitudinal study of children's speech act use. *Social development*, *5*(1), 56–84. https://doi.org/10.1111/j.1467-9507.1996.tb00072.x

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, *44*(4), 929. https://doi.org/10.1037/0012-1649.44.4.929

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. https://doi.org/10.1111/1467-8624.00045

Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., Evans, A., et al. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, *14*(2), 175–186. https://doi.org/10.1016/S0047-2484(85)80005-1

Zuberbühler, K. (2000). Referential labelling in diana monkeys. *Animal behaviour*, *59*(5), 917–927. https://doi.org/10.1006/anbe.1999.1317

Zuberbühler, K. (2001). Predator-specific alarm calls in campbell's monkeys, cercopithecus campbelli. *Behavioral Ecology and Sociobiology*, *50*, 414–422. https://doi.org/10.1007/s002650100383

Conclusion

In this dissertation, I have explored various factors that have an influence on the acquisition of new meanings in human and nonhuman great apes. In particular, in an evolutionary perspective, I have compared human and nonhuman apes' ability to rapidly learn new meanings in a novel paradigm. Focusing on human language acquisition, I have then explored experimentally the role of situational context and that of semantics on the acquisition of novel words in adults and children. I have also reviewed the effect of grammatical categories, of semantic classes and of social information on the productions of very young children. Finally, I have discussed the possible bridges that can be constructed between the early speech acts of human children and the communication pattern of nonhuman apes. In the following paragraphs, I will detail the main results of this dissertation.

Because certain pre-linguistic features of human language seem to exist in nonhuman primates (e.g. compositionality, pragmatics, see Schlenker et al., 2016) and because fast mapping is a feature that could be nonexclusive to language (Behrend et al., 2001; Markson & Bloom, 1997) and because it has been observed in domestic animals (Fugazza et al., 2021; Kaminski et al., 2004; Takagi et al., 2024), I investigated whether it was present in our close cousins. In chapter 1, I have presented a study in which two experiments were designed to compare the fast mapping of novel sound forms onto novel objects by human adults and captive nonhuman great apes. For these experiments, I designed a fast mapping eye-tracking paradigm and adapted it for use in nonhuman apes as well as human adults and infants. Event though I confirmed that humans could rapidly associate novel words to novel objects, even when using an indirect measuring method, I could not show an effect of learning in nonhuman great apes. However, this could be due to the experimental design in itself and not to the absence of a fast mapping capacity in nonhuman apes. Indeed, the apes struggled to attend to such a task for this long and did not seem very interested in the stimuli. It is also possible that the use of sounds alone as a form to associate to the novel referent was not relevant enough (see Dezecache et al., 2019, in which captive orangutans respond better to gestures than to words). If different species have different priors concerning what can be fast mapped, then we just have to find what nonhuman great apes fast map (words and objects might not be it). As with children in chapter 2, there was also a lack of social interactions or any social component in this experiment, which might be linked to their poor attention to the task. Indeed, social stimuli have been shown to enhance memory in nonhuman apes (Howard et al., 2017).

Taking into account the limitations of the eye-tracking task for the nonhuman apes, two options seem possible to me to continue to study fast mapping in this population. The first one would be to adapt the existing eye-tracking protocol to the apes' attention span and reaction time: shortening the duration of both the learning phases and test phases seems like a promising avenue. The second option would be to switch to the touchscreen method. In domestic animals, both direct and indirect methods have been successful in the observation of fast mapping: indirect methods like a habituation-dishabituation design in cats (with manual coding of gaze, see Takagi et al., 2024), and direct methods like toy naming and retrieval in dogs (Fugazza et al.,

2021; Kaminski et al., 2004). Considering the fact that our nonhuman apes struggled to stay attentive to the screen for a long period in our task, a habituation-dishabituation task, which relies on exposing a novel association for a long time during the habituation phase might not be ideal. It would also run into the same problem as our task if the ape is not attentive during any of the phases. A direct method like a touchscreen forced-choice task, while more demanding (only some highly trained dogs show fast mapping for example in Fugazza et al., 2021), seems a better idea. The nonhuman apes at Basel zoo have already been trained to use touchscreens, and have successfully completed complex tasks like touching an individual in a scene (see Brocard et al., 2024). It is also a method that can be used with both adults and children (see for example Brocard et al., 2024; Skoruppa, 2019), allowing for cross-species comparisons. Concerning the visual stimuli, a first possibility would be continuing with objects. In that case, I would suggest to extract the object from the background so that we can only see the object on the neutral background of the screen. Instead of keeping it as a still image, the object could move about during learning, to attract the apes' attention (like what can be done with children, see Piot et al., 2024). The second possibility would be to use videos (and stills during test) of animals, which might be more interesting than fruit objects.

Coming back to the humans' performance in this experiment, an interesting finding was that, while they were able to learn the novel associations in the stable condition, we did not observe a learning effect in the context-change condition. This highlights the difficulty of this task when distractions abound. This condition might have necessitated more working memory than the more simple stable condition, and/or the latter could have been more explicit in revealing the goal of the experiment and thus be more familiar of a situation for the participants. As I showed in chapter 2 with the adult pilots however, adults are able to complete the word leaning task in both conditions when no distractions are engineered and when a direct measuring method is used. The fact that infants exhibited the opposite pattern in chapter 2 (better performance in the context change condition) could be a matter of differences in experimental conditions (calm environment for the infants) or due to the fact that for them, the context-change condition was a more familiar learning situation.

Beginning to question the role of situational context on language acquisition, I have presented in chapter 2 an experimental study on 14- to 19- month-old children. I have compared children's ability to fast map novel words to novel objects in two conditions: a condition where I manipulated the background and orientation of the target object during learning and test phases so that it changed, and a condition where neither background nor orientation changed during learning and testing. I have shown that, contrary to previous studies on older children, 14- and 19-month-olds could only rapidly learn new associations in my context-change condition. This might be due to the difference in the method used in my study compared to previous ones. I used an indirect measuring method (eye-tracking), which might have involved different mechanisms than a forced-choice task. I hypothesise that the context-change condition, by mimicking the manipulations that a child could perform on an object in a real life learning-oriented situation, could facilitate the forming of associations in children's minds.

In a follow-up experiment, I have tried to find out if the context-change condition could help children extract the object from the complex background, on the assumption that children in the first experiment failed to do so in the stable condition. I designed a similar experiment with

the same context-change condition, but an easier invariant condition using grey backgrounds behind the objects. Without enough data, I was not able to draw solid conclusions, but since I did not observe a learning effect in the grey background condition, I hypothesise that the stable condition with a grey background was not particularly easy for children.

Moreover, in a third experiment, I tried to find out if words for objects were easier to rapidly learn than words for actions, in an attempt to link the noun bias to the (major) semantic classes of nouns and verbs. However, the scenes in the video stimuli that I used in this experiment seem to have been too complex for the infants, as I could not observe a learning effect in either condition. As actions are more salient, I hypothesise that the infant's failure to fast map has to do with the concept of actions being harder to learn than that of objects. Other researchers who have managed to find rapid learning of novel actions (Imai et al., 2008; Waxman et al., 2009) all used verb-like pseudowords embedded in carrier sentences, suggesting that morphosyntactic cues are, if not necessary, at least very important for infants to learn novel words for actions. The scenes I chose for my visual stimuli were very ecologically relevant, by being similar to the environment in which children could find themselves during a learning situation that is not specifically a naming event directed at them. The fact that I could not observe fast mapping in this context raises the question of the relevance of fast mapping in real-life situations that are not specific learning situations induced by the caretaker. However, it has been shown that children can rapidly learn words through overhearing (Gampe et al., 2012). It is thus possible that in complex visual situations, rapid word learning requires the presence of a little more cues to take place. These could be morphosyntactic cues, as in Waxman et al. (2009) and Imai et al. (2008), or social cues, like the presence of live human beings in Gampe et al. (2012) and Kuhl et al. (2003). As in Kuhl et al. (2003), the presence of people in the video during learning did not seem to be enough for systematic learning to occur.

Using a direct method of measuring, I have then proved that adults have no problem in rapidly associating novel words to novel objects or actions, even when those appear in complex visual scenes, and that they also perform at ceiling when the background is stable or changes during learning and test. The difference in method of testing could then also have an influence on our capacity in detecting rapid word learning. It is possible that children could have performed above chance if the second experiment had been conducted on a touchscreen instead of using an indirect method of measuring like eye-tracking.

Finally, in chapter 3, by conducting a systematic review of the literature, I have confirmed that the noun bias that can be observed in children's early productive lexicon seems to be universal. However, verbs also occupy an important place in their lexicon. Depending on the typological characteristics of the language (*noun-* or *verb-friendly*), the noun bias can be mitigated and verbs can be more easily acquired. Certain observational methods can also have an influence on the proportions of the different grammatical categories observed, with parental questionnaires being biased towards nouns. Besides, the semantic nature of certain nouns and verbs seems to be one of the causes of the noun bias. In the corpus study on the productive lexicon of young children, I have verified that nouns outnumber verbs in French-learning children, even when looking only at ecological data, and have confirmed that the dominance of nouns gradually decreases with development. Furthermore, the pattern observed for nouns and verbs seems to hold for objects and actions, which therefore reinforces the hypothesis that the noun bias is linked to

semantics.

I have also confirmed the tendency observed in the systematic review: that children's very early productions consist mainly of social terms. This finding seems highly relevant, in particular in that it reinforces the premise that social information plays a major role in early language acquisition. As I have pointed out in my review, there is a lack of systematicity in the categorisation of children's productions in research studies, particularly concerning social words. I believe it is important for future research to take that finding in consideration. A universal category for social terms could be proposed and used by researchers working on any language.

Moreover, I have once again explored the role of situational context on children's lexicon. More specifically, I have shown in the corpus study that, like what had been found in previous studies (Ogura et al., 2006; Tardif et al., 1999), nouns are overrepresented in shared-reading contexts compared to play contexts. However, more fine-grained analyses are needed to explore the role of other contexts on the frequency of verbs and nouns, as well as on the frequency of object words, action words and social terms.

In a final section, I presented additional data from the corpus study in chapter 3 on the temporal evolution of the distribution of various classes of children's speech acts. I linked such communicative acts in infants to vocal and gestural communication in nonhuman primates. I postulated that children's first communicative acts (like declarations, expressions of emotions, attracting attention and requests, for example) could be similar to the ones used by nonhuman primates, and encouraged future research to conduct comparative studies of such communicative acts between humans and nonhumans.

To conclude, during the development of children's lexicon, many factors have an influence on novel word acquisition, and on fast mapping in particular. In this dissertation I focused on the influence of grammatical categories, semantics, and visual and situational context, but many other factors are still under study. I have explored one facet of language, namely the rapid acquisition of new meanings, in nonhuman great apes, but encourage further comparative work on this subject and on other components of human language.

List of Figures

Bibliography

Abeillé, A., & Godard, D. (2021). *La grande grammaire du français*. Éditions Actes Sud.

Altınkamış, N. F., Kern, S., & Sofu, H. (2014). When context matters more than language: Verb or noun in french and turkish caregiver speech. *First language*, *34*(6), 537–550. https://doi.org/10.1177/0142723714560179

Arunachalam, S. (2013). Two-year-olds can begin to acquire verb meanings in socially impoverished contexts. *Cognition*, *129*(3), 569–573. https://doi.org/10.1016/j.cognition.2013.08.021

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development*, *7*(3), 169–184. https://doi.org/10.1080/15475441.2011.573760

Axelsson, E. L., & Horst, J. S. (2014). Contextual repetition facilitates word learning via fast mapping. *Acta Psychologica*, *152*, 95–99. https://doi.org/10.1016/j.actpsy.2014.08.002

Barrett, M. D. (1986). Early semantic representations and early word-usage. In *The development of word meaning: Progress in cognitive development research* (pp. 39–67). Springer. https://doi.org/10.1007/978-1-4612-4844-6_2

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. https://doi.org/10.1017/S0305000900004396

Bassano, D. (2007). Emergence et développement du langage: Enjeux et apports des nouvelles approches fonctionnalistes. *L'acquisition du langage et ses troubles*, 13–46.

Bassano, D., Eme, P.-E., & Champaud, C. (2005). A naturalistic study of early lexical development: General processes and inter-individual variations in French children. *First Language*, *25*(1), 67–101. https://doi.org/10.1177/0142723705049119

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. https://doi.org/10.3917/enf2.202.0151

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of child language*, *21*(1), 85–123. https://doi.org/10.1017/S0305000900008680

Behrend, D. A., Scofield, J., & Kleinknecht, E. E. (2001). Beyond fast mapping: Young children's extensions of novel words and novel facts. *Developmental Psychology*, *37*(5), 698. https://doi.org/10.1037/0012-1649.37.5.698

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of child language*, *6*(2), 183–200. https://doi.org/10.1017/S0305000900002245

Berthet, M., Coye, C., Dezecache, G., & Kuhn, J. (2023). Animal linguistics: A primer. *Biological reviews*, *98*(1), 81–98. https://doi.org/10.1111/brv.12897

Bloom, P. (2002). *How children learn the meanings of words*. MIT press.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., Pêcheux, M.-G., 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. https://doi.org/10.1111/j.1467-8624.2004.00729.x

Bourgoyne, A., & Alt, M. (2017). The effect of visual variability on the learning of academic concepts. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1568–1576. https://doi.org/10.1044/2017_JSLHR-L-16-0271

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433–436.

Brocard, S., Wilson, V. A., Berton, C., Zuberbühler, K., & Bickel, B. (2024). A universal preference for animate agents in hominids. *iScience*, *27*(6). https://doi.org/10.1016/j.isci.2024.109996

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). Glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9*(2), 378–400. https://doi.org/10.3929/ethz-b-000240890

Cameron-Faulkner, T. (2014). The development of speech acts. *Pragmatic development in first language acquisition*, 37–52.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). MIT press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

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. https://doi.org/10.1016/0885-2014(95)90008-X

Chan, C. C., Tardif, T., Chen, J., Pulverman, R. B., Zhu, L., & Meng, X. (2011). English-and chinese-learning infants map novel labels to objects and actions differently. *Developmental psychology*, *47*(5), 1459. https://doi.org/10.1037/a0024049

Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, *115*(9), 1974–1979. https://doi.org/https://doi.org/10.1073/pnas.1717572115

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of child language*, *22*(3), 497–529. https://doi.org/10.1017/S0305000900009934

Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, *3*, e1124. https://doi.org/10.7717/peerj.1124

Crockford, C., & Boesch, C. (2003). Context-specific calls in wild chimpanzees, pan troglodytes verus: Analysis of barks. *Animal behaviour*, *66*(1), 115–125. https://doi.org/10.1006/anbe.2003.2166

Curio, E., Ernst, U., & Vieth, W. (1978). Cultural transmission of enemy recognition: One function of mobbing. *Science*, *202*(4370), 899–901. https://doi.org/10.1126/science.202.4370.899

Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 892. https://doi.org/10.1037/a0035657

Déaux, E. C., Baumeyer, A., & Zuberbühler, K. (2021). Food calls enhance visual discrimination learning in chimpanzees (pan troglodytes verus). *Journal of Comparative Psychology*, *135*(3), 420. https://doi.org/10.1037/com0000286

DeKeyser, R. (2003). Implicit and explicit learning. *The handbook of second language acquisition*, 312–348. https://doi.org/10.1002/9780470756492

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of french determiners. *Journal of child language*, *35*(1), 99–127. https://doi.org/10.1017/S0305000907008276

Deshpande, A., Van Boekholt, B., & Zuberbuhler, K. (2022). Preliminary evidence for one-trial social learning of vervet monkey alarm calling. *Royal Society Open Science*, *9*(8), 210560. https://doi.org/10.1098/rsos.210560

Dezecache, G., Bourgeois, A., Bazin, C., Schlenker, P., Chemla, E., & Maille, A. (2019). Orangutans' comprehension of zoo keepers' communicative signals. *Animals*, *9*(6), 300. https://doi.org/10.3390/ani9060300

Dore, J. (1974). A pragmatic description of early language development. *Journal of psycholinguistic Research*, *3*(4), 343–350. https://doi.org/10.1007/BF01068169

Fröhlich, M., Wittig, R. M., & Pika, S. (2016). Play-solicitation gestures in chimpanzees in the wild: Flexible adjustment to social circumstances and individual matrices. *Royal Society Open Science*, *3*(8), 160278. https://doi.org/10.1098/rsos.160278

Fugazza, C., Andics, A., Magyari, L., Dror, S., Zempléni, A., & Miklósi, Á. (2021). Rapid learning of object names in dogs. *Scientific reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-81699-2

Gampe, A., Liebal, K., & Tomasello, M. (2012). Eighteen-month-olds learn novel words through overhearing. *First language*, *32*(3), 385–397. https://doi.org/10.1177/0142723711433584

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental psychology*, *40*(4), 621. https://doi.org/10.1037/0012-1649.40.4.621

Garcia, J., Kimeldorf, D. J., & Koelling, R. A. (1955). Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science*, *122*(3160), 157–158. https://doi.org/10.1126/science.122.3160.157

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report; no. 257*.

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (gorilla gorilla): Repertoire, intentionality and possible origins. *Animal cognition*, *12*, 527–546. https://doi.org/10.1007/s10071-009-0213-4

Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi-modal use of a socially directed call in bonobos. *PloS one*, *9*(1), e84738. https://doi.org/10.1371/journal.pone.0084738

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language learning and development*, *1*(1), 23–64. https://doi.org/10.1207/s15473341lld0101_4

Goldenberg, E. R., & Sandhofer, C. M. (2013). Same, varied, or both? contextual support aids young children in generalizing category labels. *Journal of Experimental Child Psychology*, *115*(1), 150–162. https://doi.org/10.1016/j.jecp.2012.11.011

Gopnik, A., & Meltzoff, A. N. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development*, 1040–1053. https://doi.org/10.2307/1130378

Halberda, J. (2006). Is this a dax which i see before me? use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive psychology*, *53*(4), 310–344. https://doi.org/10.1016/j.cogpsych.2006.04.003

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, *298*(5598), 1634–1636. https://doi.org/10.1126/science.1072702

Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language*, *15*(1), 77–94.

Havy, M., Bertoncini, J., & Nazzi, T. (2011). Word learning and phonetic processing in preschool-age children. *Journal of Experimental Child Psychology*, *108*(1), 25–43. https://doi.org/10.1016/j.jecp.2010.08.002

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child development*, 1021–1034. https://doi.org/10.2307/1130543

Hobaiter, C., Graham, K. E., & Byrne, R. W. (2022). Are ape gestures like words? outstanding issues in detecting similarities and differences between human language and ape gesture. *Philosophical Transactions of the Royal Society B*, *377*(1860), 20210301. https://doi.org/10.1098/rstb.2021.0301

Hoff, E. (2013). *Language development*. Cengage Learning.

Hollander, J., & Huette, S. (2022). Extracting blinks from continuous eye-tracking data in a mind wandering paradigm. *Consciousness and Cognition*, *100*, 103–303. https://doi.org/10.1016/j.concog.2022.103303

Hollich, G., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology*, *43*(5), 1051.

Howard, L. H., Wagner, K. E., Woodward, A. L., Ross, S. R., & Hopper, L. M. (2017). Social models enhance apes' memory for novel events. *Scientific reports*, *7*(1), 1–7. https://doi.org/10.1038/srep40926

Hu, J., Torr, J., Wei, Y., & Jiang, C. (2021). Mealtime talk as a language learning context: Australian chinese parents' language use in interactions with their preschool-aged children at the dinner table. *Early Child Development and Care*. https://doi.org/10.1080/03004430.2019.1621862

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for japanese children? *Child development*, *76*(2), 340–355. https://doi.org/10.1111/j.1467-8624.2005.00849_a.x

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child development*, *79*(4), 979–1000. https://doi.org/10.1111/j.1467-8624.2008.01171.x

Jeong, H., Sugiura, M., Sassa, Y., Wakusawa, K., Horie, K., Sato, S., & Kawashima, R. (2010). Learning second language vocabulary: Neural dissociation of situation-based learning and text-based learning. *Neuroimage*, *50*(2), 802–809. https://doi.org/10.1016/j.neuroimage.2009.12.038

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping". *Science*, *304*(5677), 1682–1683. https://doi.org/10.1126/science.1097859

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kano, F., & Call, J. (2014). Cross-species variation in gaze following and conspecific preference among great apes, human infants and adults. *Animal Behaviour*, *91*, 137–150. https://doi.org/10.1016/j.anbehav.2014.03.011

Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas (Gorilla gorilla), orangutans (Pongo abelii), and humans (Homo sapiens): Unique eye-viewing patterns in humans among hominids. *Journal of comparative psychology*, *126*(4), 388. https://doi.org/10.1037/a0029615

Kano, F., & Tomonaga, M. (2011). Species difference in the timing of gaze movement between chimpanzees and humans. *Animal cognition*, *14*, 879–892. https://doi.org/10.1007/s10071-011-0422-5

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. https://doi.org/10.1017/S0305000902005330

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, *27*(3), 227–250. https://doi.org/10.1177/0142723706075789

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. https://doi.org/10.1017/S0305000900004104

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Knox, A., Markx, J., How, E., Azis, A., Hobaiter, C., van Veen, F. J., & Morrogh-Bernard, H. (2019). Gesture use in communication between mothers and offspring in wild orangutans (pongo pygmaeus wurmbii) from the sabangau peat-swamp forest, borneo. *International Journal of Primatology*, *40*, 393–416. https://doi.org/10.1007/s10764-019-00095-w

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

Labertoniere, D., & Skoruppa, K. (2022). La composition du lexique productif dans les trois premières années de la vie : Revue systématique. *A.N.A.E.*, *181*, 689–705.

Laporte, M. N., & Zuberbühler, K. (2011). The development of a greeting signal in wild chimpanzees. *Developmental Science*, *14*(5), 1220–1234. https://doi.org/10.1111/j.1467-7687.2011.01069.x

León, J., Thiriau, C., Bodin, C., Crockford, C., & Zuberbühler, K. (2022). Acquisition of predator knowledge from alarm calls via one-trial social learning in monkeys. *Iscience*, *25*(9). https://doi.org/10.1016/j.isci.2022.104853

Lieven, E. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua*, *120*(11), 2546–2556.

Lytle, S. R., & Kuhl, P. K. (2017). Social interaction and language acquisition: Toward a neurobiological view. *The handbook of psycholinguistics*, 615–634. https://doi.org/10.1002/9781118829516.ch27

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in chinese children. *Journal of child language*, *36*(2), 405–423.

MacLeod, A. A., Sutton, A., Trudeau, N., & Thordardottir, E. (2011). The acquisition of consonants in québécois french: A cross-sectional study of pre-school aged children. *International journal of speech-language pathology*, *13*(2), 93–109. https://doi.org/10.3109/17549507.2011.487543

MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of child language*, *12*(2), 271–295. https://doi.org/10.1017/S0305000900006449

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, *92*, 199–227. https://doi.org/10.1016/0024-3841(94)90342-5

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive psychology*, *20*(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*(6619), 813–815. https://doi.org/10.1038/385813a0

Matychuk, P. (2005). The role of child-directed speech in language acquisition: A case study. *Language sciences*, *27*(3), 301–379. https://doi.org/10.1016/j.langsci.2004.04.004

McShane, J. (1980). *Learning to talk*. Cambridge University Press.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014–9019. https://doi.org/10.1073/pnas.1105040108

Mikkelsen, T., Hillier, L., Eichler, E., Zody, M., Jaffe, D., Yang, S.-P., Enard, W., Hellmann, I., Lindblad-Toh, K., Altheide, T., et al. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, *437*(7055), 69–87. https://doi.org/10.1038/nature04072

Morgenstern, A. (2006). *Un je en construction: Genèse de l'auto-désignation chez le jeune enfant*. Editions OPHRYS.

Myowa-Yamakoshi, M., Scola, C., & Hirata, S. (2012). Humans and chimpanzees attend differently to goal-directed actions. *Nature communications*, *3*(1), 1–7. https://doi.org/10.1038/ncomms1695

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the society for research in child development*, 1–135.

Niehorster, D. C., Andersson, R., & Nyström, M. (2020). Titta: A toolbox for creating psychtoolbox and psychopy experiments with tobii eye trackers. *Behavior research methods*, *52*, 1970–1979. https://doi.org/10.3758/s13428-020-01358-8

Ninio, A., Snow, C. E., Pan, B. A., & Rollins, P. R. (1994). Classifying communicative acts in children's interactions. *Journal of communication disorders*, *27*(2), 157–187. https://doi.org/10.1016/0021-9924(94)90039-6

Ninio, A., & Wheeler, P. (1986). A manual for classifying verbal communicative acts in mother-infant interaction. *Transcript analysis*, *3*(1), 1–82.

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.

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6318–6323. https://doi.org/10.1073/pnas.1300337110

Oshima-Takane, Y., Ariyama, J., Kobayashi, T., Katerelos, M., & Poulin-Dubois, D. (2011). Early verb learning in 20-month-old japanese-speaking children. *Journal of child language*, *38*(3), 455–484. https://doi.org/10.1017/S0305000910000127

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, *106*(51), 22026–22031. https://doi.org/10.1073/pnas.0908118106

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897x00366

Pepperberg, I. M., Gardiner, L. I., & Luttrell, L. J. (1999). Limited contextual vocal learning in the grey parrot (psittacus erithacus): The effect of interactive co-viewers on videotaped instruction. *Journal of Comparative Psychology*, *113*(2), 158. https://doi.org/10.1037/0735-7036.113.2.158

Pepperberg, I. M. (1990). Some cognitive capacities of an african grey parrot (psittacus erithacus). In *Advances in the study of behavior* (pp. 357–409, Vol. 19). Elsevier. https://doi.org/10.1016/S0065-3454(08)60207-7

Piot, L., Chen, H., Picaud, A., Dos Santos, M., Granjon, L., Luo, Z., To, A. W. H., Lai, R. Y., Cheung, H., & Nazzi, T. (2024). Tonal interference in word learning? a comparison of cantonese and french. *Journal of Experimental Child Psychology*, *242*, 105883. https://doi.org/10.1016/j.jecp.2024.105883

Rovee-Collier, C., & Dufault, D. (1991). Multiple contexts and memory retrieval at three months. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *24*(1), 39–49. https://doi.org/10.1002/dev.420240104

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. https://doi.org/10.1177/0142723707067545

Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of child language*, *27*(3), 561–585. https://doi.org/10.1017/S0305000900004256

Savage-Rumbaugh, S., Rumbaugh, D. M., & McDonald, K. (1985). Language learning in two species of apes. *Neuroscience & Biobehavioral Reviews*, *9*(4), 653–665. https://doi.org/10.1016/0149-7634(85)90012-0

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child development*, *69*(2), 309–320. https://doi.org/10.1111/j.1467-8624.1998.tb06190.x

Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, *122*, 109–116. https://doi.org/10.1016/j.anbehav.2016.10.003

Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinović, D., Arnold, K., Cäsar, C., Keenan, S., et al. (2016). Formal monkey linguistics. *Theoretical Linguistics*, *42*(1-2), 1–90. https://doi.org/10.1515/tl-2016-0001

Schlenker, P., Coye, C., Steinert-Threlkeld, S., Klinedinst, N., & Chemla, E. (2022). Beyond anthropocentrism in comparative cognition: Recentering animal linguistics. *Cognitive Science*, *46*(12). https://doi.org/10.1111/cogs.13220

Skoruppa, K. (2019). Novel noun and verb learning in mono-and multilingual children. *Travaux neuchâtelois de linguistique*, (71), 109–123. https://doi.org/10.26034/tranel.2019.2994

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Snedeker, J., & Gleitman, L. R. (2004). Why it is hard to label our concepts. In D. Hall & S. Waxman (Eds.), *Weaving a lexicon* (pp. 257–294). MIT Press. https://doi.org/10.7551/mitpress/7185.003.0012

Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning how to say what one means: A longitudinal study of children's speech act use. *Social development*, *5*(1), 56–84. https://doi.org/10.1111/j.1467-9507.1996.tb00072.x

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166. https://doi.org/10.1016/S0010-0277(00)00081-0

Takagi, S., Koyasu, H., Nagasawa, M., & Kikusui, T. (2024). Rapid formation of picture-word association in cats. *Scientific Reports*, *14*(1), 23091. https://doi.org/10.1038/s41598-024-74006-2

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, *44*(4), 929. https://doi.org/10.1037/0012-1649.44.4.929

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. https://doi.org/10.1111/1467-8624.00045

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental psychology*, *30*(5), 639.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development*, 1454–1463. https://doi.org/10.2307/1130423

Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., Evans, A., et al. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, *14*(2), 175–186. https://doi.org/10.1016/S0047-2484(85)80005-1

Townsend, S. W., Deschner, T., & Zuberbühler, K. (2008). Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS One*, *3*(6), e2431. https://doi.org/10.1371/journal.pone.0002431

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive psychology*, *66*(1), 126–156. https://doi.org/10.1016/j.cogpsych.2012.10.001

Twomey, K. E., Ma, L., & Westermann, G. (2018). All the right noises: Background variability helps early word learning. *Cognitive Science*, *42*, 413–438. https://doi.org/10.1111/cogs.12539

Vernes, S. C., Janik, V. M., Fitch, W. T., & Slater, P. J. (2021). Vocal learning in animals and humans. https://doi.org/10.1098/rstb.2020.0234

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, *108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in psychology*, *3*, 46. https://doi.org/10.3389/fpsyg.2012.00046

Waxman, S. R., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H.-j. (2013). Are nouns learned before verbs? infants provide insight into a long-standing debate. *Child development perspectives*, *7*(3), 155–159. https://doi.org/10.1111/cdep.12032

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67–95. https://doi.org/10.1016/j.cogpsych.2009.02.001

Wilson, V. A., Bethell, E. J., & Nawroth, C. (2023). The use of gaze to study cognition: Limitations, solutions, and applications to animal welfare. *Frontiers in Psychology*, *14*, 1147278. https://doi.org/10.3389/fpsyg.2023.1147278

Wilson, V. A., Sauppe, S., Brocard, S., Ringen, E., Daum, M. M., Wermelinger, S., Gu, N., Andrews, C., Isasi-Isasmendi, A., Bickel, B., et al. (2024). Humans and great apes visually track event roles in similar ways. *PLoS Biology*, *22*(11). https://doi.org/10.1371/journal.pbio.3002857

Wojcik, E. H., & Saffran, J. R. (2013). The ontogeny of lexical networks: Toddlers encode the relationships among referents when learning novel words. *Psychological science*, *24*(10), 1898–1905. https://doi.org/10.1177/0956797613478198

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

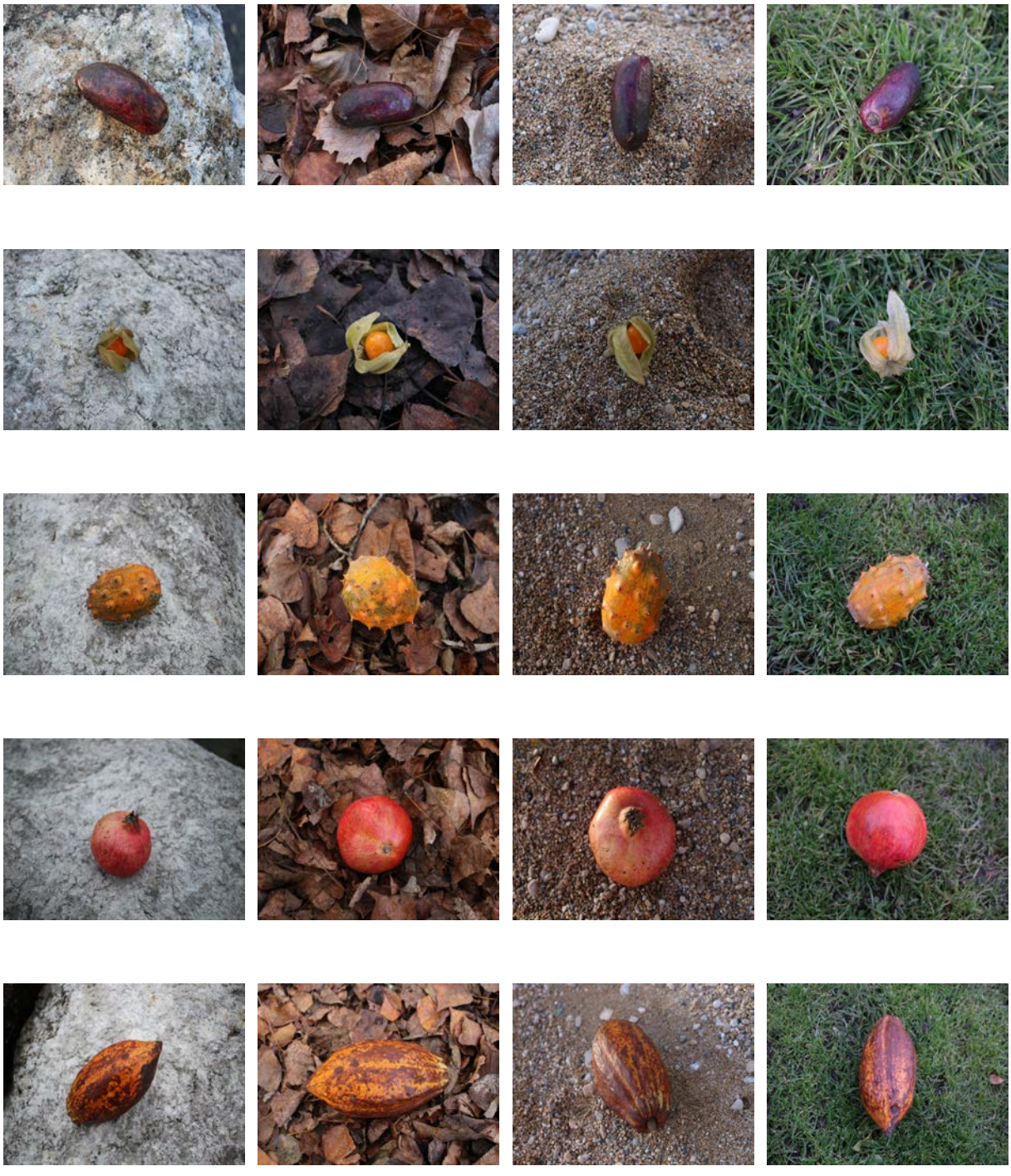
Zettersten, M., Yurovsky, D., Xu, T. L., Uner, S., Tsui, A. S. M., Schneider, R. M., Saleh, A. N., Meylan, S. C., Marchman, V. A., Mankewitz, J., et al. (2023). Peekbank: An open, large-scale repository for developmental eye-tracking data of children's word recognition. *Behavior Research Methods*, *55*(5), 2485–2500. https://doi.org/10.3758/s13428-022-01906-4

Zuberbühler, K. (2000). Referential labelling in diana monkeys. *Animal behaviour*, *59*(5), 917–927. https://doi.org/10.1006/anbe.1999.1317

Zuberbühler, K. (2001). Predator-specific alarm calls in campbell's monkeys, cercopithecus campbelli. *Behavioral Ecology and Sociobiology*, *50*, 414–422. https://doi.org/10.1007/s002650100383

A. Stimuli and Babylab setup

Visual stimuli used in the context experiments



Screencaps of visual stimuli used in the semantics experiments













Infant and their parent in the testing booth for the eye-tracking experiments at the Babylab

Specific consent about the use of this photograph in scientific reports was obtained from the parent.

B. List of speech acts for use in the corpus study

1. Answer to questions or demands

- QA Answer a question with a wh-question.
- TA Answer a limited-alternative question (ex. 'C'est de quelle couleur entre bleu, jaune et vert ?').
- AA Answer in the affirmative to yes/no question.
- AN Answer in the negative to yes/no question.
- NA Intentionally non-satisfying answer to question.
- SA Answer a wh-question with a statement.
- YA Answer a question with a yes/no question.
- AD Agree to carry out an act requested or proposed by other.
- GI Give in/ accept other's insistence or refusal.
- AL Agree to do something for the last time.

2. Questions

- QN Ask a product-question (wh-question).
- EQ Eliciting question (e.g. hmm?).
- RQ Yes/no question or suggestion about hearer's wishes and intentions.
- YQ Ask a yes/no question.
- TQ Ask a limited-alternative yes/no question .
- AQ Aggravated question expression of disapproval by restating a question.

3. Emotions

- EM Exclaim in distress pain.
- EN Express positive emotion.
- ES Express surprise.
- CM Commiserate express sympathy for hearer's distress.

- ED Exclaim in disapproval.
- ET Express enthusiasm for hearer's performance.

4. Demandes

- RP Request propose or suggest an action for hearer or for hearer and speaker.
- WS Express a wish.
- DR Dare or challenge hearer to perform an action.
- RR Request to repeat utterance.
- FP Ask for permission to carry out act.
- EC Elicit completion of word or sentence.
- EX Elicit completion of rote-learned text.

5. Declarations

- ST Make a declarative statement (comment uttered without input, spontaneous narration ...).
- YD Agree to a declaration.
- SI State intent to carry out act by speaker.
- DC Create a new state of affairs by declaration.
- DP Declare make-believe reality.
- AP Agree with proposition or proposal expressed by previous speaker.
- SC Complete statement or other utterance in compliance with request.
- AB Approve of appropriate behavior.
- MK Mark occurrence of event (thank greet apologize congratulate etc.).
- PM Praise for motor acts i.e for nonverbal behavior.
- CN Count.
- PD Promise.
- CX Complete text if so demanded.
- CR Criticize or point out error in nonverbal act.
- GR Give reason/ justify a request for an action refusal or prohibition.
- TX Read or recite written text aloud.

- TD Threaten to do.
- PA Permit hearer to perform act.

6. Refusal
   - RD Refuse to carry out an act requested or proposed by other.
   - PF Prohibit/forbid/protest hearer's performance of an act.
   - DW Disagree with proposition expressed by previous speaker.
   - RA Refuse to answer.
   - DS Disapprove scold protest disruptive behavior.
   - ND Disagree with a declaration.
   - CS Counter-suggestion/ an indirect refusal.
7. Attirer l'attention
   - AC Answer calls/ show attentiveness to communications.
   - CL Call attention to hearer by name or by substitute exclamations.
   - XA Exhibit attentiveness to hearer.
   - WD Warn of danger.
   - SS Signal to start performing an act such as running or rolling a ball.
8. Accompagnement verbal d'une action
   - TO Mark transfer of object to hearer.
   - PR Perform verbal move in game.
9. Sound / Imitation / Repetition
   - EA Elicit onomatopoeic or animal sounds.
   - EI Elicit imitation of word or sentence by modelling or by explicit command.
   - RT Repeat or imitate other's utterance.
   - CT Correct provide correct verbal form in place of erroneous one (auto-correction).
   - YY Make a word-like utterance without clear function.
   - OO Unintelligible vocalization (babillage).