Implicit Learning as a Mechanism for Syntactic Acquisition and Processing: Evidence from Syntactic Priming

Shanthi Julia Catanzariti Kumarage
Primary supervisor: Prof Evan Kidd
Supervisory panel: Prof Caroline Rowland, Dr Amy Dawel
May 2024

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

Declaration

This thesis contains original research undertaken during the Doctor of Philosophy in the School of Medicine and Psychology at the Australian National University, in affiliation with and financially supported by the Max Planck Institute of Psycholinguistics. Chapters 1 and 5 are my own work. Chapters 2 and 3 are manuscripts which have been published. Chapter 4 is in preparation for submission. The contributions of co-authors are detailed at the start of each chapter. All ideas that are not my own have been properly acknowledged and referenced.

Shanthi Julia Catanzariti Kumarage
May 2024

Acknowledgements

This research was funded by the Max Planck Society for the Advancement of Science (www.mpg.de/en) in the form of a PhD studentship in the Language Development Department at the Max Planck Institute for Psycholinguistics awarded to Shanthi Julia Catanzariti Kumarage.

I am grateful to so many people who supported me during the PhD journey and to the Max Planck Institute of Psycholinguistics for funding my study.

Firstly, I would not have finished or even started my thesis without the mentorship and support of my primary supervisor Professor Evan Kidd. I count myself very lucky to have crossed paths with you as an undergraduate curious about how language works in the brain. It is difficult to enumerate all that I am grateful for: your generosity with both time and knowledge, the invaluable insights you always have during the discussion of ideas, your wit, humour and down-to-earth nature in the often competitive world of academia, the incredibly quick return of drafts and emails, and your kindness and belief in me throughout my PhD.

I am also grateful to the members of my supervisory panel, Professor Caroline Rowland, who created the fantastic environment at the Language Development Department where I learnt so much, and Associate Professor Amy Dawel who warmly welcomed me back to ANU.

Dr Seamus Donnelly was my statistics guru, co-author and close friend. Thank you for teaching me so much of what I know about statistics, for being a bouncing board for ideas and coding woes, for our many walks for sweet treats, and for humouring me when I send through a graph I'm proud of making. Thank you to my other co-author, Dr Anton Malko, who taught me even more about coding and introduced me to the world of analysing eye-tracking data.

I am grateful for the time I was able to spend at the Language Development Department at the Max Planck Institute of Psycholinguistics. I was introduced to concepts and practices of Open Science and reproducibility that influenced my research throughout my PhD. I'm also grateful for all the incredibly clever and friendly people in the department who broadened my understanding of psycholinguistic research and made my stay in the Netherlands more enjoyable: Rebecca, Middy, Rowena, Raquel, Judit, Katja, Julia, Melis, and Yevheniy to name a few.

I owe a great deal to the members of the Language Lab; those already mentioned, and Lauren, Amanda, Katherine and Tanya. Thank you for the incredible work you did on the Canberra Longitudinal Child Language project, which gave me the data for my first PhD study. Thank you also for the best work environment I ever had, complete with cake, friendship and handstand walls. The parents and children who participated in the research showed amazing continued commitment to the project and I am extremely grateful for their time.

Thank you to the participants in my final study, including the many friends who responded to the callout. It was a wonderful feeling to have the support of my Canberra network as I came to the end of my studies. A special shoutout to Maggie, who lent her vocal talent to the cause.

No one understands the ups and downs of the thesis quite like another PhD student. Thank you to Ellie, Aseel, Caroline, Kira, Zach, Imogen, Liz, Keira and many others for sharing in the successes and failures at various points over the years.

Finally thank you to my friends and family. The fun, laughter and happy moments we shared mean a lot to me and kept me sane during the PhD journey. Thanks Mum and Dad for putting me up during a global pandemic, and all your love, support and encouragement to pursue my passions. I’m looking forward to creating more memories with all of you now that this chapter is closed.

Thesis Abstract

Learning to decode and communicate meaning from how words are combined is a challenge that children must meet in order to acquire the syntax of their language. The mechanisms of this process are hotly debated: do children have innate linguistic knowledge guiding their learning or are their innate abilities limited to learning mechanisms that infer knowledge from input? Syntactic priming offers an experimental paradigm that can test different theories of syntactic acquisition. Presenting a prime sentence of a particular syntactic structure (e.g., the passive: *the swimmer was eaten by a crocodile*) tends to increase the likelihood of participants later producing that structure over an alternative (e.g., *the cyclist was swooped by the magpie* vs the active: *the magpie swooped the cyclist*). A syntactic priming effect implies shared representation between the prime and target, illuminating the nature of the underlying syntactic representation. In addition, syntactic priming may be a short-term manifestation of a proposed mechanism of syntactic acquisition and processing: implicit error-based learning. The aim of this thesis was to investigate the contribution that research using syntactic priming can make to our understanding of mechanisms of syntactic acquisition and processing.

The first part of this thesis focuses on acquisition. It reports the first longitudinal study of syntactic priming in children aged 3;0 – 4;6 years, tracking the development of priming with and without shared lexical content between primes and participants' responses (e.g., both being *swooping* events). The developmental trajectories of abstract and lexically-dependent knowledge are key to differentiating between theories of syntactic acquisition. Abstract priming emerged early and decreased across development once the target structure had been acquired, while lexically-specific priming emerged later and increased over development. This pattern is most consistent with an implicit error-based learning account rather than lexicalist accounts where initial syntactic representations are tied to lexical items, or purely nativist accounts where priming effects are expected to be stable, like the representations they tap into.

The second study in Part 1 of this thesis synthesised the existing syntactic priming literature. A meta-analysis of syntactic priming studies in children showed that the priming effect is robust and reliable. The structural alternation under investigation and aspects of study design were identified as influences on the syntactic priming effect that researchers should consider. A key finding was that priming was larger with, but not dependent on,

shared lexical content between primes and participants' responses, supporting the findings of the longitudinal study.

The second part of this thesis explored combining syntactic priming with pupillometry, a real-time psychophysiological measure. The implicit error-based learning account proposes a cognitive architecture that is continuous from children to adults, linking syntactic acquisition in children to syntax processing in adults. It posits that prediction error leads to representational change. Pupil size provided a potential index of prediction error, allowing exploration of the mechanistic link between unexpected syntactic structure and representational change as measured via priming.

Overall, this thesis applies three lenses to syntactic priming – longitudinal research, meta-analysis, and online psychophysiological measurement – to extend the utility of the methodology, the conclusions we can draw from it and the depth of evidence for an implicit error-based learning account of syntax processing and acquisition.

Table of Contents

Chapter 1:

General Introduction

Thesis Rationale

Imagine someone telling an anecdote to their friend ending with the sentence "and then the dog was bitten by the man!" The friend knows that this sentence means the unusual occurrence *man bit dog* rather than *dog bit man*. However, it is unlikely that they could articulate their knowledge that the combination of auxiliary verb, past participle and oblique by-phrase indicates a passive structure where the subject of the sentence is the undergoer of the action and the agent of the action is expressed in the by-phrase. Nor could they point to when they learnt this or how. The story-teller also possesses this knowledge and, presuming they don't commonly associate with men with a taste for companion animals, used it to produce a sentence they have never heard before.

This example illustrates several key points. Firstly, both speaker and listener have intricate knowledge of English sentence structure and how structural information relates to functional roles of participants. That is, over the course of their development they were able to acquire complex and detailed syntactic knowledge. Secondly, language use is creative and speakers are able to produce sentences that are not simple imitations of language input. This requires abstract representations of syntax which operate over categories of words rather than words themselves. Third, there are cognitive processing mechanisms that allow the use of this knowledge *in the moment*. And finally, although we possess detailed syntactic knowledge and undertake complex syntactic processing in order to communicate, we can't explain what that knowledge looks like or how we use it in real-time. That is, it is largely an *implicit process*.

These are central questions in psycholinguistics. How is detailed complex linguistic knowledge represented and utilised by the brain? How is it acquired? And how can we study cognitive representations and processes into which people, especially children, do not have conscious insight? This thesis focuses on using syntactic priming as an experimental methodology to investigate syntactic acquisition and processing. The aims are: (i) to discriminate between theoretical accounts of these cognitive abilities, and (ii) to investigate the evidence that syntactic priming can provide when three different lenses are applied: longitudinal research, meta-analysis and pupillometry.

The first part of this thesis focuses on the acquisition of syntax. Chapter 2 reports the first longitudinal study of syntactic priming in children. It tracks the trajectory of syntactic priming effects to distinguish between the predictions of three accounts of syntax acquisition. Chapter 3 evaluates the existing developmental syntactic priming literature using meta-analysis to identify factors influencing syntactic priming and the insights they reveal into syntactic acquisition. The second part of this thesis focuses on syntactic processing. Chapter 4 combines pupillometry with syntactic priming in adults to attempt to index the cognitive processing proposed by the implicit learning account (Chang et al., 2006).

Accounts of Syntactic Acquisition

Children must acquire syntax to become skilled communicators. Once they have overcome the challenge of segmenting words from speech and mapping them to meanings, children must learn how to creatively combine words into messages. Researchers agree that mature syntactic knowledge is abstract, operating over categories of words. However, a key theoretical debate in language acquisition research is how children get to that destination. Theories contrast in whether they invoke innate linguistic knowledge or innately specified but domain-general learning mechanisms (Ambridge & Lieven, 2011).

The modern nativist tradition began with Chomsky's (1965) proposal that children come to language acquisition armed with substantial innate syntactic knowledge, termed Universal Grammar. He argued that given the "poverty of the stimulus", or the sparsity and noisiness of language input compared to the vast number of potential rules, children could not induce a complex and abstract grammar from input. Instead, Chomsky argued that children are endowed with a set of innate linguistic principles, universal to all languages, with variation across language captured by innate parameters that constrain how languages can vary. Children arrive at the syntax of their native language by 'setting' these parameters rather than constructing them from the input. Modern nativist theories differ in details but retain the core assumption that children already possess syntactic knowledge (Lidz et al., 2003; Valian, 2014; Yang, 2018). A lighter version of the nativist approach argues that children are born with some basic guiding heuristics and quickly bootstrap into abstract knowledge (Gertner et al., 2006). The commonality is that the approaches predict early abstract knowledge in children.

In contrast, *emergentist* approaches assume that children can learn syntactic knowledge using domain-general mechanisms. The mechanisms proposed differ but all assume that the bulk of linguistic knowledge is induced from the input (i.e., learnt). The *lexicalist account* assumes that children transition from lexically-specific to abstract syntactic representation (Ambridge & Lieven, 2015; Tomasello, 2003). Tomasello (2000) proposed that children first learn syntactic structures for particular words that they often hear in that structure (i.e., as *verb islands*). Once a critical mass of words in that structure is learnt, children can abstract across their separate item-based representations to form an abstract one. In comparison to nativist theories, lexicalist ones assume the late abstraction of syntactic knowledge. Another emergentist approach, which is the focus of the second part of this thesis, is the implicit error-based learning account (Chang et al., 2006), which is instantiated in a connectionist model. Under this account, syntactic knowledge is represented via connection weights within a structured network (Chang et al., 2006). The language processor predicts upcoming words and adjusts network weights whenever an error is made in order to acquire syntactic structures in language.

These three accounts of syntactic acquisition, nativist, lexicalist, and implicit learning, differ in the assumed nature of early syntactic representation and how those representations develop. In the next section of this chapter, I introduce syntactic priming as a method for tapping into children's syntactic representations. The three theories make different predictions about syntactic priming effects and how they develop. Chapter 2 tests these predictions by tracking the trajectory of abstract and lexical syntactic priming effects in a longitudinal design. Chapter 3 synthesises the developmental syntactic priming literature, including lexical, abstract and age effects using meta-analysis.

Syntactic Priming

Syntactic priming is the phenomenon whereby processing a syntactic structure facilitates the subsequent processing of that structure (Bock, 1986). For example, hearing the passive sentence *Geppetto was swallowed by the whale* increases the likelihood of later producing a passive structure like *Pinocchio was transformed by the fairy* rather than an active structure such as *the fairy transformed Pinocchio*. In an experimental setting, participants are presented with a prime sentence to read, listen to or produce, and then its influence on the processing of a target is measured (Bock, 1986). Prime sentences occur as one of two structures in a structural alternation such as the active-passive alternation or dative alternation (*Geppetto gave an apple to Pinocchio* vs *Geppetto gave Pinocchio an apple*). Targets may be a picture which can be described using either structure, a sentence stem which can be completed with either structure, or, in the case of comprehension priming, a second sentence to perceive. Priming is said to occur when participants show an increased likelihood of producing the primed structure in their target descriptions or facilitated comprehension of the primed structure in target sentences.

If a priming effect is observed, shared representation between the prime and target is assumed (Branigan & Pickering, 2017). Through the careful manipulation of the content of primes and targets, researchers have tested the nature of the representations tapped by priming. For example, if the prime uses one lexical item (e.g., *swallow*) but the target response contains a different one (*transform*), then a priming effect indicates the existence of an abstract representation of the structure (e.g., passive) that is not specific to either verb. If a priming effect is only observed when the prime and target share lexical content (e.g., *Geppetto was* ***swallowed*** *by the whale* primes *the red pill was* ***swallowed*** *by Neo* but not *Pinocchio was* ***transformed*** *by the fairy*), then only lexically-specific syntactic representations are supported. Participants' reuse of sentence structure occurs largely outside of conscious awareness, suggesting that it provides a measure of implicit representations of syntax (Bock, 1986). Syntactic priming is therefore a useful alternative to methods like acceptability judgments, which require explicit insight into what is considered grammatically correct (Branigan & Pickering, 2017). This is a particular advantage in acquisition research in developmental populations. Whereas young children have difficulty making meta-linguistic judgments (Ambridge & Rowland, 2013), they do demonstrate syntactic priming effects (Bencini & Valian, 2008). In the next section, I summarise the key topics explored by syntactic priming research in children.

Syntactic Priming in Acquisition

Theory Testing and the Need for Longitudinal Data

In children, the mental representations that syntactic priming taps into are still being acquired, thus enabling researchers to gain insight into the nature of *developing* representations. The theoretical debate between nativist and emergentist accounts of syntactic acquisition means much of the syntactic priming literature focuses on the emergence of abstract as opposed to lexically-specific syntactic knowledge (e.g., Bencini & Valian, 2008; Savage et al., 2003). Nativist accounts assume early abstraction and therefore early evidence of abstract knowledge (Messenger & Fisher, 2018) and lexicalist accounts assume late abstraction with lexically-specific knowledge preceding abstract knowledge (Tomasello, 2000). Research appears to support early over late abstraction, with evidence for priming without lexical overlap between prime and target from age 3 years (Bencini & Valian, 2008; Hsu, 2019; Shimpi et al., 2007). However, studies have typically studied priming effects at a group level rather than examining individual differences (Kidd, 2012). Task adjustments in order to observe abstract syntactic priming in the youngest children (Shimpi et al., 2007) and high dropout rates (Bencini & Valian, 2008) suggest variability in early knowledge, a feature that is best accommodated by emergentist not nativist accounts.

Regarding lexically-specific knowledge, only two studies have demonstrated priming with lexical overlap in the absence of abstract syntactic priming (Donnelly et al., 2024; Savage et al., 2003). In adults, shared lexical content increases the magnitude of the priming effect, an effect called the *lexical boost*. Beyond the lack of evidence for lexically-specific priming, syntactic priming studies in children have not consistently found evidence for the lexical boost. Whilst some researchers have found larger priming effects when lexical content is repeated (e.g., Branigan, McLean, et al., 2005; Branigan & McLean, 2016), others have found that priming effects are equivalent in abstract and lexical overlap conditions (e.g., Foltz et al., 2015; Peter et al., 2015; Rowland et al., 2012).

Moving beyond the emergence of abstract and lexical effects is research on their developmental trajectory. Chapter 2 details the predictions that nativist, lexicalist and implicit learning accounts make about how abstract priming and the lexical boost effect develop. In brief, nativist accounts assume stable priming effects whilst emergentist accounts predict varying trajectories given syntactic representations themselves undergo change. Rowland et al. (2012) were the first to investigate these trajectories, measuring syntactic priming with and without lexical overlap in 3-4 year olds, 5-6 year olds, and adults. Abstract priming was largest in the youngest age group but the lexical boost was largest in adults. Other cross-sectional studies have also compared samples of children of different ages to determine whether and how priming effects develop over relevant periods of acquisition (Donnelly et al., 2024: study 1; Garcia & Kidd, 2020; Hsu, 2019; Kholodova et al., 2023; Peter et al., 2015). However, these studies have not reached a consensus regarding whether the abstract priming effect is stable, increases or decreases over development and only studies of the dative structural alternation have investigated the developmental trajectory of the lexical boost effect (Donnelly et al., 2024; Kholodova et al., 2023; Peter et al., 2015; Rowland et al., 2012). Messenger et al. (2022) point out that in syntactic priming studies comparing samples cross-sectionally, the age range of children within a sample is typically 20 months. During a period of acquisition where language undergoes rapid development, such samples are likely to contain considerable variability, hindering the ability to detect developmental change. That is, any differences across age groups in a cross-sectional design might be obscured by the fact that there is wide variability of knowledge *within* age groups.

In sum, syntactic priming studies find that abstract syntactic knowledge emerges early and is not preceded by lexically-specific knowledge. However, variability in early abstract knowledge has not been investigated in detail and the development of abstract priming effects is unclear. The lexical boost effect is not reliably found in children and its trajectory is understudied. Chapter 2 aimed to address these gaps in the literature and the shortcomings of cross-sectional research by using a longitudinal design where children are compared to themselves at timepoints with tightly-controlled ages. The study manipulated lexical overlap between primes and targets and so could investigate the emergence and development of abstract priming and the lexical boost effect.

Synthesising the Literature

The developmental syntactic priming literature is not limited to investigations of abstract and lexically-based knowledge. Studies have also investigated how syntactic representations are constrained by or interact with semantic features. For example, researchers have investigated whether priming is affected by animacy preferences for human and animate arguments to precede inanimate ones (i.e., *the woman was hit by the car* is a more likely passive than *the car was bought by the woman*; Buckle et al., 2017; Vasilyeva & Gámez, 2015) and whether syntactic representation is shared across actional and experiential verb-types, which differ in the thematic roles assigned to arguments (i.e., in actional verbs an agent performs an action on a patient such as *agent kicks patient*, whereas experiential verbs involve an experiencer and theme such as *experiencer fears theme*; Bidgood et al., 2020; Messenger et al., 2012). Other research concerns the underlying mechanisms or architecture of the language system. For example, the implicit learning account predicts long-term priming effects so studies have investigated whether and how priming effects persist over time (Branigan & McLean, 2016; Kidd, 2012; Savage et al., 2006). If syntactic representations are shared across languages, then priming will be observed from one language to another, a prediction tested by crosslinguistic priming studies (Gámez & Vasilyeva, 2020; Vasilyeva et al., 2010; Wolleb et al., 2018). Other studies have manipulated whether children repeat or simply listen to prime sentences, addressing whether syntactic representations are shared by comprehension and production processing mechanisms (Gámez & Shimpi, 2016; Huttenlocher et al., 2004; Shimpi et al., 2007). Much of the syntactic priming literature focuses on the active-passive or dative alternations in English. However, the methodology has more recently been applied to syntactic alternations in non-European languages (Tagalog: Garcia & Kidd, 2020; Mandarin: Hsu, 2014a, 2014b, 2019).

Overall, the syntactic priming literature in children has focused on a range of theorised moderators of the effect, which reveal insights into the underlying nature of representation and processing. Quantitatively summarising research via meta-analysis achieves greater power and can reveal summary findings that aren’t apparent at the level of individual studies (Borenstein et al., 2009). Chapter 3 aimed to synthesise the breadth of syntactic priming research in children to identify summary findings that reveal insights into the acquisition of syntax, including analysing the lexical and abstract priming effects examined by Chapter 2. Together, Chapters 2 and 3 contribute novel evidence to discriminate theories of syntax acquisition by combining syntactic priming with longitudinal design and meta-analysis.

I now turn from acquisition to adult syntax processing, where syntactic priming has long been an important and widely-used experimental paradigm. Research in adults informs both how mature syntactic knowledge is represented and mechanisms underlying syntactic priming and therefore sentence production (Branigan & Pickering, 2017; Pickering & Branigan, 1998). In the next section, I introduce the proposed mechanisms of syntactic priming and then findings from syntactic priming that bear upon them.

Mechanisms of Syntactic Priming

When the sentence *Geppetto was swallowed by the whale* primes *Pinocchio was transformed by the fairy*, we can assume an abstract representation of the passive that is shared between the two sentences. However, why does encountering a structure make it more likely to be used soon after? The phenomenon reveals something about the way that we process language, which is the focus of theories about the mechanisms of syntactic priming. There are two main mechanisms proposed for syntactic priming effects: residual activation and error-based learning.

Under the residual activation account, processing the prime structure activates the syntactic representation of the structure, facilitating its use in the target (Pickering & Branigan, 1998). Additionally, links between structural representations and particular words are subject to residual activation, explaining why lexical overlap between primes and targets increases the magnitude of priming (Pickering & Branigan, 1998). However, the residual activation account predicts that priming effects should be relatively short-lived, a prediction that is inconsistent with the observation that priming effects can endure over 10 sentences (Bock et al., 2007; Bock & Griffin, 2000) or even up to a week (Kaschak, Kutta, & Schatschneider, 2011). These findings suggest priming is not solely due to transient changes to the linguistic system, but also enduring ones that result from learning. Reitter et al. (2011) proposed a model that can account for learning in syntactic priming using activation-based mechanisms. A *spreading activation* mechanism behaves similarly to Pickering and Branigan's (1998) residual activation account but each retrieval of a syntactic structure also increases its *base-level activation*, inducing long-term learning effects.

Another interpretation of syntactic priming is that it is implicit error-based learning in action (Chang et al., 2006; Dell & Chang, 2014). During the prime sentence, word-by-word prediction occurs and any prediction errors instigate changes in connection weights to reduce the likelihood of the same prediction error in future. For example, on hearing *Geppetto was…*, the model is likely to predict a present participle (*snoring, eating, carving*) in line with the bias towards agent-first sentences in English. However, when the sentence continues *Geppetto was swallowed…*, prediction error is generated and backpropagates through the network. The model adjusts its connection weights to anticipate passive sentences more frequently. The same adjustments that make predicting a passive more likely also make producing a passive more likely during description of the target because the same representations and processing underlie both.

The key difference between activation accounts and the error-based learning account is the role of prediction error in their mechanisms. Predictive processing is central to the

Chang et al. (2006) account, with learning instigated by and proportional to the prediction error produced by primes, whereas predictive processing is not implemented by activation accounts. Consequently, the representations invoked by the accounts differ. Activation accounts assume existing representations whose accessibility is altered by activation mechanisms (Roelofs' (1992) lemma representations for Pickering & Branigan (1998); and Steedman's (1999) Combinatory Categorial Grammar for Reitter et al. (2011)). However, Chang et al.'s (2006) model implicitly learns the representations themselves through learning to sequence words in ways that minimise prediction error, the same mechanism that adjusts expectations for a structure once it has been learnt. Chang et al.'s (2006) model can acquire abstract and generalisable representations of syntax due to architectural assumptions of a sequencing system based on a simple recurrent network (Elman, 1990), and restricted interaction between the sequencing system and a separate meaning system. On the other hand, activation accounts simply assume these properties of the representations they utilise (i.e., they are given). In other words, the error-based learning account makes predictions about syntactic acquisition as well as priming. The first part of this thesis considered syntactic acquisition, and the second error-based learning in adult syntactic priming. In the next section, I will summarise the literature on syntactic priming in adults with a focus on effects with implications for the mechanistic understanding of priming.

Syntactic Priming Research in Adults

Similar to the developmental syntactic priming literature, early studies in adults focused on the nature of syntactic representation. Bock's (1986) seminal study demonstrated syntactic priming without the repetition of conceptual or lexical elements. She concluded that the phenomenon therefore involved the activation or strengthening of syntactic representations that are abstract and isolable from other levels of representation such as semantics or the lexicon.

Supporting lexically-independent syntactic knowledge, priming is observed in the absence of both repeated open-class lexical items (e.g., verb overlap; Pickering & Branigan, 1998) and closed-class lexical items (i.e., *for*-datives such as *the secretary was baking a cake for her boss* prime *to*-datives such as *a cheerleader offered a seat to her friend*; Bock, 1989). However, lexical overlap between primes and targets does increase the strength of syntactic priming (Pickering & Branigan, 1998). In Mahowald et al.'s (2016) meta-analysis of syntactic priming studies in adults, the large and robust *lexical boost* effect was one of the key summary findings. Activation accounts of syntactic priming can account for the lexical boost because the links between structures and lexical items are also subject to activation effects (Pickering & Branigan, 1998; Reitter et al., 2011). The error-based learning account does not explain the lexical boost effect and Chang et al. (2006) instead attribute it to explicit memory, where the repeated word provides a cue to the primed structure. Unlike abstract priming, which can endure over long periods of time (Bock et al., 2007; Bock & Griffin, 2000; Kaschak, Kutta, & Schatschneider, 2011), the lexical boost decays quickly (Hartsuiker et al., 2008; Mahowald et al., 2016). This provides some support for a dual-mechanism account of the two effects (Chang et al., 2006; Hartsuiker et al., 2008).

Syntactic priming occurs when structures differ in word order alone, supporting priming of syntax independently of semantic information (e.g., *pulled the sweater off* vs *pulled off the sweater*: Konopka & Bock, 2009; see also, Ferreira, 2003; Hartsuiker & Westenberg, 2000). However, aspects of semantics such as the ordering of thematic roles and animacy can be primed independently of syntax (Bock et al., 1992; Chang et al., 2003; Chen et al., 2022; Ziegler & Snedeker, 2018). For example, location-theme locatives (*the farmer heaped the wagon with straw*) and theme-location locatives (*the farmer heaped straw onto the wagon*) both have the same syntactic structure (NP [V NP [P NP]$_{PP}$]$_{VP}$) but the order of the theme and location arguments can be primed (Chang et al., 2003). There is less clarity regarding the interaction between syntactic and semantic representation. In some cases, when semantic cues align with syntactic ones, priming is strengthened (Ziegler & Snedeker, 2018) but in others there is no interaction (Bock et al., 1992; Chen et al., 2022). Branigan & Pickering (2017) conclude that semantics and syntax are separate levels of representation. This is compatible with the Chang et al. (2006) account given the instantiation of a separate meaning system and with activation accounts if they act upon semantic representations.

Other syntactic priming effects relate directly to the predictability or frequency of prime input. Inverse-frequency effects describe larger priming effects being observed for less frequent structures. For example, hearing *the man was bitten by the dog* is more surprising than hearing *the dog bit the man* because passive structures are extremely infrequent relative to actives (Roland et al., 2007). Empirically, passive sentences prime passives more strongly than active sentences prime actives (Bock, 1986). This finding extends to the dative alternation (Bernolet & Hartsuiker, 2010; Jaeger & Snider, 2013; Kaschak, Kutta, & Jones, 2011), the use or omission of the English complementiser *that* (Ferreira, 2003), and relative clause attachment (Scheepers, 2003). Frequency effects are not limited to structural alternations. Notably, verbs can be biased towards appearing in one of two structures in a syntactic alternation. Hearing a structure with a verb that typically uses the opposite structure is more surprising than hearing the structure with a verb biased towards it (e.g., *the fielder threw the wicketkeeper the ball* is more surprising than *the fielder threw the ball to the wicketkeeper* because *throw* typically occurs in a prepositional not double object dative). Researchers have manipulated both prime structure and verb-bias to demonstrate *prime-surprisal effects*, where stronger priming is observed when the structure of a prime mismatches the bias of the verb (Bernolet & Hartsuiker, 2010; Jaeger & Snider, 2013; Peter et al., 2015). Inverse-frequency and verb-bias effects can be interpreted as the result of less predictable input producing larger prediction error and therefore greater representational change as per the error-based learning account (Chang et al., 2006). Reitter et al.'s (2011) model can also account for inverse-frequency effects because changes in base-level activation are smaller for more frequent structures, but it is unclear how verb-bias effects can be explained.

Accounts of syntactic priming effects have focused on priming the production of a structure. However, researchers have employed a variety of methodologies to investigate priming in comprehension. These include the interpretation of sentences (e.g., whether participants interpret *she saw the man with the telescope* as *the man had a telescope* or as *she used the telescope*; Branigan, Pickering, et al., 2005) and anticipation of syntactic structures (e.g., whether participants look at a picture of a child or of a rope first after hearing *the lifeguard tossed…*; Arai et al., 2007). Other studies have used measures of processing effort such as EEG (Tooley et al., 2014) and reading times (Traxler et al., 2014) to demonstrate facilitated processing of sentences after a prime of the same syntactic structure. Early comprehension studies frequently found syntactic priming effects in lexical overlap conditions but not abstract priming (Tooley & Traxler, 2010). However, more recent studies that are careful to use comparable structures and methods to production studies have found abstract priming, suggesting abstract representations of syntax do also operate in comprehension (Tooley, 2023; Tooley & Bock, 2014). The types of methods used in comprehension studies offer online measures of underlying processing that could be applied in production priming studies (Tooley, 2023). For example, Arai & Chang (2024) attempted to link manipulations of input to a reading time measure of prediction error and subsequently to production effects. However, their reading time measure did not yield the expected results. In Chapter 4, I introduce pupillometry: the measurement of cognitive processes and events using pupil size (Sirois & Brisson, 2014). Pupil dilation can index surprisal (Preuschoff et al., 2011) and I investigate its potential as a measure of prediction error, a key feature of error-based learning, in syntax processing during syntactic priming.

Thesis Outline

In this thesis, I investigated syntactic priming as a means to answer questions about how complex syntactic knowledge is acquired, represented and utilised. Syntactic priming is a widely used methodology in developmental and adult populations. However, at present, it is difficult to discriminate between major theories of acquisition on the basis of the syntactic priming literature. In adults, a variety of findings bear upon the underlying mechanisms of syntactic priming, but using online measures to detect processing signatures of proposed mechanisms is underexplored. The aims of this thesis were: (i) to provide evidence regarding underlying mechanisms in acquisition and processing, and (ii) to extend our understanding of the utility of the syntactic priming methodology in investigating such questions.

Chapter 2 derives and tests predictions from three accounts of syntactic acquisition: nativist, lexicalist, and implicit error-based learning, using a longitudinal design. The study manipulated prime structure and lexical overlap in a longitudinal design, measuring children's syntactic priming every six months from 36 to 54 months of age. It tracked the emergence and development of abstract priming and lexical boost effects, evidence for which is unclear in the literature and limited to cross-sectional designs. **Chapter 3** quantitatively summarises the developmental syntactic priming literature using meta-analysis. The analysis synthesised findings regarding the emergence and development of abstract and lexicalised syntactic knowledge, which are relevant to the questions in Chapter 2. Additionally, other moderators of the priming effect such as lag between prime and target, prime repetition and noun animacy were investigated, which provide additional insights into the nature and mechanisms of children's acquisition of syntax. **Chapter 4** turns from the acquisition of syntax to its processing in adulthood. It investigated evidence for the role of prediction error by attempting to measure it directly using pupillometry. Participants completed a syntactic priming study and their pupil size during prime comprehension was measured and tested as a predictor of their priming during target description. Finally, **Chapter 5** summarises the findings of Chapters 2 to 4 and discusses their implications for syntactic priming research, the implicit error-based learning account, and future research.

Publication Details

Chapters 2, 3, and 4 were prepared as manuscripts for journal publication. Chapters 2 and 3 appear here as the versions accepted for publication (before the type-setting process) and have only been edited for formatting consistency. Chapter 4 has been prepared for submission but is not published.

Chapter 2: Published in *Journal of Memory and Language*

Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring children. Journal of Memory and Language, 127, Article 104374. https://doi.org/10.1016/j.jml.2022.104374

Chapter 3: Published in *Journal of Memory and Language*

Kumarage, S., Donnelly, S., & Kidd, E. (2024). A meta-analysis of syntactic priming experiments in children. Journal of Memory and Language, 138, Article 104532. https://doi.org/10.1016/j.jml.2024.104532

Chapter 4: In preparation for submission

Kumarage, S., Malko, A., & Kidd, E. (in preparation). Indexing prediction error during syntactic priming via pupillometry.

References for Chapter 1

Ambridge, B., & Lieven, E. (2015). A constructivist account of child language acquisition. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (1st ed., pp. 478–510). Wiley. https://doi.org/10.1002/9781118346136.ch22

Ambridge, B., & Lieven, E. V. M. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press. https://doi.org/10.1017/CBO9780511975073

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *WIREs Cognitive Science*, *4*(2), 149–168. https://doi.org/10.1002/wcs.1215

Arai, M., & Chang, F. (2024). Influence of word-based prediction errors on syntactic priming. *Seijo University Economic Papers*, *243*, 97–120.

Arai, M., van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *54*(3), 218–250. https://doi.org/10.1016/j.cogpsych.2006.07.001

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, *59*(1), 97–113. https://doi.org/10.1016/j.jml.2007.12.007

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*(3), 455–461. https://doi.org/10.1016/j.cognition.2009.11.005

Bidgood, A., Pine, J. M., Rowland, C. F., & Ambridge, B. (2020). Syntactic representations are both abstract and semantically constrained: Evidence from children's and adults' comprehension and production/priming of the English passive. *Cognitive Science*, *44*(9). https://doi.org/10.1111/cogs.12892

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458. https://doi.org/10.1016/j.cognition.2006.07.003

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*(2), 177–192. https://doi.org//10.1037//0096-3455.129.2.177

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, *99*, 150–171. https://doi.org/10.1037/0033-295X.99.1.150

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Branigan, H. P., & McLean, J. F. (2016). What children learn from adults' utterances: An ephemeral lexical boost and persistent syntactic priming in adult-child dialogue. *Journal of Memory and Language*, *91*, 141–157. https://doi.org/10.1016/j.jml.2016.02.002

Branigan, H. P., McLean, J. F., & Jones, M. (2005). A blue cat or a cat that is blue? - Evidence for abstract syntax in young children's noun phrases. In A. Brugos & M. R. ClarkCotton (Eds.), *Proceedings of the 29th Annual Boston University Conferece on Language Development, Vols 1 and 2* (pp. 109–121). Cascadilla Press.

Branigan, H. P., & Pickering, M. J. (2017). Structural priming and the representation of language. *Behavioral and Brain Sciences*, *40*, Article e313. https://doi.org/10.1017/S0140525X17001212

Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 468–481. https://doi.org/10.1037/0278-7393.31.3.468

Buckle, L., Lieven, E., & Theakston, A. L. (2017). The effects of animacy and syntax on priming: A developmental study. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02246

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*(1), 29–49. https://doi.org/10.1016/S0010-0277(03)00123-9

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chen, X., Hartsuiker, R. J., Muylle, M., Sarah Slim, M., & Zhang, C. (2022). The effect of animacy on structural priming: A replication of Bock, Loebell and Morey (1992). *Journal of Memory and Language*, *127*, Article 104354. https://doi.org/10.1016/j.jml.2022.104354

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Dell, G. S., & Chang, F. (2014). The p-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634). https://doi.org/10.1098/rstb.2012.0394

Donnelly, S., Rowland, C., Chang, F., & Kidd, E. (2024). A comprehensive examination of prediction-based error as a mechanism for syntactic development: Evidence from syntactic priming. *Cognitive Science*, *48*(4), Article e13431. https://doi.org/10.1111/cogs.13431

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language*, *48*(2), 379–398. https://doi.org/10.1016/S0749-596X(02)00523-5

Foltz, A., Thiele, K., Kahsnitz, D., & Stenneken, P. (2015). Children’s syntactic-priming magnitude: Lexical factors and participant characteristics. *Journal of Child Language*, *42*(4), 932–945. https://doi.org/10.1017/S0305000914000488

Gámez, P. B., & Shimpi, P. M. (2016). Structural priming in Spanish as evidence of implicit learning. *Journal of Child Language*, *43*(1), 207–233. https://doi.org/10.1017/S0305000915000161

Gámez, P. B., & Vasilyeva, M. (2020). Shared syntactic representations in balanced bilinguals: Cross-linguistic priming with and without verb overlap. *Language Learning and Development*, *16*(1), 89–106. https://doi.org/10.1080/15475441.2019.1689822

Garcia, R., & Kidd, E. (2020). The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development*, *16*(4), 399–425. https://doi.org/10.1080/15475441.2020.1814780

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, *17*(8), 684–691. https://doi.org/10.1111/j.1467-9280.2006.01767.x

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238. https://doi.org/10.1016/j.jml.2007.07.003

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*(2), B27–B39. https://doi.org/10.1016/S0010-0277(99)00080-3

Hsu, D.-B. (2014a). Mandarin-speaking three-year-olds' demonstration of productive knowledge of syntax: Evidence from syntactic productivity and structural priming with the SVO-ba alternation. *Journal of Child Language*, *41*(5), 1115–1146. https://doi.org/10.1017/S0305000913000408

Hsu, D.-B. (2014b). Structural priming as learning: Evidence from mandarin-learning 5-year-olds. *Language Acquisition: A Journal of Developmental Linguistics*, *21*(2), 156–172. https://doi.org/10.1080/10489223.2014.884571

Hsu, D.-B. (2019). Children's adaption to input change using an abstract syntactic representation: Evidence from structural priming in Mandarin-speaking preschoolers. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02186

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*(2), 182–195. https://doi.org/10.1016/j.jml.2003.09.003

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kaschak, M. P., Kutta, T. J., & Jones, J. L. (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. *Psychonomic Bulletin & Review*, *18*(6), 1133–1139. https://doi.org/10.3758/s13423-011-0157-y

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*(3), 381–388. https://doi.org/10.3758/s13421-010-0042-3

Kholodova, A., Peter, M., Rowland, C. F., Jacob, G., & Allen, S. E. M. (2023). Abstract priming and the lexical boost effect across development in a structurally biased language. *Languages*, *8*(4), Article 264. https://doi.org/10.3390/languages8040264

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, *33*(2), 393–418. https://doi.org/10.1017/S0142716411000415

Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, *58*(1), 68–101. https://doi.org/10.1016/j.cogpsych.2008.05.002

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition*, *87*(3), 151–178. https://doi.org/10.1016/S0010-0277(02)00230-5

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Messenger, K., Branigan, H., Buckle, L., & Lindsay, L. (2022). How does syntactic priming experience support language development? In K. Messenger (Ed.), *Syntactic Priming in Language Acquisition: Representations, mechanisms and applications* (pp. 57–81). John Benjamins Publishing Company.

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children’s passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, *66*(4), 568–587. https://doi.org/10.1016/j.jml.2012.03.008

Messenger, K., & Fisher, C. (2018). Mistakes weren’t made: Three-year-olds’ comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition*, *178*, 118–132. https://doi.org/10.1016/j.cognition.2018.05.002

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Preuschoff, K., ’t Hart, M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline’s role in decision making. *Frontiers in Neuroscience*, *5*. https://doi.org/10.3389/fnins.2011.00115

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*(4), 587–637. https://doi.org/10.1111/j.1551-6709.2010.01165.x

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-f

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348–379. https://doi.org/10.1016/j.jml.2007.03.002

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63. https://doi.org/10.1016/j.cognition.2012.06.008

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557–567. https://doi.org/10.1111/1467-7687.00312

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, *2*(1), 27–49. https://doi.org/10.1207/s15473341lld0201_2

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition*, *89*(3), 179–205. https://doi.org/10.1016/S0010-0277(03)00119-7

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3- and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, *43*(6), 1334–1346. https://doi.org/10.1037/0012-1649.43.6.1334

Sirois, S., & Brisson, J. (2014). Pupillometry. *WIREs Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

Steedman, M. (1999). Connectionist sentence processing in perspective. *Cognitive Science*, *23*(4), 615–634. https://doi.org/10.1016/S0364-0213(99)00018-X

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253. https://doi.org/10.1016/S0010-0277(99)00069-4

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Harvard University Press.

Tooley, K. M. (2023). Structural priming during comprehension: A pattern from many pieces. *Psychonomic Bulletin & Review*, *30*(3), 882–896. https://doi.org/10.3758/s13423-022-02209-7

Tooley, K. M., & Bock, K. (2014). On the parity of structural persistence in language production and comprehension. *Cognition*, *132*(2), 101–136. https://doi.org/10.1016/j.cognition.2014.04.002

Tooley, K. M., Swaab, T. Y., Boudewyn, M. A., Zirnstein, M., & Traxler, M. J. (2014). Evidence for priming across intervening sentences during on-line sentence comprehension. *Language and Cognitive Processes*, *29*(3), 289–311. https://doi.org/10.1080/01690965.2013.770892

Tooley, K. M., & Traxler, M. J. (2010). Syntactic priming effects in comprehension: A critical review. *Linguistics and Language Compass*, *4*(10), 925–937. Scopus. https://doi.org/10.1111/j.1749-818X.2010.00249.x

Traxler, M. J., Tooley, K. M., & Pickering, M. J. (2014). Syntactic priming during sentence comprehension: Evidence for the lexical boost. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(4), 905–918. https://doi.org/10.1037/a0036377

Valian, V. (2014). Arguing about innateness. *Journal of Child Language*, *41*(S1), 78–92. https://doi.org/10.1017/S0305000914000336

Vasilyeva, M., & Gámez, P. B. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in syntactic priming. *Journal of Experimental Child Psychology*, *138*, 15–30. https://doi.org/10.1016/j.jecp.2015.04.009

Vasilyeva, M., Waterfall, H., Gámez, P. B., Gomez, L. E., Bowers, E., & Shimpi, P. (2010). Cross-linguistic syntactic priming in bilingual children. *Journal of Child Language*, *37*(5), 1047–1064. https://doi.org/10.1017/S0305000909990213

Wolleb, A., Sorace, A., & Westergaard, M. (2018). Exploring the role of cognitive control in syntactic processing: Evidence from cross-language priming in bilingual children. *Linguistic Approaches to Bilingualism*, *8*(5), 606–636. https://doi.org/10.1075/lab.17002.wol

Yang, C. (2018). A formalist perspective on language acquisition. *Linguistic Approaches to Bilingualism*, *8*(6), 665–706. https://doi.org/10.1075/lab.18014.yan

Ziegler, J., & Snedeker, J. (2018). How broad are thematic roles? Evidence from structural priming. *Cognition*, *179*, 221–240. https://doi.org/10.1016/j.cognition.2018.06.019

Chapter 2:

Implicit Learning of Structure across Time: A Longitudinal Investigation of Syntactic Priming in Young English-Acquiring Children

Chapter Overview

Chapter 2 is the first of two chapters in this thesis that address syntactic acquisition. The findings from syntactic priming studies in children are used to make claims about the nature of early syntactic representations and, by extension, how representations develop and the mechanisms responsible for this process. However, no previous study had tracked development longitudinally. The aim of Chapter 2 was to address this gap in the literature and to track the trajectory of two syntactic priming effects – abstract priming and the lexical boost – in order to distinguish three theories of syntactic acquisition: the nativist (e.g., Bencini & Valian, 2008), lexicalist (Tomasello, 2003), and implicit error-based learning accounts (Chang et al., 2006). This addressed the aims of the thesis by providing evidence that could discriminate mechanisms of acquisition and by demonstrating the value of using a longitudinal design with syntactic priming. The study found that abstract passive priming emerged early on average but contained substantial variability that was related to children's linguistic proficiency, in line with emergentist approaches. In children who did know the passive at 3 years old, the abstract priming effect decreased over development in line with the error-based learning account. The lexical boost did not emerge until later and increased over development, suggesting that the lexical boost is dependent on a different underlying mechanism (e.g., explicit memory). These results were most consistent with an error-based learning account of acquisition.

A Note on Terminology

Note that in this chapter, the nativist account is referred to as the Early Syntax account and the lexicalist account as the Late Syntax account, in line with the published article.

Publication Status

This manuscript has been published as follows:

Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring

children. Journal of Memory and Language, 127, Article 104374.

https://doi.org/10.1016/j.jml.2022.104374

Acknowledgements

Thank you to Lauren Morrison and Amanda Piper for task construction and data collection.

Author Contributions

- Kidd developed the research question.
- Morrison and Donnelly constructed the task.
- Morrison and Piper collected data.
- **Kumarage** transcribed and coded data.
- **Kumarage** analysed the data with advice from Donnelly.
- **Kumarage** produced the figures.
- **Kumarage** drafted the manuscript with editing provided by Donnelly and Kidd.
- **Kumarage** produced the reproducible analysis code.
- **Kumarage** drafted responses to reviewers with advice and editing from Donnelly and Kidd.

Implicit Learning of Structure across Time: A Longitudinal Investigation of Syntactic Priming in Young English-Acquiring Children

Shanthi Kumarage[1,2]
Seamus Donnelly[1] (https://orcid.org/0000-0003-3019-6633)
Evan Kidd[1, 2, 3] (https://orcid.org/0000-0003-4301-2290)

[1]Language Development Department, Max Planck Institute of Psycholinguistics, Nijmegen
[2]Research School of Psychology, The Australian National University
[3]ARC Centre of Excellence for the Dynamics of Language

Addresses for correspondence:

Shanthi Kumarage
Language Development Department
Max Planck Institute for Psycholinguistics
Nijmegen 6525XD
NETHERLANDS
Email: shanthi.kumarage@mpi.nl

Author Note: This research was supported by the Australian Research Council (CE140100041: CI Kidd). We thank all the children and families who participated, and Lauren Morrison, Amanda Piper, and Katherine Revius for research assistance.

Abstract

Theories of language acquisition vary significantly in their assumptions regarding the content of children's early syntactic representations and how they subsequently develop towards the adult state. An important methodological tool in tapping syntactic knowledge is priming. In the current paper, we report the first longitudinal investigation of syntactic priming in children, to test the competing predictions of three different theoretical accounts. A sample of 106 children completed a syntactic priming task testing the English active-passive alternation every six months from 36 months to 54 months of age. We tracked both the emergence and development of the abstract priming effect and lexical boost effect. The lexical boost effect emerged late and increased in magnitude over development, whilst the abstract priming effect emerged early and, in a subsample of participants who produced at least one passive at 36 months, decreased in magnitude over time. In addition, there was substantial variation in the emergence of abstract priming amongst our sample, which was significantly predicted by language proficiency measured six months prior. We conclude that children's representation of the passive is abstracted early, with lexically dependent priming coming online only later in development. The results are best explained by an implicit learning account of acquisition (Chang, F., Dell, G., S., & Bock, K. 2006. Becoming Syntactic. Psychological Review, 113, 234 – 272), which induces dynamic syntactic representations from the input that continue to change across developmental time.

Keywords: syntactic priming, language acquisition, syntax acquisition

Implicit Learning of Structure across Time: A Longitudinal Investigation of Syntactic Priming in Young English-Acquiring Children

The core aim of psycholinguistics is to explain the architecture and processes underlying the human capacity for language. This includes both how linguistic representations are used during the business of language production and comprehension, but also how those representations emerge and change across ontogeny. One method that is particularly useful in investigating the nature of linguistic representations is syntactic priming, the process whereby processing a specific syntactic structure increases the frequency of its use in subsequent discourse (Bock, 1986). For example, in Bock's (1986) seminal study, adult participants were more likely to produce a passive description like '*the church is being struck by lightning*' after saying a passive prime (*the referee was punched by one of the fans*) than after saying an active prime (*one of the fans punched the referee*). In the current study, we present the first longitudinal study of syntactic priming of the English active-passive alternation in monolingual children aged 3 – 4.5 years. In doing so, we address several methodological problems present in the developmental literature, enabling us to test competing theoretical possibilities concerning the emergence and development of syntactic knowledge during a period of rapid developmental change.

Syntactic priming is an ideal method to investigate syntactic development because priming effects are assumed to reflect common representations across prime and target. Since syntactic priming is observed in the absence of shared open- or closed-class lexical content, semantic content, or sentence prosody (Bock, 1989; Bock & Loebell, 1990), the nature of that shared representation is in many circumstances argued to be *abstract* syntactic knowledge, or syntactic procedures that operate over abstract categories (Branigan & Pickering, 2017). The effect is robust, at least in adults: a recent meta-analysis of production priming studies concluded that there is strong evidence for syntactic priming without influence from publication bias (Mahowald et al., 2016).

Priming effects are larger when the prime and target share the same main verb, the so-called *lexical boost effect* (e.g., *the bird was hugged by the dog* primes *the mouse was hugged by the chicken* to a greater extent than *the cat was pushed by the boy* does). Although the lexical boost effect is robustly larger in magnitude than abstract priming (Mahowald et al., 2016), it appears to behave slightly differently, suggesting it may have a different source. Notably, in comparison to abstract priming it is short-lived (Bock & Griffin, 2000; Hartsuiker et al., 2008), and is more difficult to observe in young children (Peter et al., 2015; Rowland et al., 2012; c.f. Branigan & McLean, 2016). This has led to suggestions that it might reflect a different mechanism to that responsible for abstract priming, one which develops with age (Chang et al., 2006, 2012). We next discuss mechanistic explanations of priming and the lexical boost and how they relate to the key conceptual divisions in language acquisition.

Mechanisms of Syntactic Priming Effects

Broadly speaking, there are two competing mechanistic explanations for syntactic priming phenomena. The first attributes priming to *residual activation* of shared linguistic representations between prime and target, which underlies both abstract priming and the lexical boost (Pickering & Branigan, 1998). The second account attributes abstract priming to error-based (implicit) learning (Chang, 2002; Chang et al., 2006) and assumes the lexical boost is due to separate, potentially more explicit, memory processes (Chang et al., 2012). We discuss each in turn.

Priming as Residual Activation

On the residual activation approach, processing a prime sentence activates a mental representation of its syntactic structure, thereby increasing the structure's short term accessibility (Pickering & Branigan, 1998). The model assumes that individual verbs are linked to structural 'nodes' denoting the syntactic frames in which they can occur. For example, the verb *chase* can occur in either an active or passive transitive sentence. If a speaker hears *the cat was chased by the dog*, the passive node becomes activated, increasing the likelihood that the structure will be used again. As both syntactic nodes and the connections between verbs and syntactic nodes are subject to residual activation, the model explains both the abstract priming effect and lexical boost using the same mechanism. This type of mechanism cannot fully explain priming effects, since findings suggest that priming can endure across time periods longer than residual activation is presumed to persist (Bock & Griffin, 2000; Hartsuiker et al., 2008; though even abstract priming effects decay, Bernolet et al., 2016). However, the theory continues to be influential. In particular, the assumption that priming is due to the activation of interconnected nodes of grammatical representations, and thus reflects the presence of structural knowledge, still holds (e.g. Branigan & Pickering, 2017; Hartsuiker & Bernolet, 2017). It is in this spirit that we adopt the theory and connect it to models of acquisition.

While the residual activation model makes clear predictions about the adult data, predictions about syntactic priming in children are less clear because the model assumes fully abstract representations of syntactic structure. However, whether and when young children possess such representations is a matter of ongoing debate (Ambridge & Lieven, 2011). As we are aware of no activation-based priming account which makes strong assumptions about early syntactic knowledge, we, therefore, consider how this model would behave when assuming each of two accounts of early syntactic representations: early abstractionist (Gertner et al., 2006; Lidz et al., 2003; Valian, 2014; Yang, 2018) and usage-based accounts (Ambridge & Lieven, 2015; Tomasello, 2003). As these two accounts represent opposite sides of the theoretical spectrum, we are able to map out a broad space of predictions that are plausibly consistent with the residual activation model. We derive our predictions by combining accounts, since developmental theories lack detail regarding their architectural assumptions and how these influence and change in response to processing structure (see also Branigan & Mclean, 2016), and adult theories typically assume a competent speaker and thus do not integrate developmental constraints.

On the one hand, *early abstractionist* accounts of acquisition assume that children acquire language guided by innately conferred or constrained processes that enable them to deduce a language-specific and abstract grammatical system from very early in development, a process that depends on input for configuring but not constructing abstract linguistic categories (Bencini & Valian, 2008; Gertner et al., 2006; Lidz et al., 2003; Messenger & Fisher, 2018; Valian, 2014; Yang, 2018). Thus, a prediction derived from the early abstractionist approach is that abstract (i.e., verb-independent) priming will be observed early in development, demonstrating that children possess abstract knowledge. For a language-specific structure like the passive, which is our focus here, the assumption is that children have sufficiently abstract representations of syntactic categories such as *subject* and *object* and can flexibly map them to thematic roles like *agent* and *patient*, independently of lexically-specific knowledge, once they have acquired the syntactic frame (Bencini & Valian, 2008; Messenger & Fisher, 2018). While the account does not make any specific predictions about the lexical boost, assuming that the early emerging abstract knowledge is processed within an architecture that links specific verbs to abstract structure (as is assumed in the residual activation account), we can deduce that having adult-like abstract syntactic knowledge would result in an adult-like pattern of priming across all condition types. That is, the lexical boost should emerge at the same time as the abstract priming effect and should be higher in magnitude. Finally, since abstract knowledge is present early, the early abstraction account does not predict significant changes in priming across development. For the sake of clarity and convenience, we call the early abstraction account the *Early Syntax* instantiation of the residual activation account (or RA-Early Syntax).

On the other hand, emergentist and usage-based theories of development differ from early abstractionist accounts in that they do not assume children possess early or innate syntactic representations (Ambridge & Lieven, 2015; Savage et al., 2003; Tomasello, 2003). Instead, this approach argues that children gradually abstract over item-based instances to induce an adult-like grammar using general-cognitive learning mechanisms. For instance, a child's early use of the passive in *the cat was chased by the dog* may only be indicative of children knowing the passive structure as contingently linked to the verb *chase*, and thus they could not generalise the structure beyond the verb. On these functional accounts, abstraction takes developmental time, since children must induce generalised syntactic representations from the evidence available in the input (Ambridge & Lieven, 2015, 2011; Tomasello, 2003), which gradually become more abstract with experience. Thus, the approach does not predict early abstract syntactic priming effects; since early syntactic knowledge is lexicalised, the approach predicts the initial emergence of lexically-based priming effects in the absence of abstract priming.[1] For the sake of clarity and convenience, throughout this paper we call this the *Late Syntax* instantiation of the residual activation account (or RA-Late Syntax).

Taken together, depending on the posited nature of children's early syntactic knowledge that residual activation mechanisms operate upon, accounts of syntactic acquisition make different predictions about the development of priming effects. If one assumes children's earliest syntactic knowledge is fully abstract, the prediction is that both syntactic priming and the lexical boost should be observable once children begin producing the relevant grammatical construction (in this case, the passive). However, if one assumes

---

[1] We use the term lexically-based priming for priming of syntactic structures which relies solely on lexicalised syntactic knowledge and the term lexical boost for an increase in syntactic priming when there is lexical overlap between prime and target relative to when there is not.

children gradually construct abstract representations in an item-specific manner, the prediction that the earliest priming should be fully lexicalized, such that children exhibit a lexically-based priming effect prior to exhibiting abstract syntactic priming.[2]

Priming as Error-Based (Implicit) Learning

An alternative explanation of priming is that it reflects *Implicit Learning* of grammatical structure (Bock & Griffin, 2000; Chang et al., 2006; Fine & Jaeger, 2013). The most explicit articulation of this account comes from Chang et al.'s (2006) Dual-path model, which constitutes a theory of both the acquisition of grammatical structure and adult sentence production. A key feature of the model is that it acquires input-driven, language-specific syntactic categories via error-based learning. That is, using a Simple Recurrent Network (Elman, 1990) for the sequencing (syntactic) system and a separate meaning system for semantic information, the model makes next-word predictions based on its previous experience with the language. When those predictions are incorrect, such as when the model fails to predict a passive past participle after *the cat was…*, the network weights are adjusted via error-based learning, such that the likelihood that a passive will be predicted is higher. In this respect, the model acquires structure via priming, and each new experience with a given structure alters the likelihood that the structure will be subsequently used again. Thus, hearing a passive increases the probability that a passive will be used to describe a transitive event, relative to an active.

The model explains a diverse range of phenomena in language acquisition and adult language processing, including: the décalage between the comprehension and production of transitive sentences in acquisition (Chang et al., 2006), cross-linguistic differences in language production (Chang, 2009), structure dependence in the acquisition of subject-auxiliary inversion (Fitz & Chang, 2017), and the underlying nature of N400 and P600 event-related potential effects (Fitz & Chang, 2019). Naturally, the model explains many syntactic priming phenomena pertaining to abstract priming, including the fact that abstract priming persists over long periods, which the model explains as a signature of implicit learning of

---

[2] There is an alternative account (Reitter et al. 2011) of syntactic priming that is, in some ways, broadly consistent with the residual activation account. The model includes a short-term spreading activation mechanism, which we believe would combine with developmental accounts similarly to the residual activation model. However, this account also assumes a base-level activation mechanism to explain long-term priming effects. Given the model's complexity, it is not clear how it would behave when assuming non-adult syntactic representations and therefore difficult to speculate about its developmental predictions.

structure involving changes in representations rather than their activation (Kaschak et al., 2011). However, it crucially does not simulate the lexical boost effect, which is consistent with the observation that abstract priming and the lexical boost endure across different time frames (though Reitter et al., 2011 simulate the effect in one architecture by assuming two distinct mechanisms; see also Zhang et al., 2020). Chang et al. (2006) argued that this suggests the lexical boost is attributable to a different mechanism - specifically, explicit memory processes that may be more vulnerable to rapid decay (Chang et al., 2012).

The Chang et al. (2006) model makes predictions for priming in acquisition that differ from Early and Late Syntax accounts. Most broadly, the model predicts different developmental profiles for abstract priming and the lexical boost/lexically-based priming. In contrast to Early Syntax accounts, which predict stable abstract representations of syntax over development, the Implicit Learning model predicts that less experienced speakers will have syntactic representations that are based on fewer instances and are more susceptible to input (Rowland et al., 2012). Therefore, immature systems will be subject to greater error prediction, which predicts larger abstract priming effects in less experienced speakers. There is mixed evidence for this claim: grammatically less-skilled aphasic participants show larger abstract priming effects (Hartsuiker & Kolk, 1998), but adults do not show greater abstract priming in their second language than in their native language (Mahowald et al., 2016). In children, we would expect the abstract priming effect to *decrease* over the course of syntax acquisition (Rowland et al., 2012; however, this is contingent on children having first acquired the structure, see Messenger et al., 2022). This prediction also differs to the Late Syntax account, which involves generalisation across lexically-specific representations rather than error-based learning, and therefore predicts an increasing abstract priming effect. In contrast, if the lexical boost is dependent on explicit memory, we should expect the opposite relation: the lexical boost effect should increase across development (Chang et al., 2012; Rowland et al., 2012). This is based on well-accepted findings demonstrating that implicit and explicit memory processes have different developmental schedules, with explicit/declarative processes continuing to develop throughout childhood and beyond (e.g., Finn et al., 2016; Lum et al., 2010).

Overall, then, we distinguish between three models. The different predictions regarding priming in acquisition over development derived from each model are summarised in Table 1.

Table 1

Summary of predictions for different priming models for developmental data across time

	Abstract Priming		Lexical Boost/Lexically-based	
	Emergence	Development	Emergence	Development
RA - Early Syntax	Early	No change	Early	No change
RA - Late Syntax	Late	Increase	Early	Inverse U-shape
Implicit learning	Early	Decrease	Late	Increase

There are a few things to note from Table 1. The first is that the three models all make different predictions regarding abstract priming. On our reading, they also make differing predictions regarding the lexical boost/lexically-based priming, although the specifics for the latter are less clear. Thus, the RA-Early Syntax approach should show this adult-like pattern early in development. The RA-Late Syntax approach predicts an early increase in lexically-based priming: since children's early grammatical knowledge is lexically-restricted, priming in conditions of lexical overlap should increase as children accumulate lexically-based representations and develop towards fully abstract knowledge. However, once abstract knowledge emerges there should be an increase in abstract priming relative to lexically-driven priming, and therefore a decrease in the "boost" provided by overlapping lexical content, which suggests an inverted U-shaped pattern.

Finally, we note an additional parameter that we bring to bear upon the nature and emergence of priming, and which can act as an additional source of evidence in constraining theory – individual differences (Kidd, Donnelly, et al., 2018; Kidd & Donnelly, 2020). Accounts of acquisition and processing that rely heavily on learning from the input, such as the RA-Late Syntax and Implicit Learning accounts, make straightforward predictions regarding individual variability. Notably, since children experience differences in their language exposure and have different learning rates, the emergence of the passive structure will be varied and predicted by their prior developmental states in a systematic matter. The predictions of the RA-Early Syntax approach are less clear. A straightforward reading of the approach as instantiating a traditional nativist approach that assumes continuity between the child and adults state leads to the prediction that there will be no systematic individual differences, only variability due to experimental noise (Crain et al., 2017; Crain & Thornton, 1998). However, RA-Early Syntax approaches differ on the continuity assumption, with more moderate approaches arguing that, for low frequency and language-specific structures like the passive, children may systematically vary in their knowledge and this will be linked to variables such as experience and processing ability (Messenger & Fisher, 2018).[3] Consequently, the study of individual differences provides important boundary conditions on models of acquisition.

Thus, it is important to also examine individual differences in the emergence of the passive priming effect. Such an analysis also tells us something important about what priming means in acquisition and the conditions under which it arises. There is a general assumption in the literature that lexically-independent priming is a signal of the presence of abstract syntactic knowledge (Branigan & Pickering, 2017), but as we have seen, in acquisition the existence of abstract knowledge at various ages is hotly debated. If the emergence of the priming effect is predicted by variability in language proficiency, it tells us the conditions under which priming arises in development (Kidd, 2012), and in concert with longitudinal data, how priming changes within individuals across time.

Past Empirical Work

A significant literature on syntactic priming in children exists, although to date all of these studies have either reported on one age group or have compared age groups in cross-sectional designs. These data suggest that, consistent with the RA-Early Syntax and Implicit Learning accounts, the abstract priming effect emerges early, though whether and how it develops across development is unclear (Bencini & Valian, 2008; Hsu, 2019; Peter et al., 2015; Rowland et al., 2012). Most studies testing children at 3 years, the youngest age group for which syntactic priming tasks appear achievable, have found some evidence of abstract priming of multiple structures (for passives: Bencini & Valian, 2008 and Shimpi et al., 2007, but not Savage et al., 2003; for datives: Shimpi et al., 2007; and for SVO-ba Hsu, 2019), and by 4 years, evidence for abstract priming is consistent (e.g., Huttenlocher et al., 2004; Messenger, Branigan, McLean, et al., 2012; Messenger et al., 2011). However, 3-year-olds

---

[3] Note that we discount an older nativist explanation of the late emergence of the passive – maturation (Borer & Wexler, 1987), which argued that children do not produce passives until around 6 years because of the late maturation of components of Universal Grammar. There is now sufficient evidence to discount this as a likely explanation of development across languages (for which it was intended) or, indeed, just in English.

appear to have difficulty with the task: Bencini and Valian (2008) reported that 35% (28/81) of 3-year-olds in their study could not complete the task, and Shimpi et al. (2007) only found abstract priming in 3-year-olds when they were asked to repeat, not just listen to, the prime sentence. These are both suggestive of variability in young children's ability to be primed and therefore, developmental change. Given the ambiguous evidence for abstract priming at 3 years compared to 4 years of age, the period of development between these age groups appears to be crucial for observing this change.

Evidence related to the developmental trajectories of abstract priming effects is less clear. In studies using the dative alternation, Rowland et al. (2012) found that children's abstract priming effect was larger than that of adults, whilst Peter et al. (2015) found the opposite. Hsu's (2019) investigation of the SVO-*ba* alternation in Mandarin found equivalent abstract priming across 3-, 4- and 6-year-olds. In studies using the passive alternation, Messenger, Branigan, McLean and Sorace (2012) found no difference between the size of children's and adult's priming effects but Messenger (2021) found a marginally significant decrease in abstract priming from children to adults, and while Messenger, Branigan and McLean (2012) found that 6- and 9-year-olds were equally likely to produce a passive after a passive prime, they also found that 6-year-olds produced more invalid responses with passive syntax but reversed thematic roles (e.g., producing *the chicken was hugged by the mouse* to describe a scene in which a chicken is hugging a mouse). A shortcoming of these past developmental studies is that they were cross-sectional in design; only longitudinal research can unambiguously determine the developmental trajectories of the effects.

The lexical boost effect is not consistently found in young children and appears to increase over development. Savage et al. (2003) found lexically-based priming when there was pronoun overlap in transitive clauses. Branigan et al. (2005) found evidence of a lexical boost in a study investigating noun overlap in the adjectival/relative clause alternation (e.g., *the red cat* vs *the cat that's red*), whilst Foltz et al. (2015) did not. In studies using verb rather than noun overlap, Branigan and McLean (2016) found a lexical boost effect for the active-passive alternation in 3–4-year-olds, while both Rowland et al. (2012) and Peter et al. (2015) did not find one for the dative alternation in children of the same age. Differences in power are unlikely to explain the divergent findings, since those studies with the largest samples (Peter et al., 2015; Rowland et al., 2012) did not find the effect, but it is possible that the lexical boost effect develops on structure-specific schedules. In terms of the development of the lexical boost effect, two studies from the same lab have found increases in the magnitude of the lexical boost effect for the dative alternation from younger children to older children to adults (Peter et al., 2015; Rowland et al., 2012). No study has investigated developmental changes in the lexical boost for passive sentences, and, as is the case with the abstract priming effect, no study has ever studied the lexical boost using a longitudinal design. Therefore, whether the lexical boost also increases over development for the passive, and precisely when it may emerge is unknown.

There is little research on individual differences in priming effects in children, despite Bencini and Valian's (2008) high dropout rate and Shimpi et al.'s (2007) task adjustments for 3-year-olds suggesting variability in children's ability to be primed. Kidd (2012) found that children's nonverbal ability predicted their tendency to be primed, and that vocabulary size and grammatical knowledge predicted the magnitude of the priming effect for those children who were primed. Messenger (2021) found a marginal correlation between the magnitude of children's priming effect and their vocabulary size. Children's passive production in the priming task also predicted their passive production in a post-test phase. The findings of both studies point to the possibility that differences in developmental levels, which are imperfectly related to age, lead to individual differences in priming, but further corroborating research is required.

The Need for Longitudinal Data

Many high-quality studies of syntactic priming in children have been conducted; however, the conclusions we can draw from the current literature are limited by the cross-sectional design of those studies. Messenger et al. (2022) point out that the average age range for samples of children in studies comparing age groups is 20 months. Therefore, there is likely substantial variation in children's stage of language development, which could contribute to variation in priming effects by including both children who are more likely to be primed and less likely to be primed in the same comparison group. This problem is particularly likely considering studies typically use samples between 3 and 5 years old, the period in which, in languages like English (and other European languages), knowledge of the passive develops. Existing studies cannot track changes in children's priming effects over critical periods of change. The lack of clarity over the development of the abstract priming effect or the precise timing of the emergence of the lexical boost indicates the shortcomings of cross-sectional designs. Developmental change is best investigated using designs which tightly control age and compare children to themselves over a period of time in which development occurs, that is, longitudinal designs.

The Present Study

The past empirical evidence suggests an early-emerging abstract priming effect (i.e., unambiguously by 4-years), in support of the RA-Early Syntax and Implicit Learning accounts. However, how this priming effect changes across time and whether it varies systematically across individuals are open questions. Moreover, evidence regarding the lexical boost effect is suggestive of a late-emerging developmental effect, as predicted by the Implicit Learning account, although the data on this topic are sparse and inconsistent. Even fewer studies have investigated individual variability in priming effects but those that have done so suggest systematic differences that are related to children's prior knowledge, in line with the RA-Late Syntax and Implicit Learning accounts. In order to better distinguish between these approaches, longitudinal evidence – the gold standard for developmental science – is required at key periods of development. In the current study we investigate the longitudinal development of both the abstract priming effect and the lexical boost effect in a large sample of children acquiring English as a first language. We followed them across four testing sessions, starting from 36 months, when the abstract priming effect does not appear to have emerged for all individuals, to 54 months, when it is well established. We examine the English active-passive alternation. As outlined in Table 1, we expect the following predictions from each account.

1) If the RA-Early Syntax account holds, we expect early abstract priming and lexical boost effects that remain stable across the 18-month period. Moreover, if the most strongly nativist of these perspectives hold (e.g., Crain et al., 2017), we expect minimal systematic individual variability between participants in the emergence of the passive structure.
2) If the RA-Late syntax account holds, we expect priming to be initially restricted to trials with lexical overlap, resulting in an increasing abstract priming effect across time and a progression from lexically-based priming to a decreasing lexical boost as abstract priming increases. Such accounts further predict systematic individual variability in the emergence of the passive structure.

3) If the Implicit Learning account holds, we expect to see early abstract priming that decreases across time and a late emerging lexical boost that increases across time. This account further predicts systematic individual variability in the emergence of the passive structure.

Data Availability

Our sentences, pictures and experimental lists, data and scripts are accessible on the Open Science Framework (https://osf.io/35kzm/).

Methods

Participants

The participants were taking part in the *Canberra Longitudinal Child Language* project (Kidd, Junge, et al., 2018), a longitudinal study of children's language processing and language acquisition between the ages of 9 months and 5 years. The project initially recruited 124 children, based on the following criteria: (i) full-term (at least 37 weeks gestation) babies born with a typical birth weight (>2.5kg), (ii) a predominantly monolingual English language environment (in all but two cases, no more than 20% exposure to a language other than English), and (iii) no history of medical conditions that would affect typical language development, such as repeated ear infections, visual or hearing impairment, or diagnosed developmental disabilities. The socio-economic status of the families was measured via parental education, measured on a 7-point scale: 0 = some high school, 6 = PhD. Consistent with the demographics of the city (Canberra, Australia), SES was high, with a median education of 4 (Bachelor degree) for caregiver 1 (SD = 1.12, Range = 0 : 6) and 4 for caregiver 2 (SD = 1.12, Range = 0 : 6). Ethnicity was not recorded.

Here we report on children's priming of the active-passive alternation, which was measured longitudinally when the children were aged 36 (35;29 – 37;6 months, mean 36;18), 42 (42;12 – 43;5 months, mean 42;19), 48 (48;11 – 51;27 months, mean 48;20), and 54 months (54;10 – 55;5 months, mean 54;20). By the 36-month timepoint, 19 families had dropped out of the study, six of whom withdrew because their child had been diagnosed with a developmental disorder or delay. A further two children were excluded due to suspected developmental delay, which was corroborated by their language proficiency, as measured by the Macarthur-Bates Communicative Development Inventory (Fenson et al., 2007) at 30 months, being more than 2 standard deviations below the mean for the sample

(3.77 and 2.85 standard deviations, respectively). Of the 103 children remaining, 73 completed the priming task at all four timepoints, with a sample of over 80 at each timepoint (N_{36} = 94 (52 females), N_{42} = 92 (50 females), N_{48} = 91 (50 females), N_{54} = 82 (46 females)). Additional data loss was due to missed sessions (N_{36} = 7, N_{42} = 6, N_{48} = 5, N_{54} = 14), further withdrawals from the study (N_{42} = 5, N_{48} = 7, N_{54} = 7), or inattention/inability to complete the task (N_{36} = 2).

Design

We employed a 4 (Age: 36, 42, 48 & 54 months) x 2 (prime: active vs passive) x 2 (verb match: match vs unmatched) within-participants longitudinal design.

Materials

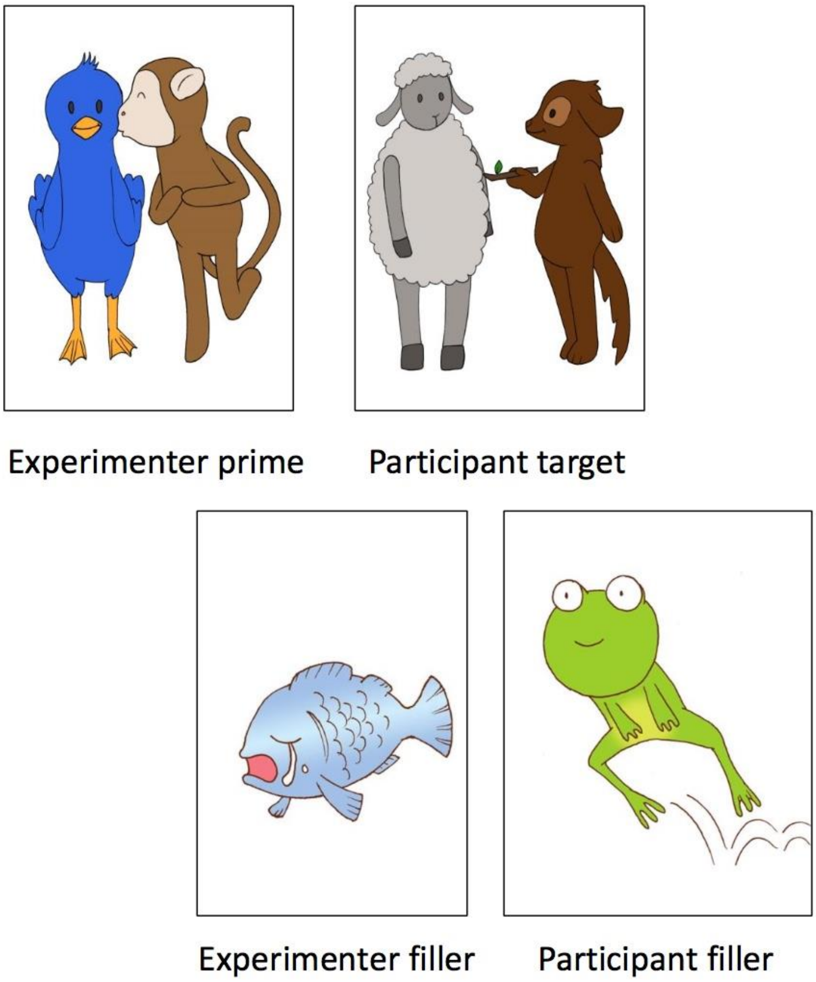
Syntactic Priming Task

The priming task was conducted using the *SNAP* (picture matching) game paradigm, a procedure which has successfully been used with young children (Branigan et al., 2005; Branigan & McLean, 2016; Branigan & Messenger, 2016). Twenty-four two-participant actions that can be described by transitive verbs were created from 6 transitive verbs (*chase, hug, kiss, poke, pull, push*) and 12 animate characters (*bird, cat, chicken, cow, dog, duck, horse, monkey, mouse, rabbit, sheep, tiger*) and then drawn by a professional artist. All verbs are highly transitive action verbs that are known to children of this age. Each has appeared in past studies of the passive and all but *poke* have been shown empirically to be understood by children in the passive construction by 3–4 years of age (see Nguyen & Pearl, 2021). The pictures were arranged into prime-target pairs across five experimental lists, each containing 12 prime-target pairs. Children therefore experienced three trials for each experimental condition (e.g. verb matched passive primes). This low number of trials was necessitated by time and attention constraints for children participating in a longer session. Across lists, each verb was counterbalanced to occur equally as often as prime and target, as passive and active prime, and in matched and unmatched verb trials. Prime descriptions for each list were present tense active and full passive sentences.

In addition, 14 one-participant actions that can be described by intransitive verbs were chosen to make up the 22 filler items, which occurred as prime-target pairs between each experimental item. For six of the filler pairs, prime and target were the same picture. Figure 1 illustrates an example sequence of the structure of each trial in the task. The task began with two practice prime-target pairs using *pull* and *feed* and one practice filler item to familiarise children with the procedure. One practice prime was active and the other passive.

Figure 1

Structure of a trial in the syntactic priming task



Note: Experimenter describes prime picture, participant describes their target picture, experimenter describes an intransitive filler, participant describes their intransitive filler.

The same materials were used at each timepoint. Children were randomly assigned to lists at each timepoint with approximately even numbers of children per list. Table 2 shows the distribution of trials presented to children across each condition and timepoint. Since each list was balanced across conditions, trials are balanced across conditions except for two experimenter errors.

Table 2

Number of trials presented per condition at each timepoint

Timepoint	Active		Passive	
	Unmatched	Verb matched	Unmatched	Verb matched
36m	283	281	282	282
42m	276	277	276	275
48m	273	273	273	273
54m	246	246	246	246
Total	1078	1077	1077	1076

MacArthur-Bates Communicative Development Inventory

Across early stages of the longitudinal study the children's primary caregiver completed the MacArthur-Bates Communicative Development Inventories (Donnelly & Kidd, 2021). At 30 months caregivers completed the Words and Sentences form (Fenson et al., 2007), which measures expressive vocabulary, grammatical complexity and mean utterance length. In the form, caregivers are asked to indicate which words and phrases their child says, and to list their 3 longest sentences (from which MLU is calculated). Following Reilly et al. (2007), some minor changes were made to some of the words in the inventory and two were removed to better reflect the Australian dialect, but otherwise the instrument was used as per the standard instructions. The instrument has excellent psychometric properties (see Fenson et al., 2007).

Procedure

At each timepoint participants completed the priming task as part of an hour-long lab visit for the larger longitudinal study. The task was audio recorded using a Zoom H2n audio recorder. Children were seated at a child-sized table and chair set with the experimenter and were told they would be playing a game of *SNAP*, a popular card game for Australian children. The experimenter and child both had a pre-ordered stack of cards. The game started with two practice trials, in which the child was familiarised with the procedure of the game. Each trial proceeded as follows: the experimenter turned over her card and described it according to a script using either an active or passive sentence for primes, or intransitive for fillers. Then the child turned over their card and was asked to describe it ('*What's* *happening on your card?*'). If the child did not use a complete sentence to describe the card, they were prompted to provide more detail (e.g. '*Tell me the whole thing*'). If children used a verb different to the one intended, they were provided feedback such as '*Yes they are running, maybe they're chasing*'. This was to reduce the number of future trials containing a 'verb error'. On six of the filler trials, the child's and experimenter's card were the same, in which case the first participant (i.e., either experimenter or child) to say '*snap*' and place their hand on the cards won. Children were always allowed to win, which aimed to increase their motivation for the game. All sessions were audio and video recorded. The task lasted approximately 10 minutes.

Coding

Children's responses were transcribed from the audio recording. Children's first complete and intended response was scored. If children corrected themselves, their corrected response was coded. If children's initial response was not a complete sentence (e.g. '*cuddles*' or '*sheep and the duck*') but they produced a full sentence after a prompt, this second sentence was scored. Sentences were coded as active, passive, or other. Sentences were coded as active if they contained an agent as the subject, a transitive verb and a patient as the object (e.g. '*the sheep is pulling the chicken*').

We chose to use a *lax* as well as *strict* coding scheme in order to capture those responses where children produced passives that did not fully match the primed model. Bencini and Valian (2008) previously used lax and strict coding schemes to allow comparison with both more generously coded child priming studies and with adult priming studies. Under our strict coding scheme, a response was coded as passive if it contained a patient in the subject position, an auxiliary, a correctly inflected transitive verb and a *by*-phrase containing the agent. A passive under the lax coding scheme required correctly assigned thematic roles, a transitive verb, and either an auxiliary or a prepositional phrase containing the agent. Therefore, truncated passives (e.g. '*he's being carried*'), omission of the auxiliary verb (e.g. '*monkey pushed by a cat*'), errors in inflection (e.g. '*a sheep is being chasing by a horse*') and errors in the prepositional phrase (e.g. '*a cow was being cuddled from a chicken*') were permitted (children also made similar mistakes in active sentences). Children sometimes produced sentences which could either be simply an error in inflecting the verb or the omission of an auxiliary and an inflection error (e.g. '*a rabbit is cuddling by a monkey*' for monkey cuddles rabbit). These were coded as other in the strict coding but passive in the lax coding.

All other sentences, including incomplete sentences, intransitives, datives, and infinitives were coded as ‘Other’. There were some notable errors in this category that may indicate early attempts at passive production. Most commonly at earlier timepoints, children occasionally inserted “by” or “being” into otherwise active sentences (e.g. ‘*a tiger is carrying by a bird*’ for tiger carries bird, and ‘*a cat is being chase a dog*’ for cat chases dog). These could indicate priming of at least surface level features of the passive, in the same way that corpus studies of adult language have shown that priming need not involve phrasal heads (e.g., Snider, 2009; see Reitter et al., 2011). We also excluded both reversal errors where the agent was the subject (e.g. ‘*a dog is being kissed by a chicken*’ for dog kisses chicken), and duplication errors where children used the same noun for both agent and patient (e.g. ‘*a monkey is being kissed by a monkey*’). Interestingly, children made reversal and duplication errors in active, intransitive, dative and infinitive sentences too. Table 3 presents the number of each of these errors made by children at each timepoint.

Table 3

Number of notable errors in children’s sentences at each timepoint

	Insertion error		Reversal error			Duplication error		
	Active	Other	Active	Passive	Other	Active	Passive	Other
36m	14		14	26 (1)	2	3	0	
42m	9		9	40 (6)	5	8	9	1
48m	6	1	3	19 (1)	1	2	8	
54m	1		2	23 (3)	1	2	4	

Note: Numbers in brackets represent the number of passives in that error category which were truncated passives.

As the study investigated the lexical boost, we also coded whether children used the verb intended by our materials or not. If the use of an incorrect verb changed the verb match condition, the trial was excluded. In verb matched trials the use of any other verb in the target sentence resulted in it no longer matching the prime verb. In contrast, children rarely produced the same verb as the prime verb in their target sentence in the unmatched verb condition.

Results

Data Loss

Trials were excluded if children did not produce a response (N_{36} = 23, N_{42} = 12, N_{48} = 3), their response was inaudible (N_{36} = 1, N_{42} = 1, N_{48} = 2), or there was an error in administering the trial (N_{42} = 3, N_{54} = 2). Trials coded *as Other* and trials where the child used a different verb to the one intended, thus changing the verb match condition, were also excluded. The distribution of trials across each timepoint is presented in Table 4 and the number of trials included for each participant is presented in Table 5. Whilst a few children produced mostly *Other* responses at each timepoint, the vast majority produced a high number of transitive responses with a median and mean of at least 9 transitive responses out of 12 trials per participant.

Table 4

Categories of children's responses at each timepoint

Timepoint	Coding	Excluded		Other		Excluded verb error		Included verb error		Remaining trials	
		N	%	*N*	%	*N*	%	*N*	%	*N*	%
36m	strict	24	2.1	250	22.2	90	8.0	143	12.7	621	55.1
	lax	24	2.1	225	20.0	94	8.3	146	12.9	639	56.7
42m	strict	16	1.5	213	19.3	47	4.3	112	10.2	716	64.9
	lax	16	1.5	166	15.0	49	4.4	118	10.7	755	68.4
48m	strict	5	0.5	130	11.9	42	3.9	108	9.9	807	73.9
	lax	5	0.5	101	9.3	43	3.9	110	10.1	833	76.3
54m	strict	2	0.2	90	9.2	20	2.0	64	6.5	808	82.1
	lax	2	0.2	78	7.9	20	2.0	68	6.9	816	82.9

Note: Trials in both the 'Included verb error' and 'Remaining trials' columns were included in analyses.

Table 5

Number of included responses per participant

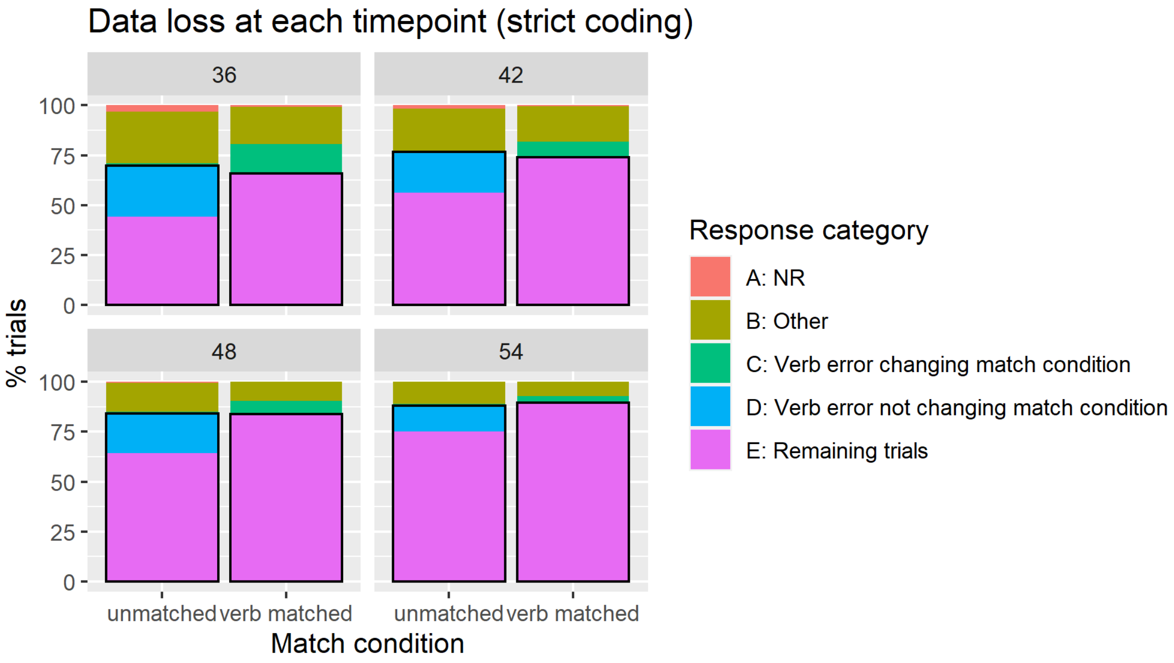
Timepoint	Coding	Transitive responses			Included trials		
		mean	median	range	mean	median	range
36m	strict	9.09	9	2 – 12	8.06	8	2 – 12
	lax	9.35	10	3 – 12	8.28	8	2 – 12
42m	strict	9.51	10	5 – 12	8.96	9	4 – 12
	lax	10.02	11	5 – 12	9.45	10	4 – 12
48m	strict	10.52	11	4 – 12	10.02	10	3 – 12
	lax	10.84	11	6 – 12	10.33	10	5 – 12
54m	strict	10.88	11	6 – 12	10.62	11	6 – 12
	lax	11.02	11	6 – 12	10.77	11	6 – 12

Note: The *Transitive responses* category includes truncated passives for the lax coding scheme. Since verb errors which changed the verb-match condition were excluded, fewer trials were included in analyses than were transitive.

Figure 2 presents the distribution of trials in Table 4 split by verb match condition and excluding administration errors and inaudible trials. In verb matched trials children made fewer verb errors than in unmatched trials and produced fewer *Other* and non-responses. Children's use of an unintended verb in verb matched trials always results in unmatched verbs, therefore changing the verb match condition and resulting in the trial being excluded. In unmatched trials, if children used an unintended verb, it was rarely the same as the prime verb and so trials rarely needed to be excluded. Therefore, at earlier timepoints, slightly more data were lost in matched than unmatched trials despite there being fewer *Other* responses in this condition. It is important to note that this data loss is non-random in a way for which our modelling does not account. The matched condition containing more transitive responses, or more successful passive productions vs incorrect attempts (captured under *Other* responses), could suggest a kind of lexical boost not captured simply by investigating the proportion of passive responses out of active and passive responses. This is important to note as we found stronger evidence for a lexical boost at the latter two timepoints, where the least data were lost. Figure 2 graphs data loss under the strict coding scheme; the pattern of data loss is identical under the lax coding scheme with only slight numerical differences.

Figure 2

Categories of children's responses at each timepoint



Note: Black outline indicates included trials; trials in both the 'Included verb error' and 'Remaining trials' categories were included in analyses.

Abstract Priming and Lexical Boost Effects

Table 6 and Figure 3 summarise the proportion of active and passive responses in each experimental condition. There appears to be a consistent abstract priming effect, with more passives produced following passive primes than active primes. In addition, there appears to be a lexical boost effect that increases over time: at 48 and 54 months, children exhibited larger priming effects for passive primes containing the same verb as the target picture.

Table 6

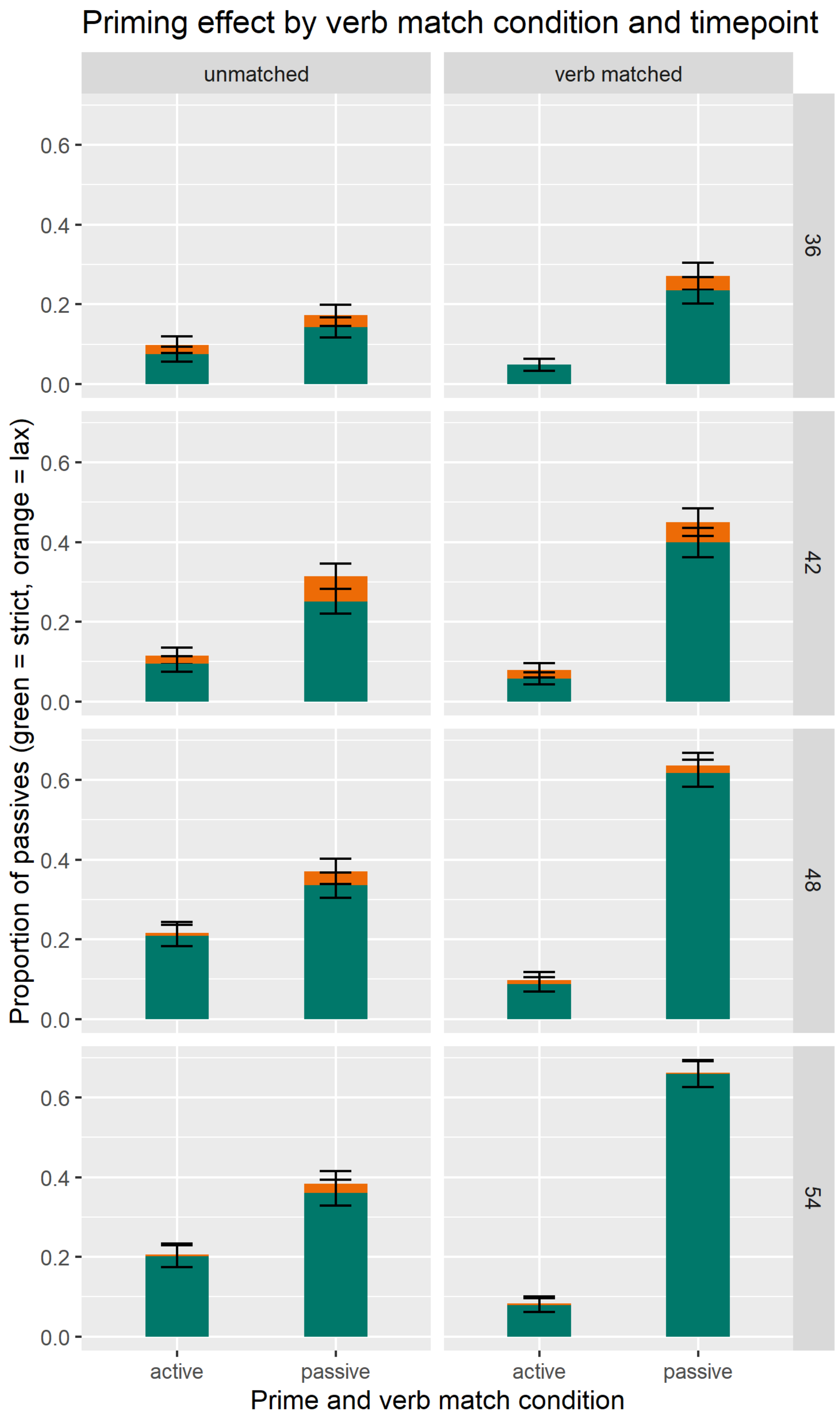
Number and proportion of passives and actives in each experimental condition

Timepoint	Prime	Verb match	Active		Passive - strict		Passive - lax	
			N	%	N	%	N	%
36m	Active	Unmatched	184	92.5 / 90.2	15	7.5	20	9.8
		Matched	197	95.2 / 95.2	10	4.8	10	4.8
	Passive	Unmatched	163	85.8 / 82.7	27	14.2	34	17.3
		Matched	124	76.5 / 72.9	38	23.5	46	27.1
42m	Active	Unmatched	201	90.5 / 88.6	21	9.5	26	11.5
		Matched	211	94.2 / 92.1	13	5.8	18	7.9
	Passive	Unmatched	146	74.9 / 68.6	49	25.1	67	31.5
		Matched	110	60.1 / 55.0	73	39.9	90	45.0
48m	Active	Unmatched	185	79.1 / 78.4	49	20.9	51	21.6
		Matched	220	91.3 / 90.2	21	8.7	24	9.8
	Passive	Unmatched	148	66.4 / 63.0	75	33.6	87	37.0
		Matched	82	38.3 / 36.4	132	61.7	143	63.6
54m	Active	Unmatched	170	79.8 / 79.4	43	20.2	44	20.6
		Matched	210	92.1 / 91.7	18	7.9	19	8.3
	Passive	Unmatched	140	63.9 / 61.7	79	36.1	87	38.3
		Matched	72	34.1 / 33.8	139	65.9	141	66.2

Note: The two percentages presented for Active responses are for the strict and lax coding schemes respectively.

Figure 3

Proportion of passive responses in each experimental condition at each timepoint



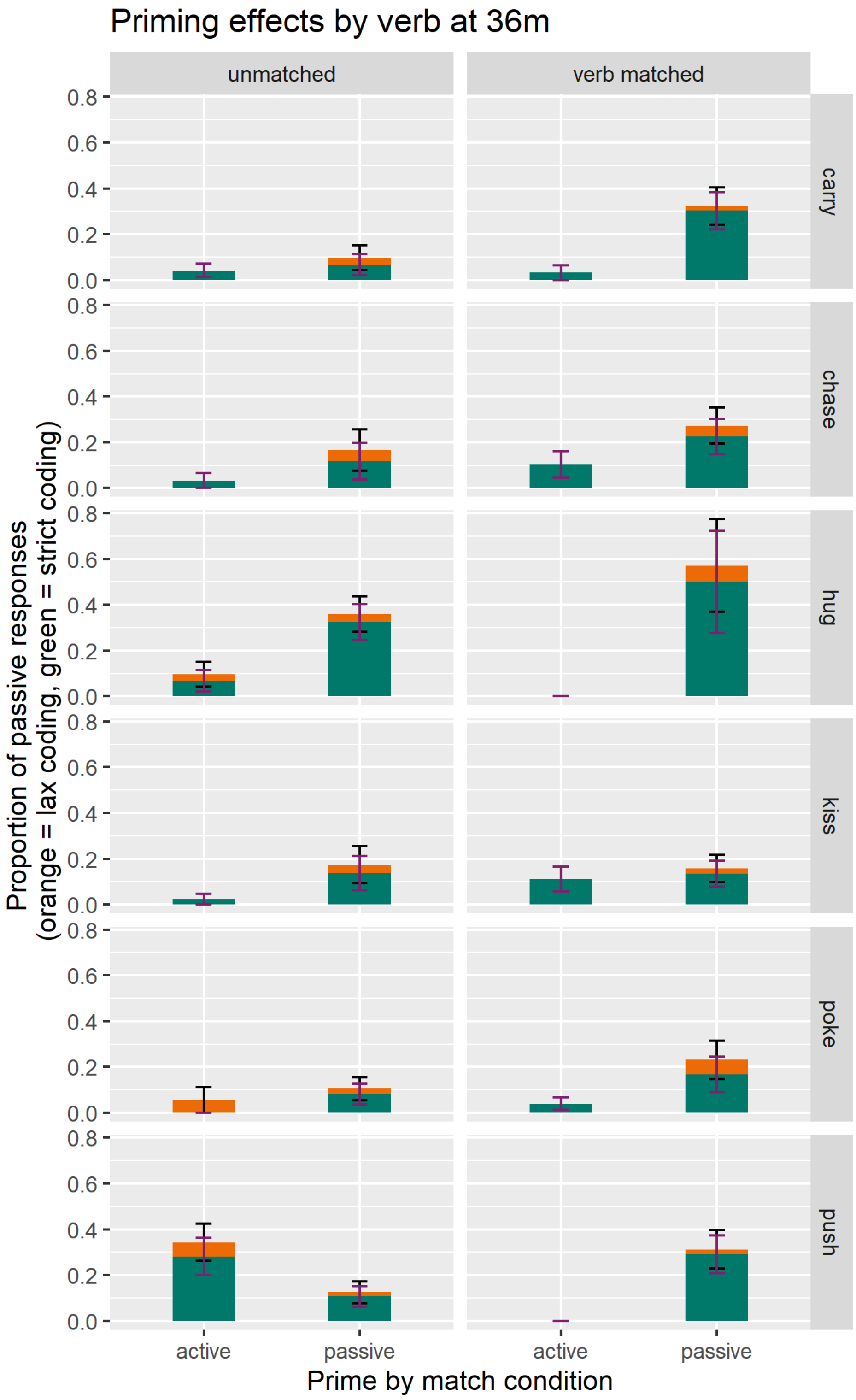
Note: Abstract priming is indicated by a larger proportion of passives in the passive than prime condition for unmatched verb trials. The lexical boost effect is indicated by a larger difference between active and passive prime conditions for verb matched vs unmatched trials.

We also plotted the priming effect by target verb to check for consistency in the priming effects by item. Since our data contains only 6 levels of item, mixed models including random slopes by item are unlikely to converge, or, under Bayesian statistics, the random effects may be estimated with particularly large credible intervals. This means our models may not adequately control for differences in the priming effects by item. This is common in acquisition studies, where the choice of items is limited to what children know and can be easily depicted, and is thus not unique to our study. Figure 4 shows that the priming effects were not consistent across verbs at 36 months. Notably, *push* behaved very differently to the other verbs. When it was the target verb, children showed the opposite pattern of results in unmatched verb trials and a reversal of this pattern in matched verb trials. We see an *anti-priming effect* in the unmatched verb condition, but a typical priming effect in matched verb trials. At later timepoints, *push* continues to behave in a similar manner, with no priming effect in the unmatched verb condition, but a typical priming effect in matched verb trials.

To explain this idiosyncrasy, we first checked whether data loss was biased by verb, perhaps skewing results for *push* due to low trial numbers (Appendix A). However, *push* trials were the most numerous, suggesting low trial numbers did not contribute to the result. We then analysed a corpus of Australian child-directed English (Kidd & Bavin, 2007) and the larger Manchester corpus (i.e., a corpus of British English; Theakston et al., 2001) to check whether *push* is passivised more frequently than our other verbs, under the assumption that *push* may prefer a passive frame (see Appendix B for details). Whilst *push* was by far the most frequent verb children heard and made up nearly half of the passives identified, as a proportion of total utterances *push* was as likely as other verbs to occur in the passive. Although the reason for this item-based effect was unclear, in light of the consistent and large discrepancy in the pattern of results, we decided to analyse the data both including *push* trials, as originally intended, and excluding them. Table 7 and Figure 5 summarise the proportion of active and passive responses in each experimental condition, excluding *push* trials. Graphically, the effects appear similar to the full set of trials, with a consistent abstract priming effect and the effect of matched verbs becoming larger over time.

Figure 4

Proportion of passive responses in each experimental condition for each verb at 36 months



Note: All verbs except *push* appear to show abstract priming, a greater proportion of passives in the passive than active priming condition for unmatched verb trials.

Table 7

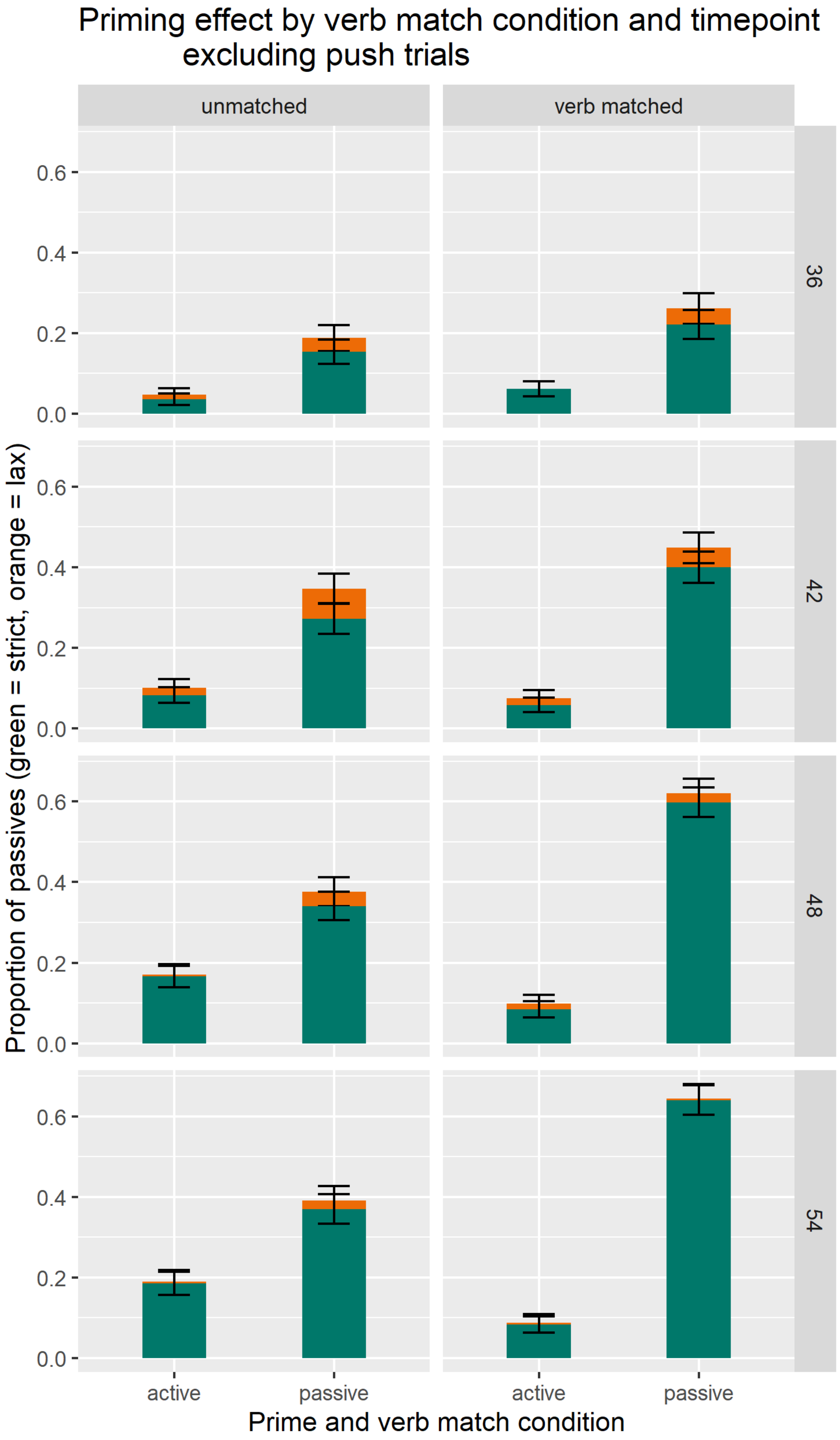
Number and proportion of passives and actives in each experimental condition excluding push *trials*

Timepoint	Prime	Verb match	Active		Passive – strict		Passive – lax	
			N	%	N	%	N	%
36m	Active	Unmatched	161	96.4/95.3	6	3.6	8	4.7
		Matched	152	93.8/93.8	10	6.2	10	6.2
	Passive	Unmatched	121	84.6/81.2	22	15.4	28	18.8
		Matched	102	77.9/73.9	29	22.1	36	26.1
42m	Active	Unmatched	177	91.7/89.9	16	8.3	20	10.2
		Matched	161	94.2/92.5	10	5.8	13	7.5
	Passive	Unmatched	107	72.8/65.2	40	27.2	57	34.8
		Matched	96	60.0/55.2	64	40.0	78	44.8
48m	Active	Unmatched	166	83.4/83.0	33	16.6	34	17.0
		Matched	174	91.6/90.2	16	8.4	19	9.9
	Passive	Unmatched	116	65.9/62.4	60	34.1	70	37.6
		Matched	72	40.2/37.9	107	59.8	118	62.1
54m	Active	Unmatched	149	81.4/81.0	34	18.6	35	19.0
		Matched	165	91.7/91.2	15	8.3	16	8.8
	Passive	Unmatched	109	63.0/60.9	64	37.0	70	39.1
		Matched	64	36.0/35.6	114	64.1	116	64.4

Note: The two percentages presented for Active responses are for the strict and lax coding schemes respectively.

Figure 5

Proportion of passive responses in each experimental condition at each timepoint, excluding push *trials*



Note: Abstract priming is indicated by a larger proportion of passives in the passive than prime condition for unmatched verb trials. The lexical boost effect is indicated by a larger difference between active and passive prime conditions for verb matched vs unmatched trials.

Cross-Sectional Models

The data were first analysed cross-sectionally, to understand the pattern of results at each timepoint. For example, whilst a longitudinal analysis may reveal that an effect increases over time, it does not differentiate between the effect emerging at a particular timepoint and the effect being present at all timepoints but becoming larger in magnitude. For the former situation, cross-sectional analyses also allow us to pinpoint when the effect emerged. In addition, the presence of an abstract priming effect at 3 years of age (our earliest timepoint) is of particular relevance to distinguishing between theories.

The production of passives was analysed as the frequency of passives out of active and passive responses (i.e., *Other* responses are excluded from the analyses). Results were analysed using mixed logistic models, which are suited to analysing binary outcome data and allow random effects for subjects and items to be accounted for in the same model (Jaeger, 2008). We used Bayesian rather than frequentist estimation due to the complexity of our models and the lower likelihood of convergence issues with Bayesian statistics (Eager & Roy, 2017). The R statistical environment was used for data analysis (verion 3.6.1.; R Core Team, 2014). The tidyverse packages (version 1.3.0.; Wickham et al., 2019) were used for data processing and visualisation and the brms package (version 2.13.0.; Bürkner, 2017) was used for statistical modelling.

We included random effects by item (target verb) and by participant. We were able to include the maximal random effects structure with random slopes for prime, verb match, and their interaction as well as correlations between random effects. The prime variable was effects coded (active: –0.5, passive: 0.5), and the verb match variable base coded (unmatched: 0, matched: 1). This allowed for an intuitive interpretation of the results with the prime effect being a simple effect for unmatched verb trials (therefore the abstract priming effect) and the match effect being a main effect across active and passive prime trials.[4] The interaction effect represents the lexical boost effect. Each model was run with 3000 iterations, 500 of them warm-up, and 4 chains. The default brms priors were used (uninformative priors). Across all the cross-sectional models for each parameter the maximum Rhat was 1.01, the minimum bulk effective sample size was 1037, and the minimum tail effective sample size was 998. The value of adapt_delta, which decreases the step sizes taken by the model, was increased from the default 0.8 closer to 1 as required to prevent divergent transitions (minimum 0.95, maximum 0.98 across all models).

Note that Bayesian statistics handle hypothesis testing differently to frequentist statistics and can in fact be interpreted more intuitively. A 95% credible interval is provided for each effect, which indicates the range of values within which the effect has a 95% chance of falling given the data. For effects of interest we provide the posterior probability for a one-sided hypothesis. The posterior probability represents the proportion of the parameter's posterior distribution that is above or below 0. A posterior probability of .95 indicates a 95% chance that the effect falls above or below 0 given the data. Rather than a binary decision about the presence or absence of an effect, we use the approach taken by Engelmann et al. (2019):

- If the posterior probability is >.95, we interpret this as strong evidence for an effect given the data.
- If the posterior probability is >.85, we interpret this as weak evidence for an effect.
- If the posterior probability is close to .5 we conclude there is no evidence for an effect.

Table 8 reports the results of the four models run at the 36-month timepoint. Despite maximal random effects being included in the model, the influence of trials where the target verb was *push* is strong enough that we observe different results depending on their inclusion. When *push* is included, there is weak evidence for the simple main effect of prime (unmatched trials only), or the abstract priming effect, whilst there is strong evidence for the interaction between prime and verb match, or the lexical boost effect, under lax coding and

---

[4] The intercept of an effect is determined by its coding. For both prime and match, the intercept is set at 0. However, 0 represents only the unmatched verb condition for the match effect, but the average of active and passive primes for the prime effect. Therefore, the effect of prime is a simple effect – at only one level of the other independent variable, whilst the effect of match is a main effect – averaging across both levels of the other independent variable.

weak evidence under strict coding. This suggests priming is driven by the verb-match condition. We see the opposite pattern of results when push is excluded, with strong evidence for abstract priming, but no evidence for the lexical boost effect.

Table 8

Results from models run at 36 months

Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-3.50	-2.87	-3.74	-3.18
	Credible interval	-4.88 | -2.38	-3.97 | -1.89	-5.37 | -2.41	-4.50 | -2.03
Prime	Estimate	1.52	1.23	2.41	2.27
	Credible interval	-0.43 | 3.83	-0.50 | 3.12	0.49 | 4.96	0.56 | 4.40
	Posterior prob.	.942†	.932†	.990*	.993*
Match	Estimate	-1.31	-1.27	-0.45	-0.55
	Credible interval	-4.05 | 0.74	-3.47 | 0.44	-3.01 | 1.48	-2.67 | 1.12
Prime*Match	Estimate	2.22	2.88	0.79	1.44
	Credible interval	-1.23 | 5.95	-0.20 | 6.37	-2.31 | 4.04	-1.24 | 4.47
	Posterior prob.	.904†	.969*	0.698	.858†

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

Table 9 reports the results from the four models run at the 42-month timepoint. The abstract priming effect now has strong evidence regardless of whether *push* trials are included, though it is numerically larger when *push* trials are excluded. There is strong evidence for the lexical boost effect when *push* trials are included but weak evidence when *push* trials are excluded.

Table 9

Results from models run at 42 months

Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-2.73	-2.14	-2.83	-2.22
	Credible interval	-3.62 | -1.98	-2.92 | -1.47	-3.90 | -1.90	-3.18 | -1.34
Prime	Estimate	1.74	1.83	2.19	2.33
	Credible interval	0.70 | 2.90	0.53 | 3.22	0.91 | 3.64	0.82 | 4.00
	Posterior prob.	.998*	.993*	.997*	.994*
Match	Estimate	-0.21	-0.14	-0.15	-0.12
	Credible interval	-1.33 | 0.76	-1.25 | 0.84	-1.70 | 1.25	-1.66 | 1.35
Prime*Match	Estimate	1.83	1.69	1.36	1.41
	Credible interval	-0.12 | 4.03	0.15 | 3.44	-1.07 | 4.04	-0.59 | 3.68
	Posterior prob.	.968*	.982*	.869†	.925†

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

Table 10 reports the results at the 48-month timepoint. Like at 42 months, there is strong evidence for the abstract priming effect regardless of the inclusion of *push* trials, and the same pattern is now evident for the lexical boost. The priming effect is numerically larger, and lexical boost effect smaller when *push* trials are excluded.

Table 10

Results from models run at 48 months

Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-1.45	-1.25	-1.58	-1.35
	Credible interval	-2.09 | -0.86	-1.85 | -0.70	-2.19 | -0.99	-1.94 | -0.81
Prime	Estimate	0.85	0.99	1.18	1.38
	Credible interval	-0.04 | 1.79	0.02 | 2.01	0.22 | 2.15	0.47 | 2.31
	Posterior prob.	.971*	.977*	.986*	.993*
Match	Estimate	0.24	0.25	0.27	0.30
	Credible interval	-0.45 | 0.90	-0.36 | 0.86	-0.58 | 1.09	-0.55 | 1.08
Prime*Match	Estimate	2.90	2.66	2.58	2.29
	Credible interval	1.30 | 4.59	1.06 | 4.30	0.87 | 4.46	0.55 | 4.24
	Posterior prob.	.999*	.997*	.995*	.989*

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

Table 11 presents results from the models run at 54 months. Again, the lexical boost effect has strong evidence across all the models and the abstract priming effect has a consistent magnitude across the models but has weak rather than strong evidence in the strict coding model excluding *push* trials. When *push* trials were excluded, the numerically larger priming effect and smaller lexical boost effect were evident, but this pattern is far less pronounced than at the earlier timepoints.

Table 11

Results from models run at 54 months

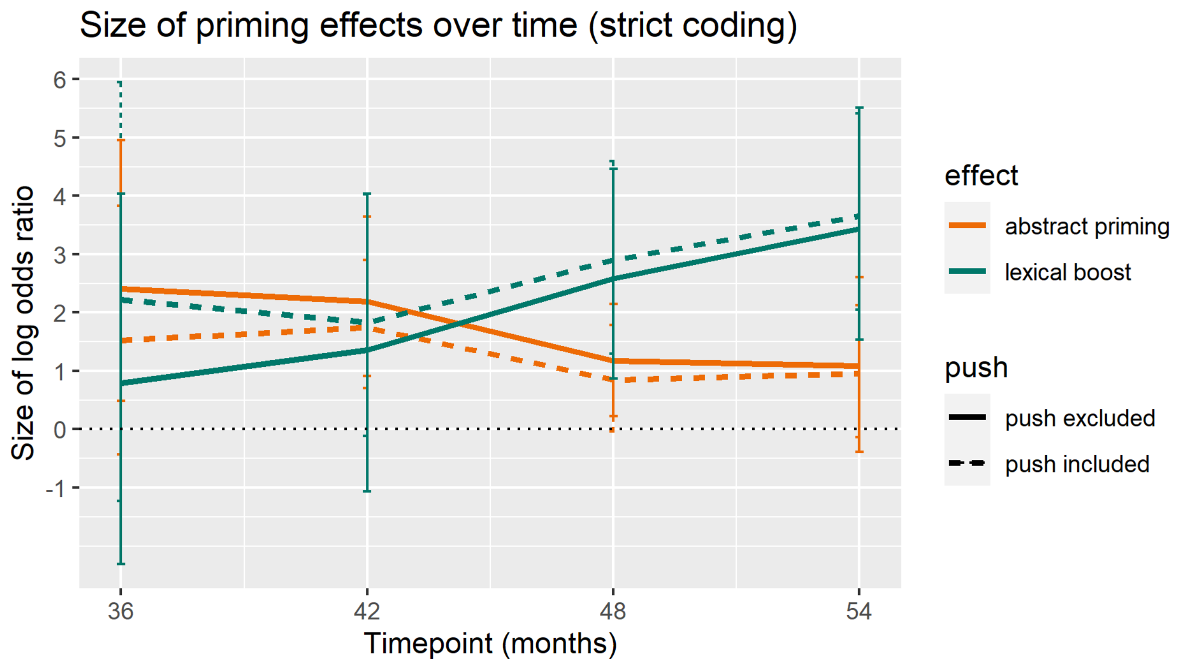
Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-1.45	-1.38	-1.47	-1.39
	Credible interval	-2.02 | -0.89	-1.97 | -0.83	-2.14 | -0.79	-2.05 | -0.75
Prime	Estimate	0.96	1.10	1.09	1.20
	Credible interval	-0.14 | 2.13	0.03 | 2.18	-0.39 | 2.61	-0.18 | 2.62
	Posterior prob.	.961*	.977*	.941†	.962*
Match	Estimate	0.44	0.33	0.52	0.41
	Credible interval	-0.33 | 1.12	-0.49 | 1.05	-0.42 | 1.43	-0.57 | 1.32
Prime*Match	Estimate	3.66	3.64	3.44	3.42
	Credible interval	2.05 | 5.41	2.09 | 5.51	1.53 | 5.51	1.51 | 5.62
	Posterior prob.	>.999*	>.999*	.998*	.998*

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

Figure 6 plots the coefficients and credible intervals of the abstract priming and lexical boost effects at each timepoint. There appears to be less certainty in the size of both effects at earlier timepoints, with larger credible intervals and greater divergence depending on the inclusion of *push* trials. Even accounting for this, there are some visible trends. In particular, when *push* trials are excluded, the size of the lexical boost effect increases over time. In contrast, the abstract priming effect appears to either slightly decrease or stay constant in magnitude over time. In order to investigate these trends over time, we combined data from all timepoints into our models.

Figure 6

Size of the abstract priming effect and lexical boost effect coefficients from cross-sectional models



Note: Error bars span the 95% credible interval for the coefficient.

Longitudinal Models

In the longitudinal models, there were two sorts of dependence within participants: dependence within participants across the four timepoints and dependence within participant for trials within a particular timepoint. As such, we included maximal random effects by participants and by participants nested within session. Note that because timepoint varies within participants, but not within sessions, random effects for time, and all its interactions were included by participant but not by session. Whilst there were also two types of dependence by items, we chose not to include more random effects by item than there were levels of the item (5 for no push models, and 6 for all verb models), since reasonable estimates would not be reached. Therefore, the by-item random effects were as for the cross-sectional models, with random intercept by item (target verb) and random slopes for prime, verb match, and their interaction. We also included correlations between random effects.

Again, the prime variable was effects-coded (–0.5, 0.5), and the verb match variable base-coded (0, 1). Time was coded as 1, 2, 3, 4 for each timepoint and then centred. We can therefore interpret our main effects as before, at the mid-point of the timepoints. However, since prime is effects coded (–0.5, 0.5) whilst verb match is base coded (0, 1), we must interpret the main effect of time for both prime conditions but only unmatched verb trials. Each model was run with 5000 iterations, 500 of them warm-up, and 4 chains. The default brms priors were used (uninformative priors). Across all longitudinal models for each parameter the maximum Rhat was 1.00, the minimum bulk effective sample size was 1319, and the minimum tail effective sample size was 1471. The value of adapt_delta, which decreases the step sizes taken by the model, was increased from the default 0.8 closer to 1 as required to prevent divergent transitions (minimum 0.99, maximum 0.995 across models).

Table 12 presents the results of the longitudinal models. Across both coding schemes and when *push* trials are excluded we see the same patterns of results. There is strong evidence for abstract priming and lexical boost effects at the midpoint of our timepoints. This is consistent with our cross-sectional models, where abstract priming reliably received strong evidence across most models and evidence for the lexical boost effect was always strong at the latter two timepoints and strong when *push* trials were included in the first two timepoints. There is strong evidence for the effect of time across all models. Due to the coding of our variables, this can be interpreted as follows: in all unmatched verb trials, the production of passives increases over time. That is, children’s overall passive production independent of prime condition increased over time. There was no support for a decrease in the abstract priming effect in models including all verbs but there was weak evidence for the decrease in models without *push* trials. There was strong evidence for an increasing lexical boost effect.

Table 12

Results from longitudinal models

Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-2.21	-1.87	-2.31	-1.97
	Credible interval	-2.77 |-1.65	-2.37 |-1.35	-2.90 |-1.70	-2.50 -1.41
Prime	Estimate	1.11	1.21	1.52	1.68
	Credible interval	0.22 | 2.06	0.20 | 2.19	0.73 | 2.32	0.92 | 2.46
	Posterior prob.	.989*	.986*	.998*	.999*
Match	Estimate	0.16	0.08	0.28	0.21
	Credible interval	-0.40 | 0.71	-0.44 | 0.57	-0.42 | 1.01	-0.42 | 0.79
Time	Estimate	0.69	0.53	0.75	0.60
	Credible interval	0.44 | 0.94	0.31 | 0.76	0.49 | 1.03	0.36 | 0.85
	Posterior prob.	>.999*	>.999*	>.999*	>.999*
Prime*Match	Estimate	2.23	2.16	1.57	1.54
	Credible interval	0.96 | 3.45	0.96 | 3.38	0.54 | 2.62	0.59 | 2.53
	Posterior prob.	.998*	.998*	.996*	.997*
Prime*Time	Estimate	-0.04	-0.00	-0.26	-0.26
	Credible interval	-0.39 | 0.32	-0.33 | 0.32	-0.67 | 0.15	-0.63 | 0.10
	Posterior prob.	.580	.505	.893†	.920†
Match*Time	Estimate	0.19	0.20	0.10	0.11
	Credible interval	-0.09 | 0.48	-0.06 | 0.47	-0.22 | 0.43	-0.18 | 0.41
Prime*Match *Time	Estimate	0.98	0.81	1.35	1.14
	Credible interval	0.41 | 1.57	0.29 | 1.35	0.70 | 2.03	0.55 | 1.75
	Posterior prob.	>.999*	>.999*	>.999*	>.999*

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

Individual Variability in Priming

We next examined variability in priming and passive production. Only 40 of 92 children produced a passive at 36 months, in comparison to 71 of 82 children at 54 months. Table 13 presents the percentage of children who produced a passive and were primed (produced a passive after a passive prime) under each coding scheme at each timepoint. In the vast majority of cases, children who produced passives were also primed. Figure 7 graphs the percentage of children, who completed all four timepoints and who were primed at each timepoint. Consistent with past studies (Bencini & Valian, 2008; Shimpi et al., 2007), it reveals substantial variability in priming in the youngest age group.

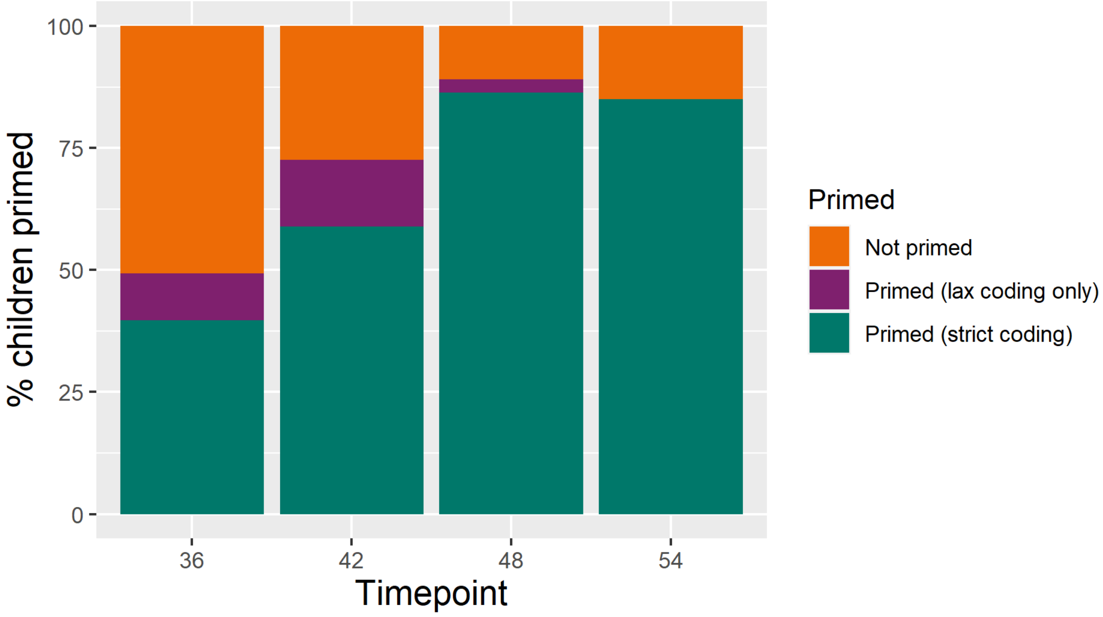
Table 13

Percentage of children who produced a passive and were primed at each timepoint

		36 months		42 months		48 months		54 months	
		N	%	*N*	%	*N*	%	*N*	%
Strict	Produced Passive	41	43.6	54	58.7	76	83.5	71	86.6
	Primed	39	41.5	51	55.4	76	83.5	68	82.9
Lax	Produced Passive	50	53.2	64	69.6	79	86.8	71	86.6
	Primed	47	50.0	63	68.5	77	84.6	68	82.9

Figure 7

Percentage of the 73 children who completed all timepoints who were primed at each timepoint



We investigated whether this variability was systematically linked to children's linguistic knowledge. The children's primary caregiver completed the Macarthur-Bates Communicative Development Inventory at 30 months, which included measures of vocabulary, grammatical complexity and mean length of utterance. These measures were intercorrelated, and so we ran a principal component analysis with promax rotation in SPSS to extract a single language proficiency component (see Appendix C for details). Table 14 presents the factor loadings of that component. It explains more than 70% of the variance of each measure individually and 65% of variance overall. Therefore, we used the extracted component as our measure of children's prior linguistic knowledge.

Table 14

Results of principal components analysis

	Factor 1
Vocabulary	.817
Complexity	.849
MLU	.747
Eigenvalue	1.95
% variance	64.91

We first analysed whether children's component scores predicted whether or not they were primed at 36 months. We ran Bayesian logistic regressions to predict children's membership category. Table 15 presents the results of these models. There was strong evidence that children's language proficiency component score predicted their tendency to produce passives and be primed (produce a passive after a passive prime) under both coding schemes.

Table 15

Results from models predicting membership category at 36 months

		Produced passive		Primed	
		Strict	Lax	Strict	Lax
Intercept	Estimate	-0.54	-0.02	-0.58	-0.14
	Credible Interval	-1.05 | -0.04	-0.52 | 0.46	-1.10 | -0.09	-0.65 | 0.36
Proficiency	Estimate	0.95	1.16	0.87	1.14
	Credible Interval	0.39 | 1.58	0.59 | 1.80	0.32 | 1.47	0.56 | 1.80
	Posterior prob.	>.999*	>.999*	.999*	>.999*

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

We then analysed whether the magnitude of the priming effect for those who were primed could be predicted by children's 30-month language proficiency. Table 16 presents the Spearman's correlations between children's language proficiency component score and their 36-month passive production and priming for only the children who produced a passive at 36 months. We observed correlations between the number of passives produced after a passive prime and the proficiency component. Under lax coding, there was a significant medium-sized correlation between priming magnitude (proportion of passives after passive primes – proportion of passives after active primes) and the proficiency component. These correlations suggest that children's language proficiency at 30 months is associated with the magnitude of their abstract priming effect.

Table 16

Spearman's correlations between proficiency and passive production and priming in subset of participants who produced passives or were primed

		1	2	3	4	5	6	7
1. Proficiency		-						
2. *N* passives produced	strict	.33^	-					
3.	lax	.38*	.92***	-				
4. *N* passives primed	strict	.41*	.86***	.73***	-			
5.	lax	.44**	.83***	.89***	.83***	-		
6. Priming magnitude	strict	.25	.33*	.25	.68***	.51***	-	
7.	lax	.31*	.35*	.38**	.62***	.65***	.94***	-

^*p*-value between .05 - .1, *correlation is significant at the *p* < .05 level, **at the p < .01 level, ***at the .001 level

We next re-ran the 36-month cross-sectional models on the subset of participants who produced a passive, with language proficiency and its interactions included as additional predictors (subsetting the children by whether they were primed made no difference to the pattern of results). Across these models for each parameter the maximum Rhat was 1.00, the minimum bulk effective sample size was 1353, and the minimum tail effective sample size was 1625. The value of adapt_delta, which decreases the step sizes taken by the model, was increased from the default 0.8 to 0.97 to prevent divergent transitions. Table 17 presents the results of these models. The evidence for an abstract priming effect was strong in all models except for the lax coding model including *push* trials, where is it was weak. This is similar to the original models where there was only strong evidence for it when *push* was excluded. Interestingly, in this more linguistically advanced subset of participants, there is strong evidence for the lexical boost effect even with *push* trials excluded. Children's language proficiency does not interact with prime to predict the magnitude of priming, but there is weak evidence for it predicting the production of passives more generally in three of the four models.

Table 17

Results from models predicting magnitude of priming at 36 months in children who produced a passive

Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-1.87	-1.74	-2.64	-2.53
	Credible interval	-3.13 | -0.83	-2.84 | -0.76	-4.59 | -1.14	-4.26 | -1.18
Prime	Estimate	1.66	1.00	3.14	2.79
	Credible interval	-0.26 | 3.95	-0.71 | 3.00	0.60 | 6.77	0.68 | 5.91
	Posterior prob.	.958*	0.878†	.991*	.994*
Match	Estimate	-1.28	-1.57	0.02	-0.60
	Credible interval	-4.71 | 1.29	-4.75 | 0.86	-3.11 | 2.69	-4.01 | 1.99
Proficiency	Estimate	0.34	0.50	0.91	1.00
	Credible interval	-0.81 | 1.54	-0.43 | 1.46	-0.79 | 2.91	-0.38 | 2.56
	Posterior prob.	.742	.879†	.868†	.933†
Prime*Match	Estimate	4.55	5.60	3.05	3.94
	Credible interval	0.21 | 10.67	1.45 | 11.71	-1.80| 9.24	-0.59 |10.02
	Posterior prob.	.981*	.997*	.884†	.956*
Prime* Proficiency	Estimate	0.12	0.48	-0.93	-0.70
	Credible interval	-1.90 | 2.09	-1.13 | 2.06	-4.16 | 1.81	-3.21 | 1.42
	Posterior prob.	.551	.738	.738	.729
Match* Proficiency	Estimate	1.22	0.35	0.72	0.09
	Credible interval	-1.21 | 4.20	-2.00 | 2.84	-2.06 | 3.88	-2.60 | 3.11
Prime*Match *Proficiency	Estimate	0.48	-0.64	1.05	0.25
	Credible interval	-4.24 | 5.48	-4.99 | 3.65	-3.98 | 6.68	-4.36 | 5.03
	Posterior prob.	.578	.624	.659	.545

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

Exploratory Analyses

Contrary to the predictions of the Chang et al. (2006) model, in our longitudinal models, the abstract priming effect did not reduce in magnitude over time. However, in the analyses excluding the idiosyncratic *push* verb, there was weak evidence toward a negative effect. In addition, we found that the age at which children first produced a passive in the task was variable. It is possible that this variability in acquisition of the passive masked a decrease in priming.

In exploratory analyses, we repeated our longitudinal analyses on the subset of children who demonstrated knowledge of the passive at 36 months because they produced at least one passive during that testing session. Table 18 presents the results of these analyses. We used the same model specifications as for the longitudinal models. Across all models for each parameter the maximum Rhat was 1.00, the minimum bulk effective sample size was 1858, and the minimum tail effective sample size was 1361. The value of adapt_delta, which decreases the step sizes taken by the model, was increased from the default 0.8 closer to 1 as required to prevent divergent transitions (minimum 0.98, maximum 0.99 across models).

Table 18

Results of longitudinal models on children who produced a passive during the first session

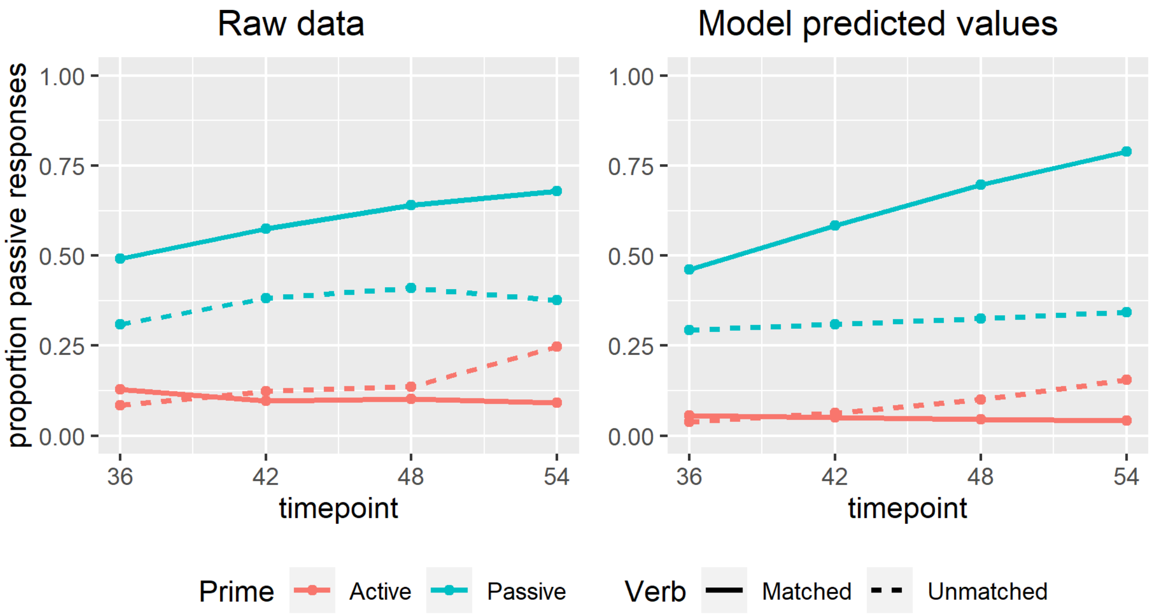
Effect		Strict	Lax	Strict – no push	Lax – no push
Intercept	Estimate	-1.39	-1.23	-1.63	-1.40
	Credible interval	-1.95 |-0.85	-1.69 |-0.79	-2.18 |-1.10	-1.88 |-0.93
Prime	Estimate	1.28	1.24	1.71	1.68
	Credible interval	0.39 | 2.21	0.47 | 2.04	0.75 | 2.65	0.95 | 2.47
	Posterior prob.	.993*	.996*	.996*	.999*
Match	Estimate	0.21	0.14	0.42	0.28
	Credible interval	-0.39 | 0.79	-0.30 | 0.57	-0.21 | 1.04	-0.25 | 0.79
Time	Estimate	0.22	0.19	0.32	0.29
	Credible interval	-0.08 | 0.51	-0.06 | 0.45	-0.01 | 0.66	0.00 | 0.58
	Posterior prob.	.922†	.933†	.972*	.977*
Prime*Match	Estimate	2.20	2.18	1.82	1.75
	Credible interval	1.15 | 3.23	1.26 | 3.08	0.57 | 3.14	0.80 | 2.80
	Posterior prob.	>.999*	>.999*	.994*	.998*
Prime*Time	Estimate	-0.13	0.01	-0.48	-0.35
	Credible interval	-0.62 | 0.35	-0.40 | 0.41	-1.03 | 0.07	-0.79 | 0.08
	Posterior prob.	.699	.480	.958*	.941†
Match*Time	Estimate	-0.01	0.09	-0.11	-0.03
	Credible interval	-0.35 | 0.34	-0.21 | 0.41	-0.51 | 0.30	-0.38 | 0.32
Prime*Match *Time	Estimate	0.64	0.51	1.14	0.92
	Credible interval	-0.06 | 1.34	-0.09 | 1.11	0.34 | 1.98	0.23 | 1.63
	Posterior prob.	.965*	.953*	.997*	.996*

Note: For all effects we report the coefficient of the effect and a non-directional 95% credible interval around this estimate. For effects of interest we report the posterior probability that the effect is > 0 in a directional hypothesis test. * indicates strong evidence that the effect is > 0 in this test. †indicates weak evidence that the effect is > 0.

As in the full sample, the evidence for the abstract priming effect, the lexical boost effect and the increase in the lexical boost effect was consistently strong. However, effects involving time were more evident once *push* is excluded from the analyses. In these analyses, there is strong rather than weak evidence for an increase in passive production over time, and strong evidence for a decreasing priming effect under strict coding, and weak evidence under lax coding. Figure 8 presents the condition means of the raw data and those predicted by the strict, no *push* model. It shows that the decrease in abstract priming but increase in passive production arises because children increase their passive production only following active primes (slope estimate = 0.56, CI_{95} = 0.10 | 1.05), whilst passives following passive primes remain stable (slope estimate = 0.09, CI_{95} = -0.29 | 0.47).

Figure 8

Predicted and actual condition means for the strict, no push analysis of children who produced a passive at 36 months



Note: these are plotted on the probability scale (the scale of the raw data) rather than the logit scale, on which the model assumes linear relationships. Plots on the logit scale are available on https://osf.io/35kzm/.

Discussion

In this paper we have presented the first longitudinal study of syntactic priming in development. A key feature of our study is that we are, therefore, able to track both the emergence and the developmental trajectory of abstract priming and lexically-based priming/the lexical boost, which enabled us to test the predictions of three different theoretical approaches that make different predictions concerning these priming effects.

These predictions and our findings are summarised in Table 19. We found an early emerging abstract priming effect in support of the RA-Early Syntax and Implicit Learning accounts but counter to the RA-Late Syntax account. This effect remained stable in magnitude across development when including all lexical items, which is only predicted by the RA-Early Syntax account. However, counter to at least some instantiations of the RA-Early Syntax account and in support of the two input-driven models, there was large variation in the onset of abstract priming, with less than 50% of 3-year-olds exhibiting priming under the strict coding scheme. This variation was systematic and meaningful, with the tendency to be primed predicted by children's language proficiency six months earlier, although the magnitude of priming was not (though the bivariate correlation between proficiency and priming magnitude was positive and significant). Moreover, when we excluded an atypical lexical item or considered only participants who exhibited a priming effect at 36 months, we found some evidence that the priming effect decreased over time. The lexical boost effect was idiosyncratic at earlier timepoints, with significance depending on a single lexical item. Excluding this item, the effect was late emerging and increased in magnitude over development, a pattern only explicitly predicted by the Implicit Learning account.

Table 19

Predictions of structural priming accounts and study findings

	Abstract Priming		Lexical Boost/Lexically-based	
	Emergence	Development	Emergence	Development
RA - Early Syntax	Early	No change	Early	No change
RA - Late Syntax	Late	Increase	Early	Inverse U-shape
Implicit learning	Early	Decrease	Late	Increase
Study findings	Early	No change/ decrease	Late	Increase

Interpretation of Results

We first note that our pattern of results was largely consistent between models that included and excluded *push* items. On the occasions where they diverge, we took the models excluding *push* to be more representative of the true priming effects, since the presence or absence of an effect in those models is not driven by a single lexical item. For the lexical boost effect, the differences between the models appeared systematic in that models including *push* were developmentally ahead of time. That is, the lexical boost effect increased in both *push* and no *push* longitudinal models, but received strong evidence at an earlier timepoint in *push* rather than no *push* cross-sectional models. For the abstract priming effect, the anti-priming effect in *push* trials at 36 months masks the size of the effect for other verbs, in turn masking the decrease in abstract priming in the subset of children who produced a passive at 36 months. Differences between strict and lax coding models were fewer, usually with the size of an effect being comparable but with strong evidence in one model and weak evidence in the other.

Abstract Priming Effect

We found an abstract priming effect at all timepoints, and an increase in overall passive production over development. In analyses of the entire data set, although overall passive production increased, the difference between the number of passives following passive primes compared to active primes remained stable over development. However, when *push,* which displayed an anti-priming effect at the first timepoint, was excluded, we found weak evidence for a decreasing abstract priming effect alongside strong evidence for an increase in overall passive production. In a set of exploratory analyses, we further examined priming effects in a sub-group of participants who produced a passive at 36 months. Our motivation for these analyses was that variation in the emergence of priming amongst children in the full sample may have masked a true decrease in priming magnitude. Children in this sample exhibited a stable abstract priming effect in models with all verbs included but decreasing effect in models excluding *push,* which eliminated the masking effect of anti-priming in *push* trials at 36 months. The strength of the evidence for the decreasing effect was stronger than in the full sample. Children in this sub-sample also exhibited an overall increase in passive production over time, though with weaker evidence than in the full sample.

An increase in overall passive production suggests learning but its association with a decreasing priming effect is, at first glance, puzzling. Therefore, one might argue that these results reflect a test-retest effect or cumulative priming effects rather than implicit learning. Children's passive production in the task increased over and above baseline passive production in natural speech (0.1%; Xiao et al., 2006). Moreover, children increased in their proportion of passives produced after *active* primes from 5.6% of utterances at 36 months to 12.8% at 54 months (strict coding; lax coding: 10.6% to 15.04%).[5] There are several possible explanations for this increase in passive production. If the increase in passives reflects a test/re-test effect, children may have learned to produce more passives in the context of the SNAP task, without changes to their linguistic knowledge more generally. If the increase reflects an increase in cumulative priming effects across filler sentences (Branigan & McLean, 2016) children's tendency to be primed by earlier passive primes in subsequent trials would increase over development. Both of these scenarios would increase passive production after active primes as well as passive primes. If passive production after passive primes had reached ceiling, this could have led to a spurious observation of a decrease in priming. However, the data are not consistent with this interpretation because the probability of producing a passive after a passive prime is far below ceiling, and indeed lower than on verb overlap trials, at all timepoints (see Figure 8). That is, it seems that what is changing is children's preference for producing passives *relative to actives*.

Such an effect suggests learning, and is consistent with the Chang et al. (2006) model. That is, as children's representation of the passive is tuned to occur more frequently, they produce more passives spontaneously after active primes and the prediction error caused by passive primes is smaller, resulting in an increase in passive production overall but a smaller effect of passive primes. Children hear few full BE passives outside of the lab, and their experience in the study across 18 months, where they hear a balanced number of actives and passives to describe transitive events, appeared to promote its use relative to the active across time, most prominently from the 48- to 54-month session. This is consistent with findings from training studies that increase children's exposure to passives. In an early study, Whitehurst, Ironside, and Goldfein (1974) modelled passive production to 4 – 5-year-old English-speaking children, which subsequently improved their production and comprehension relative to a control group. Similarly, Vasilyeva et al. (2006) increased passive sentences in 4-year-old English-speaking children's input via a two-week-long book reading intervention, which increased their production and comprehension of the passive relative to a group that heard active sentences. The difference between the current study and these

---

[5] In the full sample rather than the subset of children who produced a passive at 36 months, the proportion of passives produced after active primes increased from 4.5% to 12.4% under strict coding and 5.6% to 12.8% under lax coding from the 36- to 54-month timepoint.

past intervention studies is that we primed the active and passive within-participants, whereas the intervention studies did so between-participants. Thus, our data build upon the results of the intervention studies by showing that the relative weighting of structural options in the active-passive alternation is a property of the individual child.

Converging evidence for this interpretation comes from acquisition studies in languages that differentially weight the use of active and passive voice. Acquisition studies of languages such as Inuktitut (Allen & Crago, 1996), Ki'che' Mayan (Pye & Quixtan Poz, 1988), and Sesotho (Demuth, 1989; Kline & Demuth, 2010), where passives are structurally similar but are relatively more frequent than in European languages, show that children acquire the structure earlier and use it more frequently in their spontaneous speech.

We therefore conclude that, under conditions that take into account children's knowledge of the target structure at time 1, our results are consistent with the presence of a decrease in abstract priming over development. This is consistent with Rowland et al.'s (2012) study of the dative alternation. Additionally, Messenger (2021) found a very similar marginally significant effect in her passive priming study comparing children to adults: adults had a higher baseline rate of passive production than 3 – 4-year-olds, and their priming effect was marginally smaller than that of the children (some of whom were likely not yet primed, given our findings). Our finding is inconsistent with Peter et al. (2015), who found an increasing abstract priming effect, and those studies that have reported no developmental differences (e.g., Hsu, 2019; Messenger, Branigan, & McLean, 2012; Messenger, Branigan, McLean, et al., 2012). However, as we detail below, it is difficult to draw equivalence between longitudinal and cross-sectional designs.

The fact that there was less evidence for a decreasing abstract priming effect in the full sample than those primed at 3 years points to a lower bound on priming, such that there must be sufficient knowledge of the relevant structure in place prior to priming. This is supported by our finding that there is substantial variation in the emergence of priming, which is meaningfully linked to children's linguistic knowledge 6 months prior, and by Kidd's (2012) similar findings in older children. Thus, children need to have acquired the structure to a sufficient degree before priming can be observed. At that point priming is relatively large in magnitude but decreases across developmental time. The conflicting results from past studies may thus be attributable to variation in children's knowledge of the target structure, which is only partially related to age. We suspect the effect is small, requiring higher power than is often achieved in developmental studies, and is less easily observed in cross-sectional designs where children may be pooled with those at different developmental levels. This highlights the importance of longitudinal data, which compares children to their own past performance.

Lexically-Based Priming and the Lexical Boost Effect

Turning to priming effects on trials with lexical overlap, we did not observe lexically-based priming in the absence of an abstract priming effect, suggesting that that there was no lexically-based priming prior to abstract priming, at least for the active-passive alternation. Moreover, we found strong evidence for a 3-way interaction between prime, time and verb match, suggesting that the lexical boost effect increases over development. In the more advanced subset of children who produced a passive at 3 years of age, the lexical boost effect was strong in three of the four cross-sectional models at 36 months, whilst for the full sample it was only strong under lax coding with *push* included. This finding suggests that the lexical boost effect emerges between 3 and 4 years of age, with its idiosyncratic nature the likely reason for past inconsistent results (Branigan & McLean, 2016 vs Rowland et al., 2012; Peter et al., 2015). Our study also confirms the dissociation between abstract priming and the lexical boost effect, with the former emerging earlier on in development, and with the two effects having different developmental trajectories. This result provides crucial developmental evidence in support of the suggestion that abstract priming and the lexical boost derive from separate mechanisms (Chang et al., 2006; Hartsuiker et al., 2008; Reitter et al., 2011).

One notable yet unexpected result concerning the lexical boost was the behaviour of one verb, *push*, which showed a lexical boost effect far earlier than the other verbs. The initially verb-dependent nature of the lexical boost could also explain inconsistent findings regarding its existence in young children. However, the reason for the result is unclear: all target verbs have an early age of acquisition, and they are all action verbs and so did not have semantic differences that may have influenced passivisation (Nguyen & Pearl, 2021). We checked data loss by verb and found that *push* trials were in fact the least likely to be excluded (see Appendix A), suggesting biased data loss was not the cause. A corpus analysis revealed that *push* was more frequent than the other verbs in child-directed speech, although it was not more likely to be passivized (Appendix B). Accordingly, children may be more familiar with *push* in the passive construction simply because it is a more frequent verb.

While infrequent structures are more syntactically prime-able than frequent ones (Bock, 1986; Jaeger & Snider, 2013), it may be that more frequent lexical items are more prime-able via lexical mechanisms. Specifically, frequent verbs may produce more enduring explicit memory traces linked to structure. How this occurs is still unclear. One possibility is that the verb-specific effect relates to how lexical entrenchment (i.e., verb frequency) establishes event representations of different strengths. If *push* is more frequent, then its event structure is likely more accessible (Elman, 2009). The early lexical boost effect for *push* suggests that the event construal is more flexible for this verb; that is, children can more flexibly alternate between the agent and the patient as starting points (MacWhinney, 1977), such that hearing a prime containing *push* in the active or the passive increases the likelihood that a target event containing *push* will be construed from the perspective of the topicalized NP. Sentence construal is prime-able in adults: using eye-tracking in a visual world paradigm, Sauppe and Flecken (2021) showed an active or passive prime significantly affected whether adult Dutch-speaking participants fixated on an agent or patient in a briefly-presented transitive event following the prime. Our suggestion is that in children such an effect may vary across individual verbs, and thus influence the likelihood that the lexical boost will be observed early in development. Investigating these processes using online methodologies like eye-tracking appears to be a promising avenue of future research.

Some further features of our data and past results support our claim of a nexus between event construal and priming effects. We found biased data loss such that transitive sentences were more likely in the verb overlap condition. We are not the first to report similar effects: Gamez and Vasilyeva (2015) analysed the likelihood that 5–6-year-old L2 English learners would produce complete sentences and found a significant effect of prime type (passive primes), and prime repetition. In addition, Shimpi et al. (2007) found in their second experiment that 3-year-olds were not primed but did produce more transitives following transitive primes. Studies of priming in languages that contain multiple structures allowing the speaker to emphasise the patient have shown that these structures prime each other (e.g., Spanish: Gámez et al., 2009; Russian: Vasilyeva & Waterfall, 2012), suggesting that the prime sentence primes children to construe an event from the perspective of the patient in general, from which point children select an appropriate syntactic structure. There was even some evidence in our English-speaking data of these processes. Specifically, we had difficulty categorising some errors where children appeared to be attempting to produce structures which topicalised the patient but had not fully acquired the appropriate passive structure to do so. For example, after hearing a passive prime, one 36-month-old child described *dog kisses chicken* with “A chicken by a dog... making the chicken be happy... maybe the dog is making the chicken be happy.” Collectively, these findings show how prime sentences influence a broader spectrum of behaviour concerning scene perception and sentence construal.

Implications for Accounts of Priming and Syntax Acquisition

RA-Late Syntax Account

Our results are least compatible with the RA-Late Syntax account (e.g., Savage et al., 2003; Tomasello, 2003). The account predicts the early emergence of lexically-based priming followed by a later emergence of the abstract priming effect, when in fact we observed the opposite pattern of results. There are two points worth considering in relation to the approach, despite its poor prediction of priming effects. Firstly, the theory is mostly concerned with explaining children’s very early grammatical knowledge (e.g., Pine et al., 1998; Rowland, 2007; Tomasello, 1992, 2003), although its assumptions and mechanisms extend beyond the first sprouting of syntactic knowledge and have framed significant debates and theoretical development in the literature (e.g., Ambridge, 2019; Ambridge & Lieven, 2011; Fisher, 2002; McCauley & Christiansen, 2019; Özge et al., 2019; Thothathiri & Snedeker, 2008; Tomasello, 2000). While there is no doubt much idiosyncrasy in syntactic knowledge, such that constructional knowledge has different levels of abstraction (Goldberg, 1995, 2005), it appears that children move rapidly away from item-based syntactic knowledge of core argument relations quite early in development (Bannard et al., 2009). The challenge for the RA-Late Syntax approach is to develop a sufficiently detailed yet constrained account of syntactic development that distinguishes between item-based and abstract knowledge across different levels of developmental experience.

Secondly, we note that there are some features of our data that are broadly consistent with the RA-Late Syntax approach. The first is that there were meaningful individual differences in the emergence of children’s knowledge of the passive, which was predicted by their linguistic proficiency 6 months prior. The second is that children’s general production of the passive was more frequent across development, suggesting that children's overall knowledge of the passive strengthened across time (consistent with past research, see Marchman et al., 1991). As we will see, however, these are not unique features of the RA-Late Syntax account.

RA-Early Syntax Account

The RA-Early Syntax approach predicts a large degree of continuity between adult and child priming effects (Bencini & Valian, 2008; Valian, 2014), such that there should be early emerging abstract priming and lexical boost effects, which remain stable across development. While we did find an early emerging abstract priming effect that appeared developmentally stable across the whole cohort, the individual variability in which children had acquired productive competence with the passive masked a negative developmental effect, whereby the abstract priming effect decreased with development. This effect constrains the RA-Early Syntax approach. In particular, contra to one version of the approach that explicitly assumes full continuity between the child and adult state (e.g., Crain et al., 2017; Crain & Thornton, 1998), our data strongly suggest developmental change in the system for both the abstract priming and lexical boost effects, which varies systematically across individuals. Thus, there exist observable and measurable learning effects for language-specific structures beyond 3 years, which presumably must be attributable to both children's variable input and variability in endogenous learning mechanisms (Kidd, Donnelly, et al., 2018; Kidd & Arciuli, 2016). Amongst RA-Early Syntax approaches, these data are more consistent with accounts that assume children necessarily acquire language-specific knowledge via the input, building upon less specified innate content (e.g., Fisher et al., 2020; Messenger & Fisher, 2018). We note that these accounts currently lack detail regarding how syntactic knowledge may change or be updated in response to further experience once abstract categories emerge. The models would need to be updated to reflect the specific developmental changes we observed in our data.

Implicit Learning Account

The Implicit Learning approach of Chang et al. (2006) provides the best fit to the data, accounting for both the early emergence of abstract priming and the late emergence and increase over development of the lexical boost effect. In addition, it explicitly predicts the decrease in the abstract priming effect found in a subset of children who produced a passive at 36 months. One additional advantage is that the account is computationally implemented, and thus provides an explicit account of the system's initial conditions and its learning mechanisms. There is a general consensus amongst computational models of priming that the abstract priming effect involves a form of implicit learning (Chang et al., 2000, 2006; Reitter et al., 2011); however, the Chang model, with which our data are most consistent, explains abstract priming via error-based learning, and correctly predicts the decreasing developmental effect (see also Dell & Chang, 2014). Additionally, the Chang model correctly predicts the asymmetry in the emergence and development of the lexical boost relative to abstract priming, although the model itself does not have a mechanistic account of the lexical boost. The Chang et al. (2006) model has a further advantage in being a model of language acquisition and therefore explains empirical phenomena beyond syntactic priming.

Conclusions

In this paper, we reported on the first longitudinal study of syntactic priming in development, tracking the priming of the active-passive alternation in a large sample of English-speaking children between the ages of 3;0 and 4;6. The longitudinal design allowed us to distinguish between several accounts of the acquisition of syntactic knowledge and, importantly, whether and how that knowledge changes across time. Our use of the syntactic priming method enabled us to make explicit connections between models of acquisition and mechanistic models of adult sentence production aimed at explaining priming effects (among other effects). Overall, we found evidence for the early emergence of abstract knowledge of the passive, which both varied across individual children and changed across developmental time. We also found evidence for asynchrony in the emergence and development of the lexical boost, supporting the suggestion that abstract priming and the lexical boost emerge via different mechanisms. These data are best accommodated within Chang et al.'s (2006) connectionist model, where knowledge of structure emerges via error-based learning relatively early in development, but continues to change with language use across developmental time.

Appendix A

Data Loss by Verb

Figures A1 to A4 graph data loss at each timepoint by verb. They show that there was bias in trial numbers by verb. Especially at earlier timepoints, children often produced intransitive sentences using *run* for *chase* actions, or produced transitive sentences with *cuddle* for *hug* actions and *hold* for carry actions. However, *push* trials were least often excluded from analysis. This, and the consistency of the differing pattern of results for *push* across timepoints, suggests that this item-specific effect is reliable.

Figure A1

Data loss by verb at the 36-month timepoint

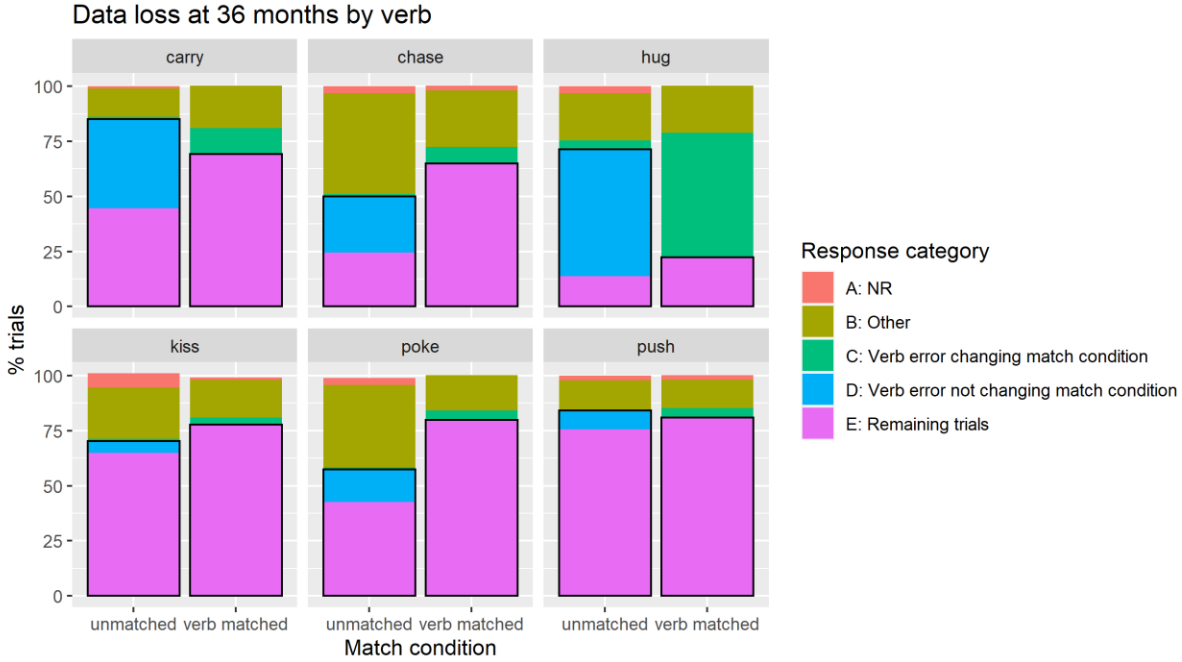


Figure A2

Data loss by verb at the 42-month timepoint

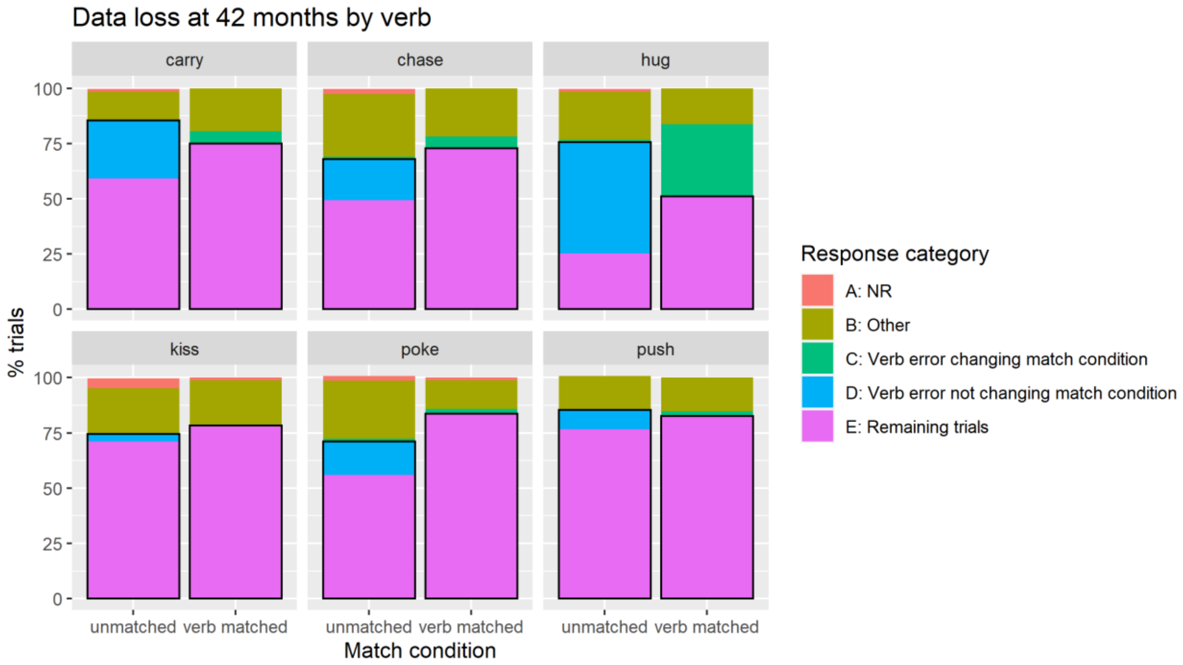


Figure A3

Data loss by verb at the 48-month timepoint

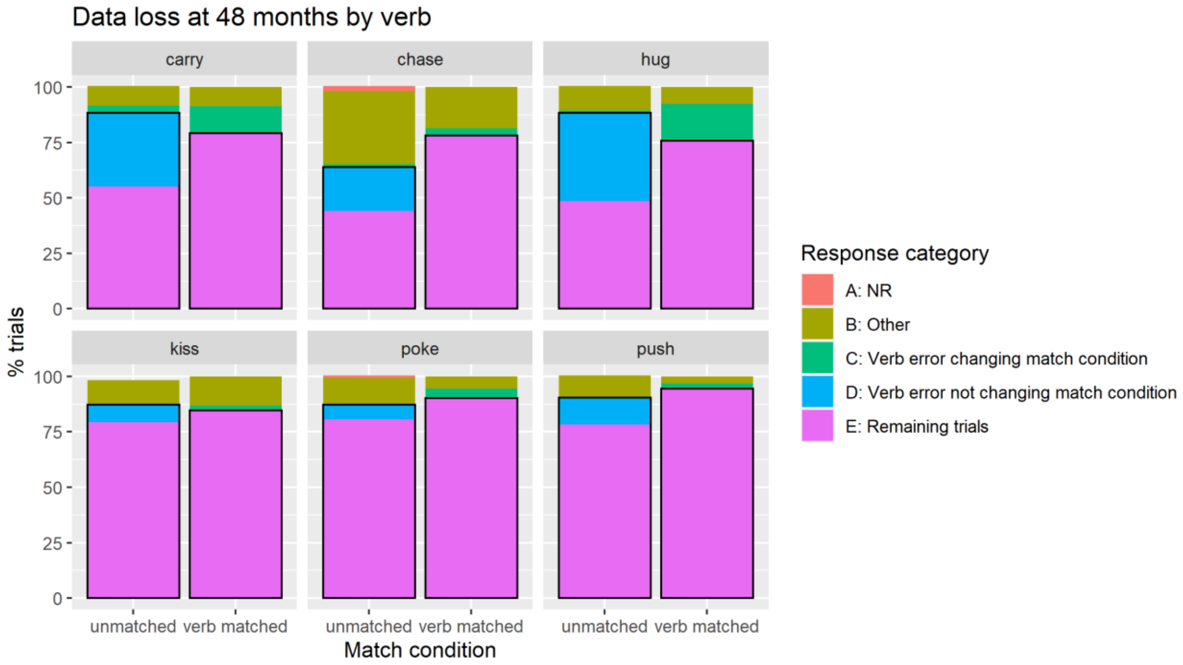
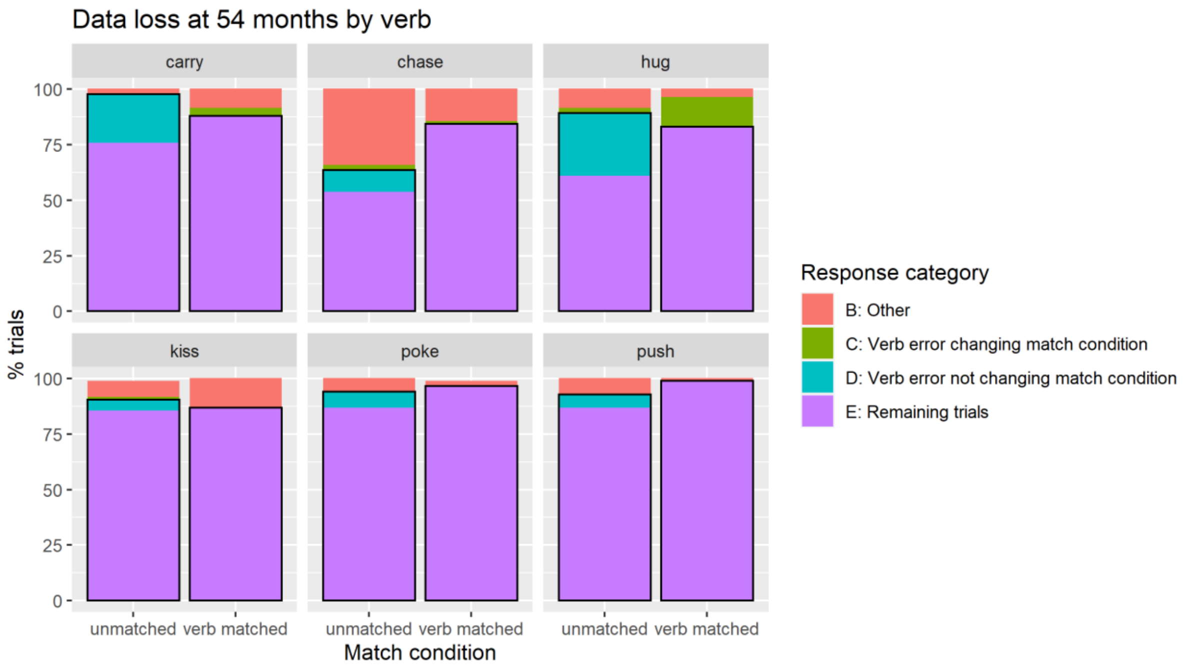


Figure A4

Data loss by verb at the 54-month timepoint



Appendix B

Corpus Analysis

We ran a corpus analysis to check the frequency of each of our verbs in child directed speech, and whether *push* occurs more frequently as a passive. We searched for utterances containing the verb lemmas, and excluded any including noun or adjectival uses. We coded the utterances as active, passive, or other. Since natural speech is less systematic than that elicited in experimental contexts, we coded transitives more generously to include utterances with transitive thematic role order if not strictly transitive syntax. Therefore questions, imperatives with subjects, infinitival structures, existential structures, sentences with modal verbs, subordinate clauses, verb arguments, relative clauses, participle phrases, and subjectless structures where the subject was clear from the context were all included in the active category if they had a SVO thematic role order. Passives included truncated passives and some subjectless structures where the subject was clear from context. The *Other* category included structures without a clear SVO order including imperatives, intransitives, gerunds, some participle phrases, relative clauses and questions with non-transitive thematic role orders, structures without a subject where that subject was not clear from the context, and phrasal verbs.

We extracted utterances containing our verbs from an Australian corpus of child directed speech (Kidd & Bavin, 2007). However, we found too few instances of our verbs, and in fact no passive instances, to perform an analysis of their frequency as passives (see Table B1). Notably, *push*, which behaved differently in the priming experiment, is by far the most frequent verb children heard but mostly as a non-transitive imperative.

Table B1

Frequency of experimental items in Australian corpus of child directed speech

Verb	*carry*	*chase*	*hug*	*kiss*	*poke*	*push*
Active	7	4	0	3	0	7
Passive	0	0	0	0	0	0
Other	1	5	0	0	1	26
Total	8	9	0	3	1	33

We ran the same analysis on the Manchester corpus from CHILDES (Theakston et al., 2001). The verbs were used similarly enough by parents to be comparable to Australian English and suitable for our purposes. Table B2 reports the outcome of this analysis and Table B3 details the 12 passive utterances found in our corpus analysis. Again, *push* was by far the most frequent verb. *Hug* was notably infrequent but occurred frequently as a noun in dative constructions which were excluded from our analyses. *Kiss* also occurred often as a noun in dative constructions but was additionally frequent as a verb.

Push was not more likely to appear as a passive than our other verbs. *Kiss* was the verb most likely to appear as a passive with 2.8% of utterances occurring in the passive and all three *full* passives using *kiss*. However, of the 12 total passive utterances, *push* was as frequent as *kiss*, with 5 passives for each verb.

Table B2

Frequency of experimental items in Manchester corpus of child directed speech

Verb	*carry*		*chase*		*hug*		*kiss*		*poke*		*Push*	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Active	114	59.1	57	87.7	9	90.0	129	72.9	19	47.5	282	46.6
Passive	1	0.5	1	1.5	0	0.0	5	2.8	0	0.0	5	0.8
Other	78	40.4	7	10.8	1	10.0	43	24.3	21	52.5	318	52.6
Total	193		65		10		177		40		605	

Table B3

Passive utterances

Verb	Utterance
carry	there's a baby being *carried* in a very special way.
chase	being *chased* again.*
kiss	**being *kissed* by a thing.*** **being *kissed* by all these creatures.*** I'm not sure you deserve to be *kissed* better because you were being silly. doesn't the polar bear like being *kissed*? **oh he's been *kissed* better by the vet, has he?**
push	the poor cow'd be *pushed* out of the way, wouldn't it ? I willn't be *pushed* down this time. I think if that's not *pushed* in you can't hear it. or was it *pushed*? yeah well that won't fit on now because it's not *pushed* down enough.

Note: Passives with a *by*-phrase in bold. The subject of * utterances was clear from context.

Appendix C

Principal Components Analysis

Children’s primary caregiver completed the Macarthur-Bates Communicative Development Inventory at 30 months, which included measures of vocabulary, grammatical complexity and mean length of utterance (MLU). Table C1 presents the correlation matrix between the variables, which were all significantly correlated at medium to large correlation sizes.

Table C1

Spearman’s correlations between predictors measured at 30 months

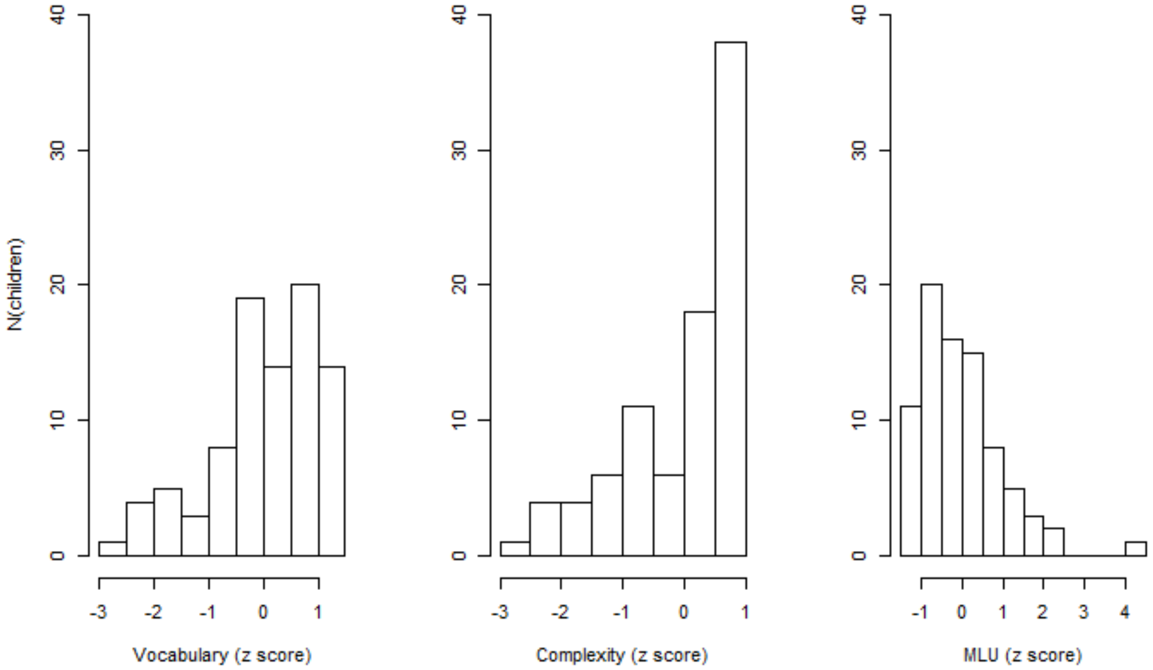
	1	2	3
1. Vocabulary	-		
2. Complexity	.56***	-	
3. MLU	.42***	.49***	-

*** Correlation is significant at the .001 level (2-tailed)

Figure C1 presents histograms for each of the three variables. All have skewed distributions, with left skew in the vocabulary and grammatical complexity measures and MLU being right skewed. Given their intercorrelation and skewed distributions, we decided to run a principal components analysis to check whether the three variables could be reduced to a single measure of language proficiency.

Figure C1

Histograms for each predictive variable



References for Chapter 2

Allen, S. E. M., & Crago, M. B. (1996). Early passive acquisition in Inuktitut*. *Journal of Child Language*, *23*(1), 129–155. https://doi.org/10.1017/S0305000900010126

Ambridge, B. (2019). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, *40*(5–6), 509–559. https://doi.org/10.1177/0142723719869731

Ambridge, B., & Lieven, E. (2015). A constructivist account of child language acquisition. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 478–510). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118346136.ch22

Ambridge, B., & Lieven, E. V. M. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press. https://doi.org/10.1017/CBO9780511975073

Bannard, C., Lieven, E., & Tomasello, M. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences*, *106*(41), 17284–17289. https://doi.org/10.1073/pnas.0905638106

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, *59*(1), 97–113. https://doi.org/10.1016/j.jml.2007.12.007

Bernolet, S., Collina, S., & Hartsuiker, R. J. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, *91*, 99–116. https://doi.org/10.1016/j.jml.2016.01.002

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, *31*(2), 163–186. https://doi.org/10.1016/0010-0277(89)90022-X

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning. *Journal of Experimental Psychology: General*, 177–192. https://doi.org/10.1037/0096-3445.129.2.177

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*(1), 1–39. https://doi.org/10.1016/0010-0277(90)90035-I

Borer, H., & Wexler, K. (1987). The maturation of syntax. In *Parameter setting* (pp. 123–172). Springer.

Branigan, H. P., & McLean, J. F. (2016). What children learn from adults' utterances: An ephemeral lexical boost and persistent syntactic priming in adult-child dialogue. *Journal of Memory and Language*, *91*, 141–157. https://doi.org/10.1016/j.jml.2016.02.002

Branigan, H. P., McLean, J. F., & Jones, M. (2005). A blue cat or a cat that is blue? - Evidence for abstract syntax in young children's noun phrases. In A. Brugos & M. R. ClarkCotton (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development, Vols 1 and 2* (pp. 109–121). Cascadilla Press.

Branigan, H. P., & Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. *Cognition*, *157*, 250–256. https://doi.org/10.1016/j.cognition.2016.09.004

Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, *40*, Article e313. https://doi.org/10.1017/S0140525X16002028

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, *26*(5), 609–651. https://doi.org/10.1207/s15516709cog2605_3

Chang, F. (2009). Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language*, *61*(3), 374–397. https://doi.org/10.1016/j.jml.2009.07.006

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–230. https://doi.org/10.1023/A:1005101313330

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass*, *6*(5), 259–278. https://doi.org/10.1002/lnc3.337

Crain, S., Koring, L., & Thornton, R. (2017). Language acquisition from a biolinguistic perspective. *Neuroscience & Biobehavioral Reviews*, *81*, 120–149. https://doi.org/10.1016/j.neubiorev.2016.09.004

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Mit Press.

Dell, G. S., & Chang, F. (2014). The p-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634). https://doi.org/10.1098/rstb.2012.0394

Demuth, K. (1989). Maturation and the Acquisition of the Sesotho Passive. *Language*, *65*(1), 56–80. https://doi.org/10.2307/414842

Donnelly, S., & Kidd, E. (2021). The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. *Child Development*, *92*(2), 609–625. https://doi.org/10.1111/cdev.13511

Eager, C., & Roy, J. (2017). *Mixed effects models are sometimes terrible*. https://arxiv.org/abs/1701.04858v1

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*(4), 547–582. https://doi.org/10.1111/j.1551-6709.2009.01023.x

Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., ... & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive psychology*, *110*, 30-69. https://doi.org/ 10.1016/j.cogpsych.2019.02.001

Fenson, L., Bates, E., Dale, P., Marchman, V. A., Reznick, J. S., & Thal, D. (2007). *MacArthur-Bates communicative development inventories.* (2nd ed.). Brookes.

Fine, A. B., & Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*(3), 578–591. https://doi.org/10.1111/cogs.12022

Finn, A. S., Kalra, P. B., Goetz, C., Leonard, J. A., Sheridan, M. A., & Gabrieli, J. D. E. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology*, *142*, 212–220. https://doi.org/10.1016/j.jecp.2015.09.027

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition*, *82*(3), 259–278.

Fisher, C., Jin, K., & Scott, R. M. (2020). The developmental origins of syntactic bootstrapping. *Topics in Cognitive Science*, *12*(1), 48–77. https://doi.org/10.1111/tops.12447

Fitz, H., & Chang, F. (2017). Meaningful questions: The acquisition of auxiliary inversion in a connectionist model of sentence production. *Cognition*, *166*, 225–250. https://doi.org/10.1016/j.cognition.2017.05.008

Fitz, H., & Chang, F. (2019). Language ERPs reflect learning through prediction error propagation. *Cognitive Psychology*, *111*, 15–52. https://doi.org/10.1016/j.cogpsych.2019.03.002

Foltz, A., Thiele, K., Kahsnitz, D., & Stenneken, P. (2015). Children's syntactic-priming magnitude: Lexical factors and participant characteristics. *Journal of Child Language*, *42*(4), 932–945. https://doi.org/10.1017/S0305000914000488

Gámez, P. B., Shimpi, P. M., Waterfall, H. R., & Huttenlocher, J. (2009). Priming a perspective in Spanish monolingual children: The use of syntactic alternatives. *Journal of Child Language*, *36*(2), 269–290. https://doi.org/10.1017/S0305000908008945

Gámez, P. B., & Vasilyeva, M. (2015). Increasing second language learners' production and comprehension of developmentally advanced syntactic forms. *Language Learning and Development*, *11*(2), 128–151. https://doi.org/10.1080/15475441.2014.894873

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, *17*(8), 684–691. https://doi.org/10.1111/j.1467-9280.2006.01767.x

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. (2005). *Constructions at work: The nature of generalization in language*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.001.0001

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning*. *Bilingualism: Language and Cognition*, *20*(2), 219–234. https://doi.org/10.1017/S1366728915000164

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238. https://doi.org/10.1016/j.jml.2007.07.003

Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic facilitation in agrammatic sentence production. *Brain and Language*, *62*(2), 221–254. https://doi.org/10.1006/brln.1997.1905

Hsu, D.-B. (2019). Children's adaption to input change using an abstract syntactic representation: Evidence from structural priming in Mandarin-speaking preschoolers. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02186

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*(2), 182–195. https://doi.org/10.1016/j.jml.2003.09.003

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*(3), 381–388. https://doi.org/10.3758/s13421-010-0042-3

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, *33*(2), 393–418. https://doi.org/10.1017/S0142716411000415

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193. https://doi.org/10.1111/cdev.12461

Kidd, E., & Bavin, E. L. (2007). Lexical and referential influences on on-line spoken language comprehension: A comparison of adults and primary-school-age children. *First Language*, *27*(1), 29–52. https://doi.org/10.1177/0142723707067437

Kidd, E., & Donnelly, S. (2020). Individual Differences in First Language Acquisition. *Annual Review of Linguistics*, *6*(1), 319–340. https://doi.org/10.1146/annurev-linguistics-011619-030326

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kidd, E., Junge, C., Spokes, T., Morrison, L., & Cutler, A. (2018). Individual differences in infant speech segmentation: Achieving the lexical shift. *Infancy*, *23*(6), 770–794. https://doi.org/10.1111/infa.12256

Kline, M., & Demuth, K. (2010). Factors facilitating implicit learning: The case of the sesotho passive. *Language Acquisition*, *17*(4), 220–234. https://doi.org/10.1080/10489223.2010.509268

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition*, *87*(3), 151–178. https://doi.org/10.1016/S0010-0277(02)00230-5

Lum, J., Kidd, E., Davis, S., & Conti-Ramsden, G. (2010). Longitudinal study of declarative and procedural memory in primary school-aged children. *Australian Journal of Psychology*, *62*(3), 139–148. https://doi.org/10.1080/00049530903150547

MacWhinney, B. (1977). Starting points. *Language*, *53*(1), 152–168. https://doi.org/10.2307/413059

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Marchman, V. A., Bates, E., Burkardt, A., & Good, A. B. (1991). Functional constraints of the acquisition of the passive: Toward a model of the competence to perform. *First Language*, *11*(31), 65–92. https://doi.org/10.1177/014272379101103104

McCauley, S. M., & Christiansen, M. H. (2019). Language learning as language use: A cross-linguistic model of child language development. *Psychological Review*, *126*(1), 1–51. https://doi.org/10.1037/rev0000126

Messenger, K. (2021). The persistence of priming: Exploring long-lasting syntactic priming effects in children and adults. *Cognitive Science*, *45*(6), e13005. https://doi.org/10.1111/cogs.13005

Messenger, K., Branigan, H., Buckle, L., & Lindsay, L. (2022). How does syntactic priming experience support language development? In K. Messenger (Ed.), *Syntactic priming* *in language development: Populations, representations and mechanisms.* John Benjamins.

Messenger, K., Branigan, H. P., & McLean, J. F. (2011). Evidence for (shared) abstract structure underlying children’s short and full passives. *Cognition*, *121*(2), 268–274. https://doi.org/10.1016/j.cognition.2011.07.003

Messenger, K., Branigan, H. P., & McLean, J. F. (2012). Is children’s acquisition of the passive a staged process? Evidence from six- and nine-year-olds’ production of passives. *Journal of Child Language*, *39*(5), 991–1016. https://doi.org/10.1017/S0305000911000377

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children’s passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, *66*(4), 568–587. https://doi.org/10.1016/j.jml.2012.03.008

Messenger, K., & Fisher, C. (2018). Mistakes weren’t made: Three-year-olds’ comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition*, *178*, 118–132. https://doi.org/10.1016/j.cognition.2018.05.002

Nguyen, E., & Pearl, L. (2021). The link between lexical semantic features and children’s comprehension of English verbal be-passives. *Language Acquisition*, *28*(4), 433–450. https://doi.org/10.1080/10489223.2021.1939354

Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition*, *183*, 152–180. https://doi.org/10.1016/j.cognition.2018.10.026

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633–651.

Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). *Comparing different models of the development of the English verb category*. *36*(4), 807–830. https://doi.org/10.1515/ling.1998.36.4.807

Pye, C., & Quixtan Poz, P. (1988). Precocious passives (and antipassives) in K’iche’ Mayan. *Proceedings of the Annual Child Language Research Forum*, *27*.

R Core Team. (2014). *R: A language and environment for statistical computing*. http://www.R-project.org/

Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., Eadie, P., Barrett, Y., & Ukoumunne, O. C. (2007). Predicting language at 2 years of age: A prospective community study. *Pediatrics*, *120*(6), e1441–e1449. https://doi.org/10.1542/peds.2007-0045

Reitter, D., Keller, F., & Moore, J. D. (2011). A Computational Cognitive Model of Syntactic Priming. *Cognitive Science*, *35*(4), 587–637. https://doi.org/10.1111/j.1551-6709.2010.01165.x

Rowland, C. F. (2007). Explaining errors in children’s questions. *Cognition*, *104*(1), 106–134. https://doi.org/10.1016/j.cognition.2006.05.011

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63. https://doi.org/10.1016/j.cognition.2012.06.008

Sauppe, S., & Flecken, M. (2021). Speaking for seeing: Sentence structure guides visual event apprehension. *Cognition*, *206*, Article 104516. https://doi.org/10.1016/j.cognition.2020.104516

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children’s linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557–567. https://doi.org/10.1111/1467-7687.00312

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3- and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, *43*(6), 1334–1346. https://doi.org/10.1037/0012-1649.43.6.1334

Snider, N. (2009). Similarity and structural priming. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *33*.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, *28*(1), 127–152. https://doi.org/10.1017/S0305000900004608

Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*, *58*(2), 188–213. https://doi.org/10.1016/j.jml.2007.06.012

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge University Press.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of-language acquisition*. Harvard University Press.

Valian, V. (2014). Arguing about innateness. *Journal of Child Language*, *41*(S1), 78–92. https://doi.org/10.1017/S0305000914000336

Vasilyeva, M., Huttenlocher, J., & Waterfall, H. (2006). Effects of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology*, *42*(1), 164–174. https://doi.org/10.1037/0012-1649.42.1.164

Vasilyeva, M., & Waterfall, H. (2012). Beyond syntactic priming: Evidence for activation of alternative syntactic structures. *Journal of Child Language*, *39*(2), 258–283. https://doi.org/10.1017/S0305000911000055

Whitehurst, G. J., Ironsmith, M., & others. (1974). Selective imitation of the passive construction through modeling. *Journal of Experimental Child Psychology*, *17*(2), 288–302.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in Contrast*, *6*(1), 109–149. https://doi.org/10.1075/lic.6.1.05xia

Yang, C. (2018). A formalist perspective on language acquisition. *Linguistic Approaches to Bilingualism*, *8*(6), 665–706. https://doi.org/10.1075/lab.18014.yan

Zhang, C., Bernolet, S., & Hartsuiker, R. J. (2020). The role of explicit memory in syntactic persistence: Effects of lexical cueing and load on sentence memory and sentence production. *PLOS ONE*, *15*(11), Article e0240909. https://doi.org/10.1371/journal.pone.0240909

Chapter 3:

A Meta-Analysis of Syntactic Priming Experiments in Children

Chapter Overview

Chapter 3 is the second chapter in Part 1 of this thesis, which focuses on syntactic acquisition. Chapter 2 aimed to address limitations of cross-sectional syntactic priming studies, which have produced conflicting evidence regarding the emergence and trajectory of priming effects, by using a longitudinal design. Another approach to resolving conflicting findings is through quantitatively summarising them via meta-analysis. Therefore, Chapter 3 presents a meta-analysis of syntactic priming studies in children. The first aim was to corroborate the findings of Chapter 2 regarding abstract priming and lexical boost effects by including lexical overlap, age and their interaction as moderators in the meta-analysis. Secondly, Chapter 3 aimed to synthesise the breadth of syntactic priming research in children to identify other summary findings that reveal insights into the acquisition of syntax. This addressed the aims of the thesis by providing summary evidence to distinguish between theories of syntactic acquisition and by demonstrating the value of meta-analysis: synthesising across studies achieves greater power and can reveal findings that aren't apparent at the level of individual studies. In support of Chapter 2's findings, lexical overlap increased the magnitude of syntactic priming but was not required to observe it. However, the results revealed no evidence for developmental change in abstract priming and the lexical boost. The finding that within- vs between-subjects design had a large impact on the magnitude of the effect implied that explicit processes are also involved in syntactic priming.

Publication Status

The following manuscript has been published as follows:

Kumarage, S., Donnelly, S., & Kidd, E. (2024). A meta-analysis of syntactic priming experiments in children. Journal of Memory and Language, 138, Article 104532. https://doi.org/10.1016/j.jml.2024.104532

Author Contributions

- **Kumarage** and Kidd developed the research question.
- **Kumarage** conducted the literature search.
- **Kumarage** extracted and coded data.
- **Kumarage** performed the meta-analysis and publication bias analysis with advice from Donnelly.
- **Kumarage** produced the figures for the meta-analysis and publication bias analysis.
- Donnelly performed the power analysis with the assistance of **Kumarage**.
- Donnelly produced the figures for the power analysis.
- **Kumarage** drafted the manuscript with editing provided by Donnelly and Kidd.
- **Kumarage** produced the reproducible analysis code for the meta-analysis and publication bias analysis.
- Donnelly produced the reproducible analysis code for the power analysis.
- **Kumarage** drafted responses to reviewers with advice and editing from Donnelly and Kidd.

A Meta-Analysis of Syntactic Priming Experiments in Children

Shanthi Kumarage[1,2] (https://orcid.org/0009-0006-3578-2737)
Seamus Donnelly[2] (https://orcid.org/ 0000-0003-3019-6633)
Evan Kidd[1,3] (https://orcid.org/0000-0003-4301-2290)

[1]Language Development Department, Max Planck Institute of Psycholinguistics, Nijmegen
[2]School of Medicine and Psychology, The Australian National University
[3]School of Languages, Literature and Linguistics, The Australian National University

Addresses for correspondence:
Shanthi Kumarage
School of Medicine and Psychology
The Australian National University
Canberra 2601
AUSTRALIA
Email: shanthi.kumarage@anu.edu.au

Author Note: This research was supported by the Max Planck Society and the Australian Research Council (CE140100041: CI Kidd).

Abstract

A substantial literature exists using the syntactic priming methodology with children to test hypotheses regarding the acquisition of syntax, under the assumption that priming effects reveal both the presence of syntactic knowledge and the underlying nature of learning mechanisms supporting the acquisition of grammar. Here we present the first meta-analysis of syntactic priming studies in children. We identified 37 eligible studies and extracted 108 effect sizes corresponding to 76 samples of 2,378 unique participants. Our analysis confirmed a medium-to-large syntactic priming effect. The overall estimate of the priming effect was a log odds ratio of 1.44 (Cohen's d = 0.80). This is equivalent to a structure that occurs 50% of the time when unprimed occurring 81% of the time when primed. Several variables moderated the magnitude of priming in children, including (i) within- or between-subjects design, (ii) lexical overlap, (iii) structural alternation investigated and, (iv) the animacy configuration of syntactic arguments. There was little evidence of publication bias in the size of the main priming effect, however, power analyses showed that, while studies typically have enough power to identify the basic priming effect, they are typically underpowered when their focus is on moderators of priming. The results provide a foundation for future research, suggesting several avenues of enquiry.

Keywords: syntactic priming, meta-analysis, language acquisition, syntax acquisition

A Meta-Analysis of Syntactic Priming Experiments in Children

Grammar is a core component of language, and thus its acquisition has long been of interest in language acquisition research, where debate about the representational nature and innateness of syntactic representations has been a key battleground. One important method in the toolkit of developmental psycholinguists is *syntactic priming*, which describes the tendency for an individual to use the same grammatical structure they have previously used or heard. For instance, if a child hears the English passive sentence *Ernie was hugged by Bert* and soon after produces another passive such as *Big Bird was surprised by Mr Snuffaluffagus*, they are said to be primed. Far from being mere imitation, priming in this instance reveals the abstract nature of grammatical representations, since the two sentences do not share overlap in their open-class lexical items (Branigan & Pickering, 2017). Furthermore, priming has been argued to be a form of implicit learning (Chang et al., 2000, 2006; Dell & Chang, 2014), and is thus potentially reflective of a *mechanism* underlying language acquisition.

Given its ability to reveal both representational knowledge and learning, the syntactic priming methodology has become a prominent one in the field (see papers in Messenger, 2022). In the current paper we present the first systematic review and meta-analysis of experimental papers that have investigated syntactic priming in developmental populations. In what follows we review the past research, highlighting the trends in the literature that we then quantitatively examine.

Past Developmental Syntactic Priming Research

One key question in developmental priming research is the abstractness of syntactic representations, tested by manipulating overlap in lexical content between prime and target trials (Branigan et al., 2005; Branigan & McLean, 2016; Foltz et al., 2015; Kumarage et al., 2022; Peter et al., 2015; Rowland et al., 2012; Savage et al., 2003). Priming in the absence of lexical overlap indicates the existence of abstract syntactic representations, whereas priming restricted to trials where prime and target share lexical content, such as a verb (i.e., *Ernie was hugged by Bert* primed *Bird Bird was hugged by Mr Snuffaluffagus*), indicates syntactic knowledge may be more item-based, therefore bearing upon a hotly debated issue in the field (e.g., Ambridge & Lieven, 2015; Fisher, 2002; Tomasello, 2000; Tomasello & Abbot-Smith, 2002). Thus, a focus on syntactic priming in development inevitably leads to the question of *when* abstract priming effects emerge, which in the literature has informed arguments over the early or late emergence of core grammatical knowledge (e.g., Bencini & Valian, 2008; Huttenlocher et al., 2004).

The distinction between abstract priming and lexically-based priming also bears upon a broader issue concerning the architecture of the language production system. In the mature adult system, where abstract knowledge can be assumed, priming effects are typically larger when the prime and target sentences share open-class lexical content. This is known as the *lexical boost*. Activation-based production architectures attribute both the abstract priming and lexical boost effects to a single mechanism driven by the residual activation of connected lexical and structural knowledge (Pickering & Branigan, 1998), and therefore predict a lexical boost to emerge in development once abstract knowledge has been acquired. Priming as residual activation also predicts that priming is a relatively fleeting event that may not lead to representational change in the linguistic system. In contrast, implicit learning accounts of priming propose that abstract priming and the lexical boost are attributable to separable implicit and explicit memory systems, respectively (Chang et al., 2006, 2012). Since explicit memory processes exhibit a relatively protracted developmental trajectory, the lexical boost is predicted to emerge after abstract priming (Chang et al., 2012; Rowland et al., 2012). Notably, implicit learning accounts also predict that abstract priming leads to representational change via the same mechanism involved in language development. The most prominent of these accounts invokes *error-based learning* (Chang et al., 2006). Specifically, during comprehension, the syntactic processor predicts upcoming input based on syntactic representations that are sensitive to frequency distributions in the input. It compares its predictions to the actual input and responds to prediction errors by updating the weights of its syntactic representations. In syntactic priming experiments, unexpected syntactic structures (e.g., low frequency structures such as the English passive) encountered in primes result in an updating of syntactic weights that increases the likelihood of their later production, thus leading to priming. In experimental terms, the implicit learning account of priming predicts priming effects to have long-term resonance (see Bock & Griffin, 2000). Several studies in the developmental literature have investigated the long-term nature of syntactic priming (Branigan & McLean, 2016; Fazekas et al., 2020; Hsu, 2019; Kidd, 2012; Messenger, 2021), with at least one study showing that it can persist for up to a month (Savage et al., 2006).

An analysis of syntactic priming across development also bears upon the question of whether and how syntactic representations change across developmental time. As a theory of syntactic acquisition, Chang et al.'s (2006) computational Dual-path model, which learns syntactic representations via implicit learning, differs from traditional nativist (Bencini & Valian, 2008; Messenger & Fisher, 2018) and traditional lexicalist (Savage et al., 2003; Tomasello, 2003) accounts of syntax acquisition. Research has compared these theories by examining the developmental trajectories of the abstract priming effect and lexical boost in both cross-sectional (abstract priming: Garcia & Kidd, 2020; Hsu, 2019; both effects: Peter et al., 2015; Rowland et al., 2012) and longitudinal designs (both effects: Kumarage et al., 2022). Notably, whereas nativist accounts predict no appreciable change in priming magnitude across development because representations are largely unaffected by frequency, lexicalist approaches predict an increase because initially lexically-specific representations become more abstract and thus more primeable with experience. Different still, priming as implicit learning predicts that priming will initially increase as children acquire the necessary representations to be primed and then decrease across development, since error-based learning is strongest when representations are weak (for more discussion see Kumarage et al., 2022).

Researchers have manipulated several other prominent variables in the syntactic priming literature. For example, manipulating whether children must repeat the prime sentence before they describe a target aims to investigate whether their syntactic representations are shared across production and comprehension (Gámez & Shimpi, 2016; Huttenlocher et al., 2004; Shimpi et al., 2007). Other studies have tested how children's syntactic representations interface with semantic information by manipulating the animacy or thematic roles of nouns in prime and target sentences (Bidgood et al., 2021; Buckle et al., 2017; Messenger, Branigan, McLean, et al., 2012; Vasilyeva & Gámez, 2015). The past literature is not limited to priming in typically-developing monolingual children. Research in clinical developmental populations has investigated whether syntactic priming is observed in these populations (e.g., children with Developmental Language Disorder (DLD), Leonard et al., 2000; children who stutter, Anderson & Conture, 2004) and whether it is associated with defining clinical features. For example, research on children with Autism Spectrum Disorder (ASD) has investigated if priming is used as a mechanism of conversational alignment and whether it differs from children without ASD (Allen et al., 2011; Hopkins et al., 2016). Other research on children with DLD has investigated whether priming, as a marker of implicit learning, is compromised in comparison to children without DLD (Garraffa et al., 2015, 2018). Crosslinguistic priming studies have investigated whether children share syntactic representations between languages (Gámez & Vasilyeva, 2020; Vasilyeva et al., 2010; Wolleb et al., 2018). For instance, does *Ernie was hugged by Bert* prime *Big Bird è stato sorpresso dal Signor Snuffaluffagus* in Italian-English bilinguals?

In summary, 20 years of research on syntactic priming in children has investigated a range of theoretical questions regarding the representation and acquisition of syntax. A past meta-analysis of syntactic priming in adults aggregated evidence to provide key summary findings: abstract syntactic priming decays slowly, lexical overlap provides a large boost to the effect, and this boost decays quickly but is more pronounced in a speaker's second language (Mahowald et al., 2016). The substantial literature in children now warrants a similar quantitative summary, which we present in this paper.

The Current Study

This study reports (i) a summary effect of syntactic priming in typically-developing developmental populations, (ii) sources of variation in the effect, (iii) an analysis of publication bias, and (iv) a power analysis. At its simplest, meta-analysis involves combining the results from many studies to produce a summary effect (Borenstein et al., 2009). In this case, studies reliably find evidence for syntactic priming in children, so we expect to find an overall effect. We report the magnitude of the summary effect computed in the meta-analysis. Additionally, we were interested in the sources of variation in the magnitude of syntactic priming. More advanced meta-analytic techniques allow the investigation of moderator variables, although researchers must be cautious not to ask more of the data than can be obtained from them (Viechtbauer, 2008). We identified which moderators can be reasonably investigated based on the studies available, including both experimentally manipulated variables and researchers' methodological choices, and investigated their influence on syntactic priming in children. We also investigated whether there is evidence for publication bias; that is, the inflation of the estimated effect due to unpublished null results missing from the sample of studies (Borenstein et al., 2009). Finally, we ran a power analysis to estimate the number of participants and items required to detect both the main priming effect and interaction effects, given the field's focus on moderators of syntactic priming. In the process, we summarise the state of the field: which questions are being investigated and how? Does how we run studies have an impact? And what do we have evidence for and where it is lacking?

Meta-Analysis

Data Availability

Our data sheet of coded and extracted information and analysis scripts are available on the Open Science Framework (OSF; https://osf.io/k6z8g/).

Literature Search

Search Strategies

The literature search was conducted using three strategies designed to identify as many syntactic priming studies on children as possible. First, a database search was conducted of PsycInfo, Scopus, and Web of Science using search terms designed to maximise the reach of the search. The record needed to contain *synta** *gramma** or *structur** within three words of *priming*, *alignment* or *persistence* and contain *child** or *develop** or *infan**. The search was conducted within English language articles (peer-reviewed journals, books, book chapters and conference proceedings) from 1986, the year of Bock's seminal paper, to February 2023. Secondly, we recorded references from the language acquisition sections of relevant reviews (Atkinson, 2022; Branigan & Pickering, 2017; Pickering & Ferreira, 2008). Finally, we searched for forward citations of the first priming studies in children using the forward citation tools on Scopus and Web of Science (Huttenlocher et al., 2004; Savage et al., 2003). The search was first conducted in November 2019 then replicated and updated in February 2023.

Search results from the three database searches were exported as *.ris* files and imported into Zotero (Corporation for Digital Scholarship, 2023), as were the forward citation search results. Backward citations from relevant reviews were manually entered into Zotero. Title and abstract screening was conducted within Zotero by tagging studies as relevant or not. Studies identified as directly related to syntactic priming were then exported and entered into the MetaLab *Decision Spreadsheet* template (Bergmann et al., 2018). Studies were then assessed against the selection criteria within this spreadsheet. The selection criteria are displayed in Table 1 and described below.

Selection Criteria

We focused on the typical development of early syntactic knowledge. Therefore, studies were limited to those testing children under 13 years of age with no history of language or other developmental disorders. Control groups from studies focusing on children with developmental or language disorders were included (e.g., Foltz et al., 2015). We only included studies reporting a syntactic priming experiment as defined by our design criteria. We excluded corpus analyses (3.a.), studies of non-syntactic outcomes or using non-syntactic primes (3.b.), elicitation and baseline production studies (3.c.), novel word or structure studies (3.d.), and training studies (3.e.).

Furthermore, we only included studies of spoken production priming, and thus excluded studies that used written production or primed comprehension (e.g., Thothathiri & Snedeker, 2008; van Beijsterveldt & van Hell, 2009). Comprehension and production processing, and the production processes for written and spoken language are likely to differ, particularly in developmental populations. Therefore, excluding these studies ensures consistency in the processing mechanisms underlying the effects. Similarly, we limited our analyses to within-language priming, and so excluded studies of cross-linguistic priming such as Vasilyeva et al. (2010) and Gámez and Vasilyeva (2020), and the crosslinguistic condition in Wolleb et al. (2018). We made this decision for the following reasons. Practically, there are only a small number of crosslinguistic priming studies in developmental populations, and including them had the potential to cloud any effects we found in the data because crosslinguistic priming inevitably involves other important variables that need to be controlled (e.g., relative proficiency of bilingual children, presence or absence of surface word order across target structures). Theoretically, positing representational overlap across languages raises several questions regarding the cognitive architecture of language in bilinguals (Hartsuiker & Bernolet, 2017; Van Dijk et al., 2022; van Gompel & Arai, 2018), which was beyond the scope of our project.

Finally, we excluded storytelling interventions aimed at increasing children's production of a structure (e.g., Vasilyeva et al., 2006). While these no doubt involve priming, they are difficult to compare to most syntactic priming studies, which investigate the effect of individual primes or a single block of primes on production, whereas interventions investigate sustained exposure to primes over longer time periods (sometimes weeks).

Table 1

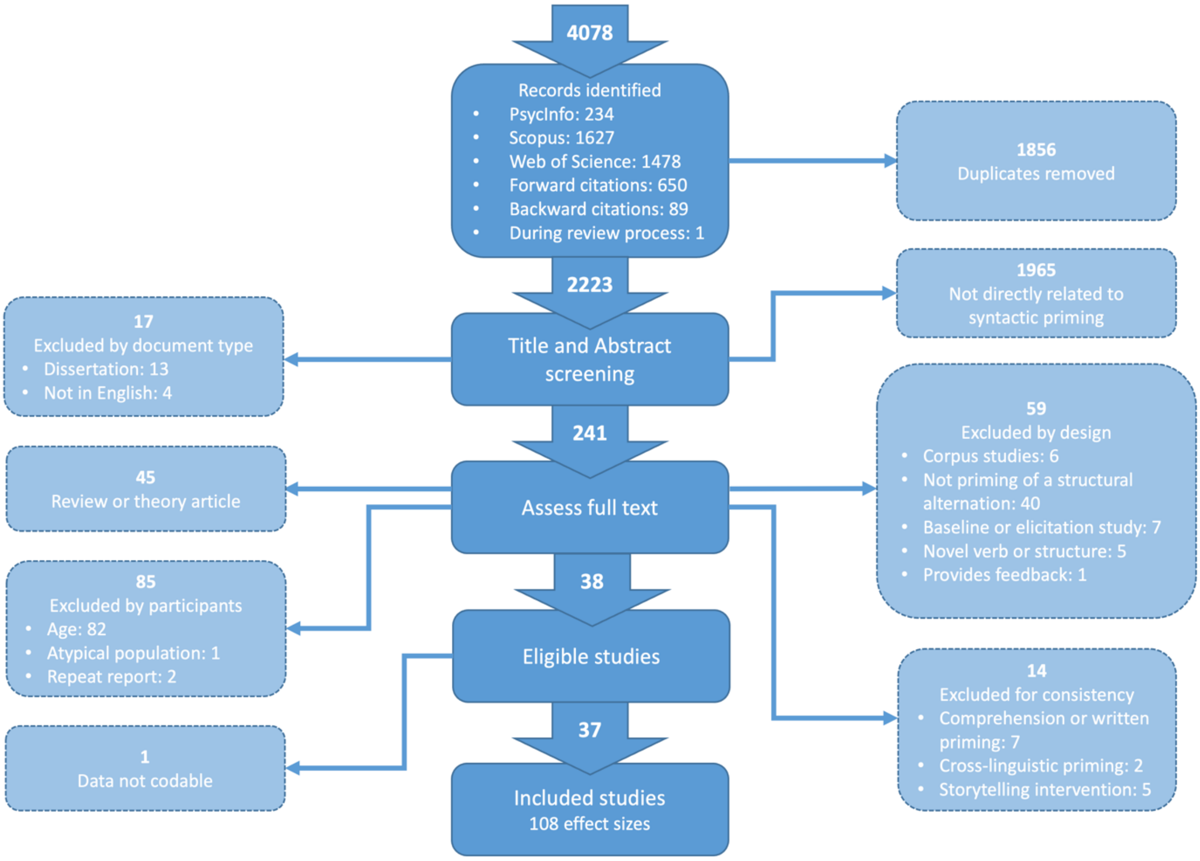
Selection criteria

Category	Criteria
1. Document type	a. Journal papers, book chapters, proceedings papers b. English language
2. Participants	a. Children under 13 years of age b. No history of developmental or language disorders i. Control groups were included c. Sample independent from any other study
3. Design	a. Must be experimental not observational i. No corpus analyses b. Must investigate priming of a structural alternation i. Dependent variable is the choice between two structures within a syntactic alternation 1. No studies investigating the effect of syntactic primes on other outcomes (e.g. lexical decision-making, grammaticality judgements, speech disfluency) 2. No morphosyntactic alternations (e.g., provision of auxiliaries) or non-alternating structures (e.g., intransitive/transitive) ii. Independent variable is the syntactic structure of prime sentences 1. No single word, rhythmic, or arithmetic primes c. Must include a baseline and primed condition i. No elicitation studies with primed condition only ii. Baseline condition could comprise unprimed targets or targets primed with an alternate structure d. Must not use novel words or structures i. Investigating existing syntactic abilities not ability to generalise e. Must not provide feedback on sentence production i. Investigating implicit priming effect not explicit learning f. Outcome variable must be *spoken* production of a sentence/structure g. Primes and targets must be produced in the same language h. Target responses must be measured after no more than one block of primes i. No storytelling interventions

As illustrated in Figure 1, following the application of selection criteria, 38 eligible studies were identified, of which 37 could be analysed. Data from one study was reported such that it could not be coded and the data were not available from the author. From the 37 included studies, 108 effect sizes were extracted. The list of included studies is available in Appendix A.

Figure 1

Flowchart showing literature search



Coding Procedures

Information was extracted separately for each experimental condition, i.e. each observation, within a study (e.g., lexical overlap vs no lexical overlap). For the primed and unprimed conditions of each observation, we extracted: the number of participants and items, the proportion and/or number of dependent (e.g. passive), alternate (e.g. active), and *other* responses. The dependent, or primed, structure was defined as the less frequently produced structure in the alternation (typically coded as 1 in the dependent variable at the study level). Table 2 describes the five structural alternations included in the meta-analysis, including the dependent and alternate structure for each. When frequencies were graphed but not reported, we estimated the number/proportion of dependent, alternate and *other* responses using WebPlotDigitizer (Rohatgi, 2022). In two cases, we calculated the frequencies from raw data available online (Fazekas et al., 2020; Garcia & Kidd, 2020). We contacted the author/s if the data was still not able to be extracted. We also recorded whether studies reported proportions or numbers of only dependent and alternate responses or also reported *other* response frequencies by condition. Note that when studies scored responses under both a lax and strict coding scheme (e.g., Bencini & Valian, 2008; Kumarage et al., 2022), we coded those scored under strict coding.

Table 2

Description of structural alternations included in meta-analysis

Structural alternation	Dependent structure	Alternate structure
Passive	Passive *Ernie was hugged by Bert*	Active *Bert hugged Ernie*
Dative	Double object dative *Elmo gave Big Bird a cat*	Prepositional dative *Elmo gave a cat to Big Bird*
Mandarin SVO/ba	ba *Xiaogou **ba** xiamao baozhu **–le*** *Little dog **ba** little cat hug-tight **PFV*** *A little dog hugged a little cat tightly*	SVO *Xiaogou baozhu **–le** xiaomao* *Little dog hug-tight **PFV** little cat* *A little dog hugged a little cat tightly*
Relative clause	Relative clause *The car that is red*	Adjective-noun *The red car*
Tagalog Symmetrical Voice (SV)-transitive	Patient-initial Agent voice *H<**um**>ahabol **ng** babae **ang** bata* *<**AV**>chase **NSBJ** woman **SBJ** child* *The child is chasing a woman* Patient voice *H<**in**>ahabol **ang** babae **ng** bata* *<**PV**>chase **SBJ** woman **NSBJ** child* *The child is chasing the woman*	Agent-initial Agent voice *H<**um**>ahabol **ang** bata **ng** babae* *<**AV**>chase **SBJ** child **NSBJ** woman* *The child is chasing a woman* Patient voice *H<**in**>ahabol **ng** bata **ang** babae* *<**PV**>chase **NSBJ** child **SBJ** woman* *The child is chasing the woman*

Note: The SVO/ba alternation occurs only in Mandarin and the Symmetrical-voice transitive alternation only in Tagalog. The three other alternations were tested in multiple languages. Example for SVO/ba adapted from Hsu (2014) and example for SV-transitive adapted from Garcia & Kidd (2020). Linguistic glosses: perfective aspect (PFV), agent voice (AV), patient voice (PV), subject (SBJ), non-subject (NSBJ).

We coded each observation for several variables. Some were experimental variables that have been proposed to influence the size of the syntactic priming effect in children and others were methodological variables that differed between the included studies. Table 3 summarises the variables that were included as moderators in our analyses, as well as language studied. It displays the levels of each variable, and how many observations fell into each level. Where observations from a study differed across a moderator that we did not analyse, they were combined into a single observation.

Table 3

Moderator variables: coding, levels and distributions

Variable	Coding	Level	N observations
Within- vs between-subjects design	0	Within-subjects	68
	1	Between-subjects	40
Baseline	0	Alternate prime	94
	1	No prime	14
Animacy	0	Controlled	57
		Animacy not relevant – Relative clause (Included as 0 not NA as otherwise automatically excluded from model)	4
	1	Uncontrolled	24
	2	Favours dependent structure	23
Prime repetition	0	Prime not repeated	76
	1	Prime repeated	32
Lexical overlap	0	No overlap	85
	1	Overlap	23
Structure	0	Passive	63
	1	Dative	22
	2	SVO-ba (Mandarin)	13
	3	Relative clause	4
	4	Symmetrical Voice Transitive (Tagalog)	6
Age		In months as z-score	*M* = 59.35 *SD* = 18.15
Lag	0	Alternating design	74
	1	Blocked design	27
	2	Long blocked design	7
Language		English	77
		German	2
		Italian	4
		Mandarin	13
		Norwegian	2
		Russian	1
		Spanish	3
		Tagalog	6

Moderators

Within- vs Between-Subjects Designs. Unlike in adult studies, developmental researchers have often used between-subjects designs, comparing primed and unprimed participants rather than conditions. Firstly, as Shimpi et al. (2007) state, requiring children to complete trials from only one condition reduces the demands of the experimental task. Secondly, researchers have raised concerns that primes from one condition will influence responses in the other (Gámez & Vasilyeva, 2015). It is common in developmental syntactic priming studies to include few or no filler items. This is another way of reducing task demands but may increase the likelihood of interference across trials. Given interference cannot occur in between-subjects designs, we expect a larger priming effect in these than in within-subjects designs. We used treatment contrasts, with within-subjects designs set as the baseline, or reference, level (0) and between-subjects as the treatment level (1). One final issue to note is that primed and unprimed conditions will be correlated in within- but not between-subjects designs. We explain how we control for this in our statistical methods section.

Baseline. Most studies compare to a baseline of responses following primes in the alternate structure, but some compare to a baseline of responses produced with no prime, typically collected in a preceding phase before priming begins. At present, there is no evidence regarding whether this difference in methodology affects the magnitude of the priming effect. However, Bencini and Valian (2008) reported that children in a no prime condition produced no passives, whilst those in the active priming condition produced two. Therefore, a larger priming effect may be expected when comparing to a no prime baseline. We used treatment contrasts, with the *alternate prime* condition as the reference level (0) and *no prime* as the treatment level (1).

Animacy. The animacy of verb arguments can favour or disprefer the dependent structure. For example, a preference to put humans in sentence-initial positions means the human-patient nonhuman-agent is the canonical passive form (e.g., *the man was bitten by the dog*; see Bock et al., 1992). In the dative, a preference for animate arguments to precede inanimate ones means the canonical double object dative involves an animate recipient and inanimate theme (e.g., *The boy gave the girl a present*). The influence of animacy on priming was demonstrated by Vasilyeva and Gámez (2015), who found that the animacy of arguments moderated abstract passive priming: sentences with an animate patient and inanimate agent were subject to greater priming. Most studies have controlled for animacy by using the same animacy configuration in primes and targets: all animate arguments (e.g., passive: Kumarage et al., 2022; e.g., dative: Rowland et al., 2012), all inanimate arguments (e.g., Savage et al., 2003) or an animacy configuration favouring the dependent structure (e.g., passive: Branigan & McLean, 2016; e.g., dative: Fazekas et al., 2020). Others have used materials with a mix of configurations in both primes and targets (e.g., Huttenlocher et al., 2004). We could not code for particular animacy configurations, given they are specific to particular structures. Instead, we created three broad categories: *controlled animacy* used either all animate or all inanimate arguments, *uncontrolled animacy* used a mixture of configurations, and *favourable animacy* used a configuration favouring the dependent structure. We set *controlled animacy* as the reference level and used treatment contrasts to compare the other two conditions to this baseline. We expect greater priming when animacy favours the dependent structure. It is not clear whether not controlling animacy will reduce or increase priming.

Prime Repetition. We coded whether studies required children to repeat the prime sentence before producing their target sentence using treatment contrasts (*prime not repeated,* 0; *prime repeated,* 1). Huttenlocher et al. (2004) found equivalent priming with and without prime repetition in 4–5-year-olds, concluding that syntactic representation is shared across comprehension and production. Shimpi et al. (2007) found no priming effect in 3-year-olds unless they repeated the prime, whereas 4-year-olds did not need to repeat the prime to demonstrate a priming effect. This suggests prime repetition may facilitate abstract priming, with the effect weakening as children develop. We therefore included an interaction between prime repetition and age in the meta-analysis.

Lexical Overlap. In adults, lexical overlap between prime and target sentences (e.g., shared verb or noun) reliably boosts priming (Mahowald et al., 2016). In children, evidence for the lexical boost is more mixed. Some studies have found a lexical boost at a young age (Branigan et al., 2005; Branigan & McLean, 2016; Savage et al., 2003), whilst others have only found the effect in older children (Kumarage et al., 2022; Rowland et al., 2012) or not at all (Foltz et al., 2015; Peter et al., 2015). The developmental trajectory of the lexical boost has implications for theories of syntactic acquisition. A lexical boost effect that increases over development is in line with the prediction that lexically-based priming reflects developing explicit memory processes rather than implicit learning (Chang et al., 2012;

Rowland et al., 2012), whilst a decreasing effect is in line with a transition from lexically-based to abstract syntactic representation (Savage et al., 2003; Tomasello, 2003). We coded whether or not there was lexical overlap between prime and target using treatment contrasts (*no lexical overlap,* 0; *lexical overlap,* 1) and included an interaction with age.

Structure. We found five structural alternations in the included studies: passive/active, double object dative/prepositional dative, Mandarin SVO/ba, relative clause/adjective-noun and patient-initial/agent-initial in Tagalog's symmetrical voice transitive (see Table 2 for details). Since syntactic priming is stronger for infrequent structures (Ferreira, 2003; Jaeger & Snider, 2013; Kaschak et al., 2011), the strength of the priming effect may vary by structure. We used the passive alternation as the reference level with which to compare other structures because it is the most researched alternation and likely to be well estimated. Given children acquire different syntactic structures at different ages, the age at which children have abstract knowledge and therefore show abstract priming of a structure is likely to vary. We therefore included an interaction between structure and age.

Age. A variety of developmental trajectories for abstract syntactic priming have been proposed. For example, Nativist accounts typically propose stable priming effects over development (Bencini & Valian, 2008; Messenger & Fisher, 2018). Error-based learning in the Chang et al. (2006) model predicts that abstract priming can decrease as children become better at predicting the prime sentences (lower error; Peter et al., 2015). However, the Chang et al. model also implicitly learns its syntactic representations from word sequences, so priming can also increase during an early developmental period as structures become more abstract, since abstract structures increase the transfer of changes on the prime to the target. This is similar to the predictions of lexicalist accounts such as Tomasello (2003). As mentioned above, age may also interact with prime repetition, lexical overlap and structure. We coded the average age in months of the sample for each observation. This value was converted to a z-score for our analyses.

Lag. The observation of priming at long distances between prime and target led to the theory that priming is a form of implicit learning (Bock & Griffin, 2000). We coded three categories of lag between prime and target: alternating design, blocked design, and long blocked design. Studies which alternate between prime and target sentences are the most common. Several studies use a blocked design, where primes are presented as a block before children describe a block of targets. Hsu (2019) found that a blocked design showed a larger priming effect than an alternating design. Relatively few studies in children have investigated the length of time for which priming endures. Only three studies contained conditions with long blocked designs with a delay of an hour (Hsu, 2019), a week (Kidd, 2012; Savage et al., 2006) or a month (Savage et al., 2006) between primes and targets. We excluded the lag 2 condition from Branigan and McLean (2016) and Garraffa et al. (2018), where two sentences intervened between prime and target. To include multiple effect observations from the same sample of participants we required a reported or estimated correlation between observations, which was not available in this case. As with animacy and structure, we dummy coded this variable, with alternating design set as the reference level.

Other Variables

We identified other potential moderator variables that, for several reasons, were not included in our analyses. Firstly, in some cases, there were too few observations of one level of the variable to make a reasonable comparison. For instance, only one study investigated priming in a second language (Gámez & Vasilyeva, 2015), ruling out a comparison between priming in a first and second language. Secondly, some variables would be suitable for a meta-analysis of a particular structure but not across structures. For example, several studies have investigated the influence of verb-type (i.e. agent-patient, experiencer-theme, theme-experiencer; Bidgood et al., 2020; Messenger, Branigan, McLean, et al., 2012) on priming of the active-passive alternation; however, this variable cannot be generalised across the other structural alternations so could not be included. Thirdly, other variables were precluded from inclusion because of confounds. Task type was confounded with structure, with stem completion typically used in studies of the dative alternation and picture description for other structures. Similarly, for priming paradigm, the *bingo* game was used almost exclusively in studies of the dative alternation, whilst studies of other structures used the *snap* game or picture description.[1] We recorded the language in which the study was conducted but could not include it as a moderator or random effect. In the case of Mandarin and Tagalog, language is confounded with structure (SVO/ba, symmetrical voice transitive), preventing its inclusion as a moderator. In addition, most studies were conducted

---

[1] We thank an anonymous reviewer for suggesting the analysis of priming paradigm. Whilst we could not include it as a moderator due to this confound we discuss its potential impacts in the discussion.

in English, with very few observations from other languages, reflecting a common bias in the field (Kidd & Garcia, 2022). A random effects structure is unlikely to be sensibly estimated in this case and so we did not include random effects by language. Finally, we could only include a limited number of moderators (as a rough rule, Borenstein et al. (2009) recommend approximately 10 observations per moderator). We coded but did not analyse the influence of: the number of confederates (as in Mahowald et al., 2016), the number and frequency of filler items, and lax vs strict coding of responses. This information is included in the Data Collection Form accessible on our OSF site.

Statistical Methods

We ran our analyses in R (version 4.3.2; R Core Team, 2022). We used *tidyverse* packages for data manipulation (version 2.0.0; Wickham et al., 2019) and the *metafor* package for meta-analysis functions (version 4.4-0; Viechtbauer, 2010). All data and code is available on our OSF site (https://osf.io/k6z8g/).

Effect Size

The first step in meta-analysis is to compute a common effect size across studies (Borenstein et al., 2009). We chose to compute a log odds ratio (LOR) as the effect size for each observation based on raw cell count data. The LOR is appropriate for binary outcome data, such as in syntactic priming studies, where participants respond with either the dependent or alternate structure (e.g., passive or active: Jaeger, 2008). In addition, studies reported a mixture of ANOVAs and logistic mixed models in their results section. These are difficult to combine due to the different statistics than can be extracted from them. Lastly, LORs taken directly from reported mixed models would not be equivalent if those models used different fixed effects or random effects structures. Calculating a LOR from cell count data, which is routinely reported in syntactic priming studies, combats these issues.

We calculated the LOR using the *escalc* function in the *metafor* package. This takes the number of responses in each condition (as in Table 4) and calculates the LOR using Equation 2[2]. This calculation cannot handle 0 scores in Table 4 so values of *a, b, c* and *d* were adjusted beforehand using Smithson and Verkuilen's (2006) approach. The proportion of

---

[2] We excluded *other* responses from our calculation of the LOR. However, Appendix C compares results including and excluding *other* responses for a subset of studies that reported the frequencies of *other* responses.

dependent responses in each condition was adjusted using Equation 1, where *N* is the number of participants in that condition. This proportion was then multiplied by *N* to calculate *a* and *c*, and (1 – P_{adjusted}) was multiplied by *N* to calculate *b* and *d*.

The *escalc* function also calculates the associated variance, or sampling variance, for each LOR effect size using Equation 3. We adjust this *escalc* calculated sampling variance for within-subjects designs and for two definitions of sample size. We next explain these issues in accounting for sampling variance, as well as our approach to accounting for two other forms of variance.

Table 4

Example cell count data

	Dependent response	Alternate response
Primed	*a*	*b*
Unprimed	*c*	*d*

$$P_{adjusted} = \frac{P(N-1) + 0.5}{N} \tag{1}$$

$$LOR = \ln\left(\frac{ad}{bc}\right) \tag{2}$$

$$V_{LOR} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \tag{3}$$

Variance: Accounting for Complex Data Structure Using Multilevel Meta-Analysis

The nature of the developmental syntactic priming literature leads to three sources of variance in effect sizes: sampling variance, between-study variance, and within-study variance. We conducted a multilevel meta-analysis to account for all three sources of variance, as detailed below.

Sampling Variance. The *escalc* function automatically calculates the LOR and its associated sampling variance, or sampling error, as per Equation 3. Sampling variance is used to weight studies, or more accurately, observations, by their precision, giving more weight to more precise observations (Borenstein et al., 2009). It is mostly influenced by sample size and reflects sampling error in the estimation of an effect (Borenstein et al., 2009).

Multiple Responses per Participant. Log odds ratios have typically been used in the meta-analysis of clinical trials where each outcome comes from a separate participant (e.g., cured vs not cured). In syntactic priming studies, participants provide multiple responses per condition. Calculating sampling variance using $N_{responses}$ ignores the dependence between responses from the same participant. However, calculating sampling variance using $N_{participants}$ dramatically reduces the power of the analysis and does not account for the true number of responses. In line with Mahowald et al. (2016), we calculated sampling variance using both $N_{responses}$ and $N_{participants}$. To do this, we followed the procedure described in the *Effect Size* section twice, first multiplying $P_{adjusted}$ by $N_{responses}$ to calculate the adjusted values of *a, b, c* and *d,* and second using $N_{participants}$. Our results section reports results from analyses using both forms of variance.

Within- vs Between-Subjects Designs. The use of a within-subjects design also affects precision: the *primed* and *unprimed* conditions are likely to be correlated due to individual participants' tendency to produce the dependent structure. In other words, comparing participants to themselves reduces error and increases precision. Sampling variance should be calculated accordingly (Morris & DeShon, 2002).

Therefore, we adjusted our sampling variance estimates for within-subjects studies using the Becker-Balagtas method (Becker & Balagtas, 1993) as described in Stedman et al. (2011). Equation 4 can be used instead of Equation 3 to calculate sampling variance if we know *p*, the correlation between production of the dependent structure in the primed and unprimed conditions (Stedman et al., 2011). The correlation, *p*, is used to calculate *s* (Equation 6), and *s* to calculate Δ (Equation 5). Then Δ and *n*, the total number of responses, can be used to calculate an adjustment to the sampling variance (Equation 4).

$$V_{LOR} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} - \frac{\Delta}{2n} \tag{4}$$

$$\Delta = n^2 \left(\frac{ns - ab}{abcd} \right) \tag{5}$$

$$s = \frac{p\sqrt{abcd} + ab}{n} \tag{6}$$

The correlation, *p*, between production of the dependent structure in the primed and unprimed conditions is not reported in syntactic priming studies. However, raw data available to the authors from the *Canberra Longitudinal Child Language* study (Donnelly et al., 2024; Kidd et al., 2018; Kumarage et al., 2022) allowed us to estimate it at $r = .35$. Appendix B details how correlations between conditions were estimated. Using $p = .35$, the adjusted sampling variance was calculated for studies with within-subjects designs. All analyses (both $N_{\text{responses}}$ and $N_{\text{participants}}$) use sampling variance corrected for within-subjects designs.

Between-Study Variance. Traditional meta-analysis assumes that each study estimates the same true effect size with differences between studies due only to sampling error (Borenstein et al., 2009). However, effects are likely to vary between studies due to factors outside of control or moderators not included in analysis (Borenstein et al., 2009). Including a random effects structure allows us to instead assume that there is distribution of true effects, which differ due to sampling error and heterogeneity in effect sizes. We therefore included random intercepts by observation nested within study. This adds the assumption that effect sizes vary both within- and between-studies due to heterogeneity, not just sampling error.

Within-Study Variance. This meta-analysis includes 108 observations from 37 studies. Ignoring the dependence between effect sizes from the same study is a common mistake in meta-analyses (e.g., 9 of the 20 most highly cited meta-analyses in the exercise science field of strength and conditioning did not account for this dependence; Kadlec et al., 2022). Treating correlated effect sizes as independent observations is problematic because it assumes they contribute independent information, which can inflate the strength of evidence for an effect and assigns more weight to studies with multiple observations (Borenstein et al., 2009). A multilevel meta-analysis allows the inclusion of multiple dependent effect sizes within a single analysis. Several types of dependence between observations were present in our data set and we describe how we accounted for each below.

Independent Groups. Multiple observations from independent samples within the same study result from manipulating moderators between-subjects. For example, comparing different age groups, or assigning children to either the prime repetition or no repetition condition. Despite having independent samples, effects from the same study are likely to be subject to the same uncontrolled variables or moderators. Specifying random effects that are nested by study allows us to account for the possibility that independent effects from the same study may be more similar than independent effects from different studies.

Shared Comparison Groups. When two or more experimental conditions are compared to the same baseline, the observations in those conditions will be correlated. We specified which group of participants within a study was primed and unprimed in each observation using the *grp1* and *grp2* arguments in *metafor'*s *vcalc* function, and the number of participants or responses in each group using *w1* and *w2*. The *vcalc* function uses this information to calculate the correlation between observations and the variance-covariance matrix used in multilevel meta-analysis (Viechtbauer, 2010).

Multiple Outcomes. Several studies manipulated lexical overlap or target structure (passive vs dative) within-subjects, leading to multiple dependent outcomes. Treating these observations as independent assumes the correlation between them is 0, despite them coming from the same sample of participants. Another common approach is to use an average effect size and variance. However, this implicitly assumes that the correlation between the effects is 1, or that they do not contribute any extra information as separate observations (for detailed discussion see Borenstein et al., 2009). It also prevents the analysis of moderators in studies where they are manipulated within-subjects, and therefore more precisely estimated. Borenstein et al. (2009) recommend using a plausible correlation over assuming an extreme of either 0 or 1.

We estimated correlations between priming with and without lexical overlap, and between passive and dative priming (see Table 5). Appendix B details how these were calculated from priming data from the *Canberra Longitudinal Child Language* project (Donnelly et al., 2024; Kidd et al., 2018; Kumarage et al., 2022). Using these estimates, *vcalc* was used to calculate a variance-covariance matrix for dependent observations, which was specified in our meta-analysis.[3]

---

[3] Two studies manipulated animacy within-subjects (Buckle et al., 2017; Vasilyeva & Gámez, 2015). Since we had no estimate of the correlation between dependent animacy conditions, these conditions were combined into a single observation per sample coded as uncontrolled animacy.

Table 5

Estimated correlations between priming conditions

Structure	Overlap	Correlation estimate
Same	Same	1
	Different	.20
Different	Same	.15
	Different	.06

Multiple Timepoints. Some studies measured priming at more than one timepoint. Branigan and Messenger (2016) conducted two priming sessions one week apart. They reported the correlation between these sessions which we were able to specify in our meta-analysis. Kumarage et al. (2022) conducted four priming sessions over 18 months. The data from this study were available to calculate correlations between the sessions, which we specified in our meta-analysis. Kidd (2012), Fazekas et al. (2020), and Savage et al. (2006) instead primed children at only one time but recorded target responses at more than one timepoint after this. The correlations between the test, posttest (immediately after a prime phase), and long posttest (one week later) timepoints from Kidd (2012) were provided by the author. The data from Fazekas (2020) were available to calculate correlations between the baseline, test, and two posttest (immediately after test) conditions. We specified the provided and calculated correlations from these two studies in our meta-analysis and used an average of these correlations as an estimate for the correlation between timepoints for Savage et al. (2006; see Appendix B). Using this estimate of .242, *vcalc* was used to calculate the variance-covariance matrix for this study, which was then specified in our meta-analysis.

Results

We report results for two different meta-analytic models because we calculated sampling variance in two ways. Using the number of responses as sample size ignores the fact that, unlike in clinical trials, participants in priming studies provide multiple responses, which will depend on an individual’s tendency to produce the target structure. Using the number of participants as the sample size is more conservative but ignores the extra information provided by collecting multiple responses from each participant. All models included random effects of observation nested within study and a variance-covariance matrix specifying reported or estimated correlations between observations.

Overall Effect Size and Heterogeneity

We first ran models without moderators to estimate an overall effect size and the heterogeneity in observations. For both models, profile-likelihood plots indicated the variance components could be estimated and DFBETA values revealed no influential observations. Residuals were normally distributed for the first model and showed slightly less variance than expected for the second.

Table 6 displays the results of these models. In both, the overall priming effect is significant with odds ratios larger than one, indicating greater odds of producing the dependent structure in the primed than unprimed condition. The magnitude of this effect is medium-to-large: when converted to odds ratios, the LORs translate to an effect of 3.6 to 4.2 times greater odds of the dependent structure in the primed than unprimed condition, or a Cohen's *d* of 0.70–0.80 (multiplying the LOR by $\frac{\sqrt{3}}{\pi}$ converts it to Cohen's *d*; Borenstein et al., 2009). Figure 2 shows a forest plot of all included effect sizes and the estimated overall effect from the first model.

Table 6

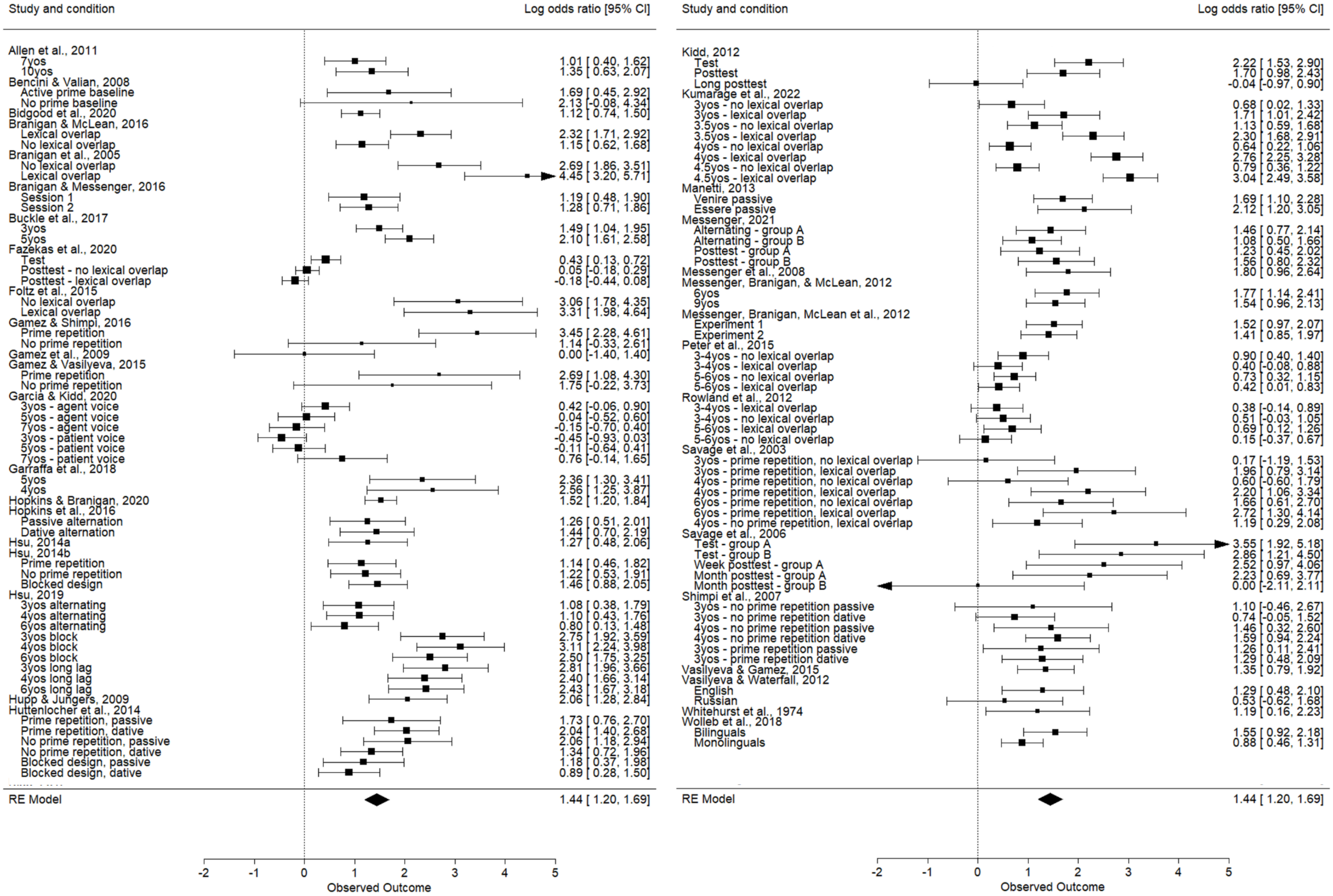
Results of multilevel meta-analytic models without moderators

	N responses		N subjects	
Estimate (*log odds ratio)*	1.44		1.27	
	1.20 | 1.69		1.01 | 1.53	
	p < .001***		p < .001***	
Odds ratio	4.24		3.55	
Cohen's d	0.80		0.70	
Q	698.15		149.23	
	df = 107		df = 107	
	p <.001***		p = .005**	
Variance	0.675		0.823	
Sampling variance	0.102	15.12%	0.504	61.22%
I^2	0.573	84.88%	0.319	38.78%
Between-study		48.27%		25.33%
Within-study		36.61%		13.45%

We also observe significant heterogeneity relative to overall variance (Q_1(107) = 698.15, p < .001; Q_2(107) = 149.23, p < .01). That is, there is significant variance not explained by sampling error. Estimates of I^2 indicate this remaining heterogeneity is substantial in the first model, at about 85% of the observed variance. In the second model, we used a more conservative estimate of sampling variance and the remaining heterogeneity reduced to 39% of observed variance. An investigation of theorised moderators of the priming effect is warranted in both cases; however, the power to detect heterogeneity differs between our two models. This difference in power can be attributed to the more and less conservative calculations of sampling variance because lower power of individual studies can reduce overall power to detect heterogeneity (Borenstein et al., 2009).

Figure 2

Forest plot for meta-analysis using N responses


Study and condition
Log odds ratio [95% CI]
Allen et al., 2011
7yos 1.01 [0.40, 1.62]
10yos 1.35 [0.63, 2.07]
Bencini & Valian, 2008
Active prime baseline 1.69 [0.45, 2.92]
No prime baseline 2.13 [-0.08, 4.34]
Bidgood et al., 2020 1.12 [0.74, 1.50]
Branigan & McLean, 2016
Lexical overlap 2.32 [1.71, 2.92]
No lexical overlap 1.15 [0.62, 1.68]
Branigan et al., 2005
No lexical overlap 2.69 [1.86, 3.51]
Lexical overlap 4.45 [3.20, 5.71]
Branigan & Messenger, 2016
Session 1 1.19 [0.48, 1.90]
Session 2 1.28 [0.71, 1.86]
Buckle et al., 2017
3yos 1.49 [1.04, 1.95]
5yos 2.10 [1.61, 2.58]
Fazekas et al., 2020
Test 0.43 [0.13, 0.72]
Posttest - no lexical overlap 0.05 [-0.18, 0.29]
Posttest - lexical overlap -0.18 [-0.44, 0.08]
Foltz et al., 2015
No lexical overlap 3.06 [1.78, 4.35]
Lexical overlap 3.31 [1.98, 4.64]
Gamez & Shimpi, 2016
Prime repetition 3.45 [2.28, 4.61]
No prime repetition 1.14 [-0.33, 2.61]
Gamez et al., 2009 0.00 [-1.40, 1.40]
Gamez & Vasilyeva, 2015
Prime repetition 2.69 [1.08, 4.30]
No prime repetition 1.75 [-0.22, 3.73]
Garcia & Kidd, 2020
3yos - agent voice 0.42 [-0.06, 0.90]
5yos - agent voice 0.04 [-0.52, 0.60]
7yos - agent voice -0.15 [-0.70, 0.40]
3yos - patient voice -0.45 [-0.93, 0.03]
5yos - patient voice -0.11 [-0.64, 0.41]
7yos - patient voice 0.76 [-0.14, 1.65]
Garraffa et al., 2018
5yos 2.36 [1.30, 3.41]
4yos 2.56 [1.25, 3.87]
Hopkins & Branigan, 2020 1.52 [1.20, 1.84]
Hopkins et al., 2016
Passive alternation 1.26 [0.51, 2.01]
Dative alternation 1.44 [0.70, 2.19]
Hsu, 2014a 1.27 [0.48, 2.06]
Hsu, 2014b
Prime repetition 1.14 [0.46, 1.82]
No prime repetition 1.22 [0.53, 1.91]
Blocked design 1.46 [0.88, 2.05]
Hsu, 2019
3yos alternating 1.08 [0.38, 1.79]
4yos alternating 1.10 [0.43, 1.76]
6yos alternating 0.80 [0.13, 1.48]
3yos block 2.75 [1.92, 3.59]
4yos block 3.11 [2.24, 3.98]
6yos block 2.50 [1.75, 3.25]
3yos long lag 2.81 [1.96, 3.66]
4yos long lag 2.40 [1.66, 3.14]
6yos long lag 2.43 [1.67, 3.18]
Hupp & Jungers, 2009 2.06 [1.28, 2.84]
Huttenlocher et al., 2014
Prime repetition, passive 1.73 [0.76, 2.70]
Prime repetition, dative 2.04 [1.40, 2.68]
No prime repetition, passive 2.06 [1.18, 2.94]
No prime repetition, dative 1.34 [0.72, 1.96]
Blocked design, passive 1.18 [0.37, 1.98]
Blocked design, dative 0.89 [0.28, 1.50]
RE Model 1.44 [1.20, 1.69]
-2 -1 0 1 2 3 4 5
Observed Outcome
Study and condition
Log odds ratio [95% CI]
Kidd, 2012
Test 2.22 [1.53, 2.90]
Posttest 1.70 [0.98, 2.43]
Long posttest -0.04 [-0.97, 0.90]
Kumarage et al., 2022
3yos - no lexical overlap 0.68 [0.02, 1.33]
3yos - lexical overlap 1.71 [1.01, 2.42]
3.5yos - no lexical overlap 1.13 [0.59, 1.68]
3.5yos - lexical overlap 2.30 [1.68, 2.91]
4yos - no lexical overlap 0.64 [0.22, 1.06]
4yos - lexical overlap 2.76 [2.25, 3.28]
4.5yos - no lexical overlap 0.79 [0.36, 1.22]
4.5yos - lexical overlap 3.04 [2.49, 3.58]
Manetti, 2013
Venire passive 1.69 [1.10, 2.28]
Essere passive 2.12 [1.20, 3.05]
Messenger, 2021
Alternating - group A 1.46 [0.77, 2.14]
Alternating - group B 1.08 [0.50, 1.66]
Posttest - group A 1.23 [0.45, 2.02]
Posttest - group B 1.56 [0.80, 2.32]
Messenger et al., 2008 1.80 [0.96, 2.64]
Messenger, Branigan, & McLean, 2012
6yos 1.77 [1.14, 2.41]
9yos 1.54 [0.96, 2.13]
Messenger, Branigan, McLean et al., 2012
Experiment 1 1.52 [0.97, 2.07]
Experiment 2 1.41 [0.85, 1.97]
Peter et al., 2015
3-4yos - no lexical overlap 0.90 [0.40, 1.40]
3-4yos - lexical overlap 0.40 [-0.08, 0.88]
5-6yos - no lexical overlap 0.73 [0.32, 1.15]
5-6yos - lexical overlap 0.42 [0.01, 0.83]
Rowland et al., 2012
3-4yos - lexical overlap 0.38 [-0.14, 0.89]
3-4yos - no lexical overlap 0.51 [-0.03, 1.05]
5-6yos - lexical overlap 0.69 [0.12, 1.26]
5-6yos - no lexical overlap 0.15 [-0.37, 0.67]
Savage et al., 2003
3yos - prime repetition, no lexical overlap 0.17 [-1.19, 1.53]
3yos - prime repetition, lexical overlap 1.96 [0.79, 3.14]
4yos - prime repetition, no lexical overlap 0.60 [-0.60, 1.79]
4yos - prime repetition, lexical overlap 2.20 [1.06, 3.34]
6yos - prime repetition, no lexical overlap 1.66 [0.61, 2.70]
6yos - prime repetition, lexical overlap 2.72 [1.30, 4.14]
4yos - no prime repetition, lexical overlap 1.19 [0.29, 2.08]
Savage et al., 2006
Test - group A 3.55 [1.92, 5.18]
Test - group B 2.86 [1.21, 4.50]
Week posttest - group A 2.52 [0.97, 4.06]
Month posttest - group A 2.23 [0.69, 3.77]
Month posttest - group B 0.00 [-2.11, 2.11]
Shimpi et al., 2007
3yos - no prime repetition passive 1.10 [-0.46, 2.67]
3yos - no prime repetition dative 0.74 [-0.05, 1.52]
4yos - no prime repetition passive 1.46 [0.32, 2.60]
4yos - no prime repetition dative 1.59 [0.94, 2.24]
3yos - prime repetition passive 1.26 [0.11, 2.41]
3yos - prime repetition dative 1.29 [0.48, 2.09]
Vasilyeva & Gamez, 2015 1.35 [0.79, 1.92]
Vasilyeva & Waterfall, 2012
English 1.29 [0.48, 2.10]
Russian 0.53 [-0.62, 1.68]
Whitehurst et al., 1974 1.19 [0.16, 2.23]
Wolleb et al., 2018
Bilinguals 1.55 [0.92, 2.18]
Monolinguals 0.88 [0.46, 1.31]
RE Model 1.44 [1.20, 1.69]
-2 -1 0 1 2 3 4 5
Observed Outcome

Moderator Analyses

We next ran models that included our identified moderator variables: design (between- vs within-subjects), baseline (alternate structure prime, no prime), animacy (uncontrolled, controlled, favourable), age, prime repetition (and interaction with age), lexical overlap (and interaction with age), structure (passive, dative, relative clause, SVO-ba, SV-Transitive; and interaction with age), and lag (alternating, blocked, long-blocked). Residuals for both models showed less variability than expected, likely due to slight overfitting after including moderators of the effect. Profile-likelihood plots showed that variance components were adequately estimated. We identified the second post-test condition of Fazekas et al. (2020) as an outlier with large influence over parameters in the model by examining DFBETAS values and running models with and without the outlier. We report results excluding this outlier.

The moderator models successfully explained remaining heterogeneity. The test of moderators was significant for both models (F(19,17) = 5.22, p < .001; F(19,17) = 4.42, p < .01), indicating that a significant portion of heterogeneity is explained by predictors in the model. This is illustrated in Figure 3, which shows that non-sampling variance is reduced after including moderators. After accounting for moderators there was significant residual heterogeneity in the $N_{responses}$ model, Q(87) = 230.79, p < .001, but not the $N_{subjects}$ model, Q(87) = 53.55, p = .998.

Figure 3

Proportion of sampling variance and non-sampling variance in models including and excluding moderator variables

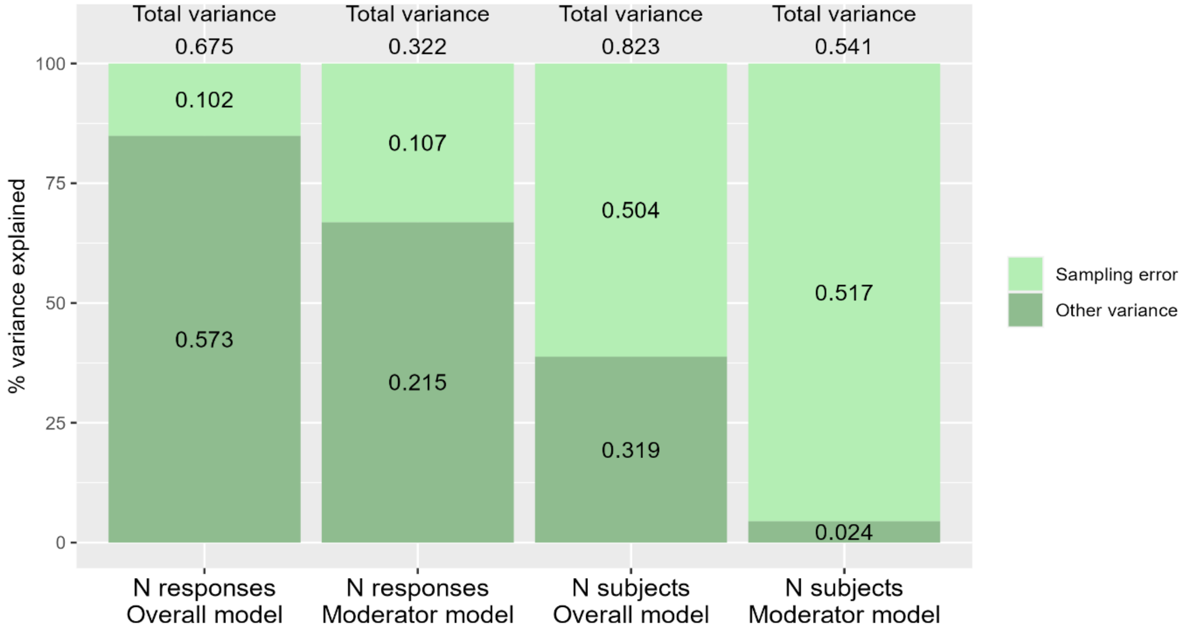


Table 7 displays the results of the moderator models. The intercept is significant in both models, indicating a significant priming effect in the baseline-coded condition of all our moderators. That is, the model predicts a significant priming effect for a within-subjects study of the passive in 59.35-month-olds, which compares to an active-primed baseline, controls for animacy, has no lexical overlap, no prime repetition and no lag between prime and target. Figure 4 shows the model-predicted LOR at treatment levels of moderators for the model using $N_{responses}$ (Figure 4a: main effects, Figure 4b: interaction effects). For example, the predicted priming effect for within-subjects studies (baseline level, 0) is shown by the Main effect (intercept), whilst the predicted priming effect for between-subjects studies (treatment level, 1) is shown under Moderators.

The priming effect was significantly larger in studies using between-subjects designs. We adjusted for the increased precision of within-subjects designs, where the primed and unprimed conditions are correlated due to individual rates of producing the dependent structure, using the Becker-Balagtas method. Therefore, a remaining difference between study designs suggests that there is a factor besides this correlation which differs between them (Morris & DeShon, 2002).

Controlling animacy did not have a significant effect on the size of the priming effect. However, using animacy configurations that favoured the dependent structure did significantly increase the size of the priming effect in the more powerful ($N_{responses}$ as sample size) model.

Studies with lexical overlap between prime and target showed larger priming effects than those with no lexical overlap. We note that including the second posttest condition of Fazekas et al. (2020) reduced the size of the lexical overlap effect for both models and produced a significant lexical overlap*age interaction in the $N_{subjects}$ model, with a decreasing lexical boost over development. This observation is unusual in combining lexical overlap with a long lag between prime and target and found no priming. The reduced priming effect is likely due to the lexical boost being short-lived (Branigan & McLean, 2016) rather than the older age of the sample. Since removing this single observation eliminates the interaction effect, we interpret the evidence to support a non-significant interaction effect.

The structural alternation significantly affected the size of the priming effect. Studies of the dative alternation reported smaller effects than those of the passive. In fact, Figure 4a shows that the model-predicted priming effect is no longer significant in dative studies, with the null effect of 0 contained within the 95% confidence interval. Studies of the relative clause reported larger effects than the passive in the more powerful model; however, there were only two studies of this structure. Similar caution should be applied when interpreting the non-significant model-predicted priming effect in observations of the Tagalog SV-Transitive alternation: there were 6 observations from 6 samples, all from the same study.

There were no effects of baseline type (no prime or alternate prime), prime repetition, lag between prime and target, or age. There was also no significant interaction between age and other moderators: Figure 4b shows that the effects of prime repetition, lexical overlap and structure were constant across age.

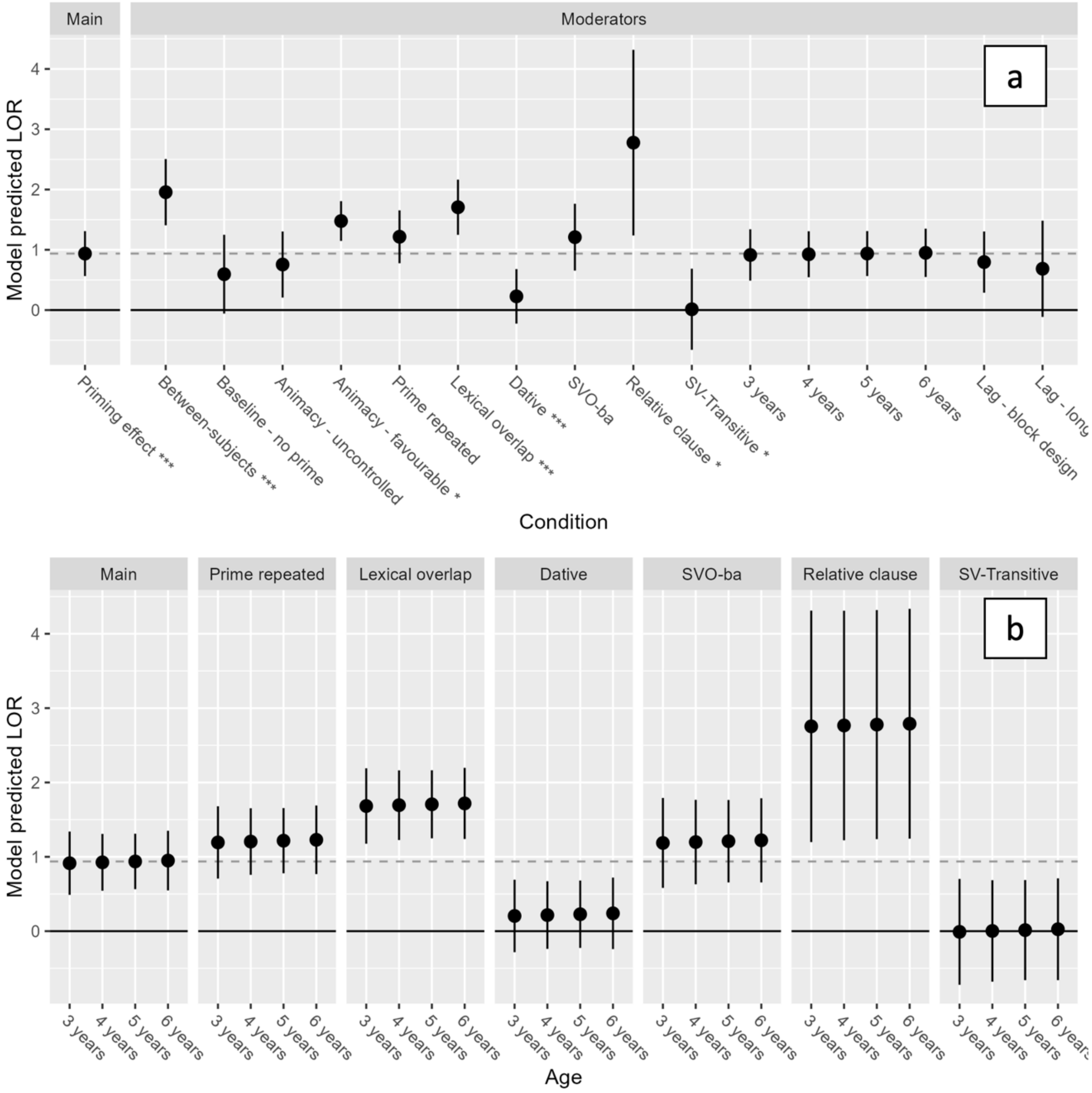
Table 7

Results of meta-analyses including moderator variables

	N responses			N subjects		
	β	CI	p	β	CI	p
Intercept	0.94	0.56 | 1.31	<.001***	1.03	0.66 | 1.39	<.001***
Between-subjects	1.02	0.57 | 1.47	<.001***	1.02	0.35 | 1.70	.003**
Baseline	-0.34	-0.88 | 0.19	.207	-0.22	-0.95 | 0.52	.563
Animacy						
Uncontrolled	-0.18	-0.72 | 0.36	.489	0.03	-0.73 | 0.78	.938
Favourable	0.54	0.06 | 1.02	.028*	0.38	-0.16 | 0.92	.159
Prime repetition	0.28	-0.07 | 0.63	.117	0.06	-0.37 | 0.49	.785
Lexical overlap	0.77	0.39 | 1.15	<.001***	0.74	0.28 | 1.19	.002**
Structure						
Dative	-0.71	-1.04 |-0.38	<.001***	-0.85	-1.26 |-0.44	<.001***
SVO-ba	0.27	-0.34 | 0.89	.364	0.16	-0.48 | 0.80	.611
RC	1.84	0.27 | 3.41	.024*	1.80	-0.91 | 4.50	.179
SV-Transitive	-0.92	-1.69 |-0.15	.021*	-1.03	-1.66 |-0.41	.003**
Age	0.02	-0.15 | 0.19	.835	0.07	-0.16 | 0.32	.566
Lag						
0 vs block	-0.14	-0.52 | 0.24	.461	-0.19	-0.76 | 0.30	.484
0 vs long	-0.25	-0.92 | 0.41	.454	-0.14	-1.23 | 0.95	.797
Prime repetition*Age	0.23	-0.17 | 0.63	.263	0.01	-0.51 | 0.50	.978
Lexical overlap*Age	-0.09	-0.58 | 0.40	.704	-0.26	-0.87 | 0.34	.389
Structure*Age						
Dative	0.03	-0.27 | 0.33	.842	0.07	-0.35 | 0.50	.735
SVO-ba	-0.20	-0.68 | 0.27	.398	-0.23	-0.89 | 0.43	.497
RC	-0.40	-3.61 | 2.81	.795	-0.22	-5.79 | 5.35	.936
SV-Transitive	0.06	-0.36 | 0.49	.764	-0.03	-0.55 | 0.49	.906

Figure 4

Size of model-predicted priming effect under different experimental conditions for (a) main effects and (b) interaction effects



Assessing Publication Bias

Studies with significant results or larger effect sizes are more likely to be published than those with smaller or null results (Dickersin, 2005; Franco et al., 2014). In meta-analyses, synthesising the results of a biased sample of studies can then lead to spurious findings. This is common in psychology, with a recent estimate that 60% of meta-analyses in psychology overestimate the evidence for an effect (Bartoš et al., 2023). For an example in psycholinguistics, the widely accepted bilingual advantage in executive functioning has more recently been attributed to publication bias (de Bruin et al., 2015; Lehtonen et al., 2018). In the syntactic priming literature, Mahowald et al. (2016) found their meta-analysis of studies in adults was not overly influenced by publication bias. We expected a similar finding if the child literature is comparable and given that in the developmental context null results can themselves be of interest (e.g., Savage et al. (2003) reported 3-year-olds were not primed in the absence of lexical overlap).

Funnel plots aid in detecting publication bias by depicting the relationship between study precision and effect size. More precise studies with smaller standard errors tend to cluster around the estimated effect size at the top of the plot. In a sample of studies without publication bias, smaller studies with larger standard errors towards the bottom of the plot will be symmetrically distributed around the estimated effect size. However, if studies suffer from publication bias, funnel plot asymmetry is observed, wherein small studies with null effects are missing from the bottom left. Funnel plot asymmetry should be assessed statistically, not just visually. We used a test conceptually similar to Egger's regression test and Peters' regression test (Egger et al., 1997; Peters et al., 2006). It is not possible to execute these tests in a multilevel meta-analysis containing dependent sampling variances[4]. Instead, we added the inverse of sample size to the meta-regression as a moderator. If study precision, as indexed by inverse sample size, significantly predicts the size of the overall effect, we can conclude there is funnel plot asymmetry. We use inverse sample size rather than the variance or standard error of the LOR because these are already mathematically dependent on the size of the LOR (Peters et al., 2006).

Figure 5 shows funnel plots for models including and excluding moderators, with a circle indicating where we may expect missing observations to be in asymmetrical plots. Funnel plots for the models without moderators show a somewhat asymmetrical distribution of effects when using number of responses as sample size, and more so when using the number of participants as sample size (Figures 5a and 5b). This is confirmed by marginally significant funnel plot asymmetry for the first model, but significant asymmetry for the second. Asymmetry can indicate publication bias, but also heterogeneity in observations (Sterne et al., 2011). In both plots, observations do not narrow around the overall estimate with increased precision. This can also indicate substantial heterogeneity in effect sizes (Sterne et al., 2011). For example, Kidd (2012) had a large sample size but

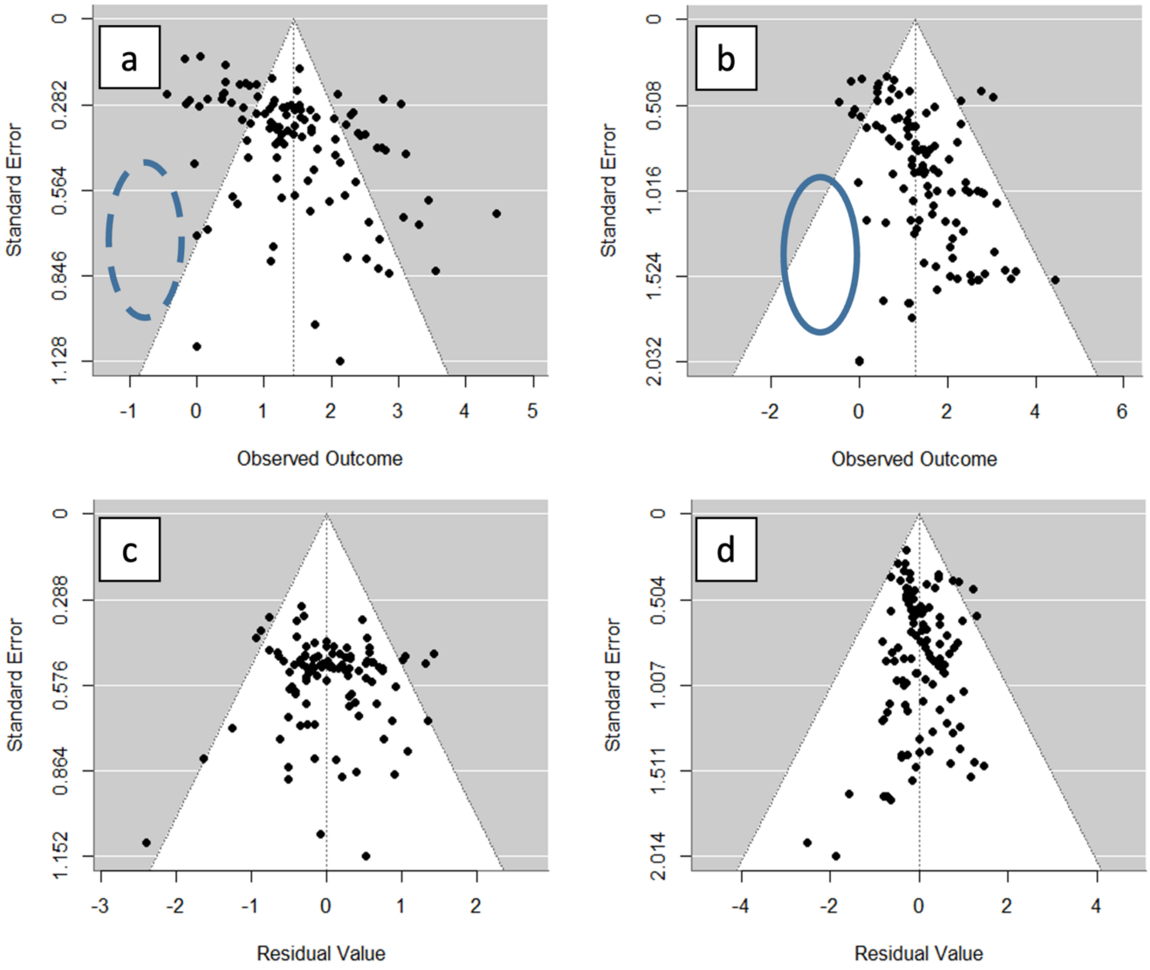
---

[4] For discussion of this issue by Wolfgang Viechtbauer, see https://stats.stackexchange.com/q/155875

reported an observation of no priming in the long post-test condition (one week after priming), Branigan et al. (2005) and Foltz et al. (2015) had small sample sizes but investigated priming in the relative clause, finding large priming effects. The manipulation of lag and the structure may be better predictors of the size of the effect than sample size in these cases.

Figure 5

Funnel plots for meta-analysis models (a) without moderators, using $N_{responses}$*, (b) without moderators, using* $N_{subjects}$*, (c) including moderators, using* $N_{responses}$*, and (d) including moderators, using* $N_{subjects}$



Funnel plots for the models including moderators plot standard error against residual value rather than observed outcome – thus taking moderators into account. Both funnel plots are fairly symmetrical (see Figures 5c and 5d) and the inverse of sample size did not significantly predict effect size in either model containing moderators. This suggests there is no evidence for publication bias in the priming literature once accounting for different manipulations in studies of different sizes.

Power Analysis

Following Mahowald et al. (2016), we conducted several simulations to estimate the power of observing a significant priming effect with and without lexical overlap at incrementally increasing values of participants and items. To make our estimates more relevant to the child language literature, we made a few different decisions than Mahowald et al. (2016). They simulated experiments in which each participant (P) saw each word (W) once, resulting in a data frame with P*W rows. A 'word' here corresponds mostly to a verb, since the majority of priming studies test argument structure alternations, and it is this sense in which we use it. Because child languages studies often use a smaller set of verbs than adult studies, it is common for participants to see each word twice or more. Thus, we simulated data where each participant saw each word twice, resulting in data frames of P*W*2 rows. However, we only simulated random effects by subject and by item (i.e., verb), and not by the interaction between subject and item. While it is very possible that this standard deviation is non-zero in this population, these effects would be very difficult to estimate given the sample sizes of most child language studies, and models estimating them would be very unlikely to converge (especially with only 2 unique values per crossing of participant and item). Following Mahowald et al. (2016), we removed 20% of observations at random to simulate missing data.

Mahowald et al. (2016) estimated random effect standard deviations from their validation model, a generalized linear mixed model fit to raw data they had available. We suspect between-participant and between-item heterogeneity in child priming studies will be larger than that typically observed adult priming studies, given that child language samples often contain wide age ranges and participants with varying levels of linguistic proficiency. We therefore considered two empirically-based random effects structures. To do so, we first collected the by-participant and by-item random intercepts and random slopes (for priming) from eight analyses reported in two recent developmental studies of priming, one focused on the passive (Kumarage et al., 2022) and one focused on the dative (Donnelly et al., 2024). We took the mean of each of these standard deviations as our average scenario and the highest standard deviation as a high variability scenario. We simulated (and estimated) uncorrelated random effects for each scenario.

Power Analyses for Detecting a Priming Effect

Given that children produce the relevant grammatical structures at very low rates, we chose a value of –2 as our intercept (corresponding to a probability of .12). For models of main effects, we simulated from the models in Equations 7 and 8.

Model without lexical overlap:

$$\Pr(\text{Structure}) = \text{logit}^{-1}(-2 + \tau_{1i} + \lambda_{1j} + \text{Prime}*(0.94 + \tau_{2i} + \lambda_{2j})) \qquad (7)$$

Model with lexical overlap:

$$\Pr(\text{Structure}) = \text{logit}^{-1}(-2 + \tau_{1i} + \lambda_{1j} + \text{Prime}*(1.71 + \tau_{2i} + \lambda_{2j})) \qquad (8)$$

We assume a priming effect of 0.94 and 1.71 respectively, which are the estimated priming effects taken from our meta-analysis including moderators. Importantly, as the active-passive alternation was the reference level, these effect sizes estimate passive priming. These simulations correspond to a within-subjects design, where prime was coded as +/– .5, τ is a matrix of by-participant random effects, and λ is a matrix of by-item random effects. Random effects values for the two scenarios (average and high variability) are shown in Table 8. Overall, then, we considered 4 scenarios (2 random effects specifications with each main effect specified above). We simulated 1000 data sets for each crossing of participant number (20, 30, 40, 60, 80, 100 and 200) and item number (6, 8, 10, 12, 18, 24, 30), fitting the above model to each data set. Note that each item corresponded to two trials, not one, since we assumed most developmental priming studies would repeat words across trials.

Table 8

Values of random effects for power simulations of the priming effect

Variability	τ_1	τ_2	λ_1	λ_2
Average	1.85	0.81	1.07	0.77
High	2.20	1.24	1.92	1.59

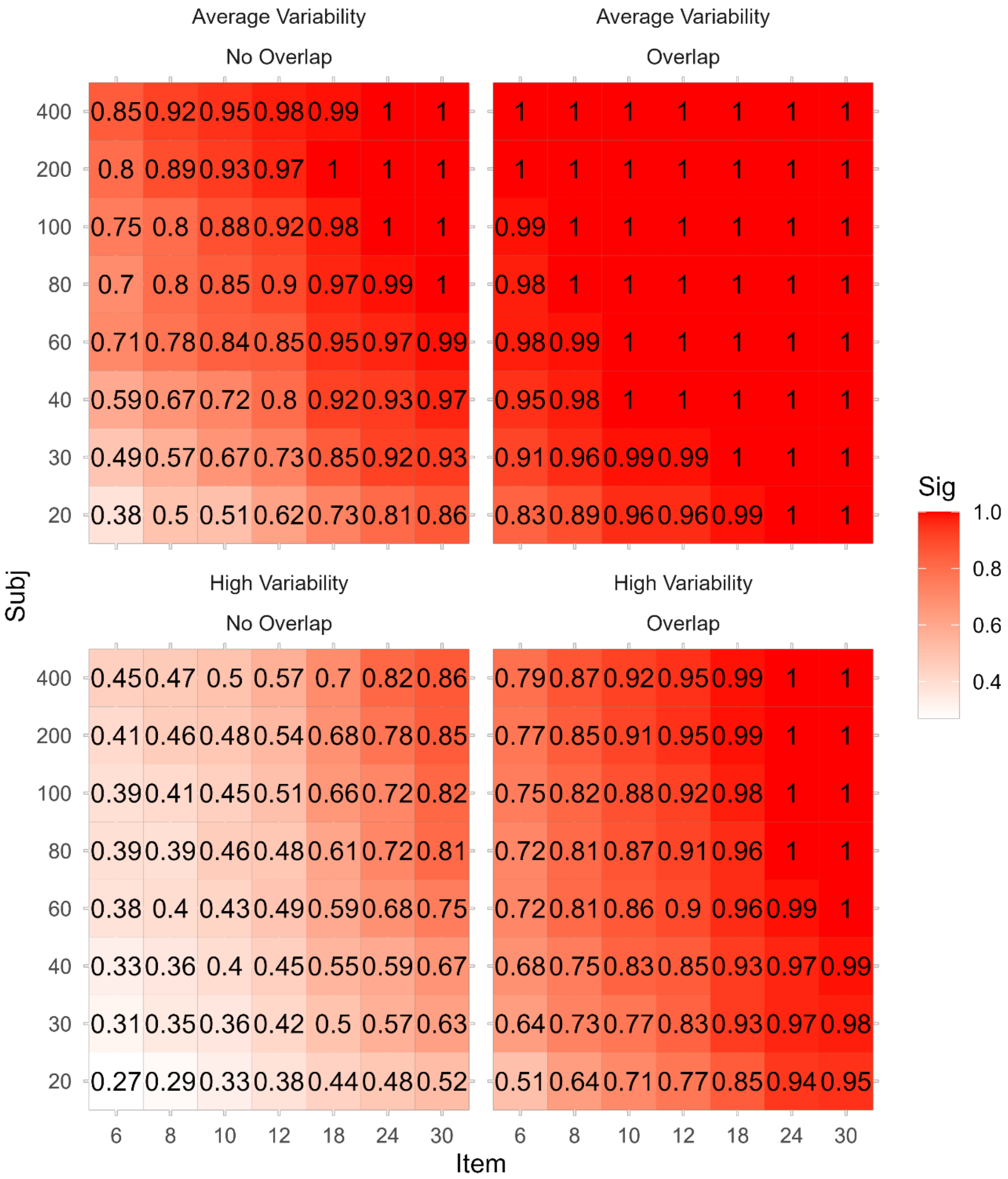
One challenge in power analysis with mixed models is handling non-convergent models. While it is advisable to consider the full random effect structure implied by the design (Barr et al., 2013), such models are often empirically unidentified, resulting in non-convergent model with unreliable parameter estimates (Bates et al., 2015). This is likely exacerbated in developmental priming research, where between-participant variability is high and mean productions of the target structure are low. Because these data were simulated, we knew a priori that the random effect structure specified in the model was the correct one, and we included all models in our main power calculations. However, we also report on the number of non-convergent models (See Appendix D).[5]

---

[5] A small number of iterations (roughly 4.5% for simulations of both priming effects and interactions) produced implausible standard errors (<.1) or, in very rare cases implausible effect sizes (> 5 or < –5 on the logit scale). Because these estimation errors would lead to significant results, we removed these iterations from power calculations. We report on the number of iterations producing implausible values for each scenario (range 0 – 16% of iterations) in Appendix D. We have also posted power calculations with these values included to file Supp 1 on the OSF site.

Figure 6

Power estimates for priming effect with and without lexical overlap at average and high levels of heterogeneity



Average Variability
No Overlap

Subj	6	8	10	12	18	24	30
400	0.85	0.92	0.95	0.98	0.99	1	1
200	0.8	0.89	0.93	0.97	1	1	1
100	0.75	0.8	0.88	0.92	0.98	1	1
80	0.7	0.8	0.85	0.9	0.97	0.99	1
60	0.71	0.78	0.84	0.85	0.95	0.97	0.99
40	0.59	0.67	0.72	0.8	0.92	0.93	0.97
30	0.49	0.57	0.67	0.73	0.85	0.92	0.93
20	0.38	0.5	0.51	0.62	0.73	0.81	0.86

Average Variability
Overlap

Subj	6	8	10	12	18	24	30
400	1	1	1	1	1	1	1
200	1	1	1	1	1	1	1
100	0.99	1	1	1	1	1	1
80	0.98	1	1	1	1	1	1
60	0.98	0.99	1	1	1	1	1
40	0.95	0.98	1	1	1	1	1
30	0.91	0.96	0.99	0.99	1	1	1
20	0.83	0.89	0.96	0.96	0.99	1	1

High Variability
No Overlap

Subj	6	8	10	12	18	24	30
400	0.45	0.47	0.5	0.57	0.7	0.82	0.86
200	0.41	0.46	0.48	0.54	0.68	0.78	0.85
100	0.39	0.41	0.45	0.51	0.66	0.72	0.82
80	0.39	0.39	0.46	0.48	0.61	0.72	0.81
60	0.38	0.4	0.43	0.49	0.59	0.68	0.75
40	0.33	0.36	0.4	0.45	0.55	0.59	0.67
30	0.31	0.35	0.36	0.42	0.5	0.57	0.63
20	0.27	0.29	0.33	0.38	0.44	0.48	0.52

High Variability
Overlap

Subj	6	8	10	12	18	24	30
400	0.79	0.87	0.92	0.95	0.99	1	1
200	0.77	0.85	0.91	0.95	0.99	1	1
100	0.75	0.82	0.88	0.92	0.98	1	1
80	0.72	0.81	0.87	0.91	0.96	1	1
60	0.72	0.81	0.86	0.9	0.96	0.99	1
40	0.68	0.75	0.83	0.85	0.93	0.97	0.99
30	0.64	0.73	0.77	0.83	0.93	0.97	0.98
20	0.51	0.64	0.71	0.77	0.85	0.94	0.95

Item

We plot the power for detecting a significant abstract priming effect (with and without overlap) at average and high levels of between-participant and between-item heterogeneity in Figure 6. The average priming study in the present meta-analysis included 49.27 subjects and 9.08 items. Our analyses suggest that this as an adequate number of participants and items to detect a significant priming effect with lexical overlap, but results in less power to observe significant effect without lexical overlap (~ .67) when between-participant and between-item variability are average. In the latter case, relatively small increases in the number of items and participants would yield power close to or above .8. However, when between-participant and between-item heterogeneity are especially high, the average sample size is close to adequately powered for effects with lexical overlap (~ .75) but considerably underpowered to detect priming effects without lexical overlap (~ .36). It will be especially challenging to reach sufficient levels of power for detecting priming effects without overlap under such high levels of variability, given the number of items needed (24 or 30 words corresponding to 48 or 60 trials). However, reasonable increases in the number of participants and items may yield sufficient power to detect a priming effect with lexical overlap. At the same time, we think these high-variability estimates are overly conservative (given the large random effect standard deviations) and might be expected in sample with (a) a very wide age range of participants or linguistic abilities (as in studies of clinical populations), and (b) constructions that have very strong verb-biases, such that some verbs consistently elicit one construction or the other. Nonetheless, our results suggest that while the average sample size in the field is likely to be sufficiently powered to observe priming effects, power estimates are affected by the size of random effects, and researchers should carefully consider the degree of between-participant and between-item heterogeneity when designing their studies.

Power Analyses for Detecting an Interaction

We next conducted simulations testing the power for detecting interactions of various magnitudes between prime and some other factor *B*. We included four random effects by random factor: a random intercept, a random slope for prime, a random slope for *B* and a random slope for the interaction. We used the average by-participant and by-item random slopes for prime from the previous simulations. As it was unclear what the random slope on the interaction should be, we divided the random slopes for the priming effect by 2, under the assumption that individual differences in the interactions would generally be smaller than individual differences in priming. We also included by-participant and by-item random effects for the main effect of *B*, though we assumed its fixed effect was 0. We simulated these random effects because *lme4* (Bates et al., 2015) estimates these by default when including random slopes for prime*B, and if their true value was 0, the model will likely not converge.

We simulated from the model in Equation 9 where prime and B were coded as .5, –.5 and M was the magnitude of the interaction effect.

Model with interaction effect:

$$\begin{gathered} Y = \text{logit}^{-1}(-2 + \tau_{1i} + \lambda_{1j} + \text{Prime}*(0.94 + \tau_{2i} + \lambda_{2j}) + (\tau_{3i} + \lambda_{3j})*B \\ + (M + \tau_{4i} + \lambda_{4j})*B*\text{Prime} \end{gathered} \tag{9}$$

$$M_{\text{small}} = 0.4;\ M_{\text{medium}} = 0.8;\ M_{\text{large}} = 1.6$$

We considered interactions of three magnitudes, 0.4, 0.8 and 1.6 on the logit scale, corresponding to a small, medium and large effect. As interactions on the logit scale can be difficult to interpret (especially since the variables for prime condition and the interaction are on different scales), we present the difference (and ratio) between priming effects at each level of B for each of the three interaction sizes considered on the probability scale (see Table 9). For each of the three simulated interaction sizes (0.4, 0.8, and 1.6), the crossed factors of Prime and B indicate the probability of producing the relevant construction. The column *Priming effect at level of B (Additive)* presents the priming effect as a difference score (e.g., Pr(Passive | Passive Prime) – Pr(Passive | Active Prime)) at each level of B. The column *Priming effect at level of B (Ratio)* presents the priming effect at each level of B as a ratio (e.g., Pr(Passive | Passive Prime)/Pr(Passive | Active Prime)). The columns *Difference in priming effect* represent the difference in priming effects across levels of B (as a difference score and a ratio). From these numbers, we can see that an interaction of 0.4 indicates a 4.4 percentage point difference in the priming effect across conditions, an interaction of 0.8 indicates a 7.9 percentage point difference in the magnitude of the priming effects across conditions, and an interaction of 1.6 indicates a 17.5 percentage point difference. The latter scenario reflects a situation where participants produce a large priming effect in one condition (being roughly 4.5 times more likely to produce the target structure after being primed by it than being primed by an alternative structure) and almost no priming effect in the other condition.

Table 9

Simulated interactions of priming effects and factor B

Interaction	*B* level	Prime level		Priming effect at level of *B*		Difference in priming effect	
		.5	-.5	Additive	Ratio	Additive	Ratio
Small (0.4)	.5	.193	.071	.122	2.72	.044	1.42
	-.5	.164	.086	.078	1.91		
Medium (0.8)	.5	.201	.065	.136	3.09	.079	1.92
	-.5	.151	.094	.057	1.61		
Large (1.6)	.5	.244	.054	.19	4.52	.175	4.00
	-.5	.127	.112	.015	1.13		

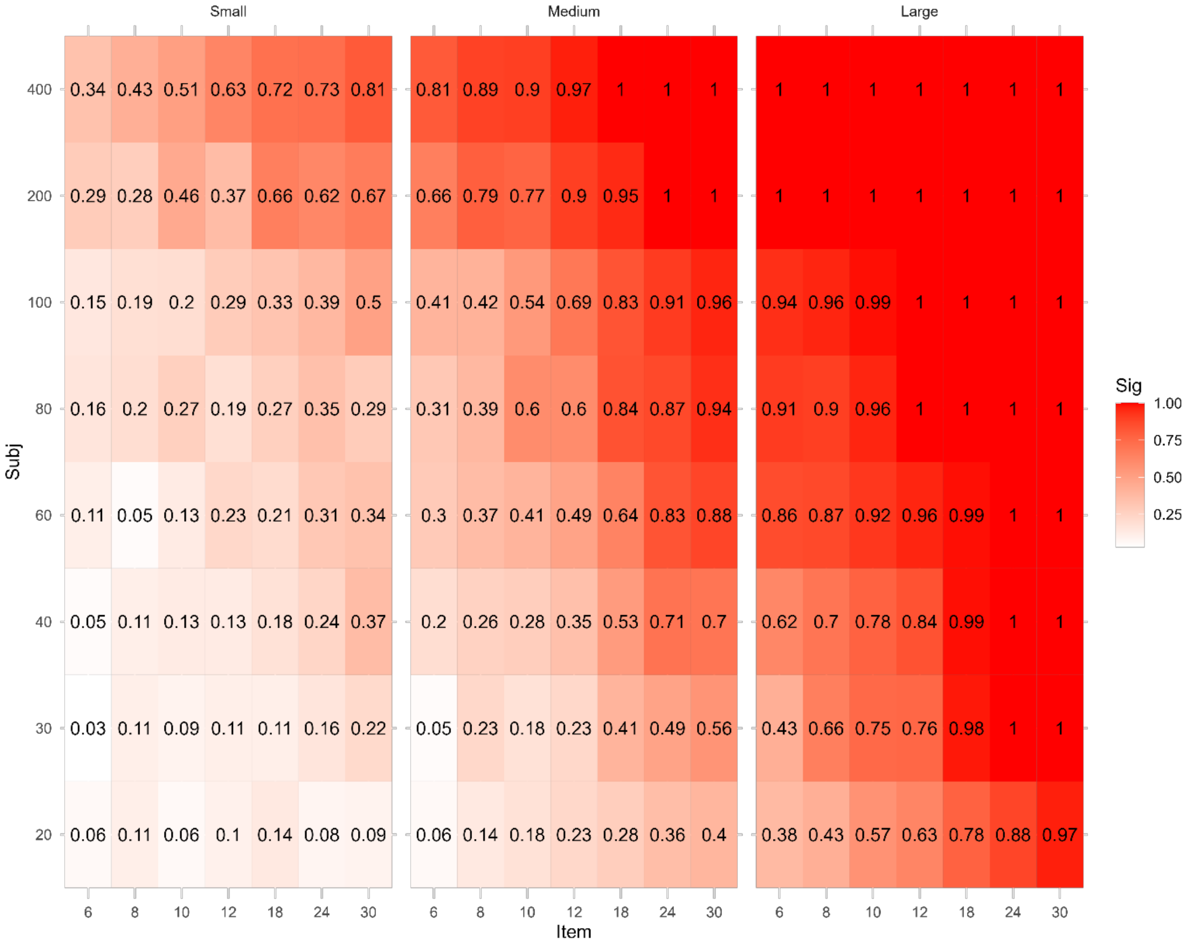
We simulated 100 data sets at each crossing of number of participants and number of items considered in the previous analyses[6], and report on the proportion of significant interactions in each. In addition, we calculated the Type M and Type S error rates for each simulation (Gelman & Carlin, 2014). In low powered studies, only extreme results are significant. As a result, samples that do reach significance in low powered studies may substantially overestimate the magnitude of the true effect or even have the wrong sign. The Type M error rate is the ratio of the average observed effect size to the true effect size amongst statistically significant studies. The Type S error rate is the proportion of statistically significant effects that have the wrong sign (Gelman & Carlin, 2014). Power, Type M and Type S Error rates for each effect size are displayed in Figures 7, 8, and 9 respectively. Note that the Type S error rate for large effect sizes was 0 at all sample sizes, so we omit this figure.

---

[6] Simulations of interactions were considerably more time-intensive than simulations for main effects, so we reduced the number of iterations.

Figure 7

Power estimates for detecting small, medium, and large interactions



Small

Subj	6	8	10	12	18	24	30
400	0.34	0.43	0.51	0.63	0.72	0.73	0.81
200	0.29	0.28	0.46	0.37	0.66	0.62	0.67
100	0.15	0.19	0.2	0.29	0.33	0.39	0.5
80	0.16	0.2	0.27	0.19	0.27	0.35	0.29
60	0.11	0.05	0.13	0.23	0.21	0.31	0.34
40	0.05	0.11	0.13	0.13	0.18	0.24	0.37
30	0.03	0.11	0.09	0.11	0.11	0.16	0.22
20	0.06	0.11	0.06	0.1	0.14	0.08	0.09

Medium

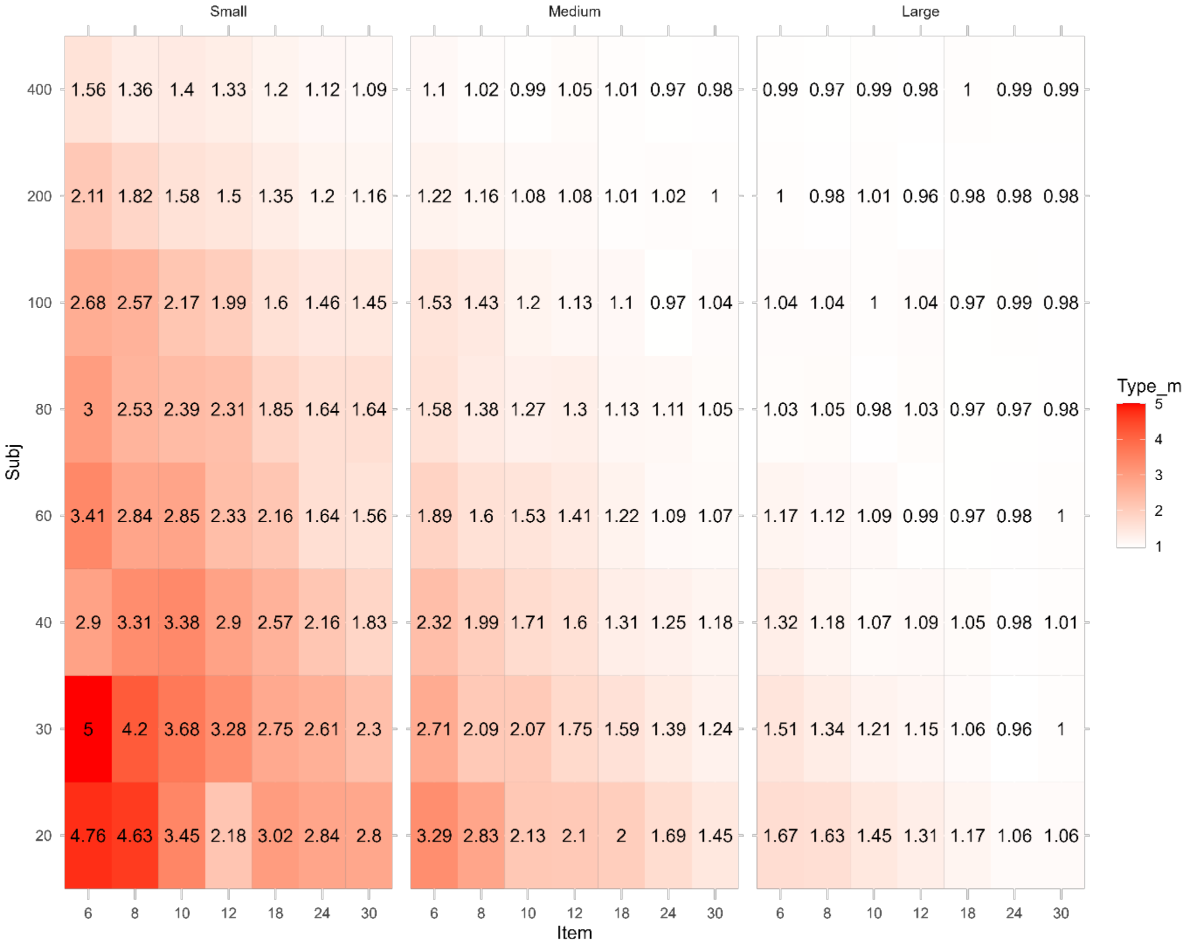
Subj	6	8	10	12	18	24	30
400	0.81	0.89	0.9	0.97	1	1	1
200	0.66	0.79	0.77	0.9	0.95	1	1
100	0.41	0.42	0.54	0.69	0.83	0.91	0.96
80	0.31	0.39	0.6	0.6	0.84	0.87	0.94
60	0.3	0.37	0.41	0.49	0.64	0.83	0.88
40	0.2	0.26	0.28	0.35	0.53	0.71	0.7
30	0.05	0.23	0.18	0.23	0.41	0.49	0.56
20	0.06	0.14	0.18	0.23	0.28	0.36	0.4

Large

Subj	6	8	10	12	18	24	30
400	1	1	1	1	1	1	1
200	1	1	1	1	1	1	1
100	0.94	0.96	0.99	1	1	1	1
80	0.91	0.9	0.96	1	1	1	1
60	0.86	0.87	0.92	0.96	0.99	1	1
40	0.62	0.7	0.78	0.84	0.99	1	1
30	0.43	0.66	0.75	0.76	0.98	1	1
20	0.38	0.43	0.57	0.63	0.78	0.88	0.97

Figure 8

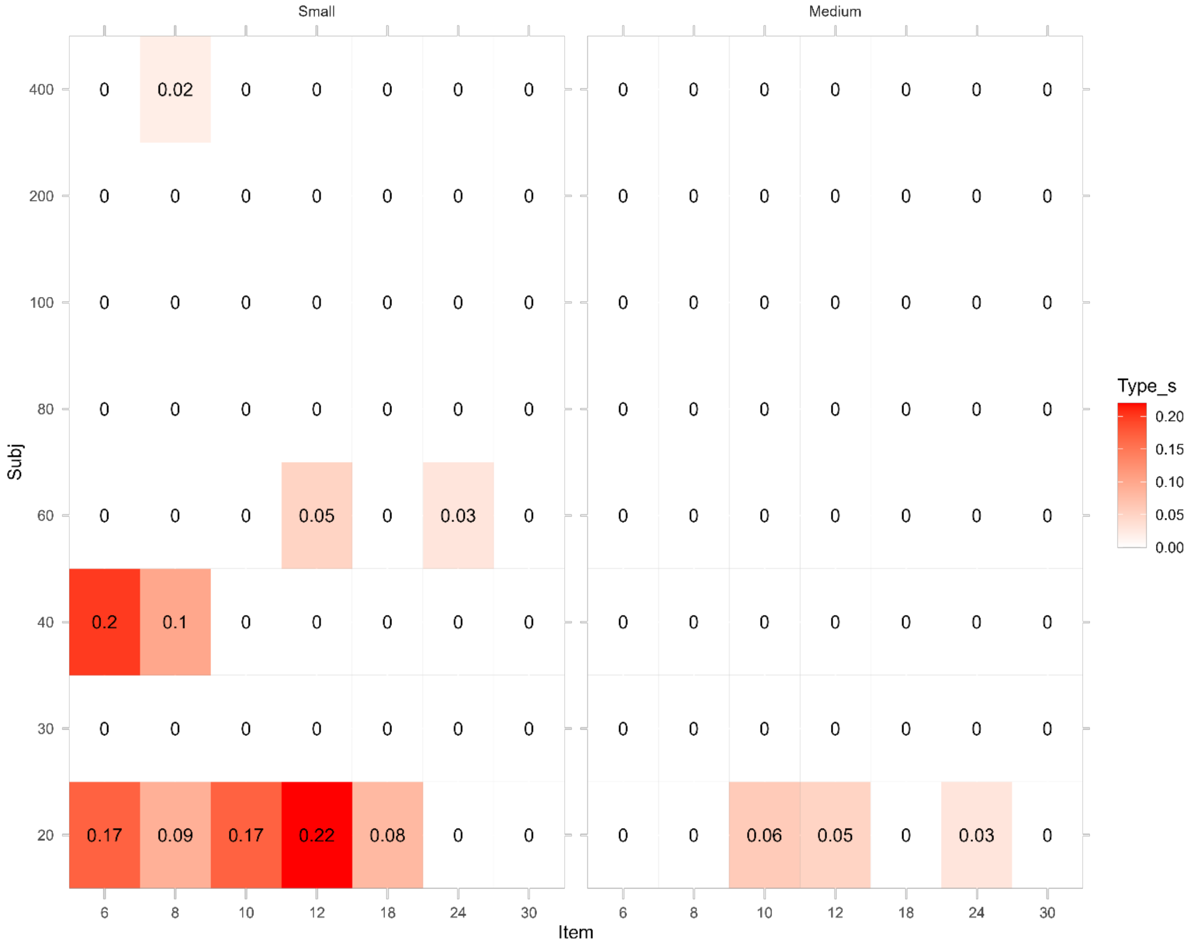
Type M errors for small, medium, and large interactions (ratio of average significant effect size to true effect size)



Subj	Small							Medium							Large						
Item	6	8	10	12	18	24	30	6	8	10	12	18	24	30	6	8	10	12	18	24	30
400	1.56	1.36	1.4	1.33	1.2	1.12	1.09	1.1	1.02	0.99	1.05	1.01	0.97	0.98	0.99	0.97	0.99	0.98	1	0.99	0.99
200	2.11	1.82	1.58	1.5	1.35	1.2	1.16	1.22	1.16	1.08	1.08	1.01	1.02	1	1	0.98	1.01	0.96	0.98	0.98	0.98
100	2.68	2.57	2.17	1.99	1.6	1.46	1.45	1.53	1.43	1.2	1.13	1.1	0.97	1.04	1.04	1.04	1	1.04	0.97	0.99	0.98
80	3	2.53	2.39	2.31	1.85	1.64	1.64	1.58	1.38	1.27	1.3	1.13	1.11	1.05	1.03	1.05	0.98	1.03	0.97	0.97	0.98
60	3.41	2.84	2.85	2.33	2.16	1.64	1.56	1.89	1.6	1.53	1.41	1.22	1.09	1.07	1.17	1.12	1.09	0.99	0.97	0.98	1
40	2.9	3.31	3.38	2.9	2.57	2.16	1.83	2.32	1.99	1.71	1.6	1.31	1.25	1.18	1.32	1.18	1.07	1.09	1.05	0.98	1.01
30	5	4.2	3.68	3.28	2.75	2.61	2.3	2.71	2.09	2.07	1.75	1.59	1.39	1.24	1.51	1.34	1.21	1.15	1.06	0.96	1
20	4.76	4.63	3.45	2.18	3.02	2.84	2.8	3.29	2.83	2.13	2.1	2	1.69	1.45	1.67	1.63	1.45	1.31	1.17	1.06	1.06

Figure 9

Type S errors for small and medium interactions (likelihood of getting an effect with the wrong sign)



The results reveal that current samples sizes are likely too small to detect interactions with sufficient power. Detection of small interactions in particular may be very resource-intensive (requiring for example 400 participants paired with 30 items and 60 trials to reach a power of .8). Detection of medium interactions may be more achievable with larger than typical sample sizes (N = 80 to 100 with 18 items and 36 trials). Detection of large interactions with adequate power will require modest increases in the number of participants and items (for example, to 60 participants or to 12 items and 24 trials). These results also reveal that statistically significant small-to-medium sized interactions may represent substantial over-estimates of the true effect sizes (for example, at current sample sizes, significant results may over-estimate true small effects by a factor of 3.31 and medium effects by a factor of 1.99) and, particularly in the case of small interactions, may reflect a non-trivial number of sign errors (for example, 10% of significant tests of small interactions observed in samples of 40 participants and 8 items had the wrong sign).

Discussion

In this paper we ran the first meta-analysis of syntactic priming studies in children. We found evidence for a medium-to-large main effect, with substantial heterogeneity but with no evidence for publication bias. However, the publication bias analysis did not apply to moderator effects and our power analyses suggest that the average developmental priming study is likely to be underpowered, especially concerning the detection of moderators. Once aggregating studies, the factors that significantly influenced the size of a study's effect were: (i) within- or between-subjects design, (ii) lexical overlap, (iii) structural alternation studied, and (iv) the animacy of verb arguments. We discuss each of these results in turn.

Size and Significance of the Effect

We found a significant overall syntactic priming effect, estimated at a log odds ratio of 1.42, or more conservatively estimated as 1.25. Converted to Cohen's *d* this is a medium-to-large effect of 0.78 or 0.69. Bergmann et al. (2018) found a median effect size of Cohen's *d* = 0.45 (range 0.12 – 1.24) across 12 meta-analyses of effects in language acquisition studies conducted with children aged 0–5 years. Our finding places the syntactic priming effect within the typical range of effect sizes in language acquisition, although priming is typically investigated in slightly older children. In comparison to adults, the effect is numerically larger than the effect reported in Mahowald et al.'s (2016) meta-analysis of syntactic priming studies in adults. After controlling for moderators (e.g. lexical overlap and between-subjects design, which both significantly increased the size of the effect), in the $N_{responses}$ model, the estimated log odds ratio was 0.94 [95% CI, 0.58, 1.29] or Cohen's *d* = 0.52. In comparison, after controlling for moderators (including lexical overlap), Mahowald et al. (2016) reported a log odds ratio of 0.52 [95% CI, 0.22, 0.82] or Cohen's *d* = 0.29 across syntactic priming studies in adults.

Moderators of the Effect

As expected, given the range of design choices and variables manipulated in syntactic priming experiments, there was significant heterogeneity in the overall syntactic priming effect. Even after including moderators in the analysis, there was significant heterogeneity in effect sizes for the $N_{responses}$ model, suggesting our list of moderators was not exhaustive.

Besides lexical overlap, the significant moderators in our meta-analysis tend not to be manipulated variables but ones that vary across studies, so despite issues with low power we remain confident these findings are not greatly affected by publication bias. In our discussion of moderators of the effect, we follow Goodman (1991) in assuming that greater insight can be gained by combining thoughtful analysis with the quantitative results of a meta-analysis. That is, when the summary result differs from the results of individual studies, we should not accept either finding without consideration of why they differ. The value of meta-analysis is not only the summary results, but also the systematic consideration of the current state of evidence in a research field.

Abstract and Lexically-Based Priming

Of particular importance to theories of syntactic acquisition is evidence regarding the emergence of abstract syntactic knowledge in comparison to lexically-based syntactic knowledge. Our findings support the early abstraction of syntax with no reliance on lexically-based representation for the most frequently studied passive structure. We found both a significant abstract priming effect (significant intercept) and lexical boost effect (significant lexical overlap effect). That is, lexical overlap increased the magnitude of a study's reported syntactic priming effect but was not necessary to observe one (cf. main priming effect to priming under lexical overlap in Figure 4a). The findings reflect a growing body of evidence that lexically-based priming does not precede abstract syntactic priming in English-speaking children (Kumarage et al., 2022; Peter et al., 2015; Rowland et al., 2012; but c.f. Savage et al., 2003), and therefore evidence against a transition from lexically-based to abstract syntactic representation (Savage et al., 2003; Tomasello, 2003). However, we also found significant effects of the structural alternation studied. For example, we did not find significant abstract priming for the dative alternation. As we later discuss, syntactic priming research is heavily skewed towards the active-passive structural alternation and generalising beyond the current state of the literature is therefore difficult.

Concerning the developmental trajectory of each effect, age did not moderate the magnitude of either abstract syntactic priming or the lexical boost (Figure 4b). Several factors limit the interpretation of this result: limited power for studies to detect an interaction effect, an inability to detect non-linear effects, and the wide age ranges of included studies. A stable lexical boost effect contrasts with accepted findings that the effect increases over development. The highest quality evidence to date comes from longitudinal data, which found an increasing lexical boost effect (Kumarage et al., 2022). Most cross-sectional data also support this finding (Peter et al., 2015; Rowland et al., 2012; cf. Savage et al., 2003). However, we found a stable effect and an almost identical estimate of the impact of lexical overlap to that found in adults (Table 7: 0.74; Mahowald et al., 2016, p.8.: 0.76). We note that although it is a key point of contention, the study of lexical influence in syntactic priming in children is at an early stage. Eight studies have manipulated lexical overlap (Branigan et al., 2005; Branigan & McLean, 2016; Fazekas et al., 2020; Foltz et al., 2015; Kumarage et al., 2022; Peter et al., 2015; Rowland et al., 2012; Savage et al., 2003), and two have investigated priming only under lexical overlap (Buckle et al., 2017; Savage et al., 2006). All but one were conducted in English. Notably, the inclusion of a single outlier (Fazekas et al., 2020: long lag lexical overlap condition) produced a significant decreasing lexical boost effect in our analysis, suggesting this interaction effect was not estimated with great certainty. We are therefore cautious in interpreting our finding that there is no change in the lexical boost effect over development.

Several factors should be considered when interpreting our finding of a developmentally stable abstract priming effect. Firstly, participants must know the structure before they can be primed (Kidd, 2012; Kumarage et al., 2022). Whilst Chang et al. (2006) predict a decreasing abstract priming effect with increasingly stable representations, they first predict an increasing priming effect as the model learns representations abstract enough to observe priming (p.261.). The linear age term in our meta-analysis would not in principle be able to detect this inverse U-shaped trajectory. However, in line with this proposal, Kumarage et al. (2022) only observed a decrease in priming from 36 to 54 months when analysing data from children who were confirmed to know the passive at 36 months (and when excluding an item effect). As further evidence, our meta-analysis did not find significant priming of datives, an alternation that takes much longer to gain productive mastery over (Donnelly et al., 2024), or priming of agent-patient order in the Tagalog Symmetrical Voice transitive, which is more complex than the transitive in European languages and thus seems to be acquired over a longer developmental trajectory (Garcia et al., 2020; Garcia & Kidd, 2020; Kidd & Garcia, 2022). Secondly, any developmental effects, but especially an increase then decrease, will be difficult to observe in samples containing a range of proficiency levels, such as those in syntactic priming studies, where age ranges are typically about 18 months (Messenger, 2022; observations included in our analysis had an average age range of 16.58 months). Indeed, cross-sectional studies have found a range of developmental trajectories (increasing: Peter et al., 2015; decreasing: Rowland et al., 2012; stable: Hsu, 2019; Messenger, 2021; Messenger, Branigan, & McLean, 2012; Messenger, Branigan, McLean, et al., 2012). Meta-analysis is essentially a cross-sectional analysis of observations of syntactic priming and so faces these same issues. Thus, any developmental trends that might exist could be obscured by differences across studies in design and participant characteristics. Overall, the developmental trajectory of priming effects is of great theoretical interest; however, the current evidence base does not provide conclusive support for any linear effects and is limited in not examining non-linear effects. We strongly encourage future research that is longitudinal and carefully considers both non-linear effects and the range of proficiency levels in the sample.

Comparing across a larger age gap, from children to adults, the implicit learning account predicts a decrease in priming magnitude (Chang et al., 2006). Interestingly, our estimate of the syntactic priming effect in children was numerically larger than that in adults (Mahowald et al., 2016). Three factors complicate the comparison between these estimates. Even when comparing across a larger age gap, if our estimate contains a mixture of developmental stages associated with varied priming magnitudes, it will not be an accurate point of comparison. Secondly, although we have compared estimates that take into account moderators of priming (e.g., greater use of between-subjects and blocked designs in children), the moderators we included and therefore the variables controlled for were not the same across the two meta-analyses. Finally, design differences remain between child and adult priming studies that could influence priming magnitude. For example, studies in children use fewer fillers and whilst the dative is the most frequently studied alternation in adults, the active-passive alternation is more frequently studied in children. Thus, although a comparison with Mahowald et al. (2016) suggests some support for a decrease in abstract priming across age, which is predicted by the implicit learning account, the two literatures differ on many different dimensions that mean a direct comparison, at this point in time, is not warranted. Future studies that compare multiple age groups and adults across multiple alternations using the same method are required.

Explicit vs Implicit Processes

An important takeaway from our findings is that study design choices can have large impacts on the magnitude of the syntactic priming effect. These impacts are likely to reflect insights into syntactic priming. Researchers' aims of reducing task demands and interference between trials have led to a higher frequency of between-subjects designs in the developmental literature than in the adult literature. Choosing to include only within-subjects designs would result in excluding a large amount of information from the literature (40 of 108 observations, see Table 3), so we chose to combine design types in our analysis. Morris and DeShon (2002) describe three main issues with combining effect size estimates from within- and between-subjects designs. Two are statistical: a common metric must be used as the effect size, and variance estimates must be calculated using design-specific methods. In order to satisfy these conditions, we used the log odds ratio and the Becker-Balagtas method to adjust variance estimates for within-subjects designs. The third issue is conceptual: do the two designs estimate the same effect? Whilst both designs answer the question of whether presenting primes increases the frequency of that structure in subsequent speech, the large difference in effect size suggests additional factors. Messenger (2021) points out that between-subjects designs model only one structure rather than providing varied input. Although child directed language has a degree of 'burstiness' (Lester et al., 2022), the kind of flooding that occurs in between-subjects priming designs does not realistically reflect the input.

We suggest that this flooding may lead to explicit processes that point children in the direction of using the structure modelled and artificially inflate the implicit priming effect. Firstly, children may invoke explicit memory. Chang et al. (2006) propose that implicit learning effects underlie long-term priming, whilst the immediate lexical boost effect relies on explicit memory. Hartsuiker et al. (2008) show evidence supporting this prediction: abstract priming was persistent, whilst the lexical boost was not. They also proposed that both implicit and explicit processes are required to explain such findings. Interestingly, in their data, the prime remaining visible on-screen after the target was presented also appeared to increase the magnitude of priming. This suggests that not only lexical overlap but other factors can act as retrieval cues, which facilitate the explicit memory mechanism and increase the syntactic priming effect size. If cues to explicit memory can facilitate priming, part of the effect is likely attributed to explicit memory. Unvaried input in between-subjects designs may act as a retrieval cue, facilitating explicit memory processes and increasing priming. Secondly, as suggested by an anonymous reviewer, such input may play a normative role, indicating to children that the structure used by the adult experimenter is the one that *should* be used in the context of the task. Given the focus on syntactic priming as implicit learning in children, we recommend the consideration of potential explicit sources of priming when designing studies. For example, the use of games such as *bingo* and *snap* makes the task more enjoyable for children, but may introduce some reliance on explicit memory for the prime structure.

Prime repetition could feasibly act as a retrieval cue in syntactic priming studies but, as in adults (Mahowald et al., 2016), was not a significant effect in the meta-analysis. Several studies of less experienced speakers have directly manipulated prime repetition and found an effect (L2 speakers: Gámez & Vasilyeva, 2015; 3yos: Shimpi et al., 2007; late acquired structure: Gámez & Shimpi, 2016). Other studies directly manipulating the variable have found no effect (Hsu, 2014; Huttenlocher et al., 2004). We included the interaction between prime repetition and age, which was not significant. Though the wide age ranges within studies make the effect more difficult to interpret, this result does not support prime repetition acting as a retrieval cue only for less experienced speakers. This finding also supports shared representation between comprehension and production (Bock et al., 2007).

Structural Alternation

Our findings regarding the structure studied raise several important points. The first regards the relative frequencies of structures in a language. Though there were only two studies of the relative clause/adjective-noun alternation, it showed a much larger priming effect than the passive. One explanation for this potential difference could be a larger difference in frequencies of adjective-noun vs relative clauses as compared to actives vs passives. That is, adjective-noun combinations are highly preferred in languages like English, whereas a relative clause structure may be more marked, even than the passive. Secondly, as previously mentioned, children must have acquired the structure to demonstrate priming of it (Kidd, 2012). The subject relative clause is early-acquired and the larger priming effect could additionally be attributed to the fact that most if not all children in the studies had sufficient mastery to display priming compared to only some in the case of the passive. At the other end of the scale, smaller and non-significant priming of the Tagalog symmetrical voice transitive may reflect the protracted acquisition of this comparatively complex set of structures (for evidence of restrictions on priming with the Tagalog symmetrical voice system see Garcia et al., 2023). Related to the acquisition of structure is the acquisition of structural *alternation*. Whereas the English active-passive alternation occurs across a broad category of transitive verbs, children must learn restrictions on which verbs participate in the dative alternation (Gropen et al., 1989). Learning such restrictions may delay the demonstration of syntactic priming. That is, even when children have acquired the double object dative for particular lexical items (e.g. "*Give me* __"), they may not have abstracted this frame, and therefore the alternation to other lexical items. The dative, in both forms, is acquired at least as early as the passive in English (Campbell & Tomasello, 2001; Marchman et al., 1991), yet we found significant abstract priming for the active-passive alternation but not the dative alternation. Therefore, an important consideration in syntactic priming studies is the age of participants and the age of acquisition of both the structure and structural alternation.

As a final point regarding the varying effect of structural alternations, we note that of the studies included in this meta-analysis 71% were conducted in English, and 58% on the active-passive alternation (see Table 3). Therefore, using findings of syntactic priming studies in children to support broad claims about syntax acquisition requires caution. For example, in English there are few structures that allow a patient-focus (e.g., the passive and object cleft), whereas languages that have more flexible word order, such as Russian and Spanish, permit the use of alternative structures that achieve a functionally equivalent outcome (e.g., the Spanish middle voice, or object fronting in Russian). These alternatives can also be primed by passives, suggesting that the locus of priming is not purely syntactic (Gámez et al., 2009; Vasilyeva & Waterfall, 2012). Overall, the dominance of English as a target language is consistent with the broader sampling bias found in the language acquisition literature (Kidd & Garcia, 2022). Given our finding that priming effects vary with structure (and language), researchers should be careful when considering how their findings from one language generalise across languages, if at all.

Other Study Design Variables

Animacy configurations moderated the priming effect, with canonical animacy patterns resulting in a higher priming effect. There has been a long-standing tendency in the psycholinguistic literature to neutralise animacy cues in experimental contexts, under the assumption that such cues are not revealing about purely syntactic processes. This tendency is not restricted to priming studies; for instance, it is common in studies of syntactic processing (e.g., relative clause comprehension, Gibson, 2000; see Kidd et al., 2007; Mak et al., 2002). Our finding that priming is higher in instances of canonical animacy configurations suggests that children's structural knowledge is best revealed when their input-based expectations about syntactic-semantic correlations for different argument roles are met.

Two other design features of the developmental priming studies did not have an influence on priming. Firstly, comparing to a no-prime baseline was uncommon and did not impact on the magnitude of priming, and neither did using a blocked rather than alternating design (c.f., Hsu, 2019). However, given that between-subjects studies showed considerably larger priming than within-subjects studies, a study which measures baseline responses, exposes participants to primes of only one structure, and then measures primed responses is subject to the issues considered by Messenger (2021) and in our section on explicit vs implicit processes.

Finally, a long lag between primes and targets did not affect the syntactic priming effect. However, there were few and heterogenous observations of priming under this condition (1 hour lag: Hsu, 2019; 1 week lag: Kidd, 2012; 1 week to 1 month lag: Savage et al., 2006). The 95% confidence interval for the predicted priming effect in this condition overlapped with both no priming effect and the main priming effect (Figure 4a). This confirms that the effect was not precisely estimated and so conclusions about long-term priming in children are preliminary.

Considering Variability

The assumptions of meta-analysis focus on moderators of the size of the effect. However, variability in the effect can also change over development. Kumarage et al. (2022) found considerable variation in 3-year-olds ability to be primed but more reliable priming at older ages. This kind of effect can tell us something about acquisition but is not well captured in a meta-analysis, such as the link between priming and proficiency. We encourage future research, particularly longitudinal research, to consider other impacts of moderators. Given the expense of longitudinal research and our power analyses, this might require multi-lab studies such as those conducted by the *Many Babies* consortium (Frank et al., 2020).

Power in the Syntactic Priming Literature

The power analysis showed that syntactic priming studies are generally underpowered, especially for the types of questions they usually aim to answer. Whilst studies typically have sufficient power to detect a priming effect under lexical overlap (because this is a large effect), for the abstract priming effect, an average study with about 40 participants and 8 verbs presented twice each has only 67% power. In a scenario with high heterogeneity between subjects and items, the power of an average study drops even further to 75% for priming with lexical overlap and 36% for abstract priming (though this may reflect a lower limit on power, given the high levels of between-participant and between-item heterogeneity assumed). Importantly, our power analyses assume a priming effect of the magnitude of priming in the active-passive alternation. Smaller priming effects would require larger samples to achieve the same power. Also important to note is that whilst the average overall sample is close to 50, average samples per cell are in reality closer to 30. This means that analyses within a cell, for example of each age group separately, have lower power to detect a priming effect.

Although some studies aim to examine whether the priming effect exists at all, many aim to test hypotheses about potential moderators of the effect. Mahowald et al. (2016) found low power and lower evidential value for interactions than the main effects in adults. Correspondingly, we found that current studies in children are seriously underpowered for this purpose, precluding the detection of all but large interaction effects. These results are consistent with recommendations from statisticians that interactions may require between 4 and 16 times more data to estimate than main effects (Gelman et al., 2020, pp. 301-304). This threatens the validity of research findings, for several reasons. Low power can result in inconclusive results and high Type II error rates, where we fail to reject the null hypothesis. Additionally, in low powered studies, significant results are not representative of the true effect size, resulting in the possibility that studies could overestimate the underlying effect – Type M error – or even find an effect in the opposite direction – Type S error. Our simulations suggest that significant results at current sample sizes may overestimate the magnitude of true effect sizes. When combined with publication bias, this is a serious problem. Meta-analysis can mitigate Type II errors by aggregating results to gain greater power, but if the sample of studies is not in fact representative, it will be aggregating exaggerated or invalid findings. We did find that, once accounting for moderators, there is no evidence for publication bias in the priming effect. However, this does not imply whether the effects of moderators themselves are subject to publication bias. This raises concerns regarding the quality of current evidence for moderator effects: given power estimates, reported significant effects may be overstated. Future research requires much higher power than currently achieved, but, encouragingly, more recent studies do contain substantially larger samples than earlier research (e.g., Fazekas et al., 2020; Garcia & Kidd, 2020; Messenger, 2021).

Researchers typically conceive of increasing power by increasing sample size. However, we are attempting to generalise across the language from a selection of items, as well as across the population from a selection of participants. Thus, another way to increase power is to increase items, which is often overlooked. In most alternations tested (excluding the N-Adj/RC alternation), the relevant number of items in a priming study is the number of target verbs presented to participants rather than the number of trials. This is because, all other things being equal, we would expect responses to vary more strongly by verb, which differ in their associations with various grammatical structures. This poses a problem in developmental studies: there are fewer verbs that are well known by children and appropriate for experimental use, especially for a structure like the dative (although for verb-specific effects in the active-passive alternation see Kumarage et al., 2022). However, if possible, we encourage researchers to consider increasing the number of verb types (vs tokens) as well as the number of participants. For the abstract priming effect, an increase from 8 to 12 unique verbs requires 40 participants for 80% power, while using 8 unique verbs requires 80 participants for 80% power (assuming average heterogeneity).

While our power analyses yielded important information about the existing evidence base, they necessarily rely on several assumptions which may by imperfect representations of reality. First, we chose an intercept value of –2 on the logit scale, which corresponds to an average probability of producing the target construction of .12. We believe this is justifiable given that children often produce the target construction at near-floor levels (indeed large numbers often produce 0 instances of the target construction; see Donnelly et al., 2024; Kumarage et al., 2022). However, it may be easier to reliably detect moderation of the priming effect if the baseline rate of production is higher. Second, consistent with the widespread use of mixed models, we assumed Gaussian random effects. In reality, this is very unlikely to be true for participants. In empirical studies, we have observed that large numbers of children often fail to produce the target structure at all, with one possible explanation being that they have not acquired it (Donnelly et al., 2024; Kumarage et al., 2022), meaning that the true distribution is probably bimodal. It is unclear what the consequences of this would be for power estimation.

Reporting Recommendations

Our meta-analysis included a wide range of studies that studied different structures and languages, and which made a range of methodological choices. Accordingly, our analysis could be considered a broad quantitative overview of the literature. Meta-analyses could be applied in the future to more specific questions. For example, meta-analyses of particular structural alternations could investigate more specific moderators like particular animacy configurations and thematic role structures in passives (Messenger, Branigan, McLean, et al., 2012; Vasilyeva & Gámez, 2015), or verb-bias effects in datives (Peter et al., 2015). To aid in this endeavour we suggest reporting the following information. Firstly, a table of cell count data separated by condition is required. That is, the number of dependent, alternate and *other* responses in the primed and unprimed condition at each level of each variable manipulated in the study (e.g., Table 1 in Branigan & McLean, 2016). We recommend reporting raw numbers instead of or in addition to proportions: proportions can be calculated from raw numbers but if excluded trials are not reported by condition, the reverse is not always true. Most researchers excluded *other* responses from their analysis but we still recommend reporting their raw frequencies. Regarding this analysis decision, see Appendix C for a comparison between meta-analyses run with and without *other* responses. Including *other* responses appears to numerically reduce the priming effect and potentially the power to detect moderators. Secondly, for within-subjects studies, we recommend reporting the correlation between dependent structure production in the primed and unprimed condition for each level of the variable manipulated in the study. We report how we calculated an estimate of this correlation in Appendix B. Thirdly, when moderators are manipulated within-subjects, reporting the correlation between priming in different experimental conditions allows dependent observations to be included in a meta-analysis. We report how we calculated an estimate for dependent observations at different levels of lexical overlap, and structure studied in Appendix B. Finally, the recent move towards making raw data accessible online allows researchers to calculate both these and other potentially useful statistics themselves (Fazekas et al., 2020; Garcia & Kidd, 2020).

Conclusion

In this paper we reported, to our knowledge, the first meta-analysis of syntactic priming in children. We found evidence of a medium-to-large priming effect, which appears to be influenced by but not dependent upon several moderating variables, including lexical and semantic factors of content words and methodological choices made by different researchers. Therefore, like in adults (Mahowald et al., 2016), syntactic priming is a robust though variable phenomenon in children. Additionally, it fares well as a reliable effect against other meta-analysis estimates in other domains of language development (Bergmann et al., 2018). These features make it an important empirical phenomenon bridging language acquisition and adult language processing. However, studies using syntactic priming to answer theoretical questions involving moderators of the effect are currently seriously underpowered, limiting the reliability of their findings. At the same time, we found that the current evidence base is limited in several ways, which prevents us from generalising from the data to acquisition in general. Notably, we found that the majority of studies focused on English and the active-passive structural alternation, suggesting that investigating a wider array of languages and structures is an important priority (Atkinson, 2022; Kidd & Garcia, 2022). Additional future directions include the careful investigation of developmental effects in priming and the nature and persistence of long-term priming effects.

Appendix A

References for Included Papers

Allen, M. L., Haywood, S., Rajendran, G., & Branigan, H. (2011). Evidence for syntactic alignment in children with autism. *Developmental Science*, *14*(3), 540–548. https://doi.org/10.1111/j.1467-7687.2010.01001.x

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, *59*(1), 97–113. https://doi.org/10.1016/j.jml.2007.12.007

Bidgood, A., Pine, J. M., Rowland, C. F., & Ambridge, B. (2020). Syntactic representations are both abstract and semantically constrained: Evidence from children's and adults' comprehension and production/priming of the English passive. *Cognitive Science*, *44*(9). https://doi.org/10.1111/cogs.12892

Branigan, H. P., & McLean, J. F. (2016). What children learn from adults' utterances: An ephemeral lexical boost and persistent syntactic priming in adult-child dialogue. *Journal of Memory and Language*, *91*, 141–157. https://doi.org/10.1016/j.jml.2016.02.002

Branigan, H. P., McLean, J. F., & Jones, M. (2005). A blue cat or a cat that is blue? - Evidence for abstract syntax in young children's noun phrases. In A. Brugos, M. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development* (pp. 109–121). Cascadilla Press.

Branigan, H. P., & Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. *Cognition*, *157*, 250–256. https://doi.org/10.1016/j.cognition.2016.09.004

Buckle, L., Lieven, E., & Theakston, A. L. (2017). The effects of animacy and syntax on priming: A developmental study. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02246

Fazekas, J., Jessop, A., Pine, J., & Rowland, C. (2020). Do children learn from their prediction mistakes? A registered report evaluating error-based theories of language acquisition. *Royal Society Open Science*, *7*(11), Article 180877. https://doi.org/10.1098/rsos.180877

Foltz, A., Thiele, K., Kahsnitz, D., & Stenneken, P. (2015). Children's syntactic-priming magnitude: Lexical factors and participant characteristics. *Journal of Child Language*, *42*(4), 932–945. https://doi.org/10.1017/S0305000914000488

Gámez, P. B., & Shimpi, P. M. (2016). Structural priming in Spanish as evidence of implicit learning. *Journal of Child Language*, *43*(1), 207–233. https://doi.org/10.1017/S0305000915000161

Gámez, P. B., Shimpi, P. M., Waterfall, H. R., & Huttenlocher, J. (2009). Priming a perspective in Spanish monolingual children: The use of syntactic alternatives. *Journal of Child Language*, *36*(2), 269–290. https://doi.org/10.1017/S0305000908008945

Gámez, P. B., & Vasilyeva, M. (2015). Increasing Second Language Learners' Production and Comprehension of Developmentally Advanced Syntactic Forms. *Language Learning and Development*, *11*(2), 128–151. https://doi.org/10.1080/15475441.2014.894873

Garcia, R., & Kidd, E. (2020). The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development*, *16*(4), 399–425. https://doi.org/10.1080/15475441.2020.1814780

Garraffa, M., Coco, M. I., & Branigan, H. P. (2018). Impaired implicit learning of syntactic structure in children with developmental language disorder: Evidence from syntactic priming. *Autism & Developmental Language Impairments*, *3*, 2396941518779939. https://doi.org/10.1177/2396941518779939

Hopkins, Z. L., & Branigan, H. P. (2020). Children show selectively increased language imitation after experiencing ostracism. *Developmental Psychology*, *56*(5), 897–911. https://doi.org/10.1037/dev0000915

Hopkins, Z., Yuill, N., & Keller, B. (2016). Children with autism align syntax in natural conversation. *Applied Psycholinguistics*, *37*(2), 347–370. https://doi.org/10.1017/S0142716414000599

Hsu, D.-B. (2014a). Mandarin-speaking three-year-olds' demonstration of productive knowledge of syntax: Evidence from syntactic productivity and structural priming with the SVO-ba alternation. *Journal of Child Language*, *41*(5), 1115–1146. https://doi.org/10.1017/S0305000913000408

Hsu, D.-B. (2014b). Structural priming as learning: Evidence from mandarin-learning 5-year-olds. *Language Acquisition: A Journal of Developmental Linguistics*, *21*(2), 156–172. https://doi.org/10.1080/10489223.2014.884571

Hsu, D.-B. (2019). Children's adaption to input change using an abstract syntactic representation: Evidence from structural priming in Mandarin-speaking preschoolers. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02186

Hupp, J. M., & Jungers, M. K. (2009). Speech priming: An examination of rate and syntactic persistence in preschoolers. *British Journal of Developmental Psychology*, *27*(2), 495–504. https://doi.org/10.1348/026151008X345988

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*(2), 182–195. https://doi.org/10.1016/j.jml.2003.09.003

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, *33*(2), 393–418. https://doi.org/10.1017/S0142716411000415

Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring children. *Journal of Memory and Language*, *127*, Article 104374. https://doi.org/10.1016/j.jml.2022.104374

Manetti, C. (2013). On the production of passives in Italian: Evidence from an elicited production task and a syntactic priming study with preschool children. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *Proceedings of the 37th Boston University conference on language development online supplement* (pp. 1–16). Boston University.

Messenger, K. (2021). The persistence of priming: Exploring long-lasting syntactic priming effects in children and adults. *Cognitive Science*, *45*(6), Article e13005. https://doi.org/10.1111/cogs.13005

Messenger, K., Branigan, H., McLean, J., & Sorace, A. (2008). English-speaking children's early passives: Evidence from syntactic priming. In H. Chan, H. Jacob, & E. Kapia (Eds.), *Proceedings of the 32nd Annual Boston University Conference on Language Development, Vols 1 and 2* (pp. 275–286). Cascadilla Press.

Messenger, K., Branigan, H. P., & McLean, J. F. (2012). Is children's acquisition of the passive a staged process? Evidence from six- and nine-year-olds' production of passives. *Journal of Child Language*, *39*(5), 991–1016. https://doi.org/10.1017/S0305000911000377

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, *66*(4), 568–587. https://doi.org/10.1016/j.jml.2012.03.008

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63. https://doi.org/10.1016/j.cognition.2012.06.008

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557–567. https://doi.org/10.1111/1467-7687.00312

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, *2*(1), 27–49. https://doi.org/10.1207/s15473341lld0201_2

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3- and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, *43*(6), 1334–1346. https://doi.org/10.1037/0012-1649.43.6.1334

Vasilyeva, M., & Gámez, P. B. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in syntactic priming. *Journal of Experimental Child Psychology*, *138*, 15–30. https://doi.org/10.1016/j.jecp.2015.04.009

Vasilyeva, M., & Waterfall, H. (2012). Beyond syntactic priming: Evidence for activation of alternative syntactic structures. *Journal of Child Language*, *39*(2), 258–283. https://doi.org/10.1017/S0305000911000055

Whitehurst, G. J., Ironsmith, M., & Goldfein, M. (1974). Selective imitation of the passive construction through modeling. *Journal of Experimental Child Psychology*, *17*(2), 288–302.

Wolleb, A., Sorace, A., & Westergaard, M. (2018). Exploring the role of cognitive control in syntactic processing: Evidence from cross-language priming in bilingual children. *Linguistic Approaches to Bilingualism*, *8*(5), 606–636. https://doi.org/10.1075/lab.17002.wol

Appendix B

Calculating Correlations

We provide the analysis code for calculating correlations on our OSF page. To estimate required correlations, we mostly used raw data from the *Canberra Longitudinal Child Language* study (Donnelly et al., 2024; Kidd et al., 2018; Kumarage et al., 2022). This consisted of longitudinal priming data for the passive alternation with and without lexical overlap at 4 timepoints (36m, 42m, 48m, and 54m) and for the dative alternation with and without lexical overlap at 3 timepoints (42m, 48m, 54m). Compared to other studies in the literature, the study has a large sample. Over 100 children participated in the study, although not all provide data at every timepoint in the priming tasks.

Correlation between Primed and Unprimed Conditions for Within-Subjects Studies

We calculated the proportion of dependent responses (either passive or double object dative) produced by each participant at each timepoint, for each structural alternation in the lexical overlap and abstract priming conditions. We then calculated the correlation between production in the primed and unprimed conditions for each condition at each timepoint (see Table B1). We used the average of these 14 correlation coefficients as the estimated correlation between dependent structure production in the primed and unprimed condition.

Table B1

Correlations between primed and unprimed conditions in the Canberra Longitudinal Child Language study

Condition			Correlation
36m	Passive	Lexical overlap	.40
		No lexical overlap	.17
42m	Passive	Lexical overlap	.07
		No lexical overlap	.38
	Dative	Lexical overlap	.39
		No lexical overlap	.30
48m	Passive	Lexical overlap	.24
		No lexical overlap	.45
	Dative	Lexical overlap	.53
		No lexical overlap	.61
54m	Passive	Lexical overlap	.19
		No lexical overlap	.41
	Dative	Lexical overlap	.49
		No lexical overlap	.27
Average			.35

Correlation between Lexical Overlap Conditions

We subtracted the proportion of dependent responses in the unprimed condition from the proportion in the primed condition to calculate a priming effect for each participant in each lexical overlap condition, for each structural alternation at each timepoint. We then calculated the correlation between priming with and without lexical overlap for each structural alternation at each timepoint (see Table B2). We used the average of these seven correlation coefficients as our estimate.

Table B2

Correlations between lexical overlap conditions in the Canberra Longitudinal Child Language study

Condition		Correlation
36m	Passive	.06
42m	Passive	.13
	Dative	.21
48m	Passive	.19
	Dative	.25
54m	Passive	.36
	Dative	.17
Average		.20

Correlation between Passive and Dative Priming

We again used the proportion of dependent responses in the unprimed condition subtracted from the proportion in the primed condition to calculate a priming effect for each participant in each condition. The correlation between this priming effect in the passive and dative alternation was calculated at both levels of lexical overlap for the three timepoints where both alternations were tested (see Table B3). We used the average of these six coefficients as our estimate.

Table B3

Correlations between passive and dative priming in the Canberra Longitudinal Child Language study

Condition		Correlation
Lexical overlap	42m	.24
	48m	.17
	54m	.24
No lexical overlap	42m	.15
	48m	.21
	54m	-.14
Average		.15

Correlation between Passive and Dative Priming under Different Lexical Overlap Conditions

Using the calculated priming effects for each participant, we calculated the correlation between priming at different levels of both structure and lexical overlap (see Table B4). We used the average of these six coefficients as our estimate.

Table B4

Correlations between passive and dative priming under different lexical overlap conditions in the Canberra Longitudinal Child Language study

Condition		Correlation
Dative overlap* passive abstract	42m	.11
	48m	.32
	54m	.01
Passive overlap* dative abstract	42m	-.07
	48m	.09
	54m	-.10
Average		.06

Correlation between Timepoints

The correlation estimate required by *vcalc()* in *metafor* is of the correlation between passive production at adjacent timepoints (e.g. baseline-test correlation; test-posttest correlation) rather than of the correlation between priming (e.g. correlation between baseline-test and baseline-posttest priming effects). Table B5 displays the appropriate correlations from two studies. Kidd (2012) provided correlations between passive production in different test blocks. Fazekas et al. (2020) made their raw data publicly accessible so we could calculate the proportion of double object datives produced by each participant in each test block and calculate correlations. We excluded the posttest 2 condition from Fazekas et al. (2020) because it is a lexical overlap condition unlike the other conditions in the study. We used the average of these five coefficients as our estimate.

Table B5

Correlations between production of the dependent structure in different test blocks

Condition		Correlation
Kidd (2012)	Baseline-test	.07
	Test-posttest	.48
	Posttest-long posttest	.01
Fazekas et al. (2020)	Baseline-test	.54
	Test-posttest	.65
Average		.35

Appendix C

Analyses Including and Excluding *Other* Responses

Not all productions in syntactic priming experiments can be classified as the dependent or alternate structure in the alternation being studied. “Other” responses can make up a large proportion of responses (e.g., over 50% of responses from 3 year olds in Shimpi et al., 2007) or a very small proportion in stem-completion studies (e.g. 0-5% in Garcia & Kidd, 2020 and Rowland et al., 2012). Researchers can choose to include *other* responses in the denominator when calculating the proportion of dependent responses or exclude *other* responses and restrict analyses to only structures within the alternation being studied (see Bencini & Valian, 2008 for discussion). A subset of 31 studies reported the frequency of *other* responses in each condition. We ran analyses for the 94 observations from these studies to compare the analysis choice of including vs excluding *other* responses. The analysis script is available on our OSF page. Table C1 shows that including other responses numerically reduces the overall effect. In addition, including *other* responses results in less heterogeneity between effect sizes. Although the amount of sampling variance is similar, the amount attributed to within- and between-study heterogeneity is reduced.

Table C1

Results of overall multilevel meta-analytic models including and excluding other responses

	N responses		N subjects	
	Other excluded	Other included	Other excluded	Other included
Estimate	1.42 1.16 | 1.68 p < .001***	1.22 1.00 | 1.44 p < .001***	1.27 0.98 | 1.57 p < .001***	1.07 0.81 | 1.34 p < .001***
Odds ratio	4.12	3.38	3.57	2.92
Cohen’s d	0.78	0.67	0.70	0.59
Q	615.72 df = 93 p < .001***	539.66 df = 93 p < .001***	135.34 df = 93 p < .01**	104.20 df = 93 p = .201
Variance	0.673	0.524	0.893	0.776
Sampling	0.108	0.094	0.533	0.544
I^2	0.565	0.430	0.360	0.232

Table C2 shows results for the models using N responses to calculate variance and including moderators. We report results excluding the lexical overlap with lag condition from Fazekas et al. (2020) and the dative priming condition from Hopkins et al. (2016), both of which were identified as outliers with large effects on the results. Regardless of *other* response inclusion, the moderators accounted for a significant proportion of variance (*other* excluded: F(18,12) = 5.91, p < .01**; *other* included: F(18,12) = 5.99, p < .01***) but did not eliminate residual heterogeneity (*other* excluded: Q(73) = 153.26, p < .001***; *other* included: Q(73) = 155.33, p < .001***). Compared to the full sample of studies, additional moderators significantly predict the priming effect in the subset of studies that reported *other* responses. However, we focus on the difference between including and excluding *other* responses. The results are very similar but there is a trend towards numerically smaller model coefficients and less significant effects for moderators when including *other* responses. Therefore, there may be less power to detect moderators of priming when including *other* responses. This may be related to there being less total variance to explain when including *other* responses (Table C1).

Table C2

Results of N responses moderator models including and excluding other responses

	Excluding other responses			Including other responses		
	β	CI	p	β	CI	p
Intercept	0.87	0.48 | 1.26	<.001***	0.78	0.46 | 1.10	<.001***
Between-subjects	1.13	0.67 | 1.58	<.001***	0.85	0.47 | 1.24	<.001***
Baseline	-0.42	-0.94 | 0.10	.112	-0.23	-0.67 | 0.20	.287
Animacy						
Uncontrolled	-0.30	-0.89 | 0.29	.288	-0.22	-0.69 | 0.26	.343
Favourable	0.52	0.00 | 1.03	.049*	0.38	-0.05 | 0.81	.076^
Prime repetition	0.51	0.12 | 0.89	.011*	0.42	0.09 | 0.74	.013*
Lexical overlap	1.13	0.69 | 1.57	<.001***	0.97	0.56 | 1.39	<.001***
Structure						
Dative	-0.84	-1.23 |-0.45	<.001***	-0.77	-1.10 |-0.44	<.001***
SVO-ba	0.53	-0.24 | 1.31	.158	0.44	-0.17 | 1.05	.141
RC	1.78	0.37 | 3.19	.018*	1.71	0.45 | 2.97	.012*
SV-Transitive	-0.86	-1.65 |-0.06	.037*	-0.77	-1.39 |-0.15	.019*
Age	0.03	-0.14 | 0.19	.744	0.07	-0.07 | 0.21	.336
Lag						
0 vs block	-0.15	-0.52 | 0.23	.443	-0.07	-0.40 | 0.25	.651
0 vs long	-0.42	-1.06 | 0.23	.200	-0.23	-0.80 | 0.35	.438
Prime repetition*Age	0.49	0.06 | 0.93	.026*	0.45	0.08 | 0.83	.018*
Lexical overlap*Age	0.14	-0.42 | 0.71	.620	0.20	-0.33 | 0.74	.452
Structure*Age						
Dative	-0.46	-0.91 |-0.01	.046*	-0.52	-0.92 |-0.11	.014*
SVO-ba	-0.16	-0.62 | 0.29	.478	-0.22	-0.63 | 0.20	.302
SV-Transitive	0.05	-0.34 | 0.43	.809	0.01	-0.36 | 0.38	.957

Table C3 shows results for the models using N subjects to calculate variance and including moderators. For the N subjects models, the same observation from Fazekas et al. (2020) was excluded but the Hopkins et al. (2016) observation no longer had large impacts on the results so was left in. In both models, the moderators accounted for a significant proportion of variance (*other* excluded: F(18,12) = 5.26, p < .01**; *other* included: F(18,12) = 4.15, p < .01**), with no significant residual heterogeneity (*other* excluded: Q(74) = 35.64, p = 1.000; *other* included: Q(74) = 33.13, p = 1.000). The results of these models should be interpreted with caution: although profile likelihood plots showed clear peaks in estimating variance components, both were estimated at zero. Nevertheless, for most effects, we also see a pattern of numerically smaller effects when including *other* responses.

Table C3

Results of N subjects moderator models including and excluding other responses

	Excluding other responses			Including other responses		
	β	CI	p	β	CI	p
Intercept	0.89	0.51 | 1.27	<.001***	0.72	0.34 | 1.10	.002**
Between-subjects	1.15	0.45 | 1.85	.002**	1.04	0.33 | 1.75	.005**
Baseline	-0.22	-0.97 | 0.52	.552	-0.12	-0.89 | 0.64	.749
Animacy						
Uncontrolled	0.13	-0.70 | 0.96	.738	-0.27	-1.11 | 0.58	.507
Favourable	0.41	-0.21 | 1.02	.175	-0.04	-0.59 | 0.66	.898
Prime repetition	0.15	-0.38 | 0.69	.570	0.11	-0.43 | 0.65	.682
Lexical overlap	1.37	0.77 | 1.96	<.001***	1.26	0.67 | 1.86	<.001***
Structure						
Dative	-0.97	-1.44 | -0.50	<.001***	-1.07	-1.55 | -0.60	<.001***
SVO-ba	0.36	-0.43 | 1.14	.345	-0.07	-0.84 | 0.70	.844
RC	1.68	-0.77 | 4.12	.161	1.63	-0.75 | 4.01	.162
SV-Transitive	-0.89	-1.52 | -0.26	.009**	-0.74	-1.37 | -0.12	.024*
Age	0.07	-0.17 | 0.31	.545	0.24	-0.00 | 0.48	.053^
Lag						
0 vs block	-0.23	-0.77 | 0.32	.412	-0.17	-0.72 | 0.38	.533
0 vs long	-0.36	-1.47 | 0.76	.525	-0.12	-1.21 | 0.97	.824
Prime repetition*Age	0.14	-0.43 | 0.71	.623	0.31	-0.27 | 0.88	.289
Lexical overlap*Age	0.12	-0.65 | 0.88	.764	0.16	-0.61 | 0.93	.682
Structure*Age						
Dative	-0.01	-0.47 | 0.45	.960	0.05	-0.41 | 0.52	.820
SVO-ba	-0.21	-0.91 | 0.49	.549	-0.29	-0.96 | 0.37	.384
SV-Transitive	-0.04	-0.55 | 0.47	.880	-0.18	-0.69 | 0.33	.479

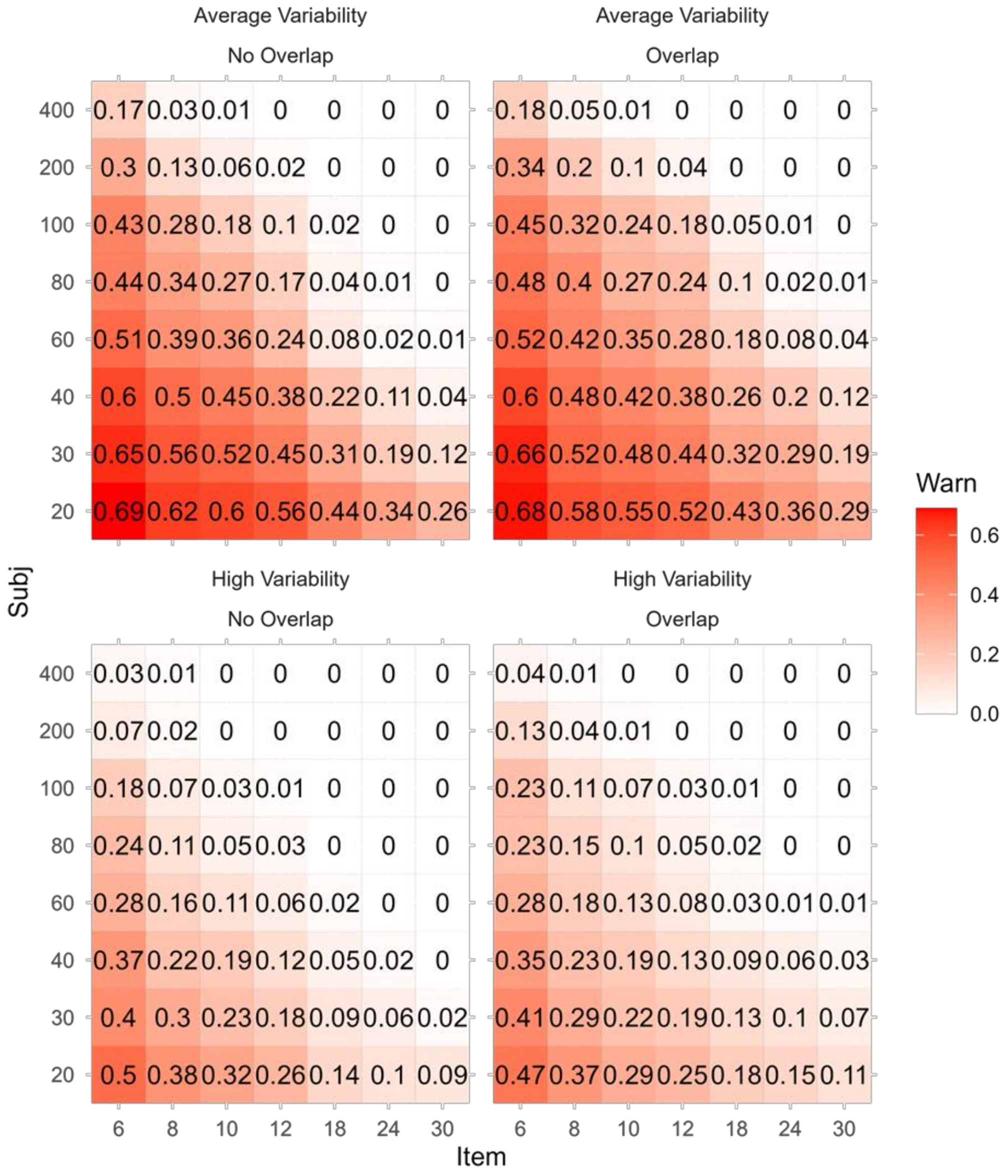
Appendix D

Additional Power Analysis Figures

Non-convergent models were included in power analyses because the full random effects structure was known a priori. Figure D1 reports the proportion of non-convergent models in power analyses for the main priming effects. Figure D2 reports the same proportion for power analyses of interaction effects. However, models with implausible estimates were excluded from power analyses because they produce significant results which may inflate estimates of power. Figures D3 and D4 report the proportion of models excluded for this reason for main and interaction effects respectively.

Figure D1

Proportions of models producing warning messages for power analyses of priming effect



Average Variability
No Overlap

	6	8	10	12	18	24	30
400	0.17	0.03	0.01	0	0	0	0
200	0.3	0.13	0.06	0.02	0	0	0
100	0.43	0.28	0.18	0.1	0.02	0	0
80	0.44	0.34	0.27	0.17	0.04	0.01	0
60	0.51	0.39	0.36	0.24	0.08	0.02	0.01
40	0.6	0.5	0.45	0.38	0.22	0.11	0.04
30	0.65	0.56	0.52	0.45	0.31	0.19	0.12
20	0.69	0.62	0.6	0.56	0.44	0.34	0.26

Average Variability
Overlap

	6	8	10	12	18	24	30
400	0.18	0.05	0.01	0	0	0	0
200	0.34	0.2	0.1	0.04	0	0	0
100	0.45	0.32	0.24	0.18	0.05	0.01	0
80	0.48	0.4	0.27	0.24	0.1	0.02	0.01
60	0.52	0.42	0.35	0.28	0.18	0.08	0.04
40	0.6	0.48	0.42	0.38	0.26	0.2	0.12
30	0.66	0.52	0.48	0.44	0.32	0.29	0.19
20	0.68	0.58	0.55	0.52	0.43	0.36	0.29

High Variability
No Overlap

	6	8	10	12	18	24	30
400	0.03	0.01	0	0	0	0	0
200	0.07	0.02	0	0	0	0	0
100	0.18	0.07	0.03	0.01	0	0	0
80	0.24	0.11	0.05	0.03	0	0	0
60	0.28	0.16	0.11	0.06	0.02	0	0
40	0.37	0.22	0.19	0.12	0.05	0.02	0
30	0.4	0.3	0.23	0.18	0.09	0.06	0.02
20	0.5	0.38	0.32	0.26	0.14	0.1	0.09

High Variability
Overlap

	6	8	10	12	18	24	30
400	0.04	0.01	0	0	0	0	0
200	0.13	0.04	0.01	0	0	0	0
100	0.23	0.11	0.07	0.03	0.01	0	0
80	0.23	0.15	0.1	0.05	0.02	0	0
60	0.28	0.18	0.13	0.08	0.03	0.01	0.01
40	0.35	0.23	0.19	0.13	0.09	0.06	0.03
30	0.41	0.29	0.22	0.19	0.13	0.1	0.07
20	0.47	0.37	0.29	0.25	0.18	0.15	0.11

Figure D2

Proportion of models producing warning messages in power analyses of interactions

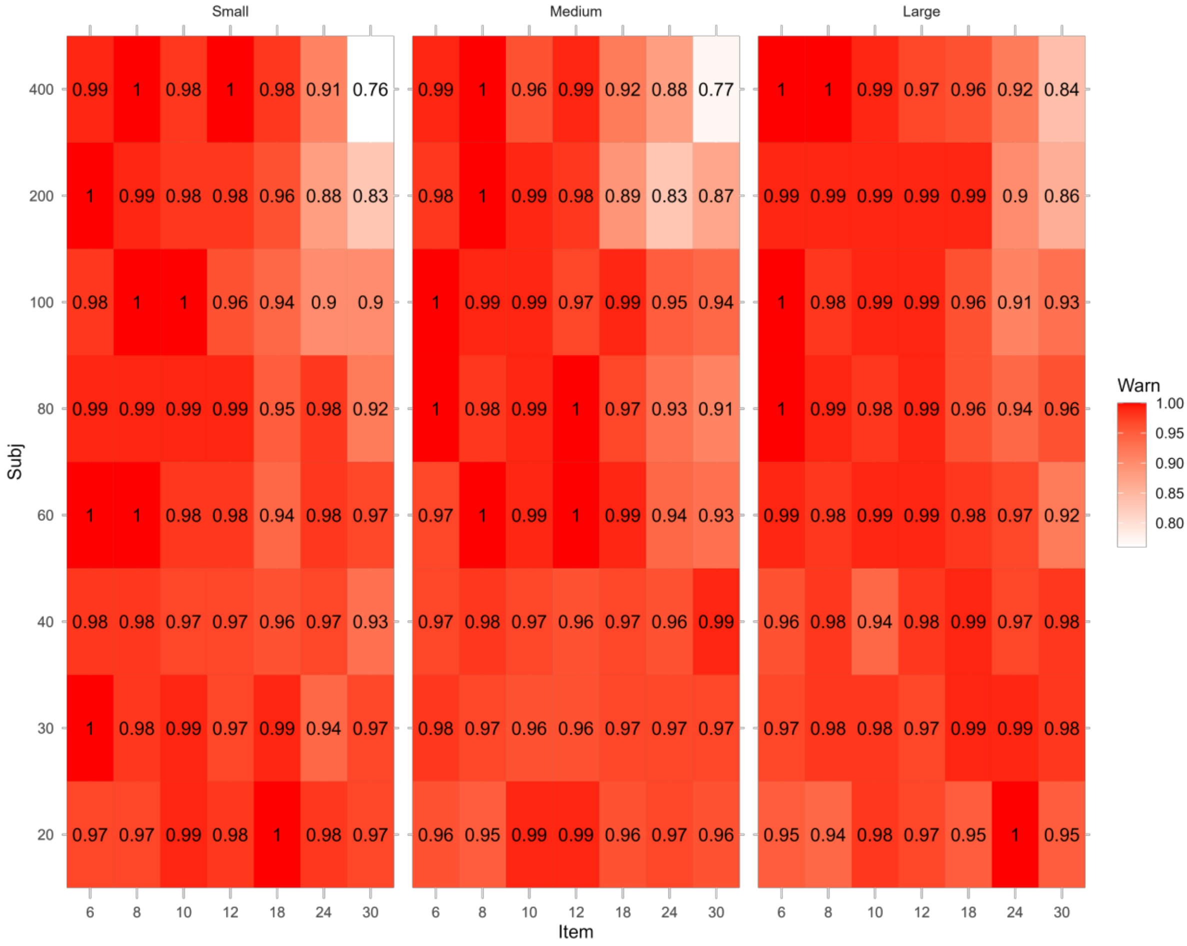
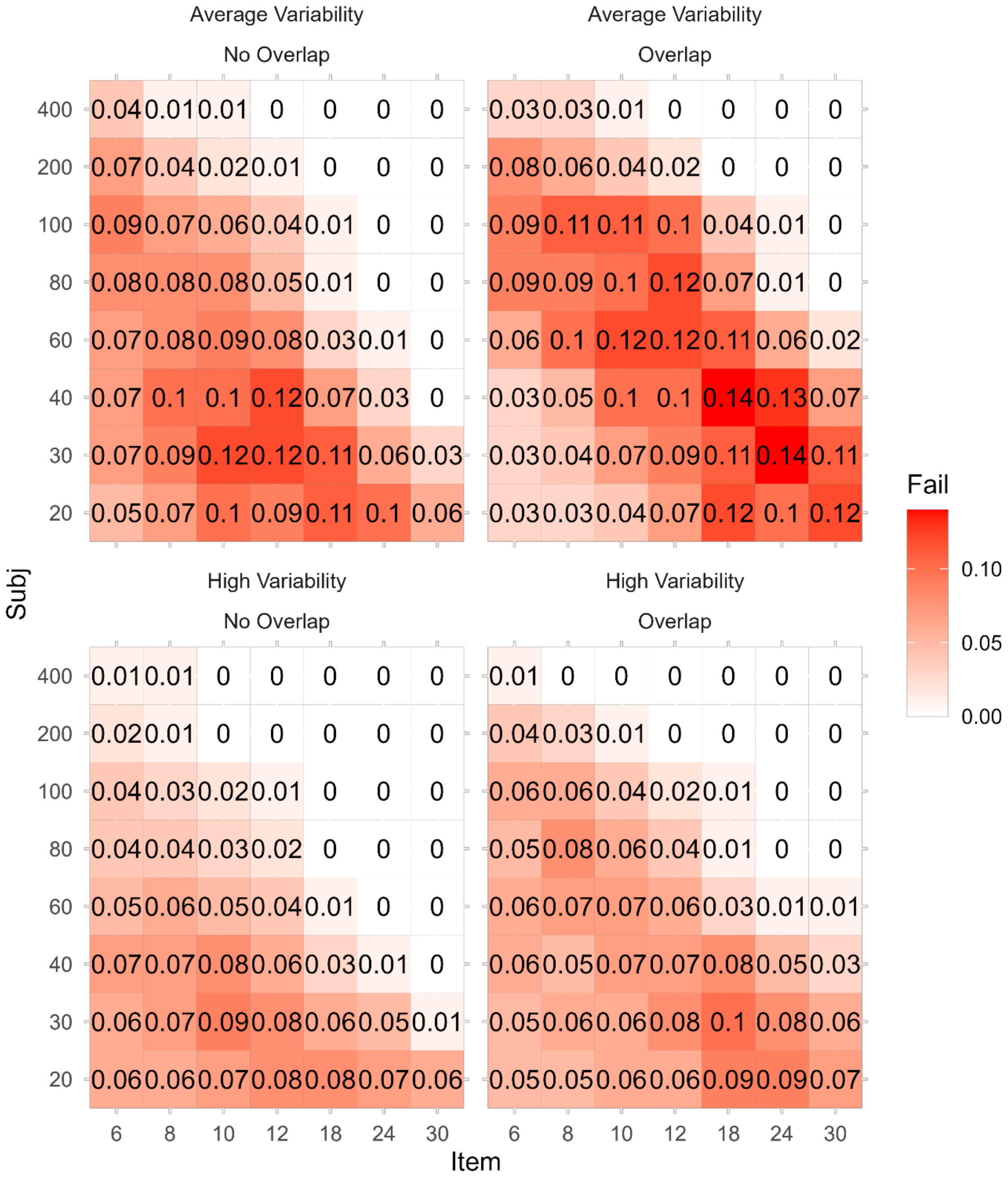


Figure D3

Proportions of models producing implausibly small standard errors (< .1) or large effect sizes (>5 or < –5) for power analyses of priming effect



Average Variability
No Overlap

Subj	6	8	10	12	18	24	30
400	0.04	0.01	0.01	0	0	0	0
200	0.07	0.04	0.02	0.01	0	0	0
100	0.09	0.07	0.06	0.04	0.01	0	0
80	0.08	0.08	0.08	0.05	0.01	0	0
60	0.07	0.08	0.09	0.08	0.03	0.01	0
40	0.07	0.1	0.1	0.12	0.07	0.03	0
30	0.07	0.09	0.12	0.12	0.11	0.06	0.03
20	0.05	0.07	0.1	0.09	0.11	0.1	0.06

Average Variability
Overlap

Subj	6	8	10	12	18	24	30
400	0.03	0.03	0.01	0	0	0	0
200	0.08	0.06	0.04	0.02	0	0	0
100	0.09	0.11	0.11	0.1	0.04	0.01	0
80	0.09	0.09	0.1	0.12	0.07	0.01	0
60	0.06	0.1	0.12	0.12	0.11	0.06	0.02
40	0.03	0.05	0.1	0.1	0.14	0.13	0.07
30	0.03	0.04	0.07	0.09	0.11	0.14	0.11
20	0.03	0.03	0.04	0.07	0.12	0.1	0.12

High Variability
No Overlap

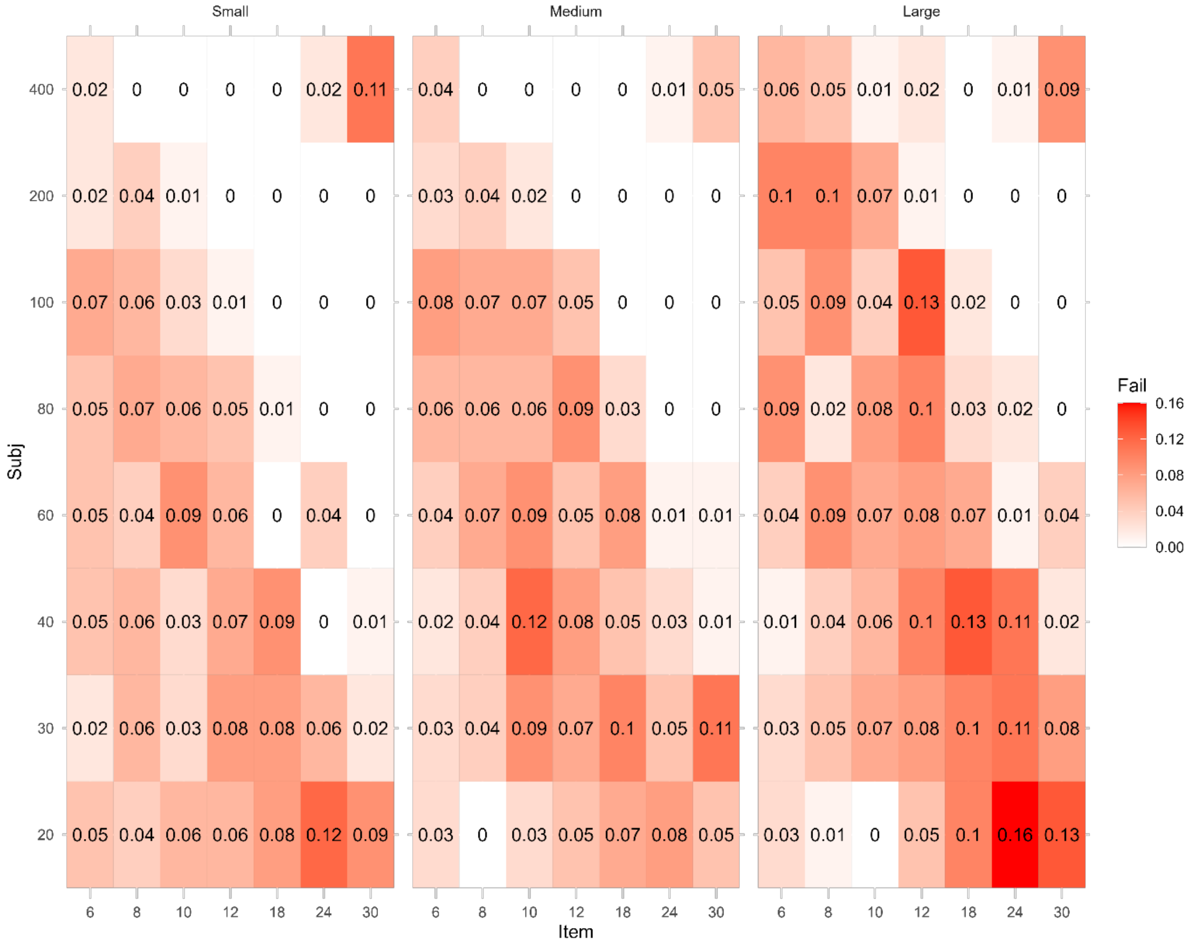
Subj	6	8	10	12	18	24	30
400	0.01	0.01	0	0	0	0	0
200	0.02	0.01	0	0	0	0	0
100	0.04	0.03	0.02	0.01	0	0	0
80	0.04	0.04	0.03	0.02	0	0	0
60	0.05	0.06	0.05	0.04	0.01	0	0
40	0.07	0.07	0.08	0.06	0.03	0.01	0
30	0.06	0.07	0.09	0.08	0.06	0.05	0.01
20	0.06	0.06	0.07	0.08	0.08	0.07	0.06

High Variability
Overlap

Subj	6	8	10	12	18	24	30
400	0.01	0	0	0	0	0	0
200	0.04	0.03	0.01	0	0	0	0
100	0.06	0.06	0.04	0.02	0.01	0	0
80	0.05	0.08	0.06	0.04	0.01	0	0
60	0.06	0.07	0.07	0.06	0.03	0.01	0.01
40	0.06	0.05	0.07	0.07	0.08	0.05	0.03
30	0.05	0.06	0.06	0.08	0.1	0.08	0.06
20	0.05	0.05	0.06	0.06	0.09	0.09	0.07

Figure D4

Proportion of models producing implausibly small standard errors or implausibly large effect sizes for power analyses of interactions



References for Chapter 3

Allen, M. L., Haywood, S., Rajendran, G., & Branigan, H. (2011). Evidence for syntactic alignment in children with autism. *Developmental Science*, *14*(3), 540–548. https://doi.org/10.1111/j.1467-7687.2010.01001.x

Ambridge, B., & Lieven, E. (2015). A Constructivist Account of Child Language Acquisition. In *The Handbook of Language Emergence* (pp. 478–510). John Wiley & Sons. https://doi.org/10.1002/9781118346136.ch22

Anderson, J. D., & Conture, E. G. (2004). Sentence-Structure Priming in Young Children Who Do and Do Not Stutter. *Journal of Speech, Language, and Hearing Research*, *47*(3), 552–571. https://doi.org/10.1044/1092-4388%282004/043%29

Atkinson, E. (2022). Sticking to what we know: Methodological limitations to generalizability. In K. Messenger (Ed.), *Syntactic priming in language acquisition: Representations, mechanisms and applications* (Vol. 31, pp. 83–106). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.31.05atk

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bartoš, F., Maier, M., Shanks, D. R., Stanley, T. D., Sladekova, M., & Wagenmakers, E.-J. (2023). Meta-analyses in psychology often overestimate evidence for and size of effects. *Royal Society Open Science*, *10*(7), 230224. https://doi.org/10.1098/rsos.230224

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious Mixed Models*. arXiv. https://doi.org/10.48550/arXiv.1506.04967

Becker, M. P., & Balagtas, C. C. (1993). Marginal Modeling of Binary Cross-Over Data. *Biometrics*, *49*(4), 997. https://doi.org/10.2307/2532242

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, *59*(1), 97–113. https://doi.org/10.1016/j.jml.2007.12.007

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting Replicability in Developmental Research Through Meta-analyses: Insights From Language Acquisition Research. *Child Development*, *89*(6), 1996–2009. https://doi.org/10.1111/cdev.13079

Bidgood, A., Pine, J. M., Rowland, C. F., & Ambridge, B. (2020). Syntactic representations are both abstract and semantically constrained: Evidence from children's and adults' comprehension and production/priming of the English passive. *Cognitive Science*, *44*(9). https://doi.org/10.1111/cogs.12892

Bidgood, A., Pine, J., Rowland, C., Sala, G., Freudenthal, D., & Ambridge, B. (2021). Verb argument structure overgeneralisations for the English intransitive and transitive constructions: Grammaticality judgments and production priming. *Language and Cognition: An Interdisciplinary Journal of Language and Cognitive Science*, *13*(3), 397–437. https://doi.org/10.1017/langcog.2021.8

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458. https://doi.org/10.1016/j.cognition.2006.07.003

Bock, K., & Griffin, Z. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology. General*, *129*(2), 177–192. https://doi.org/10.1037//0096-3445.129.2.177

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, *99*, 150–171. https://doi.org/10.1037/0033-295X.99.1.150

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Branigan, H. P., & McLean, J. F. (2016). What children learn from adults' utterances: An ephemeral lexical boost and persistent syntactic priming in adult-child dialogue. *Journal of Memory and Language*, *91*, 141–157. https://doi.org/10.1016/j.jml.2016.02.002

Branigan, H. P., McLean, J. F., & Jones, M. (2005). A blue cat or a cat that is blue? - Evidence for abstract syntax in young children's noun phrases. In A. Brugos, M. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development* (pp. 109–121). Cascadilla Press.

Branigan, H. P., & Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. *Cognition*, *157*, 250–256. https://doi.org/10.1016/j.cognition.2016.09.004

Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, *40*, Article e313. https://doi.org.10.1017/S0140525X16002028

Buckle, L., Lieven, E., & Theakston, A. L. (2017). The effects of animacy and syntax on priming: A developmental study. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02246

Campbell, A. L., & Tomasello, M. (2001). The acquisition of English dative constructions. *Applied Psycholinguistics*, *22*(2), 253–267. https://doi.org.10.1017/S0142716401002065

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–230. https://doi.org/10.1023/A:1005101313330

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass*, *6*(5), 259–278. https://doi.org/10.1002/lnc3.337

Corporation for Digital Scholarship. (2023). *Zotero* (6.0.26) [Computer software]. https://www.zotero.org

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive Advantage in Bilingualism: An Example of Publication Bias? *Psychological Science*, *26*(1), 99–107. https://doi.org/10.1177/0956797614557866

Dell, G. S., & Chang, F. (2014). The p-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634). https://doi.org/10.1098/rstb.2012.0394

Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In H. Rothstein, A. J. Sutton, & M. Borenstein, *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 9–33). Wiley Online Library. https://doi.org/10.1002/0470870168

Donnelly, S., Rowland, C., Chang, F., & Kidd, E. (2024). A comprehensive examination of prediction-based error as a mechanism for syntactic development: Evidence from syntactic priming. *Cognitive Science, 48*(4), Article e13431. https://doi.org/10.1111/cogs.13431

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fazekas, J., Jessop, A., Pine, J., & Rowland, C. (2020). Do children learn from their prediction mistakes? A registered report evaluating error-based theories of language acquisition. *Royal Society Open Science*, *7*(11), Article 180877. https://doi.org/10.1098/rsos.180877

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language*, *48*(2), 379–398. https://doi.org/10.1016/S0749-596X(02)00523-5

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition*, *82*(3), 259–278. https://doi.org/10.1016/S0010-0277(01)00159-7

Foltz, A., Thiele, K., Kahsnitz, D., & Stenneken, P. (2015). Children’s syntactic-priming magnitude: Lexical factors and participant characteristics. *Journal of Child Language*, *42*(4), 932–945. https://doi.org/10.1017/S0305000914000488

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Frank, M. C., Alcock, K. J., Arias-Trejo, N., Aschersleben, G., Baldwin, D., Barbu, S., Bergelson, E., Bergmann, C., Black, A. K., Blything, R., Böhland, M. P., Bolitho, P., Borovsky, A., Brady, S. M., Braun, B., Brown, A., Byers-Heinlein, K., Campbell, L. E., Cashon, C., … Soderstrom, M. (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24–52. https://doi.org/10.1177/2515245919900809

Gámez, P. B., & Shimpi, P. M. (2016). Structural priming in Spanish as evidence of implicit learning. *Journal of Child Language*, *43*(1), 207–233. https://doi.org/10.1017/S0305000915000161

Gámez, P. B., Shimpi, P. M., Waterfall, H. R., & Huttenlocher, J. (2009). Priming a perspective in Spanish monolingual children: The use of syntactic alternatives. *Journal of Child Language*, *36*(2), 269–290. https://doi.org/10.1017/S0305000908008945

Gámez, P. B., & Vasilyeva, M. (2015). Increasing Second Language Learners' Production and Comprehension of Developmentally Advanced Syntactic Forms. *Language Learning and Development*, *11*(2), 128–151. https://doi.org/10.1080/15475441.2014.894873

Gámez, P. B., & Vasilyeva, M. (2020). Shared syntactic representations in balanced bilinguals: Cross-linguistic priming with and without verb overlap. *Language Learning and Development*, *16*(1), 89–106. https://doi.org/10.1080/15475441.2019.1689822

Garcia, R., & Kidd, E. (2020). The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development*, *16*(4), 399–425. https://doi.org/10.1080/15475441.2020.1814780

Garcia, R., Roeser, J., & Höhle, B. (2020). Children's online use of word order and morphosyntactic markers in Tagalog thematic role assignment: An eye-tracking study. *Journal of Child Language*, *47*(3), 533–555. https://doi.org/10.1017/S0305000919000618

Garcia, R., Roeser, J., & Kidd, E. (2023). Finding your voice: Voice-specific effects in Tagalog reveal the limits of word order priming. *Cognition*, *236*, Article 105424. https://doi.org/10.1016/j.cognition.2023.105424

Garraffa, M., Coco, M. I., & Branigan, H. P. (2015). Effects of immediate and cumulative syntactic experience in language impairment: Evidence from priming of subject relatives in children with SLI. *Language Learning and Development*, *11*(1), 18–40. https://doi.org/10.1080/15475441.2013.876277

Garraffa, M., Coco, M. I., & Branigan, H. P. (2018). Impaired implicit learning of syntactic structure in children with developmental language disorder: Evidence from syntactic priming. *Autism & Developmental Language Impairments*, *3*, Article 2396941518779939. https://doi.org/10.1177/2396941518779939

Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and other stories*. Cambridge University Press.

Gibson, E. (2000). The Dependency Locality Theory. In A. P. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium* (pp. 95–126). MIT Press. https://doi.org/10.7551/mitpress/3654.003.0008

Goodman, S. N. (1991). Have you ever meta-analysis you didn't like? *Annals of Internal Medicine*, *114*(3), 244–246. https://doi.org/10.7326/0003-4819-114-3-244

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The Learnability and Acquisition of the Dative Alternation in English. *Language*, *65*(2), 203–257. https://doi.org/10.2307/415332

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, *20*(2), 219–234. https://doi.org/10.1017/S1366728915000164

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238. https://doi.org/10.1016/j.jml.2007.07.003

Hopkins, Z., Yuill, N., & Keller, B. (2016). Children with autism align syntax in natural conversation. *Applied Psycholinguistics*, *37*(2), 347–370. https://doi.org/10.1017/S0142716414000599

Hsu, D.-B. (2014). Structural priming as learning: Evidence from mandarin-learning 5-year-olds. *Language Acquisition: A Journal of Developmental Linguistics*, *21*(2), 156–172. https://doi.org/10.1080/10489223.2014.884571

Hsu, D.-B. (2019). Children's adaption to input change using an abstract syntactic representation: Evidence from structural priming in Mandarin-speaking preschoolers. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02186

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*(2), 182–195. https://doi.org/10.1016/j.jml.2003.09.003

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kadlec, D., Sainani, K. L., & Nimphius, S. (2022). With Great Power Comes Great Responsibility: Common Errors in Meta-Analyses and Meta-Regressions in Strength & Conditioning Research. *Sports Medicine*, *53*(2), 313–325. https://doi.org/10.1007/s40279-022-01766-0

Kaschak, M. P., Kutta, T. J., & Jones, J. L. (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. *Psychonomic Bulletin & Review*, *18*(6), 1133–1139. https://doi.org/10.3758/s13423-011-0157-y

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, *33*(2), 393–418. https://doi.org/10.1017/S0142716411000415

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860–897. https://doi.org/10.1080/01690960601155284

Kidd, E., & Garcia, R. (2022). How diverse is child language acquisition research? *First Language*, *42*(6), 703–735. https://doi.org/10.1177/01427237211066405

Kidd, E., Junge, C., Spokes, T., Morrison, L., & Cutler, A. (2018). Individual Differences in Infant Speech Segmentation: Achieving the Lexical Shift. *Infancy*, *23*(6), 770–794. https://doi.org/10.1111/infa.12256

Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring children. *Journal of Memory and Language*, *127*, Article 104374. https://doi.org/10.1016/j.jml.2022.104374

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394–425. https://doi.org/10.1037/bul0000142

Leonard, L. B., Miller, C. A., Grela, B., Holland, A. L., Gerber, E., & Petucci, M. (2000). Production operations contribute to the grammatical morpheme limitations of children with specific language impairment. *Special Issue: Disorders of Language and Memory: Implications for Cognitive Theory.*, *43*(2), 362–378. https://doi.org/10.1006/jmla.1999.2689

Lester, N. A., Moran, S., Küntay, A. C., Allen, S. E. M., Pfeiler, B., & Stoll, S. (2022). Detecting structured repetition in child-surrounding speech: Evidence from maximally diverse languages. *Cognition*, *221*, 104986. https://doi.org/10.1016/j.cognition.2021.104986

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The Influence of Animacy on Relative Clause Processing. *Journal of Memory and Language*, *47*(1), 50–68. https://doi.org/10.1006/jmla.2001.2837

Marchman, V. A., Bates, E., Burkardt, A., & Good, A. B. (1991). Functional constraints of the acquisition of the passive: Toward a model of the competence to perform. *First Language*, *11*(31), 65–92. https://doi.org.10.1177/014272379101103104

Messenger, K. (2021). The persistence of priming: Exploring long-lasting syntactic priming effects in children and adults. *Cognitive Science*, *45*(6), Article e13005. https://doi.org/10.1111/cogs.13005

Messenger, K. (Ed.). (2022). *Syntactic priming in language acquisition: Representations, mechanisms and applications* (Vol. 31). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.31.01mes

Messenger, K., Branigan, H. P., & McLean, J. F. (2012). Is children's acquisition of the passive a staged process? Evidence from six- and nine-year-olds' production of passives. *Journal of Child Language*, *39*(5), 991–1016. https://doi.org/10.1017/S0305000911000377

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, *66*(4), 568–587. https://doi.org/10.1016/j.jml.2012.03.008

Messenger, K., & Fisher, C. (2018). Mistakes weren't made: Three-year-olds' comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition*, *178*, 118–132. https://doi.org/10.1016/j.cognition.2018.05.002

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105–125. https://doi.org/10.1037/1082-989X.7.1.105

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *Journal of the American Medical Association*, *295*(6), 676–680. https://doi.org/10.1001/jama.295.6.676

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459. https://doi.org/10.1037/0033-2909.134.3.427

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rohatgi, A. (2022). *Webplotdigitizer: Version 4.6*. https://automeris.io/WebPlotDigitizer

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63. https://doi.org/10.1016/j.cognition.2012.06.008

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children’s linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557–567. https://doi.org/10.1111/1467-7687.00312

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, *2*(1), 27–49. https://doi.org/10.1207/s15473341lld0201_2

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3- and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, *43*(6), 1334–1346. https://doi.org/10.1037/0012-1649.43.6.1334

Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, *11*(1), 54–71. https://doi.org/10.1037/1082-989X.11.1.54

Stedman, F. M. R., Curtin, F., Elbourne, D. R., Kesselheim, A. S., & Brookhart, M. A. (2011). Meta-analyses involving cross-over trials: Methodological issues. *International Journal of Epidemiology*, *40*(6), 1732–1734. https://doi.org/10.1093/ije/dyp345

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal*, *343*, d4002. https://doi.org/10.1136/bmj.d4002

Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*, *58*(2), 188–213. https://doi.org/10.1016/j.jml.2007.06.012

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253. https://doi.org/10.1016/S0010-0277(99)00069-4

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Harvard University Press.

Tomasello, M., & Abbot-Smith, K. (2002). A tale of two theories: Response to Fisher. *Cognition*, *83*(2), 207–214. https://doi.org/10.1016/S0010-0277(01)00172-X

van Beijsterveldt, L. M., & van Hell, J. G. (2009). Structural priming of adjective-noun structures in hearing and deaf children. *Journal of Experimental Child Psychology*, *104*(2), 179–196. https://doi.org/10.1016/j.jecp.2009.05.002

Van Dijk, C., Van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2022). Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language*, *49*(5), 897–929. https://doi.org/10.1017/S0305000921000337 van Gompel, R. P. G., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, *21*(3), 448–455. https://doi.org/10.1017/S1366728917000542

Vasilyeva, M., & Gámez, P. B. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in syntactic priming. *Journal of Experimental Child Psychology*, *138*, 15–30. https://doi.org/10.1016/j.jecp.2015.04.009

Vasilyeva, M., Huttenlocher, J., & Waterfall, H. (2006). Effects of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology*, *42*(1), 164–174. https://doi.org/10.1037/0012-1649.42.1.164

Vasilyeva, M., & Waterfall, H. (2012). Beyond syntactic priming: Evidence for activation of alternative syntactic structures. *Journal of Child Language*, *39*(2), 258–283. https://doi.org/10.1017/S0305000911000055

Vasilyeva, M., Waterfall, H., Gámez, P. B., Gomez, L. E., Bowers, E., & Shimpi, P. (2010). Cross-linguistic syntactic priming in bilingual children. *Journal of Child Language*, *37*(5), 1047–1064. https://doi.org/10.1017/S0305000909990213

Viechtbauer, W. (2008). Accounting for heterogeneity via random-effects models and moderator analyses in meta-analysis. *Zeitschrift Für Psychologie/Journal of Psychology*, *215*(2), 104. https://doi.org/10.1027/0044-3409.215.2.104

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wolleb, A., Sorace, A., & Westergaard, M. (2018). Exploring the role of cognitive control in syntactic processing: Evidence from cross-language priming in bilingual children. *Linguistic Approaches to Bilingualism*, *8*(5), 606–636. https://doi.org/10.1075/lab.17002.wol

Chapter 4:

Indexing Prediction Error during Syntactic Priming via Pupillometry

Chapter Overview

Chapter 4 makes up the second part of this thesis, which turns from syntactic acquisition to processing mechanisms during syntactic priming in adults. Chapter 2 supported the error-based learning account as a mechanism of syntactic acquisition but Chapter 3 found that explicit processes can also influence syntactic priming. This suggested the value in being able to measure error-based learning more directly. The central assumption of the Chang et al. (2006) account is learning through predictive processing, with prediction error driving representational change. Pupil dilation has been associated with surprisal in other domains (e.g., Preuschoff et al., 2011). Therefore, Chapter 4 aimed to develop an online implicit measure of prediction error in syntactic processing by pioneering the use of pupillometry during syntactic priming. The error-based learning mechanism operates throughout the lifespan and studies in adults have linked manipulations of predictability to greater syntactic priming (Bernolet & Hartsuiker, 2010; Jaeger & Snider, 2013). Since combining pupillometry with syntactic priming had not previously been attempted, I conducted the study with adult participants rather than children. Pupil size was larger for passive than active primes, in line with larger prediction error for infrequent structures. Trial-by-trial pupil size also weakly predicted syntactic priming. Chapter 4 addressed the aims of this thesis by developing a real-time measure of an underlying mechanism of syntactic acquisition and processing that could provide complementary evidence to production data in syntactic priming experiments. It also demonstrated the value of incorporating real-time measures in production priming designs.

Publication Status

The following manuscript is in preparation for submission:

Kumarage, S., Malko, A., & Kidd, E. (in preparation). Indexing prediction error during syntactic priming via pupillometry.

Acknowledgements

Thank you to Rowena Garcia for contributing experimental materials and Jet Bustamante for producing new materials. Thank you to Maggie Otto for lending her voice to the experimental audio.

Author Contributions

- **Kumarage** and Kidd developed the research question.
- Malko wrote code to convert pictures and audio to videos.
- **Kumarage** designed and programmed the experiment.
- Malko wrote code to validate the experimental programming.
- **Kumarage** conducted data collection.
- **Kumarage** preprocessed and analysed data with advice from Malko.
- **Kumarage** drafted the manuscript with editing provided by Malko and Kidd.

Indexing Prediction Error During Syntactic Priming via Pupillometry

Shanthi Kumarage[1,2]

Anton Malko[3]

Evan Kidd[3]

[1]Language Development Department, Max Planck Institute of Psycholinguistics, Nijmegen

[2]School of Medicine and Psychology, The Australian National University

[3]School of Languages, Literature and Linguistics, The Australian National University

Abstract

Prediction is argued to be a key feature of human cognition, including in syntactic processing. The discrepancy between predicted and actual input, or prediction error, has been linked to dynamic changes in syntactic representations in theoretical models of language processing. This mechanism is termed *error-based learning*. Evidence from syntactic priming research supports error-based learning accounts; however, measuring prediction error itself has not been a research focus. Here we present a study exploring the use of pupillometry as a measure of prediction error during syntactic priming. We found that the pupil response distinguished passive and active prime sentences, with a larger response to the more complex and less expected passive structure. While individual differences in the response did not predict participants' propensity to be primed in target trials, trial-by-trial variation in the response weakly predicted priming. We conclude that the pupil response is not only sensitive to syntactic complexity, but there is some evidence that it also reflects the error signal during sentence comprehension, with the magnitude of this error signal predicting the magnitude of adjustment to mental representations measured by changes in production during syntactic priming. The results provide a foundation for future research using higher-powered samples and investigating other manipulations of prediction error.

Indexing Prediction Error during Syntactic Priming via Pupillometry

One of the biggest challenges in psycholinguistic research is identifying the mechanisms by which humans acquire and process language. Although early theoretical approaches were dominated by domain-specific proposals hypothesising language-specific mechanisms (following arguments made by Chomsky, 1959, 1965; and others, Fodor, 1983), the nature of language as a human skill instantiated in neural structures that evolved and were co-opted for language and communication implies that neurally plausible domain-general learning mechanisms at least play some determining role. In the current paper, we consider one such mechanism: learning via *prediction error*. Predictive processing is argued to confer adaptive advantages, allowing the brain to develop, store and use complex models of the environment rather than simple associations between sensory input and internal states (Clark, 2013). Learning via prediction error is plausibly instantiated in neural networks via recurrent connections and the backpropagation of error (Elman, 1990). That is, sensory input can be compared to previous predictive output and any mismatch converted to an error signal that adjusts the predictive model.

Prediction error plays a central role in one prominent computational model of language. Chang et al. (2006) proposed a dual-path connectionist model of language production and syntactic acquisition. The model comprises a meaning system, which encodes concepts and their roles within a message, and a separate sequencing system, which is a simple recurrent network that learns syntactic representations that allow it to correctly sequence words. The model makes next-word predictions during sentence comprehension and compares its predictions to the input. In the case of a mismatch, prediction error backpropagates through the system, and weights within the network are adjusted to reduce the likelihood of the error in future predictions, thus updating the model's syntactic representations. These same syntactic representations are used during sentence production, wherein the prediction mechanisms are used to incrementally output words to produce a grammatical sentence that accurately communicates the intended meaning. Therefore, the theory explains the acquisition of syntax through error-based learning and syntactic processing in the services of production. Importantly, the theory connects language acquisition to adult language use: the error-based learning mechanism responsible for the learning of syntactic knowledge continues to operate in adulthood, albeit with a lower learning rate, suggesting that language use continues to modify syntactic representations across the lifespan via implicit learning (Dell & Chang, 2014).

Error-Based Learning in Syntactic Priming

One way in which such learning has been quantified is by a phenomenon known as *syntactic priming*. Syntactic priming is the tendency to persist in the use of a structure after previously hearing or using it (Bock, 1986). For example, hearing a passive prime, *the swimmer was eaten by the crocodile*, increases the likelihood of later describing a target using a passive, *the cyclist was swooped by the magpie*, rather than an active, *the magpie swooped the cyclist*. Syntactic priming is argued to reflect the output of error-based learning (Chang et al., 2006), tapping into underlying syntactic representations (Branigan & Pickering, 2017). Priming independent of shared lexical content (both content words: Bock, 1986, and functional words: Bock, 1989; Ferreira, 2003; Hartsuiker et al., 1999), thematic structure (Bock et al., 1992), and prosody (Bock & Loebell, 1990) suggests that priming occurs at the level of representation of abstract syntactic structure (but see Ziegler et al., 2019). Further, observations of enduring priming effects imply that the mechanism of syntactic priming involves a form of implicit learning such as error-based learning (Bock & Griffin, 2000; Kaschak, Kutta, & Schatschneider, 2011). Under the Chang et al. (2006) account, next word prediction occurs during prime sentences. One layer of the model receives both the previously predicted word and perceived word, with the difference between them comprising the error signal. For example, hearing the past participle rather than a main verb in a transitive sentence indicates that the sentence is the less frequent passive structure (*the cyclist was* ***swooped*** *by the magpie*). The network uses the error signal to adjust network weights, increasing the likelihood that a passive will be used in the future (thus leading to priming).

Therefore, a key prediction of error-based learning accounts, such as Chang et al. (2006), is that more surprising input produces greater prediction error and therefore greater adjustment of syntactic representations. Syntactic priming research provides evidence for this proposal. For example, priming effects are larger for less frequent structures than more frequent ones. Of the prepositional dative and double object dative, whichever is less frequent is primed more strongly (the double object in Dutch: Bernolet & Hartsuiker, 2010; and the prepositional object in English: Jaeger & Snider, 2013; Kaschak, Kutta, & Jones, 2011). This *inverse-frequency effect* extends to the active-passive alternation (Bock, 1986), the mention or omission of the English *that* complementiser (Ferreira, 2003), and relative clause attachment (Scheepers, 2003). Another way of manipulating the predictability of primes is by leveraging verb-biases. Hearing *Bob threw Wendy a ball* should be more surprising than hearing *Bob gave Wendy a present* because *throw* is biased to the prepositional object dative whilst *give* is biased to the double object dative. Accordingly, researchers have observed *prime-surprisal effects*: stronger priming effects when the structure of a prime mismatches with the structure preferred by the verb in the prime sentence (Bernolet & Hartsuiker, 2010; Jaeger & Snider, 2013; Peter et al., 2015). Similarly, Fazekas et al. (2020) found that presenting input with an equal frequency of double object or prepositional object datives but where only one structure appeared in surprising contexts shifted participants' production preferences towards the structure they heard in surprising contexts.

Another prediction of error-based learning accounts is that participants' expectations are not static, but change based on input. In syntactic priming, the effects of primes should accumulate, with additional encounters with a syntactic structure resulting in compounding updates in expectations and representations. Indeed, several studies of syntactic priming in comprehension have observed *expectation adaptation*, where processing deficits associated with temporarily ambiguous structures (garden-path sentences) diminish as a function of the number of structures previously encountered (Farmer et al., 2014; Fine et al., 2013; Fine & Jaeger, 2016). There is also some evidence of the cumulative influence of primes on the likelihood of producing syntactic structure (Bernolet et al., 2016; Jaeger & Snider, 2013; Kaschak, Kutta, & Jones, 2011).

Inverse-frequency, prime-surprisal and cumulative priming effects provide strong but indirect evidence for the role of prediction error in syntactic processing. The challenge in cognitive research is getting measurable indices of cognitive processes like surprisal. More direct evidence would consist of manipulating prime surprisal to demonstrate effects on an index of prediction error and linking that measure of prediction error to priming. To our knowledge, only one study has attempted to do so. Arai and Chang (2024) investigated priming of the active-passive alternation in Japanese. They utilised the ambiguous case-marker *ni*, which marks both dative case in a sentence like *the boy talked* ***to his friend*** and the oblique argument in a passive: *the boy was hit* ***by his friend***. Since Japanese is a verb-final language, the ambiguity is not resolved until the verb. The words *boy-NOM* ***friend-ni*** could be interpreted as either *the boy* ***to his friend*** __ or *the boy was* ***by his friend*** __ until the verb is encountered. Arai and Chang (2024) manipulated participants' expectations of encountering a passive through the content of fillers, which either biased the interpretation of preverbal arguments towards a dative case-marked noun (*to his friend*) or not. Participants who were biased to a dative interpretation showed longer reading times on the verb in passive primes than those who were not, an indicator that they experienced greater prediction error. However, although priming was also larger when passives were less expected, using reading time on the prime verb as a measure of prediction error did not predict participants' priming on the target.

One problem with using reading time, as measured via self-paced reading, as a measure of prediction error, is that it is likely to involve many interacting processes that may obscure reliable measurement of prediction error at the level of the individual (Frinsel & Christiansen, 2024). Another is that participants get faster at self-paced reading in general over the course of an experiment so both reduced prediction error due to expectation adaptation and the time spent on the task contribute to faster reading times (Prasad & Linzen, 2021). In this study we explore a promising potential measure of prediction error that is more implicit and does not require deliberative action on behalf of the participant: pupil diameter.

Pupillometry

Pupil diameter is strongly associated with ambient light levels but also shows small but detectable changes due to cognitive processing (Sirois & Brisson, 2014). Specifically, pupil diameter indexes emotional arousal, mental effort, top-down processes, and surprisal (Hepach & Westermann, 2016). This is attributed to activation in the locus coeruleus, which, in addition to being tightly linked with pupil dilation, is involved in a wide range of basic functions such as attention, arousal, and memory retrieval (Sirois & Brisson, 2014). Accordingly, pupillometry has been applied to a variety of questions concerning linguistic processing. Research on word recognition under different levels of noise, speech intelligibility, and speech rates has found that larger pupil size indexes cognitive effort under different listening conditions (Koch & Janse, 2016; Kramer et al., 2013; Kuchinsky et al., 2013; Zekveld et al., 2010). Pupil dilation is also greater when the prosody of a sentence is incongruent to its information focus or syntactic structure (Engelhardt et al., 2010; Zellin et al., 2011; but see Aydın & Uzun, 2023 for contradictory results) or when a word is semantically incongruent (Demberg & Sayeed, 2016; Häuser et al., 2018). In syntactic processing, pupil dilation has been observed in response to syntactic violations such as case marking (Aydın & Uzun, 2023) and gender agreement (Demberg & Sayeed, 2016), but also to variations in syntactic complexity. Schluroff (1982) ranked a range of structures based on their Yngve depth (degree of left-branching; Yngve, 1960) and found that the magnitude of pupil dilation correlated with their syntactic complexity. Similarly, Stanners et al. (1972) found a complexity effect for structures with the same surface structure but which differed in thematic role assignment (*they are eager to please* vs *they are easy to please*). Since these early studies, complexity effects have been found for several structures including object vs subject relative clauses in both written (Demberg & Sayeed, 2016; Just & Carpenter, 1993) and auditory comprehension (Demberg & Sayeed, 2016; Piquado et al., 2010), *wh*-phrases vs *whether* clauses (Just & Carpenter, 1993), negative and passive vs affirmative and active sentences (Beatty, 1982), and SVO vs OSV word order in Danish (Wendt et al., 2016).

Thus far, research in language processing has typically interpreted the pupillary response as a measure of cognitive load or mental effort. However, many findings could also be interpreted as pupil size indexing prediction error or surprisal. Incongruencies in prosody (Engelhardt et al., 2010; Zellin et al., 2011) and semantics (Demberg & Sayeed, 2016; Häuser et al., 2018) represent violations of expectations. The same is true for syntactic violations. In Aydın & Uzun's (2023) study, Turkish speakers were presented with SVO transitive sentences (*the boy-NOM painted the desk*), in which the sentence-final object noun phrase occurred in either the grammatical accusative case (*desk-ACC*) or the unexpected and ungrammatical dative case (*desk-DAT*). Similarly, in their 1st and 4th experiments Demberg and Sayeed (2016) presented German speakers with sentences where the final noun was expected given the gender marking of the previous determiner and adjective (*Simone had a-MASC horrible-MASC dream-MASC*) or unexpected and mismatching (*Simone had a-FEM horrible-FEM dream-MASC*). In other studies, participants' accumulated syntactic knowledge would lead them to predict canonical or more frequent word orders over non-canonical ones. For instance, since Danish is an SVO language, speakers presumably expect this more frequent pattern in comparison to OSV (Wendt et al., 2016). For relative clauses, the well documented subject advantage means that English speakers expect a subject relative clause over an object relative clause (Just & Carpenter, 1993; Piquado et al., 2010). The suggestion is that pupil dilation may be a fairly implicit index of prediction error.

While this link has not been made explicitly in psycholinguistic research, pupil dilation has been argued to index prediction error in gambling tasks (Preuschoff et al., 2011) and sequence or cue learning (Rutar et al., 2023; Zhang et al., 2019). In these studies, unexpected outcomes or formalised measures of the prediction error produced by them (Preuschoff et al., 2011) are associated with larger pupil dilation. These results accord with the proposal that the neurotransmitter noradrenaline signals errors in judging uncertainty in a similar manner to how dopamine signals reward (Preuschoff et al., 2011). Given the involvement of the locus coeruleus in both the noradrenergic system and pupillary control, signalling of prediction error is argued to also result in pupil dilation. Particularly relevant here is the interpretation of these results within the context of predictive processing and model updating (Rutar et al., 2023).

The Current Study

The present study aimed to determine whether prediction error in syntactic processing can be directly measured by combining the syntactic priming and pupillometric methodologies. We test this using the active-passive alternation in English. The English passive is non-canonical in both frequency and thematic role order and is therefore a highly unpredictable structure compared to the active. Participants were presented with primes in the active or passive structure followed by a target picture of a transitive action to describe. We measured pupil diameter during prime sentences and recorded participants' productions in target sentences. We expected that passive primes would be associated with both a larger pupil response (in line with syntactic complexity effects, e.g., Just & Carpenter, 1993) and with greater production of passives (in line with syntactic priming effects, e.g., Bock, 1986). The key test is whether greater prediction error (i.e., the expected pupil dilation difference between actives and passives) predicts participants' likelihood to be primed, as predicted by error-based learning accounts of syntactic processing (Chang et al., 2006). We investigated both individual differences in prediction error (mean pupil dilation difference per participant) and pupil dilation on each trial as predictors of syntactic priming. If adaptation effects (e.g., Fine et al., 2013) operate in our study, the latter dynamic measure of prediction error is more likely to capture the effect because it can reflect updates to expectations throughout the task.

Methods

Participants

Eighty individuals who spoke Australian English as a first language, were aged 18-35, with no history of developmental or acquired language disorder, and normal or corrected-to-normal vision were recruited from the Australian National University community. An additional three participants were tested but were not included because they did not meet the eligibility criteria. Participants received a 1-hour course credit or AUD$15 as compensation for their participation. The final sample (*N* = 80: F = 53, M = 23, NB = 3, undisclosed = 1) ranged in age from 18 to 33 years (*M* = 22.55, *SD* = 4.29). All participants used English for the majority of their interactions in an average week (*M* = 98.1%, *SD* = 5.6), with 33 reporting knowing or using a language other than English. Sample size was estimated according to Mahowald et al.'s (2016) power analysis for syntactic priming effects. A sample of 80 participants has sufficient power (>80%) to detect a large interaction effect (equivalent to the size of the lexical boost effect).[1]

Materials and Design

Materials consisted of 32 prime pictures based on 8 transitive verbs (*bite, catch, carry, kick, pinch, push, kiss, lick*) and 32 target pictures based on a different set of 8 transitive verbs (*chase, drag, hit, prick, punch, shoot, tickle, feed*). To control for animacy effects, both the agent and patient in each picture were drawn from a pool of 16 animate characters (*bear, cat, chicken, cow, dog, duck, elephant, frog, goat, horse, lion, monkey, mouse, pig, rabbit, turtle*), with each character occurring as agent and patient equally often. In addition, there were 128 filler pictures, depicting noun phrases (*a red apple*), prepositional phrases (*the cups are on the chair*), intransitives (*the cow is sleeping*; *the dog is strong*), and datives (*the cow gives the chicken a present*). Most pictures were taken from Garcia et al. (2021, 2023), with some additional pictures drawn by the same artist to fulfil the requirements of this study. Experimental pictures were 800 x 450 pixels. All pictures

---

[1] We note that we used Bayesian rather than frequentist statistics as in Mahowald et al.'s (2016) power simulations. Bayesian and frequentist models do typically produce very similar results unless using informative priors, which we did not. Sample sizes that result in low power in frequentist models will result in wide credible intervals in Bayesian models, which can be interpreted similarly.

were resized to 1920 x 1200 by adding extra white background. Experimental pictures were standardised to a mean luminance of 249.38 in the HSV colourspace (scale 0-255) using the *lumMatch* function from the *SHINE_color* MATLAB toolbox (Dal Ben, 2021; Willenbockel et al., 2010). We recorded audio descriptions of each prime and half the filler pictures by a female native speaker of Australian English. For experimental items both an active (*the pig is catching the cat*) and a passive description (*the cat is being caught by the pig*) were recorded. These descriptions were recorded in the present progressive form to avoid an adjectival interpretation. The recorded description did not match the picture for 19 filler items (e.g., *the fork is on the table* for a picture depicting *the scissors are in the box*).

The task consisted of 32 prime-target pairs, with prime (active or passive structure) manipulated within-subjects. No more than two primes of the same structure occurred in a row. Between 2 and 6 filler items intervened between each prime-target pair. Half the fillers were prime trials and half were target trials, with pseudorandom ordering of trials such that no more than three trials of the same type occurred in a row. Variable trial type and spacing of prime-target pairs intended to mask the aims of the study from participants. As shown in Figure 1, we constructed 16 experimental lists counterbalancing the verb pairings in prime-target pairs, prime structure, and the direction of the action in prime and target pictures (LR or RL).

Figure 1

Counterbalancing procedures for experimental lists

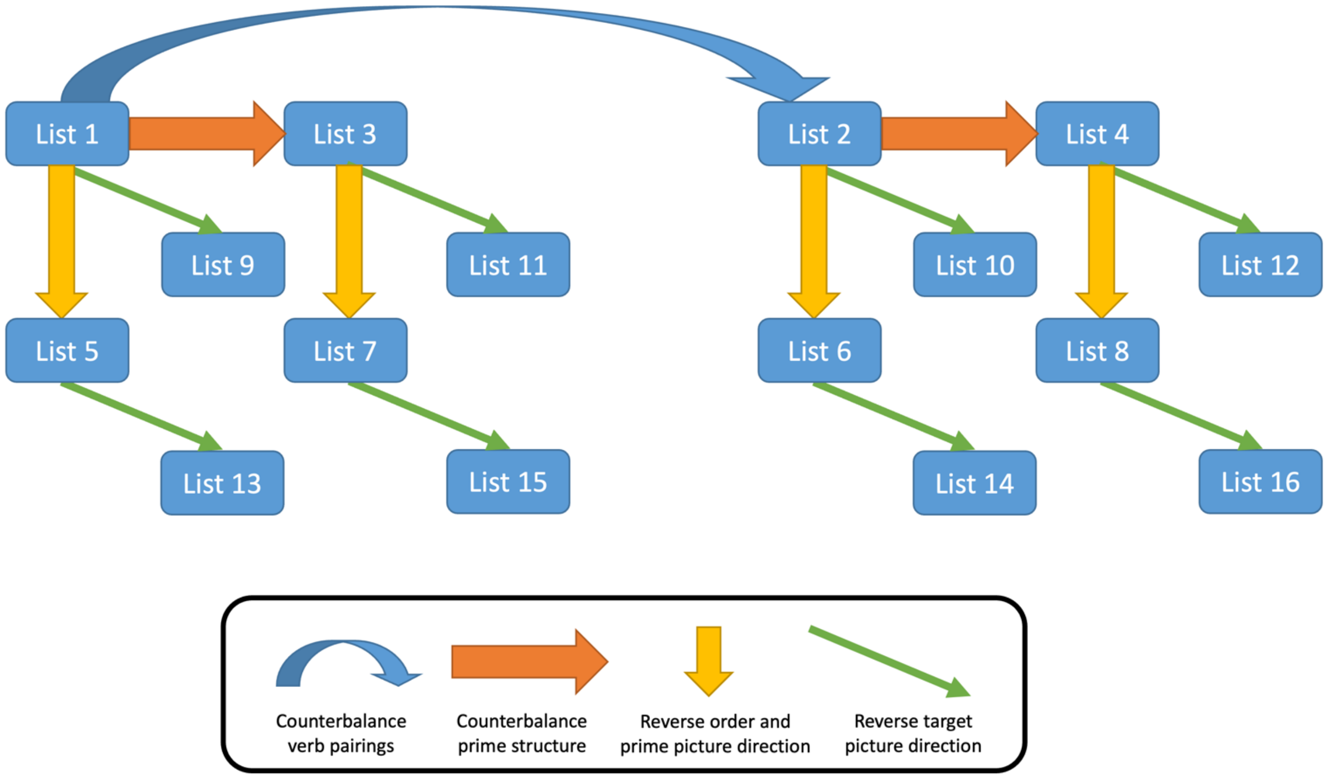


Figure 2 depicts the time course of a prime target pair in the experiment. Both prime and target trials were preceded by a 2000ms fixation symbol accompanied by a 250ms beep, indicating to participants the start of a new trial and the type of trial. This also allowed extra time for the pupil to return to baseline following the previous trial (Mathôt & Vilotijević, 2022). The picture then appeared on screen.

In prime trials, we presented the picture for 500ms before sentence onset to allow event apprehension to occur (Griffin & Bock, 2000). Therefore, participants could identify the event as transitive and form an expectation that the agent would occur first. We expected prediction error for passive primes to be induced in the very first part of the sentence, on hearing the first noun phrase. The peak pupil dilation typically occurs 1-1.5s from the point of difficulty (Just & Carpenter, 1993; Tromp et al., 2016). Recorded sentences varied from 1414ms to 2147ms and were followed by 2500ms of silence, which allowed sufficient time for the pupillary response to be observed. Additionally, the picture remained on screen during this time so that the pupil response could be measured without changes in luminance caused by switching to a fixation symbol. Because the pupil response is task-evoked (see Zellin et al., 2011, p. 136), we introduced a picture-verification task on each prime trial. When a question symbol appeared after picture offset, participants responded by pressing a green button if the picture and audio description matched and a red button if they did not. The task automatically proceeded to the next trial after the participant responded. As participants needed to wait until the question symbol appeared to make a response, we minimised any cognitive processing associated with actually making the response (e.g. motor planning etc) during the measurement of pupil diameter. About 20% of filler prime trials had mismatching audio descriptions to maintain attention to the task. On average, participants answered these items correctly 97% of the time (*M* = 0.97, *SD* = 0.06, range: 0.68 – 1.00) confirming their attention to the task.

Target trials more closely resembled typical syntactic priming experiments. The picture appeared after the fixation symbol and remained on screen until the participant had described it and the experimenter progressed the task to the next trial.

Figure 2

Structure of trials in the experiment

Apparatus and Procedure

Participants were seated in a dimly lit room (30 lux) and first completed a demographic questionnaire. Participants were introduced to the three fixation symbols for listening, answering questions, and speaking (see Figure 2) and instructed to respond accordingly. On seeing the listening symbol, they were instructed to listen carefully to the upcoming sentence and pay attention to whether it matched the accompanying picture. When the question symbol appeared, participants responded with whether the picture matched (green button) or didn't match (red button) on a keypad. On seeing the speaking symbol, participants were instructed to describe the upcoming picture. The instructions encouraged participants to respond with full sentences where possible and to remember the button locations rather than looking down to respond. Participants completed six practice items, then three blocks of trials.

Stimuli were presented using Tobii Pro Lab (version 1.207.44884) software on the T60 eyetracker, which measured pupil diameter in mm at 60Hz. Participants completed a 9 point calibration and validation procedure at the beginning of each of the three blocks of trials (mean validation accuracy 0.51 degrees). A Zoom H2n audio recording device was connected to the computer so the Tobii software automatically audio-recorded the task.

Transcription and Coding

Participants' responses were transcribed from the audio recording and scored as *active*, *passive* or *other*. If participants produced more than one sentence, only the first complete sentence was coded. If participants corrected themselves before producing a full sentence, the corrected form of the utterance was scored. Responses were scored as active if they contained an agent in the subject position, an appropriate transitive verb, and a patient in the object position and could be expressed in the alternate passive structure (e.g., *the elephant feeds the lion*). Passive responses needed to contain a patient in the subject position, an auxiliary verb (*was, got*) and appropriate transitive main verb, an agent in a by-phrase and be expressible in the alternate active structure (e.g., *the goat is being chased by the horse*). Transitive responses where the participant repeated the verb or a noun contained in the prime sentence were excluded. Other responses consisted of all other sentence forms, including datives (*the horse is feeding the duck some food*), intransitives (*the goat and the horse are walking*), and irreversible phrasal verbs (*a mouse running away from a chicken*). Overall, 90.2% of participants' sentences could be coded as Active or Passive.

Results

Pupillometry

Our main aim was to see whether prediction error can be indexed by pupil dilation and if this predicts priming. Therefore, we first analysed the time course of pupil dilation for active primes and passive primes to determine whether the pupil response does distinguish them.

Pupil Data Preprocessing

We performed data preprocessing and our analyses in R (version 4.2.1; R Core Team, 2022). The Tobii T60 eyetracker measures pupil size for both eyes. As recommended by Sirois and Brisson (2014), pupil size was regressed from each eye onto the other, thus imputing values missing from only one eye. Then, average pupil size across both eyes was calculated. Average pupil size was passed through an 11 sample Hanning filter using the *filter_data* function from the *PupillometryR* package (version 0.0.5; Forbes, 2020) to provide a smooth signal for blink detection. Blinks were detected using a velocity filter similar to the method described by Mathôt (2013). Due to the 60Hz sampling frequency, the data typically showed the fast negative velocity before blinks but not the rapid increase in velocity following blinks. Therefore, we counted blinks as any velocity larger than -0.02mm change in the filtered data preceding a period of track loss. Blinks were extended to 50ms on either side of the gap with no imputation of missing values (since generalised additive mixed models can handle missing data: van Rij et al., 2019). Filtered data was used for deciding which samples to remove during blink detection but we returned to the raw data for further preprocessing and analysis. We applied a velocity filter using a cut off of 0.15mm change in pupil size between each sample, removing one sample either side as well. Visual inspection of each trial for each participant showed that obvious outliers were removed without the exclusion of steep curves that made up the pupil response. Gaze position was used as a control variable in our analyses as the pupil foreshortening effect can influence the measured pupil size (Brisson et al., 2013). Any samples where gaze position was outside the area of the screen were removed. After removing blinks, velocity outliers and gaze position outliers, trials with more than 25% of data removed in preprocessing and/or missing due to track loss were excluded. Baseline pupil size was calculated as the average pupil size in the 250ms preceding sentence onset (i.e. 250-500ms after picture onset). Trials where the baseline could not be calculated were excluded, as were trials where the baseline was more than two standard deviations away from the participant’s mean baseline. The baseline was subtracted from pupil size to calculate baseline-corrected pupil diameter. Trials were cut to 3767ms, the length of the shortest trial. In total 78.8% of trials (*N* = 2016) were included after excluding trials due to recording failure (*N* = 9), high percentage of missing or removed data (*N* = 416), and missing or improbable baseline values (*N* = 119).

Syntactic Complexity Effect

Data were first analysed to determine whether the pupil response differed between active and passive prime sentences. It is recommended to model the time course of pupil dilation rather than analysing the peak or latency of the response (van Rij et al., 2019;

Wieling, 2018). We fit a generalised additive mixed model (GAMM) following the recommendations of van Rij et al. (2019). We ran our analyses using the *mgcv* (version 1.9.1; Wood, 2011) and *itsadug* (version 2.4.1; van Rij et al., 2022) R packages. GAMMs allow the modelling of non-linear relationships, such as that between time and pupil size, the inclusion of methods for controlling autocorrelation between samples in time-series data, and the inclusion of random effects for items and subjects.

The model estimated the effect of prime condition (*active*: 0 or *passive*: 1) as an ordered factor. This includes an intercept term, a reference smooth (for the *active* condition) and a difference smooth (for the *passive* compared to *active* condition). We controlled for gaze position by including a non-linear interaction between the X and Y coordinates of gaze position. We could not include random smooths for each trial for each participant due to the high computational demands of doing so. We instead included random slopes and intercepts for each trial, and random smooths for participant and item (van Rij et al., 2019). Residuals were not normally distributed but we did not fit the model with a scaled *t*-distribution because doing so lowered the correlation between fitted values and the data and substantially increased the size of the median absolute residuals, indicating worsened model fit. Residuals showed autocorrelation and so an AR-1 model was added to the GAMM using a value of $\rho = .947$ (van Rij et al., 2019).[2] An autocorrelation function (ACF) plot showed reduced autocorrelation. Although adding the AR-1 model slightly lowered the correlation between fitted values and the data it did not increase the size of residuals; we therefore retained this model.

Interpreting significance in GAMMs requires a combination of visualisation and interpreting model summary statistics. Figure 3a plots the model estimated baselined pupil diameter for active and passive prime sentences. The pupil response appears larger following passive than active primes. Figure 3b plots the difference smooth between passive and active primes. The pupillary response to passive sentences is significantly larger than the response to active sentences from 1115 to 3767ms after sentence-onset. Table 1 reports the

---

[2] The model syntax is provided below:

```
model <- bam(pupil ~ OFcondition + s(time) + s(time, by=OFcondition, k=20)
      + s(gazeX, gazeY) + s(time, ID, bs='fs', m=1)
      + s(time, prime_item, bs='fs', m=1) + s(event, bs='re')
      + s(time, event, bs='re'), data=data, discrete=TRUE,
       rho = .947, AR.start = data$start.event)
```

summary of the final model. The parametric effect for Condition shows a significant overall difference in pupil size between active and passive sentences ($\beta = 0.02$, $t = 4.11$, $p < .001$). For smooths, the estimated degrees of freedom (edf) indicates how wiggly the regression line is, with lower edfs indicating a smoother line. The reference degrees of freedom (Ref.df) indicates the degrees of freedom associated with the F-test of significance. A significant *p*-value indicates the regression line is significantly different from zero. Because we explicitly modelled the difference between conditions (Time:Condition), we can conclude that the time course of the pupil response is significantly different between active and passive sentences ($F(18.46) = 17.52$, $p < .001$).

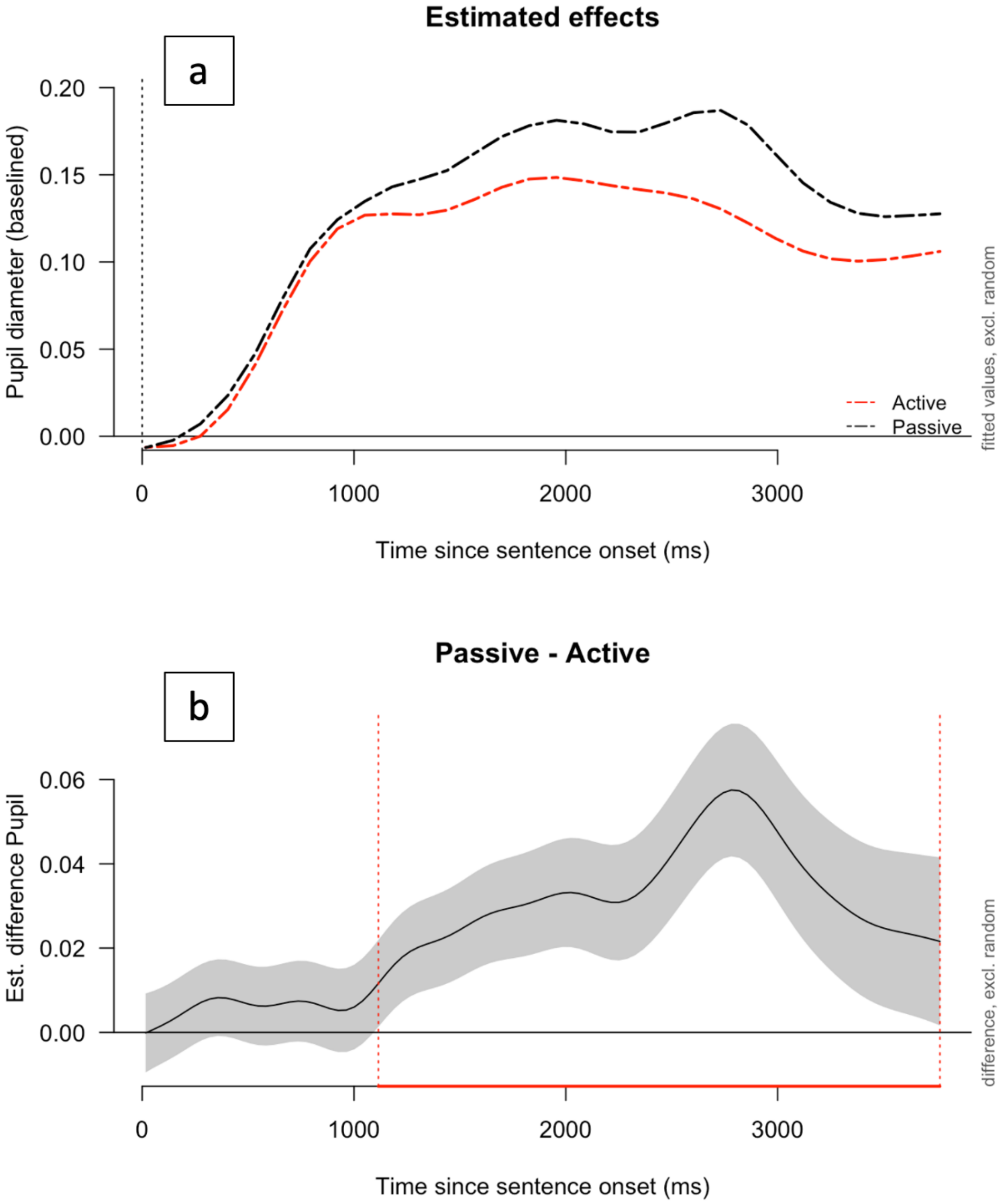
Table 1

Summary of the final model

	β	SE	*t*	*p*
Parametric coefficients				
Intercept	0.11	0.01	8.93	< .001
Condition	0.02	0.00	4.11	< .001
	edf	Ref.df	F	*p*
Smooth terms (fixed effects)				
Time	8.67	8.72	25.56	< .001
Time:Condition	16.69	18.46	17.52	< .001
GazeX,GazeY	28.31	28.95	78.50	< .001
Smooth terms (random effects)				
Time, Participant	668.22	719.00	1305.45	< .001
Time, Item	257.43	287.00	506.70	< .001
Trial	1567.29	2014.00	174.83	< .001
Time, Trial	1738.93	2014.00	216.74	< .001

Figure 3

Model-predicted effects on pupillary time course including (a) estimated effects for active and passive primes and (b) estimated difference smooth between conditions



Syntactic Priming

We next confirmed that we found a syntactic priming effect. Table 2 summarises the proportion of active, passive and other responses that participants produced in each experimental condition. *Other* responses were excluded from our analyses. We ran our priming analyses using the *brms* R package (version 2.20.4; Bürkner, 2017). We ran a Bayesian mixed-effects logistic model including the maximal random effects structure, with random intercepts and slopes for prime by participant and by target item nested in target verb.[3] The prime variable was dummy coded (active: 0, passive: 1). The model was run with 3000 iterations, 1000 of them warm-up, and 10 chains using default uninformative priors. The value of *adapt_delta*, which decreases the step sizes taken by the model, was increased to 0.95 to prevent divergent transitions. Convergence diagnostics indicated reliable convergence and estimates of posteriors: the maximum Rhat was 1.00, the minimum bulk effective sample size was 2331 and the minimum tail effective sample size was 4699. Table 3 reports the model summary statistics. The credible interval indicates the range of values within which the effect has a 95% chance of falling given the data. The posterior probability indicates the chance that an effect falls above or below zero given the data. We follow our previous approach in interpreting a posterior probability of >.95 as strong evidence for an effect given the data, of >.85 as weak evidence for an effect, and of close to .5 as no evidence for an effect (see Engelmann et al., 2019; Kumarage et al., 2022). As illustrated in Figure 4, participants produced more passive responses following passive primes than active primes. The results of the model in Table 3 indicate there is strong evidence for this priming effect with the effect having a 99.9% chance of being greater than 0 given the data.

Table 2

Number and proportion of active, passive and other responses in each condition

Prime	Active		Passive		Other	
	N	%	N	%	N	%
Active	1119	87.4	41	3.2	120	9.4
Passive	1009	78.8	141	11.0	130	10.2

---

[3] The model syntax is provided below:

```
brm(formula = response ~ prime + (prime|ID) + (prime|target_verb/target_item),
      family = bernoulli(link = "logit"),
      warmup = 1000, iter = 3000, chains = 10, cores = 10, seed = 57, init = 0,
      control = list(adapt_delta = 0.95), data = data)
```

Figure 4

Proportion of passive responses out of active and passive responses in each condition

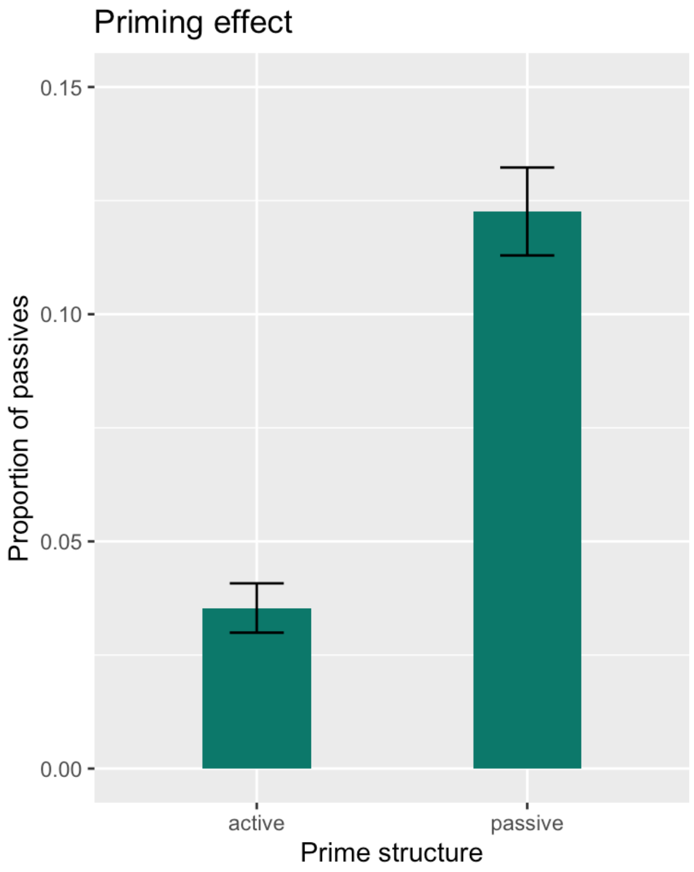


Table 3

Model summary of Bayesian mixed-effects logistic model

Effect	Estimate	95% credible interval	Posterior probability
Intercept	-4.09	-4.86 | -3.43	> .999
Prime	1.44	0.71 | 2.22	.999

Note: The posterior probability that an effect is smaller or larger than zero is calculated in a directional (one-tailed) hypothesis test.

Predicting Syntactic Priming Using Pupillometry

Our final and key hypothesis was that prediction error in processing the prime, as indexed by pupillometry, would predict participants' production of passives after passive primes. The GAMM analysis confirmed that there was a larger pupil response for passive than active primes and we also replicated the syntactic priming effect. We tested whether there was a relationship between pupil dilation and priming in two ways: (i) at the level of the participant, averaged across items, and (ii) at the level of the participant on a trial-by-trial basis. The former is a more conservative approach to the measurement of prediction error and mathematically assumes a stable level of difference in error as elicited by active and passive primes across the experiment. The latter allows pupil size to vary across individual primes across the course of the experiment. Since this is the first study to attempt to link prediction error indexed by pupillometry to priming, we thought it prudent to explore both options. However, if participants' expectations and therefore prediction error for the passive change during the task, then the trial-level predictor is more likely to capture a relationship between pupil dilation and syntactic priming.

Participant-Level Predictor

For each participant, we extracted the mean model-predicted pupil size during active and passive primes and calculated the difference score. We used model-predicted rather than measured pupil size as the model controlled for gaze position. We ran the syntactic priming model again including the z-scored measure of prediction error and its interaction with prime as predictors.[4] We included the full random effects structure, with an intercept and slope for prime by participants, and an intercept and slopes for prime, prediction error, and their interaction by items nested in verbs. The model convergence and posterior estimates were reliable: maximum Rhat = 1.00, minimum bulk effective sample size = 1952, minimum tail effective sample size = 3316. Table 4 reports the summary statistics from the model. Strong evidence for a priming effect remained, with participants producing more passives after passive primes than active primes. However, there was no evidence for prediction error as a predictor of either participants' overall passive production (main effect) or their propensity to be primed (interaction effect).

---

[4] The model syntax used for both the participant-level and trial-level predictors (error term) is provided below:

```
brm(formula = response ~ prime*error + (prime|ID) +
(prime*error|target_verb/target_item),
      family = bernoulli(link = "logit"),
      warmup = 1000, iter = 3000, chains = 10, cores = 10, seed = 57, init = 0,
      control = list(adapt_delta = 0.96), data = data)
```

Table 4

Model summary of model predicting syntactic priming with participant-level prediction error

Effect	Estimate	95% credible interval	Posterior probability
Intercept	-4.19	-4.99 | -3.52	> .999
Prime	1.47	0.69 | 2.28	> .999
Prediction error	-0.11	-0.75 | 0.51	.629
Prime*Prediction error	-0.08	-0.68 | 0.54	.603

Note: The posterior probability that an effect is smaller or larger than zero is calculated in a directional (one-tailed) hypothesis test.

Trial-Level Predictor

For each trial, we extracted the mean model-predicted pupil size. Unlike the participant-level predictor, this operationalisation of prediction error does not provide a measure of comparison between active and passive prime sentences. However, it has the advantage of being at the level of trial rather than participant. We included the z-scored measure of prediction error and its interaction with prime as predictors and the maximal random effects structure. The value of *adapt_delta* was increased from 0.95 to 0.96 to prevent divergent transitions. The model convergence and posterior estimates were reliable: maximum Rhat = 1.00, minimum bulk effective sample size = 2482, minimum tail effective sample size = 6240. Table 5 reports the summary statistics from the model. There was strong evidence for a priming effect and weak evidence for prediction error predicting the likelihood of being primed. Figure 5 plots the model-predicted production of passives at various values of prediction error following active and passive primes. The greater prediction error following an active prime, the higher active production (i.e., in Figure 5, greater values of pupil size leads to decreased likelihood of passive production). Passive production is also higher at larger pupil sizes following passive primes but additionally shows a trend towards being higher at particularly small pupil sizes. That is, smaller prediction error following a passive prime is associated with a slightly increased likelihood of passive production.

Table 5

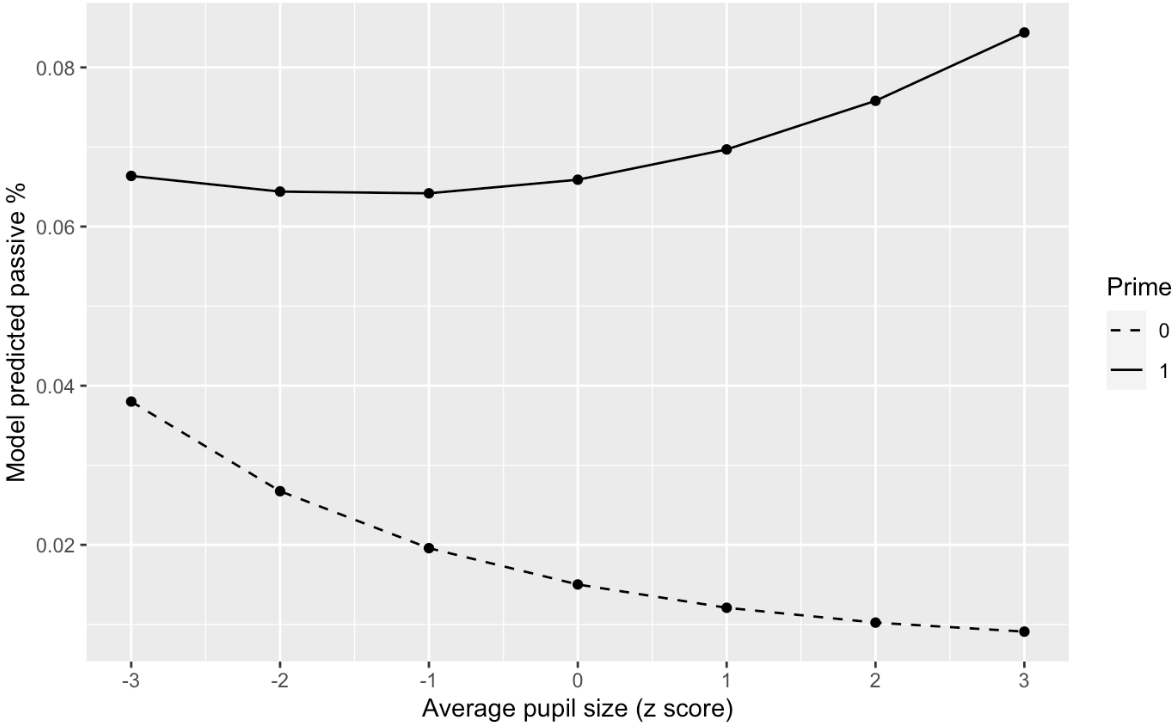
Model summary of model predicting syntactic priming with trial-level prediction error

Effect	Estimate	95% credible interval	Posterior probability
Intercept	-4.27	-5.19 | -3.52	> .999
Prime	1.56	0.72 | 2.48	> .999
Prediction error	-0.26	-0.73 | 0.20	.125
Prime*Prediction error	0.31	-0.21 | 0.85	.879

Note: The posterior probability that an effect is smaller or larger than zero is calculated in a directional (one-tailed) hypothesis test.

Figure 5

Model-fitted proportion of passive responses following active and passive primes at various levels of prediction error as measured by pupil size



Discussion

In this paper we aimed to index prediction error using pupillometry to predict syntactic priming in line with the error-based learning mechanism proposed by Chang et al. (2006). We found that the pupillary response was larger during passive than active primes, replicating syntactic complexity effects. We also replicated the syntactic priming effect, with participants producing more passive responses after passive than active primes. Using pupil size as a measure of prediction error, we found no evidence for participants' overall difference in pupil response predicting syntactic priming, but we found weak evidence that the trial-level size of the pupil response predicted syntactic priming.

Our first hypothesis was that the more complex and less frequent passive structure would be associated with a larger pupillary response in line with previously observed syntactic complexity effects (Beatty, 1982; Demberg & Sayeed, 2016; Just & Carpenter, 1993; Piquado et al., 2010). Indeed, we found that there was a significantly larger response to passive than active structures from 1115 to 3767ms after sentence-onset (Figure 3b). This is consistent with past studies that have typically found a peak of the pupil response 1000ms following the point of difficulty (Beatty, 1982: 1200ms; Demberg & Sayeed, 2016: 750-1250ms; Just & Carpenter, 1993: 1000-1500ms). Our study also found that the pupil response has a long latency: the difference between conditions is not evident until about 1100ms after sentence onset and peaks at about 2800ms, which is after sentence offset. Similar to Just and Carpenter (1993), the pupil response peaks earlier for the less complex structure. While passive sentences were longer than the active ones, previous findings suggest that the larger pupil response cannot be attributed merely to sentence length. Piquado et al. (2010) found a larger pupil size for recalling 12- vs 9-word sentences. However, the syntactic complexity of subject vs object relative clauses varies only with the order of words (*the gambler that signalled the dealer revealed the card* vs *the gambler that the dealer signalled revealed the card*) and they found a syntactic complexity effect in both the 9-word and 12-word conditions. Another example is Stanners et al.'s (1972) finding that *they are easy to please* requires greater cognitive effort to process than *they are eager to please*, which is the exact same length. We conclude that either the passive requires greater cognitive resources to process, resulting in a larger pupil response attributed to cognitive effort, or its low frequency in English results in prediction error during processing that is reflected in the pupillary response. Within a predictive processing account such as Chang et al. (2006), prediction error and cognitive effort are tightly linked because low frequency structures do not enjoy the facilitated processing of predictable ones and trigger additional cognitive activity in model updating. However, under more traditional approaches, cognitive effort can instead be explained by factors such as the requirement for argument movement in passives compared to actives (Frazier, 1987).

If the pupillary response indexes prediction error rather than cognitive effort, then the time course of the pupil response may provide insight into the time course of prediction error. We designed the study to include 500ms for picture apprehension before sentence onset in prime trials, under the assumption that participants would extract the gist of the picture in that period and expect the agent of the action to come first. In Chang et al.'s (2006) dual-path model, prediction during comprehension can occur under two circumstances: a *situated* event, wherein message content can be inferred from context, and a *messageless* event, wherein message content is unknown. If participants extract the gist of the picture, that is concepts, event roles and their bindings such as ELEPHANT-AGENT, FROG-PATIENT, BITE-ACTION[5], this constitutes a *situated* event. The sequencing system can use the message information to aid next-word predictions, in this case predicting *elephant* not *frog* to follow the determiner *the* in the initial noun phrase because the system knows that *elephant* is the agent of the action, and the network weights reflect the strong agent-first bias in English. On the other hand, in a *messageless* event, the sequencing system makes predictions based purely on its knowledge of word classes, syntactic categories, and their ordering. The emergence of a significant effect within about 1100ms following sentence-onset suggests that participants may have experienced prediction error early in the sentence, in line with processing a situated event and expecting an agent-first sentence. However, latencies from previous studies are not directly comparable because they do not use analysis techniques that reveal periods of significance but instead report the latency of the *peak* of the pupil response. The peak of the pupil response occurred later in the sentence, which would instead indicate prediction error when participants expect a main verb rather than auxiliary verb following the first noun phrase (*the elephant is* ***biting*** *the frog* vs *the frog is* ***being*** *bitten by the elephant*). This could align with either the processing of a messageless event, or passive morphology compounding the error previously encountered on hearing a patient-initial noun phrase.

Our second hypothesis that we would replicate a syntactic priming effect was also confirmed. As expected, participants produced more passives after passive than active primes. Therefore, we could test the key hypothesis of the study: syntactic priming is driven

---

[5] Note that we have used traditional thematic roles for simplicity. The dual-path model uses XYZ roles where the Y role corresponds to intransitive agents and transitive patients, X to agents, causers and stimuli, and Z to goals, locations and recipients.

by prediction error which can be indexed by pupillometry. Chang et al. (2006) propose that an error-based learning mechanism underlies syntactic priming. During prime comprehension, the syntactic processor makes next-word predictions and compares these to the actual input. The error associated with predicting exact lexical items (e.g., *elephant* vs *mammoth*) does not differ between active and passive prime sentences. However, in an active sentence (*the elephant is biting the frog*), the sequence of word classes and syntactic categories is associated with minimal prediction error, whereas when participants hear the first noun in a situated passive or the auxiliary verb in a messageless passive (*the <u>frog</u> is* ***being*** *bitten by the elephant*) and implicitly compare it to the predicted agent noun or main verb respectively, a large error signal is produced. This error backpropagates through the system, which adjusts the network weights that comprise syntactic representations and increases the likelihood of a passive structure. This translates to a greater likelihood of a passive response to the target item when the same representations are used for production. The greater the prediction error, the greater adjustment in network weights and therefore increase in passive production. The error-based learning account therefore predicts that if the larger pupil response to passive sentences indexes prediction error, then larger pupil size should predict participants' syntactic priming.

We found weak evidence for this proposal. First, taking an individual differences approach to operationalising prediction error, participants who least expect the passive should experience greater prediction error, adjustments to syntactic representations, and therefore syntactic priming. This prediction was not supported: participants' mean difference in pupil dilation between passive and active primes did not appear to index individual differences in the propensity to be primed. However, this operationalisation assumes that participants' syntactic representations do not change over the course of the experiment and their tendency to predict and produce passives is maintained throughout, counter to findings of expectation adaptation and cumulative syntactic priming effects (e.g., Fine et al., 2013; Jaeger & Snider, 2013). If we take the weak evidence for an effect, pupil size does appear to index trial-by-trial prediction error *within* participants. The more fine-grained measure of prediction error – average pupil size on each trial – weakly predicted syntactic priming. Figure 5 suggests that multiple factors could underlie this effect. Firstly, both actives and passives are more likely to be produced following a prime in that structure which produced particularly large prediction error. However, passive production also appears to be slightly more likely when passive primes are not particularly surprising. This could indicate experiment-level changes in representations and expectations: when participants increase their predictions of passives, they are both less surprising and more likely to be produced. This suggests evidence in favour of the proposition that pupil size represents a dynamic online measure of prediction error.

There are two factors that may have attenuated the strength of the effect resulting in only weak evidence. Firstly, similar to EEG, where electrodes detect the summation of electrical activity from the cortex, the pupil response is influenced by multiple cognitive processes (Hepach & Westermann, 2016). Prediction error is but one of the mental processes and events comprising the pupil response to prime sentence comprehension. Subtracting the pupil response to active primes from the pupil response to passive primes serves to isolate prediction error from other cognitive processing but the trial-based prediction error operationalisation was not a relative measure. Thus, although the participant-level operationalisation controls for factors external to prediction error, its assumption that prediction error is stably evoked by different structures may be incorrect. The item-level operationalisation assumes variation across trials, but not controlling for other factors may have introduced noise to the signal. Secondly, we may have required higher power to conclusively detect an effect, especially if the pupil response is not solely reflective of prediction error. According to Mahowald et al.'s (2016) power analysis for syntactic priming effects, our sample of 80 participants and 32 items has sufficient power (> 80%) to detect an interaction effect of the size of the lexical boost effect. The increase in syntactic priming for primes and targets with shared lexical content is a particularly large and reliable effect and much smaller effects are of theoretical interest. For an effect half the size, the power of our sample size to detect an effect is reduced to 50-60%. In this context, finding weak evidence for an effect is suggestive of a true effect, although future studies with higher power are required to confirm this. An additional point to note is that experimental paradigms that reliably produce group-level effects, such as syntactic priming, can require substantially larger sample sizes to observe individual variation (Hedge et al., 2018). Therefore, the power of our sample may be even lower for detecting individual differences effects. Whilst this poses an additional problem for the participant-level predictor, the larger issue remains the assumption of stable expectations.

Comparing our findings to those of Arai and Chang (2024) raises important methodological considerations. They measured prediction error using reading time on the final verb in active and passive primes in Japanese. While both reading time and syntactic priming increased for the condition where passives were least expected, reading time did not predict syntactic priming as expected. There does appear to be alignment between reading time measures and pupil dilation. Demberg and Sayeed (2016) used both measures in three self-paced reading experiments and found similar results in detecting grammatical violations, semantic violations and differentiating the processing of subject and object relative clauses. However, if we accept the weak evidence in our study, then pupillometry appears to be a more promising measure of prediction error than reading times. Self-paced reading tasks require participants to explicitly progress to the next word and so reading time may be influenced by extraneous factors related to deliberate action. Importantly, self-paced reading is not a reliable individual differences measure (Frinsel & Christiansen, 2024). In contrast to self-paced reading, pupillometry is an inherently implicit measure that can be measured during auditory rather than written sentence comprehension. It may more precisely capture prediction error as a result.

We chose a structural alternation which reliably produces large priming effects as a test case for combining pupillometry with syntactic priming. However, previous research has utilised a variety of manipulations of prime content that are intended to increase prediction error and do increase syntactic priming (e.g., manipulating expectations: Arai & Chang, 2024; verb-bias effects: Bernolet & Hartsuiker, 2010; inverse-frequency effects: Jaeger & Snider, 2013). Exploiting different structures and experimental designs comes with different advantages. For example, we compared the infrequent passive structure to the canonical active, but other structures may offer a larger frequency differential. There is preliminary evidence in developmental studies that the relative clause and adjective-noun alternation produces stronger priming than the active-passive alternation (Kumarage et al., 2024). If the prediction error associated with a relative clause (*the apple that is red*) relative to an adjective-noun structure (*the red apple*) is larger than when comparing passives to actives, the effect of pupil size predicting syntactic priming may be more easily observed. One drawback is that both these alternations involve one very basic and one marked structure, which may lead to conscious priming effects. The dative alternation enables cross-linguistic investigations and investigations of verb-biases that could control for this. The double object dative is less frequent in Dutch, whilst the prepositional object dative is less frequent in English. If the pupillary response to primes of the same structure varies according to the structure's distribution within the language and predicts priming, this could tie pupil size more definitively to prediction error than cognitive effort. Similarly, if the pupil response to primes of the same dative structure differs for verbs biased towards or against the structure it cannot be attributed to differences in syntactic complexity. Our results suggested some evidence for dynamic changes to expectations and syntactic representations during the experiment. Jaeger and Snider (2013) used an experimental design that lends itself well to investigating this issue. All participants were presented with an equal number of prepositional object and double object datives, but half saw primes in block order (e.g., all double objects then all prepositional objects), while the other half saw primes in alternating order. In the blocked condition, prediction error should show attenuation over the first block, then a sharp increase and attenuation again over the second block, whereas in the alternating condition prediction error should remain more stable. Pupil size could be compared at points in the experiment where prediction error is expected to be larger and smaller and correlated with syntactic priming at these times.

Future research may also address the limitations of this study. Firstly, studies likely require larger sample sizes than we were able to collect in order to definitively detect effects, especially under an individual differences approach. A second limitation was the uncertainty in deciding on how to operationalise prediction error from the pupil response. While GAMMs allow sophisticated modelling of the time course of the pupil response, it is less clear how to extract an equally sophisticated trial-by-trial measure of prediction error from that model. Given the GAMM analysis showed an overall larger pupil response for passive sentences, the average pupil size during the prime is a reasonable but coarse operationalisation. We did not have the computational power to include random smooths for each trial for each participant. Extracting these random effects from a model that did include them is one option for a more precise measure. Finally, finding an association between syntactic priming and pupil size using a trial-level but not participant-level predictor is suggestive of adaptation effects however analyses incorporating trial number are required to confirm this.

In conclusion, we found a larger pupil response elicited by passive when compared to active sentences, in line with previous syntactic complexity effects (e.g., Just & Carpenter, 1993). This finding indicates that the passive structure either requires greater cognitive effort to process than the active or is more unexpected and generates a larger error signal. Individual differences in the average pupil response difference did not predict participants' syntactic priming. However, a more fine-grained measure – prediction error on a trial-by-trial basis – weakly predicted syntactic priming. In light of power estimates for detecting interactions with syntactic priming, this supports the interpretation that pupil response, at least partially, reflects prediction error in syntactic processing and predicts syntactic priming effects, in line with the Chang et al. (2006) account. In light of the difficulty of measuring and operationalising cognitive processes, this finding opens a range of possible applications within research that investigates error-based learning and the links between expectations, frequencies, prediction error and syntactic representation and processing.

References for Chapter 4

Arai, M., & Chang, F. (2024). Influence of word-based prediction errors on syntactic priming. *Seijo University Economic Papers*, *243*, 97–120.

Aydın, Ö., & Uzun, İ. P. (2023). Pupil dilation response to prosody and syntax during auditory sentence processing. *Journal of Psycholinguistic Research*, *52*(1), 153–177. https://doi.org/10.1007/s10936-021-09830-y

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. https://doi.org/10.1037/0033-2909.91.2.276

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*(3), 455–461. https://doi.org/10.1016/j.cognition.2009.11.005

Bernolet, S., Collina, S., & Hartsuiker, R. J. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, *91*, 99–116. https://doi.org/10.1016/j.jml.2016.01.002

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, *31*(2), 163–186. https://doi.org/10.1016/0010-0277(89)90022-X

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*(2), 177–192. https://doi.org/10.1037//0096-3455.129.2.177

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*(1), 1–39. https://doi.org/10.1016/0010-0277(90)90035-I

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, *99*, 150–171. https://doi.org/10.1037/0033-295X.99.1.150

Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, *40*, Article e313. https://doi.org.10.1017/S0140525X16002028

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, *45*(4), 1322–1331. https://doi.org/10.3758/s13428-013-0327-0

Bürkner, P.-C. (2017). brms: An R package for bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chomsky, N. (1959). Review of Skinner's verbal behavior. *Language*, *35*, 26–58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Dal Ben, R. (2021). *SHINE_color: Controlling low-level properties of colorful images*. PsyArXiv. https://doi.org/10.31234/osf.io/fec6x

Dell, G. S., & Chang, F. (2014). The p-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634). https://doi.org/10.1098/rstb.2012.0394

Demberg, V., & Sayeed, A. (2016). The frequency of rapid pupil dilations as a measure of linguistic processing difficulty. *PLOS ONE*, *11*(1), Article e0146194. https://doi.org/10.1371/journal.pone.0146194

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, *63*(4), 639–645. https://doi.org/10.1080/17470210903469864

Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., Theakston, A., & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive Psychology*, *110*, 30–69. https://doi.org/10.1016/j.cogpsych.2019.02.001

Farmer, T. A., Fine, A. B., Yan, S., Cheimariou, S., & Jaeger, T. F. (2014). Error-driven adaptation of higher-level expectations during reading. *Proceedings of the Annual Meeting of the Cognitive Science Society, 36*, 2181–2186.

Fazekas, J., Jessop, A., Pine, J., & Rowland, C. (2020). Do children learn from their prediction mistakes? A registered report evaluating error-based theories of language acquisition. *Royal Society Open Science*, *7*(11), Article 180877. https://doi.org/10.1098/rsos.180877

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language*, *48*(2), 379–398. https://doi.org/10.1016/S0749-596X(02)00523-5

Fine, A. B., & Jaeger, T. F. (2016). The role of verb repetition in cumulative structural priming in comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(9), 1362–1376. https://doi.org/10.1037/xlm0000236

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10), Article e77661. https://doi.org/10.1371/journal.pone.0077661

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

Forbes, S. H. (2020). PupillometryR: An R package for preparing and analysing pupillometry data. *Journal of Open Source Software*, *5*(50), Article 2285. https://doi.org/10.21105/joss.02285

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Erlbaum.

Frinsel, F. F., & Christiansen, M. H. (2024). Capturing individual differences in sentence processing: How reliable is the self-paced reading task? *Behavior Research Methods*. https://doi.org/10.3758/s13428-024-02355-x

Garcia, R., Garrido Rodriguez, G., & Kidd, E. (2021). Developmental effects in the online use of morphosyntactic cues in sentence processing: Evidence from Tagalog. *Cognition*, *216*, Article 104859. https://doi.org/10.1016/j.cognition.2021.104859

Garcia, R., Roeser, J., & Kidd, E. (2023). Finding your voice: Voice-specific effects in Tagalog reveal the limits of word order priming. *Cognition*, *236*, Article 105424. https://doi.org/10.1016/j.cognition.2023.105424

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279. https://doi.org/10.1111/1467-9280.00255

Hartsuiker, R. J., Kolk, H. H. J., & Huiskamp, P. (1999). Priming word order in sentence production. *The Quarterly Journal of Experimental Psychology Section A*, *52*(1), 129–147. https://doi.org/10.1080/713755798

Häuser, K. I., Demberg, V., & Kray, J. (2018). Surprisal modulates dual-task performance in older adults: Pupillometry shows age-related trade-offs in task performance and time-course of language processing. *Psychology and Aging*, *33*(8), 1168–1180. https://doi.org/10.1037/pag0000316

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, *17*(3), 359–377. https://doi.org/10.1080/15248372.2015.1135801

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *47*(2), 310–339. https://doi.org/10.1037/h0078820

Kaschak, M. P., Kutta, T. J., & Jones, J. L. (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. *Psychonomic Bulletin & Review*, *18*(6), 1133–1139. https://doi.org/10.3758/s13423-011-0157-y

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*(3), 381–388. https://doi.org/10.3758/s13421-010-0042-3

Koch, X., & Janse, E. (2016). Speech rate effects on the processing of conversational speech across the adult life span. *The Journal of the Acoustical Society of America*, *139*(4), 1618–1636. https://doi.org/10.1121/1.4944032

Kramer, S. E., Lorens, A., Coninx, F., Zekveld, A. A., Piotrowska, A., & Skarzynski, H. (2013). Processing load during listening: The influence of task characteristics on the pupil response. *Language and Cognitive Processes*, *28*(4), 426–442. https://doi.org/10.1080/01690965.2011.642267

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology*, *50*(1), 23–34. https://doi.org/10.1111/j.1469-8986.2012.01477.x

Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring children. *Journal of Memory and Language*, *127*, Article 104374. https://doi.org/10.1016/j.jml.2022.104374

Kumarage, S., Donnelly, S., & Kidd, E. (2024). A meta-analysis of syntactic priming experiments in children. *Journal of Memory and Language, 138,* Article 104532. https://doi.org/10.1016/j.jml.2024.104532

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Mathôt, S. (2013). *A simple way to reconstruct pupil size during eye blinks*. https://doi.org/doi:10.6084/m9.figshare.688001

Mathôt, S., & Vilotijević, A. (2022). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. *Behavior Research Methods*, *55*(6), 3055–3077. https://doi.org/10.3758/s13428-022-01957-7

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560–569. https://doi.org/10.1111/j.1469-8986.2009.00947.x

Prasad, G., & Linzen, T. (2021). Rapid syntactic adaptation in self-paced reading: Detectable, but only with many participants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(7), 1156–1172. https://doi.org/10.1037/xlm0001046

Preuschoff, K., ’t Hart, M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline’s role in decision making. *Frontiers in Neuroscience*, *5*. https://doi.org/10.3389/fnins.2011.00115

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rutar, D., Colizoli, O., Selen, L., Spieß, L., Kwisthout, J., & Hunnius, S. (2023). Differentiating between Bayesian parameter learning and structure learning based on behavioural and pupil measures. *PLOS ONE*, *18*(2), Article e0270619. https://doi.org/10.1371/journal.pone.0270619

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition*, *89*(3), 179–205. https://doi.org/10.1016/S0010-0277(03)00119-7

Schluroff, M. (1982). Pupil responses to grammatical complexity of sentences. *Brain and Language*, *17*(1), 133–145. https://doi.org/10.1016/0093-934X(82)90010-4

Sirois, S., & Brisson, J. (2014). Pupillometry. *WIREs Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

Stanners, R. F., Headley, D. B., & Clark, W. R. (1972). The pupillary response to sentences: Influences of listening set and deep structure. *Journal of Verbal Learning and Verbal Behavior*, *11*(2), 257–263. https://doi.org/10.1016/S0022-5371(72)80086-0

Tromp, J., Hagoort, P., & Meyer, A. S. (2016). Pupillometry reveals increased pupil size during indirect request comprehension. *Quarterly Journal of Experimental Psychology*, *69*(6), 1093–1108. https://doi.org/10.1080/17470218.2015.1065282

van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., & Wood, S. N. (2019). Analyzing the Time Course of Pupillometric Data. *Trends in Hearing*, *23*, 1–22. https://doi.org/10.1177/2331216519832483

van Rij, J., Wieling, M., Baayen, R. H., & Rijn, H. van. (2022). *itsadug: Interpreting time series and autocorrelated data using GAMMs*.

Wendt, D., Dau, T., & Hjortkjær, J. (2016). Impact of background noise and sentence complexity on processing demands during sentence comprehension. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00345

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *73*(1), 3–36.

Yngve, V. H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, *104*(5), 444–466.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear & Hearing*, *31*(4), 480–490. https://doi.org/10.1097/AUD.0b013e3181d4f251

Zellin, M., Pannekamp, A., Toepel, U., & Van Der Meer, E. (2011). In the eye of the listener: Pupil dilation elucidates discourse processing. *International Journal of Psychophysiology*, *81*(3), 133–141. https://doi.org/10.1016/j.ijpsycho.2011.05.009

Zhang, F., Jaffe-Dax, S., Wilson, R. C., & Emberson, L. L. (2019). Prediction in infants and adults: A pupillometry study. *Developmental Science*, *22*(4), Article e12780. https://doi.org/10.1111/desc.12780

Ziegler, J., Bencini, G., Goldberg, A., & Snedeker, J. (2019). How abstract is syntax? Evidence from structural priming. *Cognition*, *193*, Article 104045. https://doi.org/10.1016/j.cognition.2019.104045

Chapter 5:

General Discussion

In this thesis, I used syntactic priming in combination with a longitudinal design, meta-analysis, and pupillometry to investigate the mechanisms underlying two related cognitive abilities: syntactic acquisition and syntax processing. The aims of the thesis were twofold: (i) to provide evidence regarding underlying mechanisms in acquisition and processing, and (ii) to extend our understanding of the utility of the syntactic priming methodology in investigating such questions. In this chapter, I summarise the main findings, discuss their implications for both aims, and suggest future directions for further research.

Summary of Findings

In **Chapter 2**, I investigated the potential mechanisms underlying syntactic acquisition using syntactic priming. Using data from the Canberra Longitudinal Child Language Project (Kidd et al., 2018), I analysed the emergence and trajectory of children's abstract priming and lexical boost effects over a key period of development, from 36 – 54 months. I derived and tested predictions from three theories of syntactic acquisition: the nativist, lexicalist, and implicit learning approaches, which assume that abstract priming and lexical boost effects emerge and develop in different ways.

The abstract priming effect emerged at the earliest timepoint (36 months), indicating early abstract knowledge of the passive. The lexical boost effect emerged between 42 – 48 months, suggesting it is driven by a later emerging mechanism different to abstract priming. In line with its late emergence, the magnitude of the lexical boost effect significantly increased over development. There was weak evidence that the abstract priming effect decreased over development, though this evidence was strong in an exploratory analysis on a more homogenous subsample of children who produced a passive at 36 months. There was substantial variation in children's priming effects at the earliest timepoint and their linguistic proficiency predicted their likelihood of being primed. Overall, the findings of Chapter 2 best supported the implicit learning account, which predicts early abstraction but decreasing abstract priming as representations become less susceptible to change, and attributes the lexical boost effect to explicit memory, which is still developing in young children. There were two qualifications to this support for the account. Firstly, priming effects for one of the lexical items tested differed substantially from the others and we interpret the results excluding this item. Secondly, the decreasing priming effect had strong evidence in exploratory analyses but weak evidence in the full sample.

Abstract priming and lexical boost effects have been widely studied by developmental researchers, without leading to consensus on their emergence and development. Chapter 2 addressed the gap by contributing the first longitudinal evidence on this topic; however, another approach is to synthesise evidence across studies. In **Chapter 3**, I performed a meta-analysis of syntactic priming studies in children. Using advanced meta-analytic techniques, I was able to include dependent effect sizes from studies to investigate moderators of the main effect (i.e., abstract syntactic priming). Additionally, a variety of other moderators were included to determine whether any further insights could be gained into both the acquisition of syntax and the use of the syntactic priming methodology in children. A publication bias analysis and power analysis evaluated the quality of the evidence from syntactic priming studies.

In support of Chapter 2's finding that syntax is abstracted early, I found that lexical overlap increases but is not required to observe syntactic priming. However, the trajectory of abstract priming and the lexical boost could not be reliably determined. The finding that between-subjects designs induce larger priming effects than within-subjects ones suggested the involvement of explicit processes in syntactic priming studies. Priming was larger under canonical animacy configurations. The strength of enduring priming effects could not be precisely estimated. Priming effects varied substantially by the structural alternation under investigation, highlighting the low generalisability of current findings given the bias towards studies in English on passive priming. Another limitation of the current literature was that although the abstract priming effect is robust to publication bias, studies are underpowered to detect interaction effects.

Chapter 2's findings best supported the implicit learning account. Notably, the finding that abstract syntactic priming decreased over development supports a decrease in the prediction error that children experience on encountering passives as they experience more examples and increase their predictions of them. However, Chapter 2 only measured the outcome of increased expectations of the passive via production. **Chapter 4** aimed to directly measure prediction error to determine its role in syntactic priming using pupil dilation, which has been associated with prediction error in other contexts. The error-based learning mechanism is posited to operate throughout the lifetime, so I chose to first test the methodology in adults. I manipulated prime sentence structure (active or passive) and measured pupil size during prime comprehension and coded the structure of participants' target responses. I hypothesised that pupil size would be larger in response to passive than active primes and that larger pupil size would predict greater syntactic priming.

As predicted, pupil size was significantly larger during passive than active prime sentences and I replicated a syntactic priming effect. Although a participant-level relative measure of prediction error did not predict syntactic priming, a trial-level measure of pupil size did weakly predict syntactic priming. This constituted promising but preliminary evidence that prediction error and its role in the error-based learning of syntax can be indexed by pupillometry.

Implications of Main Findings

Evidence for Error-Based Learning

The first aim of this thesis was to examine mechanisms underlying syntactic acquisition and processing. Chang et al.'s (2006) error-based implicit learning account provides an account of both children's syntactic acquisition and syntactic priming in adults. The key assumption of the account, which distinguishes it from activation-based accounts of syntactic priming and nativist and lexicalist accounts of syntactic acquisition, is that learning occurs via predictive processing. Empirically, the model predicts that the magnitude of prediction error, which can vary developmentally or by structure and context, determines the magnitude of syntactic priming. In this section I discuss how key findings from Chapters 2 and 4 contribute evidence of this link to the existing literature.

In adults, the connection between prediction error and syntactic priming has been investigated by manipulating the predictability of primes. Support for error-based learning comes from inverse-frequency effects, where priming effects are larger for less frequent structures (e.g., Jaeger & Snider, 2013), and prime-surprisal effects, where priming of a structure is larger when the prime verb typically prefers the other structure in an alternation (e.g., Bernolet & Hartsuiker, 2010). Chapter 4 provided further support for the link between the magnitude of prediction error, representational change, and syntactic priming. An index of prediction error demonstrated the predicted response to the manipulation of prime structure and weakly predicted syntactic priming, which was a promising result given limited power.

In children, evidence for the role of prediction error from inverse-frequency and prime-surprisal effects is less conclusive. Inverse-frequency effects have not been explicitly investigated, with mixed evidence from studies that have run analyses of both active and passive priming (larger passive priming: Gámez & Shimpi, 2016; Gámez & Vasilyeva, 2015; equivalent priming: Messenger et al., 2012). There is also conflicting evidence for prime-surprisal effects. Peter et al.'s (2015) finding of prime surprisal effects based on dative verb-biases was not replicated by Fazekas et al. (2020) or Donnelly et al. (2024), although Fazekas et al. (2020) did find that children's post-test productions shifted towards the structure they heard only in surprising contexts.

However, the error-based learning account also predicts variations in prediction error developmentally. Younger children who have acquired a structure but represent it weakly are susceptible to greater adjustment in those representations than those who have encountered more instances of the structure (Rowland et al., 2012). Chapter 2 found weak evidence of this decreasing effect in the full sample of children and strong evidence in those who had definitely acquired the passive by 3 years. Decreasing abstract priming effects have also been observed in two cross-sectional studies of dative priming (German: Kholodova et al., 2023; English: Rowland et al., 2012). Unlike the error-based learning account, neither the lexicalist nor nativist accounts predict a decreasing priming effect at any point in development. One caveat to note is that if we consider the corresponding increase in passive productions in Chapter 2 as evidence of cumulative syntactic priming (i.e., where the influence of primes persists over filler trials; Branigan & McLean, 2016), then this finding instead represents an increase in syntactic priming over development. This would be consistent with the suggestion that children learn to predict rather than predicting to learn language (Rabagliati et al., 2016; Gambi et al., 2018; but see Gambi et al., 2021 where predictive ability was associated with greater subsequent vocabulary development). As discussed next, another explanation that is consistent with error-based learning is that children's baseline expectation for passives within the task increased over the course of the study.

One final point raised by this thesis regarding the link between prediction error and priming is the impact of baseline frequency. When a structure is infrequent, its baseline production may increase over the course of an experiment or study. For example, the baseline (i.e., unprimed) production of the passive measured after a period of priming is typically higher than prior to priming (e.g., Kidd, 2012; Messenger, 2021). Though this is evidence of implicit learning, it may muddy the expected association between prediction error and priming. For example, in Chapter 4, when the pupil response to a passive prime was particularly small, indicating minimal prediction error, participants appeared slightly more likely to produce a passive than at medium pupil sizes. This could reflect later trials where passives were less surprising but also more likely to be produced in general. Similarly, but over a longer time scale in Chapter 2, children increased their passive productions over the course of the study, alongside a less obvious decrease in the priming of the structure.

Overall, this thesis adds to existing evidence for a role of predictive processing and error-based learning in the processing and acquisition of syntax. However, the dynamic nature of expectations and representations of structure should be considered in interpreting results. Other key findings of this thesis bear upon two other predictions of the error-based learning account: that syntax is abstracted early and that the lexical boost is explained by an explicit memory mechanism. I discuss these in the next two sections.

Early Abstraction or Lexical-Specificity?

The abstractness of early syntactic representations is a focus of debate in developmental research, including the syntactic priming literature (Ambridge & Lieven, 2011). Chapter 2 found that passive representations are abstracted early, with lexical overlap not required to observe priming even at the earliest timepoint (36 months). Chapter 3 corroborated this finding across (mostly passive) syntactic priming studies, though in older children on average. However, ruling out lexicalist acquisition mechanisms may be premature given that they are actually invoked by Chang et al. (2006), and given findings from Donnelly et al. (2024) and the item-specific effects observed in Chapter 2.

The Chang et al. (2006) model learns to sequence words in ways that minimise prediction error. Through this mechanism it can both *acquire* abstract representations of a structure and subsequently *adjust* the likelihood of its prediction and production. Whilst priming research has focused on the developmental trajectory implied by the latter stage (Rowland et al., 2012), Chang et al. (2006) state that lexical-specificity similar to lexicalist proposals is consistent with the earlier stage. Although the model predicts abstract knowledge early in development, until it has acquired abstract enough representations for changes to transfer between prime and target, syntactic priming gradually increases. This nuance can explain a key finding of Chapter 2: abstract priming decreased but this was best observed in a subsample of children who produced a passive at the earliest timepoint, confirming they had acquired the structure. If abstract priming shows a non-linear trajectory, then decreases in priming from children with an abstract representation will be offset by increases from children still developing one. However, although the decreasing priming effect that is proposed following the acquisition of abstract knowledge was observed, the period of lexically-specific knowledge preceding this was not.

One explanation is that although 3 years is considered young in the syntactic priming literature, enough children in the sample may have acquired abstract knowledge of the passive by this age such that lexically-specific priming effects could not be observed. For example, the diary data from which Tomasello (1992) developed the hypothesis that children's early syntactic knowledge is dependent on specific lexical items covered a period of acquisition from 15 to 24 months (see also Israel et al., 2000). Although the passive is usually assumed to have a delayed developmental trajectory, Chapter 3's finding that there was no significant dative priming across studies suggests we may observe an even later trajectory for priming effects in this structural alternation. Donnelly et al. (2024) analysed complementary data from the Canberra Longitudinal Child Language Project (Kidd et al., 2018), which measured dative priming at the later three timepoints in Chapter 2 (42, 48, and 54 months). They found that priming at 42 and 48 months was dependent on lexical overlap, with abstract priming emerging at 54 months, without a lexical boost. This finding supports early lexical-specificity, with a later developmental trajectory for dative than passive priming.

Both Donnelly et al.'s (2024) findings and item-specific effects in Chapter 2 link distributional features of children's language input to their mental representations. Donnelly et al. (2024) suggest that the degree of lexical-specificity is related to features of the input. In datives, the restricted set of verbs that participate in the dative alternation and the dominance of one lexical item (*give*) in children's input lead to later abstraction than for passives. In Chapter 2, I found item-based priming effects for *push*, the test verb that was most frequent in children's input. Unlike the other verbs, *push* was not primed by other verbs: priming was only evident in the verb overlap condition. This suggests a role for lexically-based knowledge alongside abstract representation, perhaps for particularly frequent lexical items. This is consistent with lexicalist conceptualisations of the end point of acquisition (Tomasello, 1998) and compatible with the Chang et al. (2006) model, which is able to acquire both abstract representations and lexical restrictions on structure, such as verb preferences (Twomey et al., 2014).

So far, I have discussed how well the findings of this thesis, in combination with the existing literature, corroborate predictions of Chang et al.'s (2006) account. The role of prediction error is well supported, as is early abstraction in acquisition, though it likely exists alongside lexically-based knowledge. In the next section, I consider a theoretical proposition rather than a prediction of the computationally-instantiated model: the involvement of a separate explicit memory mechanism, which accounts for lexical boost effects.

Explicit Processing in Syntactic Priming

Error-based learning does not appear to be the only mechanism involved in syntactic priming. Explicit processes have been invoked to explain lexical boost effects since they cannot be explained by the error-based learning account (Chang et al., 2006, 2012). In support of this dual-mechanism proposition, lexical boost effects decay rapidly compared to abstract priming effects (Branigan & McLean, 2016; Hartsuiker et al., 2008; Mahowald et al., 2016). However, Bernolet et al. (2016) demonstrated that even the abstract priming effect decays in strength when sentences intervene between prime and target, concluding that explicit memory processes are additionally involved in this effect at short time lags. Short-term abstract priming cannot be completely explained by explicit memory (see findings that syntactic priming is observed in the presence of amnesia: Ferreira et al., 2008; Heyselaar et al., 2017; Yan et al., 2018), but researchers should consider the underlying mechanisms their study's design is likely to recruit when interpreting syntactic priming results (Tooley, 2023). A key finding of Chapter 3 was that the choice of within- or between-subjects design had a large impact on the magnitude of syntactic priming. A simple explanation is that children are simply exposed to more instances of the structure in a between-subjects than within-subjects design. However, in adults, who are not subject to the same limitations on experiment length as children, cumulative priming – priming in studies with between-subjects blocked designs – is similarly stronger than priming in alternating designs (Mahowald et al., 2016). Instead, I attributed this result to the presentation of only one structure in between-subjects designs providing a cue that is available to conscious awareness and explicit memory processes rather than implicit error-based learning. Repetition of a single structure, particularly an infrequent one, could be salient enough to lead to conscious awareness of a "rule" for its use in the task context and an explicit memory of the structure. If syntactic priming is intended to index implicit error-based learning, then designing studies to avoid the recruitment of other processes is important (Tooley, 2023). A simple recommendation is to use lag 2 designs since two intervening sentences are enough to eliminate lexical boost effects that are attributed to explicit memory (Branigan & McLean, 2016; Hartsuiker et al., 2008). Another direction is the development of implicit measures that tap into processing mechanisms more directly, such as the pupil size index of prediction error explored in Chapter 4. Arai and Chang (2024) did not find the expected results for a reading time measure. However, the comprehension priming literature offers other real-time methodologies, such as EEG and eye-tracking, which could be measured during prime comprehension and correlated with the production of primed structures during targets (Tooley, 2023).

A theoretical implication of the involvement of both mechanisms in abstract priming effects is that the interplay between error-based learning and memory-based processes requires further attention. Unlike error-based learning, the underlying nature of the explicit memory mechanism is underspecified, including its developmental trajectory. In explaining lexical boost effects, one suggestion is that open-class lexical overlap acts as an episodic cue to the structure of the prime sentences (Bernolet et al., 2016; Hartsuiker et al., 2008). Thus, following a prime like *Grover gave Oscar a book*, either the request to use *give* in a target or the conceptual structure of a giving event (e.g., as invoked by an experimental picture) cues the speaker to use a double object dative. However, comparing the results of Chapter 2 with those of Donnelly et al. (2024) suggests that this process is fragile in children, and that they must first have abstract knowledge of a particular structure to benefit from the cue provided by lexical overlap. Chapter 2 found an increasing lexical boost effect that emerged later than abstract priming at some point between 42 and 48 months of age. Donnelly et al. (2024) found that priming was initially lexically-specific but at 54 months children demonstrated abstract priming. Despite doing so for passives, children did not display a lexical boost effect for datives at 54 months. Most cross-sectional studies have found increasing lexical boost effects for the dative over development (Kholodova et al., 2023; Peter et al., 2015; Rowland et al., 2012 but see Donnelly et al., 2024). This suggests that the memory-driven lexical boost only comes online sometime after abstract representations have been acquired. A detailed account of the explicit memory mechanism needs to explain this delay. Future studies could manipulate the use of explicit memory via cueing in priming tasks and compare the magnitude of syntactic priming effects between conditions in children who have different levels of structural knowledge (ideally, abstract versus lexically-based). Other challenges for an explicit memory account are that it is typically invoked to explain lexical boost findings (Chang et al., 2006; 2012), but one study found that people with amnesia demonstrated an intact lexical boost effect despite declarative memory impairments (Yan et al., 2018). Additionally, in order to explain its proposed influence on short-term abstract priming effects (Bernolet et al., 2016), explicit memory of the ordering of abstract syntactic categories rather than an episodic memory of the prime sentence would need to be possible.

To summarise the implications of this thesis so far: the evidence suggests that multiple mechanisms are at play in syntactic priming and acquisition. Error-based learning is a convincing candidate for long-term priming effects, as well as contributor to short-term priming, and can accommodate both abstract and lexically-specific representation at various stages of syntactic acquisition. There is evidence for an additional explicit memory mechanism that explains lexical boost effects and the remaining portion of short-term priming. However, a detailed account of this mechanism, how it interacts with error-based learning and how it develops is not clear. In the next section, I will explore the limitations and future directions of the research in this thesis.

Limitations and Future Directions

The second aim of this thesis was to extend the utility of syntactic priming in evaluating questions about syntactic acquisition and processing through a combination of longitudinal design, meta-analysis and pupillometry. This section discusses the advantages and limitations of each and potential avenues of future research.

Chapter 2 highlighted the value of longitudinal research in development. In a longitudinal design, age varies within- rather than between-subjects. This increases the precision of estimating developmental effects because individuals are held constant across timepoints. In contrast, cross-sectional designs are unable to compare individuals to themselves and must assume that each sample is representative of a particular developmental stage, an assumption that may not be valid in syntactic priming research. Chapter 2 demonstrated substantial variation in children's syntactic priming effects in a sample of children whose age differed by at most one month. However, samples in cross-sectional syntactic priming studies typically contain a 20-month age range (Messenger et al., 2022) and likely represent children at a variety of developmental stages. Consistent with comparing heterogenous samples of children, cross-sectional research in syntactic priming has identified varying developmental trajectories for abstract priming effects even when investigating the same structure (i.e., dative; decreasing: Kholodova et al., 2023; Rowland et al., 2012; increasing: Peter et al., 2015; stable: Donnelly et al., 2024: study 1). Additionally, non-linear effects cannot be observed unless tracking the emergence and development of priming effects *within* individuals (or unless developmental stage is tied closely with chronological age). The *Early Abstraction or Lexical Specificity* section discussed a potential non-linear developmental trajectory for abstract priming. Because of its longitudinal design, Chapter 2 was able to isolate the developmental trajectory for participants who had already acquired the passive at 36 months to provide some support for this proposal. At present only the transitive and dative alternations have been studied longitudinally, with contrasting patterns of results (c.f. Chapter 2 and Donnelly et al., 2024). Future longitudinal research covering different periods of development could clarify whether a trajectory from lexically-specific to abstract priming to abstract priming with lexical boost is supported, or whether different structures display different trajectories, which is also plausible given the variable priming effects observed in Chapter 3. In this vein, extending longitudinal research to a variety of structures and languages would be extremely valuable for our understanding of syntactic acquisition. Longitudinal research is extremely resource-intensive so another option is to apply some of the principles to cross-sectional designs. For example, recruiting samples with narrow age ranges may allow more fine-grained investigations of developmental effects.

As the first meta-analysis of the field, Chapter 3 was able to identify key summary effects and limitations of the developmental syntactic priming literature. Crucially, the effect of within- vs between-subjects design would not have been identified at the study-level and points to the involvement of explicit processes in priming. As discussed previously, future research could further explore the role of explicit memory as an underlying mechanism of short-term priming effects. This finding also points to a limitation of Chapters 2 and 4.

Chapter 2 employed a *snap* game method, which increases children's engagement in the task but may also introduce explicit memory mechanisms. Chapter 4 measured priming without intervening sentences; adding a lag between prime and target could have eliminated short-term priming effects not attributed to error-based learning. The suggestion is that the effects observed in the two studies may not have been fully implicit. These are elements of experimental design that future research should consider.

The role of prediction error in syntactic processing is a key proposal of Chang et al.'s (2006) account and the focus of much syntactic priming research. Chapter 4 provided promising but preliminary evidence that prediction error can be indexed by pupil dilation. The value of implicitly measuring processing during prime sentences and predicting syntactic priming is beginning to be understood (Arai & Chang, 2024; Tooley, 2023). Replicating this result would allow the application of this method to a range of questions. For example, semantic representations also appear to be subject to priming, sometimes separately (Bock et al., 1992; Chen et al., 2022) and sometimes in interaction with syntactic structure (Vasilyeva & Gámez, 2015). Given the support for prediction error initiating weight changes in syntactic representations, whether pupil dilation surprisal effects are observed for semantic representations may disentangle whether they are encompassed by the network weights subject to error-based learning. Another future direction is extending the use of pupillometry to index prediction error to children. Inverse-frequency effects and prime-surprisal effects have conflicting evidence in children. Donnelly et al. (2024) point out that verb-biases must be acquired before they can induce prime-surprisal. Using an implicit pupil size measure could confirm whether primes are in fact producing the expected prediction error.

Finally, as with all meta-analyses, Chapter 3 is limited by the existing shortcomings of the literature. In addition to developmental syntactic priming studies being underpowered to detect interaction effects, which are often the focus of the research, it is difficult to generalise findings beyond the active-passive alternation in English. Chapter 3 found that it is the most commonly studied alternation in children, and in adults it is the second most studied after the dative alternation (Mahowald et al., 2016). A limitation of both Chapter 2 and Chapter 4 is that they reflect this widespread bias. This is problematic because the English passive is a structure where various syntactic and semantic features align. Chapter 3 found that animacy bindings interact with syntactic priming. Further, there is evidence that semantic features alone can be primed, such as the order of thematic roles in sentences with identical constituent structure (e.g., spray-load locatives: Chang et al., 2003), and the content emphasis of structures with different syntax (Vernice et al., 2012). Passives can prime alternative structures in languages with multiple structures that emphasise the patient (Gámez et al., 2009; Vasilyeva & Waterfall, 2012), but in English patient-emphasis and passive syntax are almost completely correlated. Accordingly, Ziegler et al. (2019) claim that passive priming in English could be almost entirely attributed to priming of semantic features. This raises issues of generalisability. The findings of Chapter 2 and Donnelly et al. (2024) diverge in ways that raise important questions about acquisition. Features of other languages offer a diversity of phenomena to be explained and advantages for experimental design. For example, symmetrical voice in Tagalog can disentangle syntactic role from thematic role and ordering (Garcia et al., 2023; Garcia & Kidd, 2020) and priming of syntactic alternatives in Russian and Spanish can disentangle information focus from syntactic structure (Gámez et al., 2009; Vasilyeva & Waterfall, 2012). Given the interactions between syntactic priming and features of syntactic structures that have already been observed, we can expect priming effects to vary on this basis rather than aligning with the findings for a single structure. Only by understanding the pattern of syntactic priming effects over structures and languages can we determine what is general in syntactic acquisition and representation, and where the nuances lie.

Conclusion

This thesis aimed to contribute evidence about the underlying mechanisms of syntactic acquisition and processing and the utility of syntactic priming research. Following predictions from Chang et al.'s (2006) computational model, I demonstrated that there is evidence for error-based learning as a mechanism underlying priming, but that this must be considered alongside a short-term explicit memory mechanism. In acquisition, despite support for the error-based learning account, there is a role for lexically-based knowledge. I showed that using syntactic priming with a longitudinal design, meta-analysis and online implicit measures such as pupillometry reveals insights that cannot be investigated otherwise. Future directions for research include clarifying the trajectory of priming effects early in development, investigating how underlying processing mechanisms interact, researching how prediction error effects manifest alongside developing representations in children, and broadening the evidence base for syntactic mechanisms to diverse languages and structures.

References for Chapter 5

Ambridge, B., & Lieven, E. V. M. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press. https://doi.org/10.1017/CBO9780511975073

Arai, M., & Chang, F. (2024). Influence of word-based prediction errors on syntactic priming. *Seijo University Economic Papers*, *243*, 97–120.

Bernolet, S., Collina, S., & Hartsuiker, R. J. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, *91*, 99–116. https://doi.org/10.1016/j.jml.2016.01.002

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*(3), 455–461. https://doi.org/10.1016/j.cognition.2009.11.005

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, *99*, 150–171. https://doi.org/10.1037/0033-295X.99.1.150

Branigan, H. P., & McLean, J. F. (2016). What children learn from adults' utterances: An ephemeral lexical boost and persistent syntactic priming in adult-child dialogue. *Journal of Memory and Language*, *91*, 141–157. https://doi.org/10.1016/j.jml.2016.02.002

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*(1), 29–49. https://doi.org/10.1016/S0010-0277(03)00123-9

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chen, X., Hartsuiker, R. J., Muylle, M., Sarah Slim, M., & Zhang, C. (2022). The effect of animacy on structural Priming: A replication of Bock, Loebell and Morey (1992). *Journal of Memory and Language*, *127*, Article 104354. https://doi.org/10.1016/j.jml.2022.104354

Donnelly, S., Rowland, C., Chang, F., & Kidd, E. (2024). A comprehensive examination of prediction-based error as a mechanism for syntactic development: Evidence from syntactic priming. *Cognitive Science*, *48*(4), Article e13431. https://doi.org/10.1111/cogs.13431

Fazekas, J., Jessop, A., Pine, J., & Rowland, C. (2020). Do children learn from their prediction mistakes? A registered report evaluating error-based theories of language acquisition. *Royal Society Open Science*, *7*(11), Article 180877. https://doi.org/10.1098/rsos.180877

Ferreira, V. S., Bock, K., Wilson, M. P., & Cohen, N. J. (2008). Memory for syntax despite amnesia. *Psychological Science*, *19*(9), 940–946. https://doi.org/10.1111/j.1467-9280.2008.02180.x

Gambi, C., Gorrie, F., Pickering, M. J., & Rabagliati, H. (2018). The development of linguistic prediction: Predictions of sound and meaning in 2- to 5-year-olds. *Journal of Experimental Child Psychology*, *173*, 351–370. https://doi.org/10.1016/j.jecp.2018.04.012

Gambi, C., Jindal, P., Sharpe, S., Pickering, M. J., & Rabagliati, H. (2021). The relation between preschoolers' vocabulary development and their ability to predict and recognize words. *Child Development*, *92*(3), 1048–1066. https://doi.org/10.1111/cdev.13465

Gámez, P. B., & Shimpi, P. M. (2016). Structural priming in Spanish as evidence of implicit learning. *Journal of Child Language*, *43*(1), 207–233. https://doi.org/10.1017/S0305000915000161

Gámez, P. B., Shimpi, P. M., Waterfall, H. R., & Huttenlocher, J. (2009). Priming a perspective in Spanish monolingual children: The use of syntactic alternatives. *Journal of Child Language*, *36*(2), 269–290. https://doi.org/10.1017/S0305000908008945

Gámez, P. B., & Vasilyeva, M. (2015). Increasing Second Language Learners' Production and Comprehension of Developmentally Advanced Syntactic Forms. *Language Learning and Development*, *11*(2), 128–151. https://doi.org/10.1080/15475441.2014.894873

Garcia, R., & Kidd, E. (2020). The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development*, *16*(4), 399–425. https://doi.org/10.1080/15475441.2020.1814780

Garcia, R., Roeser, J., & Kidd, E. (2023). Finding your voice: Voice-specific effects in Tagalog reveal the limits of word order priming. *Cognition*, *236*, Article 105424. https://doi.org/10.1016/j.cognition.2023.105424

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238. https://doi.org/10.1016/j.jml.2007.07.003

Heyselaar, E., Segaert, K., Walvoort, S. J. W., Kessels, R. P. C., & Hagoort, P. (2017). The role of nondeclarative memory in the skill for language: Evidence from syntactic priming in patients with amnesia. *Neuropsychologia*, *101*, 97–105. https://doi.org/10.1016/j.neuropsychologia.2017.04.033

Hsu, D.-B. (2019). Children's adaption to input change using an abstract syntactic representation: Evidence from structural priming in Mandarin-speaking preschoolers. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02186

Israel, M., Johnson, C., & Brooks, P. (2000). From states to events: The acquisition of English passive participles. *Cognitive Linguistics*, *11*(1/2), 103–129. https://doi.org/10.1515/cogl.2001.005

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kholodova, A., Peter, M., Rowland, C. F., Jacob, G., & Allen, S. E. M. (2023). Abstract priming and the lexical boost effect across development in a structurally biased language. *Languages*, *8*(4), Article 264. https://doi.org/10.3390/languages8040264

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, *33*(2), 393–418. https://doi.org/10.1017/S0142716411000415

Kidd, E., Junge, C., Spokes, T., Morrison, L., & Cutler, A. (2018). Individual Differences in Infant Speech Segmentation: Achieving the Lexical Shift. *Infancy*, *23*(6), 770–794. https://doi.org/10.1111/infa.12256

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Messenger, K. (2021). The persistence of priming: Exploring long-lasting syntactic priming effects in children and adults. *Cognitive Science*, *45*(6), Article e13005. https://doi.org/10.1111/cogs.13005

Messenger, K., Branigan, H., Buckle, L., & Lindsay, L. (2022). How does syntactic priming experience support language development? In K. Messenger (Ed.), *Syntactic Priming in Language Acquisition: Representations, mechanisms and applications* (pp. 57–81). John Benjamins Publishing Company.

Messenger, K., Branigan, H. P., & McLean, J. F. (2012). Is children's acquisition of the passive a staged process? Evidence from six- and nine-year-olds' production of passives. *Journal of Child Language*, *39*(5), 991–1016. https://doi.org/10.1017/S0305000911000377

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience*, *31*(1), 94–105. Scopus. https://doi.org/10.1080/23273798.2015.1077979

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63. https://doi.org/10.1016/j.cognition.2012.06.008

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge University Press.

Tomasello, M. (1998). *The new psychology of language: Cognitive and functional approaches to language structure*. Erlbaum.

Tooley, K. M. (2023). Structural priming during comprehension: A pattern from many pieces. *Psychonomic Bulletin & Review*, *30*(3), 882–896. https://doi.org/10.3758/s13423-022-02209-7

Twomey, K. E., Chang, F., & Ambridge, B. (2014). Do as I say, not as I do: A lexical distributional account of English locative verb class acquisition. *Cognitive Psychology*, *73*, 41–71. https://doi.org/10.1016/j.cogpsych.2014.05.001

Vasilyeva, M., & Gámez, P. B. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in syntactic priming. *Journal of Experimental Child Psychology*, *138*, 15–30. https://doi.org/10.1016/j.jecp.2015.04.009

Vasilyeva, M., & Waterfall, H. (2012). Beyond syntactic priming: Evidence for activation of alternative syntactic structures. *Journal of Child Language*, *39*(2), 258–283. https://doi.org/10.1017/S0305000911000055

Vernice, M., Pickering, M. J., & Hartsuiker, R. J. (2012). Thematic emphasis in language production. *Language and Cognitive Processes*, *27*(5), 631–664. https://doi.org/10.1080/01690965.2011.572468

Yan, H., Martin, R. C., & Slevc, L. R. (2018). Lexical overlap increases syntactic priming in aphasia independently of short-term memory abilities: Evidence against the explicit memory account of the lexical boost. *Journal of Neurolinguistics*, *48*, 76–89. https://doi.org/10.1016/j.jneuroling.2017.12.005

Ziegler, J., Bencini, G., Goldberg, A., & Snedeker, J. (2019). How abstract is syntax? Evidence from structural priming. *Cognition*, *193*, Article 104045. https://doi.org/10.1016/j.cognition.2019.104045