2023

What Relationships Exist Between Words In The Lexical-Semantic Systems Of Toddlers?

Fitzpatrick, Nadine

https://pearl.plymouth.ac.uk/handle/10026.1/20623

http://dx.doi.org/10.24382/2766
University of Plymouth

What Relationships Exist Between Words In The Lexical-Semantic Systems Of Toddlers?

by

Nadine Fitzpatrick

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

July 2022

Copyright Statement

Acknowledgements

I am extremely thankful to Professor Caroline Floccia who has been a fantastic supervisor, offering support and guidance at all stages of the PhD. Her direct approach and honesty have pushed me to develop my ideas and my work in new and creative ways.

I am grateful to the School of Psychology for funding my PhD, with a special thank you to Professor Chris Mitchell (PG Coordinator at the time) who found the extra funds that enabled me to do the PhD. The tech office has also been extremely supportive, and I could not have analysed my data without the bespoke coding platform they developed for me, with particular thanks to Anthony Mee.

I am also grateful to my colleagues at the BabyLab, especially the placement students, who helped with the laborious task of hand-coding eye movement data from the online experiments.

A special thank you must be given to the families who participated in the research as without them, none of this would have been possible.

A final thank you to my friends and family, for showing interest in my research, particularly my husband, Clayton (and dog Luna!), who endured me talking about it constantly and who was always patient and supportive.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was financed with the aid of a studentship from the University of Plymouth.

Presentations at Conferences:

Fitzpatrick, N., & Floccia, C. (2022, June 9th - 11th). *A Resource of Word Associations in 3-year-olds which are Not Captured by Adult Associative Norms* [Poster Presentation]. Workshop on Infant Language Development (WILD), San Sebastian, Spain.

Fitzpatrick, N., & Floccia, C. (2022, August 24th – 26th). *Testing Infants Online using the Gorilla Experiment Builder* [Poster Presentation]. Lancaster Conference on Infant and Child Development (LCICD), Lancaster, UK.

Word count of main body of thesis: 62, 891

Signed

Date 04/07/2022

What Relationships Exist Between Words In The Lexical-Semantic Systems Of Toddlers?

Nadine Fitzpatrick

Abstract

Investigating how infants first establish relationships between words is a necessary step towards understanding the qualitative shift children make to an organised and complex interconnected network of semantic relationships which characterises a mature, adult lexical-semantic system. Since little is known about the word-word associations in infants that establish this network of meanings (Arias-Trejo & Plunkett, 2009), this thesis sought to, first, document the word associations (WA)s that young monolingual and bilingual children produce and then compare these to adult WAs. A concurrent aim was to establish a database of child-specific WAs as a resource for future studies. Second, to understand how a network of meaning establishes in different groups during infancy, an online semantic priming paradigm was developed due to the Covid-19 pandemic. The aim was to see how words are organised in the emergent lexical-semantic system by replicating in-lab findings and extending these to explore different infant groups. In parallel, this paradigm was used to validate the WAs found in monolingual and bilingual children.

Findings from Chapter 1 revealed that children share some of the WAs that adults exhibit in a mature lexical-semantic system. However, a large number of WAs shared by children were not represented in the WA norms of adults. This could indicate that adult norms under-represent the associations of children, as they might not capture the unique developmental stage and life experience of 3-year-olds. This research presents a resource of child-specific associated word pair stimuli for future studies.

Findings from Chapter 2 indicate that lexical-semantic links might be more robust in the lexical-semantic system of a 3-year-old when they capture associative meaning compared to taxonomic meaning. Furthermore, running infant studies online can replicate in-lab findings, though it remains unclear if this is only true of certain paradigms.

List Of Contents

List Of Tables

List Of Figures

1 Introduction

Every word we know and use in daily language comprehension and production has a mental representation which is stored in the mental lexicon, a construct best conceptualised as a mental dictionary (Jarema & Libben, 2007; Oldfield, 1966). Each entry in the mental lexicon contains information about the speaker's knowledge of a word's meaning, pronunciation, and its syntactic function (Fay & Cutler, 1977).

The organising principle of this lexicon uses networks to link words which are similar in sound, meaning, and use (i.e. the contexts in which a word is used and the words with which it frequently co-occurs), in an inter-connected system known as the lexical-semantic system. The lexical-semantic system's name derives from its dual nature: a network of connections between individual words (the lexical aspect); and a system which encompasses our representation of concepts and categories within these concepts (the semantic element) (Collins & Loftus, 1975; Collins & Quillian, 1969; Oldfield, 1966; Smith et al., 1974). Through its organisation, the lexical-semantic system facilitates quick and efficient retrieval of words for automatic speech comprehension and production. This happens through the activation of other words that are similar in meaning, sound, or which co-occur frequently together, so as to facilitate the processing of upcoming words in the transient linguistic stream (Collins & Loftus, 1975; Collins & Quillian, 1969; Oldfield, 1966; Smith et al., 1974). At the same time, the interconnected nature of the lexical-semantic system can inhibit word retrieval due to competition between similar words and sounds, in turn slowing down retrieval of the intended word (discussed in more detail, below).

In the adult literature, semantic priming studies have largely uncovered what we now know about lexical-semantic structure. A semantic priming study involves presenting word pairs which are either related or unrelated in meaning. When a target word (e.g. *cat*) is preceded by a word which is semantically related to the target (e.g. *dog*), it results in faster word recognition/response times compared to primes which are unrelated (e.g. *plate* - *cat*) and is widely accepted as evidence for the inter-connected organisation of words in the mental lexicon (Meyer & Schvaneveldt, 1971).

Another task frequently used to explore the organisation of words in the mental lexicon and concepts in the mind is a word association or 'free' association task. Free association tasks have been employed in psycholinguistic research for over a century already (e.g. Jung, 1910) and involve a participant saying or writing the first word that comes to mind on hearing or reading a cue word (Nelson et al., 2000). Free association is thought to index the mapping of lexical knowledge. The way in which words are associated provides information about the organisation of the mental lexicon and how this organisation affects performance in certain tasks involving memory and verbal response (Comesaña et al., 2014).

While the lexical-semantic system in adults has been explored in great depth to date, its developmental course in very young children has been stymied by methodological restrictions that come with testing young participants (Wojcik, 2017). With advancements in technology, however, procedures for testing online language comprehension are now viable for testing young, pre-verbal infants which enables us to investigate how mature lexical-semantic networks are formed: from initial emergence, through development, to maturation. This is not a trivial trajectory to map, because even in the monolingual child, the time-course and organisation of the lexicon does not seem to be a miniature version of the adult (Chow et al., 2017).

The Intermodal Preferential Looking paradigm (IPL, Golinkoff et al., 1987) has commonly been used to study language development in infants, providing evidence to suggest that the emerging infant lexical-semantic system does differ from the adult system. In this task, a word is spoken or played auditorily while the infant is presented with two visual stimuli on a screen. Preferential looking to the target visual stimulus is thought to index word recognition of the spoken word. Only by the age of two does the infant system begin to resemble the adult system in terms of a sensitivity to the phonological and semantic similarity between words, rather than encoding words in isolation (Chow et al., 2017).

During the first year of life, concepts are still being developed in an infant's mind, and words start to become associated with their referents (Wojcik, 2017). By the second year of life, lexical-semantic relations begin to develop. This is a crucial developmental phase for the lexical-semantic system, as word comprehension and the learning of new words are directly impacted by it. For example, how the lexical-semantic system is organised can influence the learning of new words (Wojcik, 2017). This happens when newly-learned words activate related words and concepts, leading to better consolidation of the new word into the lexicon (Borovsky et al., 2010). Understanding more about this system offers insight into language development and more broadly, how knowledge is represented in the mental lexicon.

The organisation of language in the mental lexicon is made more complex when considering infants who are learning two languages. A monolingual will initially learn one word per concept, but a bilingual will inevitably learn two: a word from each of their languages. For a bilingual, two words representing the same concept can share meaning, but not phonology (i.e. translation equivalents), or they can share meaning and some phonological overlap (i.e. cognates- e.g. *cat* in English and *kat* in Dutch). Cognates present a unique opportunity to investigate how the development of a bilingual's lexical-semantic system might differ to that of a monolingual. While a bilingual generally has two words for the same concept, in the case of cognates, overlapping lexical representations exist due the phonetic similarity of both words, so it is of empirical interest to explore if there is a qualitative difference in a cognate's representation in the mental lexicon, compared to words which are not cognates.

While a significant amount of monolingual and bilingual infant research has focused on the learning of individual words and their mental representations, semantic meaning has received less attention, and more is needed to assist our understanding of the development of lexical-semantic networks (Wojcik, 2017). For the few studies addressing development of the lexical-semantic network in infants, there is a particular concession that is of concern. In the experimental design of these studies, there seems to be an over-reliance on adult word association norms to inform experimental stimuli selection for use in infant populations. If the infant lexical-semantic system is not a miniature version of the adult system, then it cannot be assumed that the words adults connect in their semantic system (i.e. the word associations they have) will be mimicked in infants. While the authors of studies relying on adult associative norms acknowledge this as a compromise, the rationale behind their employment lies in the absence of a normed set of word associations for infants, or at the very least, associative norms for older children. Consequently, it must be questioned whether using adult associative norms has the potential to confound findings in infant studies by exaggerating or underestimating semantic priming effects. There is a clear need for a set of infant word association norms so that studies addressing research questions based on the development of the immature lexical-semantic system, are using word associations that have been validated as existing in the infant lexicon, rather than in a mature adult lexical-semantic system. The availability of a normed set of infant word associations would benefit not only the psycholinguistic research community, but potentially have value for clinical applications. Understanding word associations in typically developing children could prove insightful to children with a language or developmental delay. This is certainly true of adult norms, which provide a baseline from which to assess whether a certain population might differ from the norm (Nelson et al., 2004).

The overarching aim of this thesis is to better understand the development and organisation of the lexical-semantic system in young children and to consider how this system, in different infant populations, develops into the mature lexical-semantic system found in adults. This thesis uses two methodological approaches to explore our key aims: first, a database of infant word associations is compiled of 3-year-olds' language production in a free association task (see **Chapter 1:** Word Associations). A typically-developing child has already experienced a large vocabulary growth spurt and is able to understand hundreds of words by 24 months of age (Hamilton et al., 2000). At the age of three, a child will have sufficient knowledge and experience of words and word combinations to attempt a language production task, such as a free association task. No research to date has explicitly asked children this young what words they associate with common nouns, rather this has been assumed and the only evidence of word associations comes from children older than four years of age (Newman, 1970; Wojcik & Kandhadai, 2020). The results of this study are important to validate whether the word combinations often used in infant semantic studies, but taken from adult word associative norms, truly do represent the word associations in the infant lexicon.

Second, we explore whether the lexicon is already interconnected at both the semantic and lexical levels by using a semantic priming paradigm on different infant populations, and we seek to validate the findings from Chapter 1 through this paradigm (see **Chapter 2:** Semantic Priming). Monolinguals consistently show a semantic sensitivity, demonstrated by the effect of semantic priming in an Intermodal Preferential Looking (IPL) task at 24 months (see Arias-Trejo & Plunkett, 2009, 2013), with bilinguals mimicking this behaviour slightly later, at 30 months (Jardak & Byers-Heinlein, 2019). Taking the older of the two age ranges, monolinguals and bilinguals will be tested in semantic priming studies, to first, replicate prior findings using an adapted version of the primed IPL task in an online modality which was developed due to the Covid-19 pandemic; second, to test the effect of cognates on word recognition and semantic priming to observe for any qualitative difference in access to phonologically and semantically overlapping words in two languages, which might help understand how lexical-semantic organisation is affected by knowledge of more than one language; and finally, to compare word pairs generated in the infant word association task (Chapter 1) to word pairs commonly taken from adult associative norms for use as closely related prime-target words in infant experiments.

The remainder of this chapter outlines key concepts inherent to the lexical-semantic system and presents an overview of the literature on monolingual and bilingual infant lexical-semantic development in comparison to what is known about the adult population.[1]

1.1 The Mature Lexical-Semantic System

1.1.1 Monolingual Adults

What we know about a word is not restricted to its inherent meaning, but it extends to how a word relates and can be used in conjunction with other words. This interconnectedness exists in a semantic system which also represents knowledge about events, which is integral to language and thinking more generally, as well as cognitive functioning (Johnson-Laird, 1987) .

One of the fundamental questions in language development is how linguistic knowledge is acquired and organised. Of equal importance is understanding the mechanisms which enable retrieval of these words when required for speech comprehension or production. Nowadays, there is a general consensus that language processes belong to one of three types of representations: concepts (semantic), words (lexical), or sounds (phonological) (Caramazza, 1997; Dell, 1986; Levelt et al., 1999; Roelofs, 1992, 2003). The semantic level relates to the meaning of a word, whereas the lexical level is where the word form is represented, which connects to the semantic features to which the word can be attributed, and also has information regarding word class (e.g. noun) and other syntactic features

[1] An overview of the literature on word associations and associative norms will be provided in Chapter 1 and an overview of findings from semantic priming studies and the IPL paradigm will be presented in Chapter 2.

(Levelt et al., 1999). These first two levels of representation are shared for language comprehension and production, but phonological representations differ for each modality (Levelt, 1999). Sound forms for spoken word production and spoken word comprehension have output and input levels, respectively.

To illustrate, fluency in speech production is enabled through successful retrieval of the intended words: a process which is dependent on the availability of semantic and phonological information (Seiger-Gardner & Schwartz, 2008). A speaker must correctly select a word which matches their intended concept then access its abstract lexical representation (lemma) in a process called lexical selection. These words are accessed and retrieved from the mental lexicon which are then placed in a structural frame in which morphological, phonological, and phonetic information is retrieved ready for articulation (Dell, 1986; Levelt et al., 1999). By contrast, spoken word comprehension requires matching the phonological input with the lexical form to successfully identify its meaning (Hillis et al., 1990; Shelton & Caramazza, 1999). Thus, the order in which lexical and semantic representations are processed is the opposite for spoken word comprehension and production, with the latter moving from the semantic to the lexical and then the phonological level.

However, both of these processes activate not only the intended word but also words similar in sound or meaning, or related to the intended word, to differing levels of activation through spreading activation in the language system (Collins & Loftus, 1975). For example, when a word such as *boat* is heard, this activates semantically-related words (e.g. *sail*, *sea*) and phonologically-related items (e.g. *bow*, *float*). In an adult population, the connections in these semantic and phonological networks are robust, enabling their measurement during spoken recognition and production, using eye-movement techniques (Bergelson & Aslin, 2017), and neuro-imaging (Kutas & Hillyard, 1980). Evidence for the co-activation of multiple candidates during language processing comes from eye-tracking during semantic priming studies[2] in the adult literature (Meyer & Schriefers, 1991). Competition occurs between these possible candidates for selection (Dell, 1986; Levelt et al., 1999), which is due to the co-activation of lexical representation in language production, and the co-activation of semantic features in language comprehension. This subsequently can interfere with how efficiently language comprehension and production is performed. For example, in behavioural, language production studies, presenting semantically-related words early in word production inhibits responses and studies presenting phonologically-related words late in word production find facilitative retrieval of a response (Schriefers et al., 1990).

Similarly, using neuroimaging techniques such as event-related potential (ERP) measures of the electroencephalogram (EEG), which is the trace of voltage over time, semantic processing can be captured. Voltage fields are caused by active neurons and the fluctuation of large networks of neurons can be measured by using electrodes on the scalp. The ERP measures positive and negative waves of activity, which when time-locked to stimulus onset, is identifiable according to valence and latency. Kutas and Hillyard (1980) first discovered the N400, which is a negative wave of neural activity approximately 400ms after stimulus onset. The N400 is sensitive to semantic processing such that a semantically unrelated word preceding another word results in a large N400 component. This indicates the activity of word comprehension and integration of semantic information in context

[2] See the introduction to Chapter 2 of this thesis for a literature review of adult priming studies.

(McNamara, 2005). Conversely, when a semantically-related word precedes a word, the N400 is smaller or not present at all (Bentin et al., 1985; Rugg, 1985). This is because the integration of a semantically-related word into context is less effortful than words which are unrelated (Chwilla et al., 2000). The N400 component has indexed semantic relatedness (Federmeier & Kutas, 1999), congruency (Kutas & Hillyard, 1980), and category organisation (Heinze et al., 1998).

Since this thesis focuses predominantly on the semantic level of representation, we will now differentiate types of semantic meaning, before moving on to discuss the lexical-semantic system in other populations.

In the most simplistic sense, the term ‘semantic’ refers to meaning which can be shared between concepts. Take *table* and *bed* as examples. In a western context, both share similarities in meaning as both are large pieces of furniture traditionally constructed out of wood or other materials. Semantic relativity can be demarcated into sub-categories with the two most empirically studied being first, a taxonomic relationship, which refers to items that can be categorised through shared semantic/perceptual properties, within a hierarchy e.g. animals, mammals, dog breeds; and the second most studied semantic associate is an associative/thematic relationship, which refers to meaning formed through personal experience due to the co-occurrence of words in time and space for a given individual. In the above example of *table* and *bed*, the semantic relationship is taxonomic whereas *table* and *chair* share taxonomic and associative meaning as they are both items of furniture and co-occur in contexts with a high rate of frequency.

To understand how semantic memory is structured, the adult literature has explored whether semantic priming derives from the associative or taxonomic relations between words (McRae & Boisvert, 1998; Perea & Rosa, 2002). Results indicate that both taxonomic (McRae & Boisvert, 1998) and associative (Ferrand & New, 2004) relationships have a priming effect. Adult studies have shown that words both taxonomically and associatively related (e.g. *chair* - *table*) show the most reliable effect of semantic priming and this phenomenon is commonly referred to as a 'priming boost' (McRae & Boisvert, 1998; Perea & Rosa, 2002). The distinction between associative and taxonomic semantic relativity is especially pertinent in the context of infants since research indicates that both taxonomic and associative/thematic relationships contribute to the formation of semantic networks (Arias-Trejo & Plunkett, 2009, 2013; Styles & Plunkett, 2009).

1.1.2 Bilingual Adults

A bilingual population is a unique context in which to explore the representation of word meaning. This is because a bilingual speaker will have two lexical representations for any given concept, compared to a monolingual speaker who will just have one. Of empirical interest, then, is how a bilingual speaker organises these two labels for the same concept in the mental lexicon, and whether a bilingual's two languages interact in the lexical-semantic system in a facilitatory or inhibitory sense, at the semantic or lexical level (Bilson et al., 2015).

To explore questions regarding the structure of the language system and the mechanisms underpinning language processing in a bilingual population, two types of words are of particular interest: translation equivalents and cognates. Translation equivalents are words which represent the same concept, i.e. they are direct translations of one another in any two languages, but they have different phonology (e.g. *car* and *auto*- car in Dutch), therefore sound different. Cognates, on the other hand, are words which share the same meaning and have very similar phonology (e.g. *foot* and *voet*- foot in Dutch).

Adult bilinguals have been shown to automatically activate both languages in parallel (Kroll et al., 2006; Thierry & Wu, 2007) when use of only one language is required. This evidence can be found in behavioural studies that demonstrate that cognates facilitate processing during language production, compared to processing control words (Costa et al., 2000). This facilitation effect is evidenced in bilingual processing of cognates due to the semantic and phonological activation in both languages which is thought to speed up language production. In the case of translation equivalents in the bilingual adult literature, evidence for semantic priming during language comprehension is stronger across languages for translation equivalents than it is for semantically related words in the two languages, but which are not translation equivalents (Schoonbaert et al., 2009).

A question that pervades the bilingual literature is whether language - and more specifically lexical - processing in a bilingual is language-specific or non-specific (Marian & Spivey, 2003a). That is, the question is whether interaction occurs between the two languages during lexical processing or not. Empirical findings indicate that activation is initially not language specific when moving from the conceptual to the lexical system (Costa & Santesteban, 2004; Hermans et al., 1998). It is not clear, however, whether lexical selection itself is language-specific or not (Lemhöfer et al., 2018). Those supporting the latter (Abutalebi & Green, 2007; Green, 1998; Hermans et al., 1998) propose that lexical representations in both languages compete for selection. Those supporting the former

(Colomé, 2001; Costa et al., 1999) argue that lexical competition is restricted to one language. Most existing research into the language-specificity of bilingual lexical selection focuses on the adult bilingual population, however; some researchers have sought to explore the same phenomenon in young bilinguals (e.g. DeAnda et al., 2016) to ascertain whether a child's "representational system is best conceptualized as separated by language or integrated, as in adults" (Persici et al., 2019, p. 104).

Nonetheless, a widely accepted view is that language processing is slower for bilinguals compared to monolinguals (de Groot et al., 2002; Gollan et al., 2011) since bilinguals are constantly managing two languages simultaneously. Under experimental conditions, bilinguals exhibit slower word recognition compared to monolinguals, demonstrated in a number of task types (Martin et al., 2012). One reason for longer processing during bilingual word recognition might relate to implicit naming. There is empirical evidence that monolingual infants (Mani et al., 2012; Mani & Plunkett, 2010, 2011) and adults (Meyer & Damian, 2007) implicitly name objects on seeing a visual display in an experimental setting. In the case of a bilingual, implicit naming might involve the co-activation of a translation equivalent as well as the label in the target language. Evidence for bilingual adults activating both the L1 and L2 label for an object has not been consistently demonstrated and might be influenced by the language of immersion and language of testing. There is some evidence of dual activation when testing bilinguals immersed in the L1 and testing the L1 (Von Holzen & Mani, 2014; Weber & Cutler, 2004), but not if testing the L2 (Weber & Cutler, 2004). Similarly, when testing bilinguals immersed in the L2, there is evidence for activation of both labels for an object when testing the L1 (Spivey & Marian, 1999, Marian & Spivey, 2003a, Marian &

Spivey, 2003b), but not always (Wu & Thierry, 2011). Consequently, it is clear that a consistent replicable finding is not yet attainable, and it remains to be determined if the bilingual infant population would demonstrate the same pattern of findings.

Considering the evidence mentioned so far, it suggests that a bilingual's languages do interact at various levels of processing, with potential for both a facilitatory and an inhibitory effect. It would seem that different language combinations and different periods of language exposure might modulate the organisation of a bilingual's representations in the mental lexicon and understanding how these representations initially form in an immature system and evolve into a mature system is important to inform our understanding of the developmental trajectory between the two systems.

1.2 The Immature Lexical-Semantic System

1.2.1 Monolingual Children

There has been significant progress in understanding semantic organisation and its neural basis in adults, yet its origins and developmental trajectory are less clear. Early research in this field began with explicit questioning of children and their knowledge of word meanings (Gelman & Wellman, 1991; Keil, 1983), helping to lay the foundations of our current understanding. Techniques to explore implicit knowledge in infants came much later (Meints et al., 1999), but acquiring reliable and replicable evidence for the origins of semantic knowledge is complex.

Little is known about the word-word associations in infants that establish the network of meanings in the mature semantic system (Arias-Trejo & Plunkett, 2009). To understand how children acquire meaning requires us to understand how they acquire concepts, and how different concepts relate to one another. Either words are learnt and understood in isolation and become connected in a network of relationships within the mental lexicon, which develops into a mature system; or infants are born with this network already in place and integrate newly-learned words into it (Arias-Trejo & Plunkett, 2013).

Research into the foundations of the infant lexicon has established that infants comprehend common nouns by 12 months, and from as young as 6 months (Bergelson & Aslin, 2017; Delle Luche et al., 2014; Tincoff & Jusczyk, 1999). This word-world knowledge accelerates rapidly by the second year of life in both receptive and productive abilities (Reznick & Goldfield, 1992), establishing a mental lexicon with many lexical entries. By 24 months, there is evidence for semantic sensitivity to the relationships between words (Bergelson & Aslin, 2017). More specifically, in the case of implicit semantic priming[3], evidence in behavioural (Arias-Trejo & Plunkett, 2009; Willits et al., 2013) and electrophysiological studies (Rämä et al., 2013; Torkildsen et al., 2007) on infants is relatively robust at 24 months of age, with some evidence suggesting the establishment of semantic relationship already at 18 months in an event-related potential (ERP) study (Sirri & Rämä, 2015), in children with large vocabularies (Rämä et al., 2013), and in a head-turn preference procedure (Delle Luche et al., 2014).

Another technique to explore the evolution of the lexical-semantic system and investigate spoken word recognition in very young children is the Intermodal Preferential Looking Paradigm (IPL, Golinkoff et al., 1987). In this task, a word is spoken or played auditorily while

[3] A literature review of semantic priming in infants is presented in Chapter 2.

the infant is presented with two visual stimuli on a screen. Preferential looking to the target visual stimulus is thought to index word recognition of the spoken word. The IPL paradigm measures the tendency a listener has to look at a visual display when a component from the display is referred to: either explicitly or implicitly. It therefore does not necessitate an overt response and results show that a target object is fixated in the absence of any task requirement.

The reason listeners fixate a target picture might be in order to relate the visual to the auditory input (Altmann & Kamide, 2007; Huettig et al., 2011). Mapping the auditory to the visual world is a strategy embraced in everyday situations e.g. when learning a new skill and following a demonstration. The two often complement one another making it beneficial for this dual processing. Matching auditory information to the visual environment is usually achieved by directing one's attention and eye gaze to the relevant objects in this environment (Huettig et al., 2011). This enables the listener to recognise an object and access related information which might facilitate cognition or speech processing (Malpass & Meyer, 2010).

When young participants perform an IPL task investigating semantic relationships, longer looking times to target pictures relative to neutral trials indexes a priming effect of semantic-relatedness (Floccia et al., 2020). Arias-Trejo and Plunkett (2009) tested 18-month-old and 21-month-old infants in an IPL task using associatively and taxonomically related, or unrelated word pairs. Semantic priming was found in the 21-month-old group, but not at 18 months. While previous research on priming has found phonological priming in this younger, 18-month-old group (Mani & Plunkett, 2010), it would seem that semantic knowledge is not implicit until 21 months. In a later study, Arias-Trejo and Plunkett (2013) tested word pairs that were associatively related, taxonomically related, or unrelated: thus differentiating between the two main types of semantic relatedness as a primary focus. Two groups of participants were tested: aged 21 months and 24 months. Both associative and taxonomic priming was found in the older, 24-month-old group, but not in the 21-month-olds. The authors extrapolated that a "priming boost" may have caused an effect of semantic priming in 21-month-olds in an earlier study (Arias-Trejo & Plunkett, 2009) to be measurable, but not in this study, owing to the word pairs not being associatively and taxonomically related as in the earlier study. No effect of semantic priming in the 21-month-old group in the later study (Arias-Trejo & Plunkett, 2013) was possibly due to relatedness being separated out into associative or taxonomic word pairs. These findings align with a semantic priming effect found at 24 months, but not at 18 months (Styles & Plunkett, 2009).

These findings might suggest that before 21 months, priming effects do not exist. However, in another behavioural paradigm, Delle Luche et al. (2014) used related (e.g. animals) and unrelated (e.g. clothes and food) word lists in a head-turn preference procedure[4] experiment and found that on average, infants listened longer to related lists of words, compared to unrelated lists of words. They took this as evidence for an 18-month-old's sensitivity to the semantic relationship between words in the mental lexicon. This is supported by evidence from brain imaging (e.g., Rämä et al., 2013).

In the case of verbs, 2-year-olds demonstrate semantic knowledge by anticipating the subject of the verb according to semantic relatedness. Mani and Huettig (2012) showed that on hearing the verb *eat*, 2-year-olds were more likely to fixate an image of something edible

[4] In this procedure, a child's attention is directed to the left or right side of a booth using flashing lights before auditory stimuli are played from the left/right side. The child's attention, indicated by their head movement to the auditory stimulus is measured.

while listening to the sentence containing the verb and its subject, before hearing the referent. Toddlers with a larger expressive vocabulary were more reliable in their prediction of the upcoming noun than those with a small vocabulary.

When learning new words, 2-year-olds perform better when the word in question belongs to an already-established category in the child's lexicon (Borovsky et al., 2016). Therefore if a child knows a larger number of words that can be categorised as a type of food, learning another word in this category will be easy relative to learning a word belonging to a category in which there are few exemplars in the child's lexicon (DeAnda et al., 2016). Thus, 2-year-olds encode relatedness between category members which facilitates word learning (DeAnda et al., 2016).

To sum up, it seems that the lexical-semantic system develops around the age of two with infants understanding the meaning of individual words by 12 months, but only appreciating the relationship in meaning between words in the second year of life, amounting to robust evidence at 24 months of age. Based on findings from priming studies, both taxonomic and associative relationships contribute to structuring a semantic network in the emerging lexicon. Furthermore, 2-year-olds evidence semantic sensitivity to verbs and utilise category membership to facilitate word learning.

While it still remains unclear how the lexical-semantic system establishes during infancy, and how new words are integrated into this system, advancements in computational modelling have enabled predictions to be tested about language acquisition, based on existing empirical findings. Language-processing computational modelling uses a controlled environment in which variables can be manipulated to predict their effect on processing, which can be used to inform theories about how language systems work (Shook & Marian, 2013). For example, Hills et al. (2009) used longitudinal analyses of 130 nouns that children have usually learnt by 30 months to model the structure of the emerging lexicon. They modelled development using networks of connections based on feature overlap and based on associations, and found the latter to be a better growth model, but argue that perceptual and functional feature overlap structure the infant lexicon, much like the taxonomic groups found in the adult lexicon. In another study, Hills et al. (2010) analysed vocabulary development using CDI scores from parents and found that words were preferentially acquired based on adult associative norms.

This modelling data converges with behavioural evidence to suggest that taxonomic and associative relationships help structure the infant lexical-semantic system.

1.2.2 Bilingual Children

The development of meaning in the bilingual mental lexicon has received minimal attention to date, with little known about the nature of semantic development in these learners (DeAnda et al., 2016). Whether a bilingual child maps the same developmental trajectory as a monolingual child when establishing meaning in words and creating connections between different words, is still open to debate. Bilinguals do reach language milestones at the same rate as their monolingual counterparts (Werker & Byers-Heinlein, 2008), which might suggest comparable semantic development in the two populations. However, Bilson et al. (2015) demonstrated a protracted naming of concepts in bilinguals relative to monolinguals, and it remains unclear if the reason bilingual children take longer to name a concept is due to the distributed exposure to its label, across two languages (Bilson et al., 2015). Some research suggests a language delay in bilingual children due to this split linguistic input from two, rather than one language (Hoff et al., 2012). However, while the act of naming may be delayed, bilingual infants seem to develop conceptual knowledge at the same rate as monolingual children, but the way words are learnt seems to differ (Bilson et al., 2015). These differences may have a qualitative effect on the way semantic networks develop in bilingual infants and it is important to disambiguate such differences to establish whether theories of language acquisition, which are grounded in monolingual findings, are generalisable to other infant populations.

In the case of translation equivalents, as a bilingual child's vocabulary grows, the number of translation equivalents known increases proportionately (Legacy et al., 2016), with evidence from computational modelling to suggest that learning a new word is facilitated by knowing its translation equivalent (Bilson et al., 2015). Furthermore, according to the Lexicon Structure Hypothesis (Byers-Heinlein & Werker, 2013), if a bilingual knows many translation equivalents receptively and productively, their lexical-semantic system will have a richer network of semantic connections compared to a monolingual, especially during development. Thus, this highlights the importance of investigating the effect of monolingual versus bilingual experience on lexical-semantic development.

While there has been a marked representation in the literature for English-speaking, monolingual infants on semantic network development, there has been little investigation into other infant populations (Jardak & Byers-Heinlein, 2019).

The development of the lexical-semantic system in a bilingual infant is complicated by the presence of two labels for every concept. So, how exactly does a bilingual infant organise these two labels for the same concept in their mental lexicon? Being able to answer this question would be informative to the field of bilingual and monolingual language development alike, to help better understand the early formation of semantic networks in different populations of learners. However, the few studies that have explored semantic development in bilingual infants do not always converge in their findings, for example, when testing translation equivalents and when one of the languages is more dominant than the other.

In one study, Poulin-Dubois et al. (2018) tested 22-month-old French-English bilingual children using a touch-screen, word identification task. Translation equivalents were recognised faster than non-translation equivalents in both languages: the dominant and non-dominant language.

Floccia et al. (2020) found no effect of language dominance, similar to Poulin-Dubois et al. (2018), but the finding of a translation equivalent processing advantage by Poulin-Dubois and colleagues, was not mirrored by Floccia et al. (2020). Floccia et al. (2020) tested 27-month-old bilinguals, who speak English and one other language, in two experiments investigating translation equivalent priming and cross-linguistic priming. There was an effect of priming in both experiments with no effect of language dominance or language distance (i.e. how similar or different the two known languages are). Time-course analysis showed translation equivalent priming and cross-linguistic priming demonstrated similar word recognition patterns. The authors propose that the processing of words to concepts in the bilingual infant lexicon mimics the adult lexicon, but that learning two languages simultaneously from a young age likely integrates lexical representation from both languages into one system, which may explain the lack of an advantage for processing translation equivalents over cross-linguistic semantic priming, and the absence of language dominance and distance effects.

Unlike the previous two studies, Singh (2013) did find an effect of language dominance when testing forward and backward semantic priming in 30-month-old Mandarin-English bilinguals. Forward priming refers to the prime word being in the dominant language (L1) and the target word being in the non-dominant language (L2), whereas a prime in the non-dominant language and a target in the dominant language, is referred to as backward priming. Singh (2013) found forward semantic priming but not backward priming. This mirrors findings in the adult literature (Basnight-Brown & Altarriba, 2007). Singh (2013) attributed this to greater word familiarity in the dominant language since more familiar words might have processing privileges compared to less familiar words (Mills et al., 2005).

Von Holzen et al. (2019) studied the effect of cognates on word recognition in thirty-one German-English bilinguals (18-53 months) and twenty-three German monolinguals (23-47 months). They specifically investigated the amount of overlap in cognates and found that phonological similarity (comparing 0- to 3- feature changes between concrete nouns) facilitated word recognition for the L2 but not L1 in bilingual toddlers. In contrast, when testing the processing of 'false friends' i.e. words which are similar in form but have different meanings in the two languages (e.g. English: *bald*; Spanish: *balde* which means 'bucket' rather than the English meaning of: having no hair) during language comprehension, findings reveal that comprehension is impeded in school-aged bilingual (Italian-French/German) children due to an overlap in form but a difference in meaning (Persici et al., 2019).

Thus, the effect of language and uniquely bilingual word types, such as translation equivalents, on processing semantic information remains unclear in a bilingual infant population. Cognates seem to facilitate processing, but there is scope for further studies to contribute to the findings in an infant population. Consequently, the organisation of these representations is yet to be explained by the predictions made by theories of language acquisition. We now consider some of these theories in more detail.

The most influential models of developmental bilingual word processing are: BLINCS - the Bilingual Language Interaction Network for Comprehension of Speech (Shook & Marian, 2013); DevLex-II (Zhao & Li, 2010, 2013); PRIMIR (Curtin et al., 2011); and BIA-d (Grainger et al., 2010). These bilingual models are based on findings from the monolingual literature but are not conceptualised merely by incorporating an additional language into a monolingual system, rather, they are able to capture the interaction between the two languages in the bilingual lexicon (Shook & Marian, 2013).

This thesis will focus on the two most recent models: BLINCS (Shook & Marian, 2013) and DevLex-II (Zhao & Li, 2010, 2013), which specifically address and attempt to explain the processing of cognates and translation equivalents.

BLINCS (Shook & Marian, 2013) attempts to simulate bilingual spoken language comprehension using features of distributed and localist models of language processing. The model uses an interconnected network of self-organising maps, which make use of unsupervised learning algorithms (Kohonen, 1995). The BLINCS model predicts the underlying architecture of the bilingual language system and the interactions within the system, modelled using English and Spanish. It can account for an advantage of cognate activation, leading to increased or faster activation, due to the close proximity of cognates in phono-lexical space with additional overlap at the semantic level. The authors tested how cognates are processed in the model by comparing them to false cognates (e.g. *arena* and *arena*- sand in Spanish); words without a translation in the model; and words with a translation in the model (i.e. translation equivalents). The BLINCS model predicted a higher level of activation for cognates than the other word types. This might suggest an advantage of learning cognates over translation equivalents, on processing. The implication of using cognates in a semantic priming study might result in facilitation when using a cognate as a prime or target, but could create inhibition if used as a distractor.

The second model of interest, the DevLex-II (Zhao & Li, 2010, 2013), is a self-organising neural network model which aims to simulate the developmental pattern of semantic priming both cross-linguistically and through translation. Using Chinese and English, the model simulates spreading activation according to the distance between bilingual words in semantic space. The DevLex-II model mirrors empirical findings that demonstrate a stronger translation priming effect than a semantic priming effect and shows that additional learning creates stronger connections between translation equivalents. Spreading activation from a prime word to its translation equivalent occurs via strong lateral connections which may account for the larger effect of translation priming, compared to cross-linguistic priming. The authors suggest that the better a bilingual understands the meaning of a word, the less they will experience competition from other lexical items, which explains why there may be greater confusion or competition from L2 words, when the L2 is a less dominant language in a bilingual's lexicon. This suggests that translation equivalents could contribute to stronger lexical connections in the developing semantic system. In a priming study, this should result in faster processing of words known in both languages.

1.3 Thesis Overview

The aim of this thesis is to investigate the development of the immature lexical-semantic system. It has been organised into two main chapters with each chapter reflecting the experimental approach chosen to study the lexical-semantic system in infants. Chapter 1 uses a free association task on 34- to 42-month-olds to establish whether adult associative norms accurately capture the word associations of a typical 3-year-old, and thus assess whether the current approach taken to select stimuli in an infant experiment on semantics (i.e. using the word associations found in the adult literature), particularly when investigating the nature and development of meaning in the immature lexicon, is valid. In Experiment 1, the most common word associations in 3-year-olds are documented, and the associative strength is calculated for frequently occurring word association pairs in this population. These observations are compared to adult associative norms and their corresponding associative strengths. Experiment 2 changes the modality of the task to an online format to validate the findings from the at-home format in Experiment 1. From Experiment 1 and 2, the most frequently-occurring word association pairs are tabulated with their associative strengths, to be in a comparable format to adult associative norms, so that they might act as a resource of related word pairs in children for future studies. Experiment 1 and 2 also consider whether vocabulary size modulates the word associations produced, especially when looking at adult norms and the infant associations occurring with greatest frequency. Experiment 3 continues the online methodology but is administered on a different population of the same age, namely a bilingual infant population. Bilingual word associations are compared to the newly established monolingual child associations from Experiments 1 and 2. Additionally, and specific to a bilingual population, the nature of each response is evaluated to determine whether the word association response a bilingual gives, reflects their knowledge of the word in one or two languages (i.e. translation equivalents), and whether the words chosen can be considered cognates or not.

Chapter 2 aims to replicate a semantic priming effect found at 30 months in monolinguals and bilinguals and extend these findings to address new research questions. More specifically, the aim is to see whether a priming effect is robust enough to be found in a variation of the traditional, in-lab, eye-tracking measures, in an online modality of the IPL, which was prompted by the 2020/21 Covid-19 pandemic. Specific to English-Dutch/German bilinguals is whether a cognate will inhibit or facilitate word recognition, particularly when the cognate acts as prime and target, as one of the two, or neither; and the implications of this for understanding how monolingual and bilingual lexical-semantic systems might differ. Therefore, Experiment 4 explores the validity of an online adaptation of the infant inter-modal preferential looking paradigm by replicating a simple word recognition task in 24-month-olds. Experiment 5 is an online semantic priming IPL task at 30 months in monolinguals. Experiment 5 uses the findings from Chapter 1 (i.e. word associations unique to children) to compare the priming strength in word pairs which: i.) have been taken from adult associative norms and replicated in a free association task on infants (Experiments 1 + 2); ii.) have been taken from adult associative norms but not replicated in a free association task on infants (Experiments 1 + 2); iii.) have been presented in Chapter 1 of this thesis, as unique children's associations, but not represented in adult associative norms.

Experiment 6 tests the same age monolinguals using a more common stimulus selection process (i.e., using taxonomic categories and adult norms to inform stimuli selection). Experiment 7 tests a bilingual population of the same age, for which Experiment 6 is a control. Experiment 8 and 9 aim to replicate findings in Experiment 6, adjusting the age for Experiment 8 (36-39 months vs. 30 months) and modality for Experiment 9 (in-lab vs. online).

2 Chapter 1: Word Associations

2.1 Literature Review

Free association or word association tasks have been employed in various areas of psychological research for over a century (Fitzpatrick et al., 2013). To study word associations (WA)s, the most common task is a free association task. This task instructs a participant to name or write the first word they think of in response to a cue word, usually to be done as quickly as possible.

Word association tasks were originally used to explore behaviours or as a psychodiagnostic tool in early studies (Jung, 1910), but there was a shift from clinical applications to linguistic domains in the 1950s. This new focus did not look specifically at the individual responses participants gave to cue words, rather the patterns of responses generated from large groups of participants (Moss et al., 1996). The focus of these large-scale studies was to collect association norms (Palermo & Jenkins, 1964) to understand the structure of semantic memory.

Network models of semantic memory (Collins & Loftus, 1975) represent concepts in an interconnected network of nodes. Spreading activation occurs between related concepts in such a system so that when one concept is activated, like the cue in a WA task, this activates other nodes related to the concept, such as the responses generated to the cue word. A common opinion is that these WAs represent the links in the network (de Groot, 1989) and by knowing the types of responses (e.g. taxonomic or associative), it can reveal the types of links between concepts in semantic memory (Moss et al., 1996). The strength of the links between concepts might be represented by the associative strength of a word pair (Kiss et al., 1973), which is the proportion of respondents generating a given response to a cue word. A free association task is considered a reliable measure of word connection strength (Palermo & Jenkins, 1964). So, in studies collating associative norms, a free association response is an indicator of the probability that one word will cue another in the absence of further contextual clues (Nelson et al., 2004).

Conceptual links are not the only factor affecting associative strength. The frequent co-occurrence of words such as *cat-dog* are thought to contribute to the associative strength in addition to their category membership, which means co-ordinates such as *cat-horse* would have a lower associative strength to *cat-dog* as the words might belong to the same semantic category, but they do not occur frequently together in everyday language (Moss et al., 1996).

In the lingusitc domain, WAs have more recently been used to investigate the organisation of the mental lexicon and how this organisation affects performance in certain tasks involving memory and verbal response (Comesaña et al., 2014). Through our experience of the world, associative structures form, linking word representations together in the mental lexicon. The shared lexical experience of many people is represented by this associative structure and the way in which words are associated provides information about the organisation of the mental lexicon (Nelson et al., 2000). When one word readily cues another, the links between the two are believed to have a strong connection in memory (Nelson et al., 2000). This makes the study of WAs a useful tool for investigating meaning and internal representations related to language (De Deyne et al., 2019).

2.1.1 Word Association Studies in Adults

Studies investigating infant semantic development often draw stimuli from, and reference the work of, three key adult associative norms studies: the Edinburgh Associative Thesaurus (Kiss et al., 1973), the Birkbeck Word Association Norms (Moss et al., 1996), and the University of South Florida free association norms (Nelson et al., 2004).

Kiss et al. (1975; 1973) collected WAs between 1968 and 1971 from 100 British, 17-22-year-olds for the Edinburgh Associative Thesaurus. There are 8,400 cues with 100 responses per cue. Cue words were taken from Kent and Rosanoff (1910). This resource is no longer readily available, but it has been transformed into an RDF dataset (Resource Description Framework- a model for data interchange on the Web) (Hees et al., 2016) more recently, though it is by far the oldest of the WA norms sets referenced in infant studies.

Moss and Older (1996) compiled the Birkbeck Word Association Norms from the associative responses to 2464 words, organized into fourteen tests, over seven years. Participants were between 17 and 45, living in the UK. Each cue word was allocated to 41-50 British-English participants and each participant responded to 50-387 cue words, with some participants completing more than one test session. Cues cover many grammatical classes, semantic categories, and word frequencies. Ten of the fourteen tests were completed by reading the cues in a booklet and writing the responses, and four tests had the cues delivered auditorily, with responses either typed on a computer, or written in a booklet. The authors counted the instances of a response and calculated the percentage of participants producing the same response. They refer to the percentage value as the associative strength between cue and response. The WAs are presented in a very accessible format, organised alphabetically, however, they do not seem to have been normed, despite being referred to as ‘norms’.

In the University of South Florida free association norms, Nelson et al. (2004) reported the WAs of more than 6000 participants to 5,019 cues. A total of 149 participants responded to 100-200 words on average, which generated 72,000 word pairs. Participants were asked to write the first (single) word that came to mind that was "meaningfully related or strongly associated to the presented cue word" (Nelson et al., 2004). Data collection started forty years before its publication according to De Deyne et al. (2019). The research has been cited 1,900+ times and is the most commonly used resource in English (De Deyne et al., 2019). Nelson et al. (2004) provide a variety of metrics, organised into appendices (e.g. Appendix B contains the 5,019 normed words with associative and frequency information), which is far more detailed than other studies (Moss et al., 1996).

A more recent adult study is the English Small World of Words project (SWOW-EN) (De Deyne et al., 2019) which compiled a new English WA dataset, collected between 2011 and 2018. The study tested 12,000 cue words from previous studies: the Semantic Priming Project of Hutchison et al. (2013); the University of South Florida norms (Nelson et al., 2004); the English word modality norms (Lynott & Connell, 2013); and semantic feature production norms (McRae et al., 2005). Over 90,000 participants were crowd-sourced using social media, e-mail, and university sites. The sample included over-16-year-olds who were predominantly American-English and British-English speakers. A unique feature of this task was its web-based data collection technique. In the experimental task, a word appeared on screen and a participant was asked to respond with the first three words they could think of connected to the cue word. Participants entered their associations into three text fields in the order in which they thought of the three responses. A participant responded to between 14 and 18 cue words and could stop responding on any given cue by clicking on the “No More Responses” button or “Unknown Word” if the meaning of the word was not known to the participant.

De Deyne et al. (2019) compared their findings to the University of South Florida free association norms (Nelson et al., 2004) and Edinburgh Associative Thesaurus (Kiss et al., 1973) and noticed small differences which they attribute to the difference in sample size between studies, the repeated vs. single response nature of the task, and differences in task instructions (e.g. first thing that comes to mind versus the first ‘meaningful’ thing). They also noted that their sample is less homogenous, probably due to demographic differences between the studies. Both Nelson et al. (2004) and Kiss et al. (1973) tested college-age students, compared to De Deyne et al. (2019) who tested a broader age range from 16 years and older. The authors predicted a higher correlation between the WAs of American college students and the University of South Florida free association norms (Nelson et al., 2004) than their own norms, as this was the participant demographic sampled in Nelson et al.’s (2004) study. This comparison clearly highlights the effect of task administration and participant sample attributes on the types of WAs produced.

Fitzpatrick et al. (2013) investigated the methodologies used in a number of influential adult WA studies. Published studies with the highest number of citations (Harzing, 2007) were compared. The authors concluded that there was “little commonality of approach” making it very difficult for a researcher to decide on the most suitable design and execution of a WA task and its subsequent data handling. One example of high variability is the way in which WA responses are categorised and analysed (Fitzpatrick et al., 2013). The relationship between cue and response followed Saussurian definitions in early studies, whereby syntagmatic responses (e.g. park-play) reflect the co-occurrence of words in text, and paradigmatic responses (e.g. park-playground) are those which are interchangeable without a significant change in the grammaticality or meaning of the sentence. While some studies retain this type of coding system, the definition of syntagmatic and paradigmatic responses is not always consistent. Another shortcoming of many studies according to Fitzpatrick et al. (2013) is the lack of rationale for methodological and procedural decisions, though earlier studies, such as Entwisle et al. (1964), seem to have done this more successfully.

In response to the inconsistencies found in their analysis of existing WA studies, Fitzpatrick et al. (2013) devised a WA on task that they conducted on twin 16-year-olds and twin over-65-year-olds (*N*= 48 twins per group). One hundred cue words were randomly selected from the 2000-3000 band of most frequently used English words in the British National Corpus. Participants were given a booklet of the 100 cue words and asked to write their responses in spaces provided. The instructions given were to write the first word they thought of for a given cue. Participants had ten minutes to complete the task. The authors created normed lists from the responses given by each age group (*N*= 96 per group) and found differences in the associated responses of 16-year-olds and over-65-year-olds. According to Fitzpatrick et al. (2013) the age-related differences might stem from the vocabulary preferences of the two age groups or due to changes related to ageing. If the latter is true, then the young group's responses should evolve to resemble the old group's associations over time. If the former is true then as the younger group ages, their associations will remain the same, and if a new cohort of 16-year-olds were tested, they would produce unique norms. The authors propose a third possibility suggesting an interaction of age and generation. In this scenario, neither of the norms lists from their study would reflect the WAs of the younger group when they turn 65.

Consequently, Fitzpatrick et al. (2013) caution against using normed lists such as the South Florida Association Norms (1998) to compare responses of a target population against (e.g. in stereotypy analysis), as it fails to acknowledge the characteristics of a cohort, such as generational differences, which might influence how a group responds. Not enough is known about the variables affecting WA behaviour and so the authors endorse the creation of a norms list from the study population itself, rather than reliance on published norms. A population-specific list will reflect the characteristics of those tested. This will enable better identification of differences within a population and across populations too.

To summarise, much work has been done on adult WAs, with the studies mentioned here representing the widely-cited studies in the infant literature, but by no means the only WA studies in the adult literature. There seems to be huge variation in how the WA task is administered and its data handled in adults, but a commonality in all three studies mentioned here is the requirement for a participant to read and write, rendering the procedure incompatible for testing young children unless it is done with an experimenter or parent/care-giver. The literature on children’s WA studies will prove instrumental in informing the experimental protocol for a free association task on young participants and so the next section shifts to the studies that have previously been administered on young participants.

2.1.2 Word Association Studies in Children

There have been very few WA studies conducted on children compared to adult studies, with most testing children aged 4 years and older. This is despite its suitability as a task for naturalistically investigating early lexical-semantic networks, specifically, the implicit activation of related words during word processing (Wojcik & Kandhadai, 2020).

An area of particular interest in free association research in children is investigating the occurrence of a developmental shift referred to as the 'syntagmatic- paradigmatic' shift (White, 1985). Findings suggest that until six years of age, children's responses to a WA association task are based on syntagmatic links (Brown & Berko, 1960; Entwisle et al., 1964; Ervin, 1961), but after this age, up until eleven years, children responses become paradigmatic in nature (Newman, 1970). Most researchers now agree that this developmental shift occurs in the first years of schooling (Wojcik & Kandhadai, 2020).

Paradigmatic responses to a WA task indicate a more developed semantic system, thus are more common in adult associated responses. It is believed that a higher level of cognitive processing is behind this type of response, which involves processes such as conceptual and lexical reorganisation (Nelson, 1977). Thus, as children develop cognitively and linguistically, it is thought that the types of WAs they produce will become more adult-like, and paradigmatic in nature. Paradigmatic knowledge helps structure semantic networks and in the retrieval of semantic knowledge, which develops as a child increases their vocabulary and word knowledge (Sheng et al., 2006). However, according to Wojcik and Kandhadai (2020), the belief that young children only produce syntagmatic responses in a WA task is inaccurate and is due to the lack of data on young children's associations. In fact, syntagmatic and paradigmatic relationships between words have been observed in infants at 24 months (Arias-Trejo & Plunkett, 2013) in experiments testing comprehension, with some evidence suggesting the existence of paradigmatic relations as young as 6 months (Bergelson & Aslin, 2017).

To explore the developmental trajectory of paradigmatic relations in children, Wojcik and Kandhadai (2020) conducted a free association task on sixty 3-8-year-olds (*M*= 4.85, *SD*= 1.27). They also tested a group of adults for comparison (*N*= 60). A total of 65 cue words were used (nouns= 25) and eight order lists were created, all 32 or 33 words in length. Two experimenters demonstrated the task to a participant, who was tested individually. A participant had three practice trials and an experimenter gave positive feedback after a child's responses. During the test phase, however, no feedback was provided. Children's responses were audio recorded as an experimenter made a written record. Responses were considered 'unusable' if the prime was repeated or if the child said "I don't know" or failed to respond. Other categorisation of response included: syntagmatic, paradigmatic, both, 'not X' (negation), rhyme, or 'other'. The proportion of paradigmatic responses was calculated as a function of age, and the percentage of associations and idiosyncratic responses were analysed, as well as the associative strength of the top associate for each prime. When using age as a categorical variable for analyses, the following distinction was made in children: 'older' children were classified as 6-8 years (*N*= 17) and 'young' children as 3-5 years (*N*= 43).

Wojcik and Kandhadai (2020) found clear evidence of paradigmatic responses in 'young' children, with a higher proportion of this response type in 'older' children, and a higher proportion still in adults (though the interaction was not significant). This is in line with previous research suggesting a paradigmatic shift in school-age children. Taken as a continuous variable, age predicted the proportion of paradigmatic responses with a linear trend as age increases. This might point to the paradigmatic shift beginning before children enter the school system. 'Younger' children produced a higher rate of idiosyncratic responses compared to 'older' children. This indicates more variability in related responses in younger children, leading the authors to surmise that lexical-semantic networks in early childhood "are more idiosyncratic".

Overall, this study points to the existence of paradigmatic relations in children as young as 3 years, though proportionally at this age, responses are more likely to be syntagmatic. However, caution must be taken before generalising this assumption as the methodology used only asked participants to give one response per cue, and while a child's first response might not have been paradigmatic, this does not indicate an absence of a paradigmatic link in their wider network of associations (Wojcik & Kandhadai, 2020). Another point to note, which is mirrored in other recent free association studies testing children, is the relatively small sample tested (e.g. Cronin, 2002: *N*= 59; Sheng et al., 2006: *N*= 24; Wojcik & Kandhadai, 2020: *N*= 60). Further empirical evidence is needed to corroborate findings before generalising about any patterns of WAs, or specific types of associated cue-response pairs, produced by children as young as 3 years.

There are much larger-scale WA studies in children, but these were conducted over fifty years ago already (Entwisle, 1966). One such study was conducted in 1963 by Koff (1965), who tested 8 to 12-year-olds (*N*= 147) on a list of 51 words to compare children's associative responses with responses collected in one of the first infant studies on WAs (Woodrow & Lowell, 1916- testing children aged 9-12, *N*= 1000). Koff was interested in investigating any changes in children's WAs, especially changes that might demonstrate an effect of mass culture on thinking. The test was administered by the teacher in the children's classroom, first with an oral example to demonstrate the task, followed by the teacher calling out the 51 words in turn, while each child wrote down their written responses on paper. The instruction was: "Write the very first word that comes to your minds". Koff found a significant difference in primary responses in children from 1916 to 1963, but when compared to adult responses given in 1954 (Jenkins & Russell, 1960), there was not a large difference between the responses given by children and adults. This differs to Woodrow and Lowell's (1916) finding of a large discrepancy between children and adults. Koff (1965) concluded that a cumulative effect on WAs can be attributed to changes in culture.

Palermo and Jenkins (1964) tested participants from fourth grade children up to university level students (*N*= 250 boys and 250 girls in grades 4-8, 10, 12; *N*= 500 male and 500 female university students) on a set of 200 words derived from Kent and Rosanoff's (1910) list of 100 nouns and adjectives, plus an extra list of 100 high frequency verbs, pronouns etc. which were included to broaden the range of word class represented. The test took a written format that was administered to fourth and fifth grade children in schools of middle and upper socioeconomic status. A pilot study had revealed that children from elementary schools with low socioeconomic status (SES) did not have the adequate reading and writing skills for the test. Other grades were tested in a range of SES schools. Participants were given the following instructions:

> "You are to write next to each word the first word that it makes you think of. It doesn't make any difference what word you write as long as the word on the paper makes you think of it. There are no right or wrong answers. The purpose of the test is just to see how quickly words will come to your mind."

This was followed by a booklet with the stimulus words. Participants were instructed to write the associated word that came to mind next to each word in the list. On completion, participants had to raise their hands and could continue answering until everyone had finished.

The same authors (Palermo, 1971; Palermo & Jenkins, 1966) also collected data on children in grade 1-4 (*N*= 50 boys and 50 girls per grade) using oral presentation and responses for a set of 100 words. Words were taken from a previous study (Palermo & Jenkins, 1964) selected to exclude homophones, represent most grammatical classes, and likely to elicit superordinates and contrasts. Each participant was tested individually in a room in their school. The task was described as a word game and each child was told to give the first word he/she thought of. Five words (*window*, *cold*, *have*, *soldier*, *high*) were used as practice. The test words were presented in the same order for each participant and the experimenter recorded each response after it was given.

A participant's responses were excluded if they did not meet the following criteria:

> "they had more than 10% nonword responses, omissions, phrases, repetitions of the stimulus word, or the children responded during the task by naming objects in the room without apparent reference to the stimulus words."

These were also used as exclusion criteria in Palermo and Jenkin's (1964) study.

Palermo (1971) argues that the oral mode of administration, at least when comparing fourth grade children to a previous study (Palermo & Jenkins, 1964), has an effect on the type of response produced, with a larger percentage of participants responding with the most popular responses than when instructed to give written responses.

Newman (1970) conducted two studies which have not been cited as frequently as the others mentioned above, but the methodology used differs slightly, which is why it has been included here. Study 1 tested 34 kindergarten children (54-66 months) on 14 words (*store*, *children*, *egg*, *kitten*, *pig*, *rabbit*, *I*, *wish*, *see*, *cut*, *smile*, *up*, *still*, *no*) with a different order presented for each child. Half of the participants were asked to respond using the Continued Associations (CA) method and half using the Continued Sentence Associations (CSA) method. The instructions were the same for both groups:

"When I say a word I want you to tell me the first thing that you think of."

The procedure differed for each group only in the examples given:

CA:

"If I say TRUCK you might say CAR. If I ask 'What else does TRUCK make you think of?' you might say FAST. .."

CSA:

"If I say TRUCK you might say 'A truck is bigger than a car.' If I ask 'What else does TRUCK make you think of?' you might say 'A truck goes real fast on a road. . ."

Multiple examples were given until the participant felt ready to proceed.

The instructions then remained the same for both groups:

"What do you think of when I say the word…? What else do you think of when I say the word…?"

A time limit of 60 seconds was set using a stopwatch. The question and stimulus word were repeated 4-6 times in this time window.

Newman (1970) compared the two methods and found that the second, sentence association task, was a more natural task type for children of this age, compared to the first, more standard version of a WA task. Newman suggests that the CSA method accommodates syntagmatic responses better than the CA task, which is thought to be a more natural associative response type at a young age. Further, it allows children to offer multiword responses which has been noted as a commonality in children of 4-5yrs engaged in associative word tasks (see Entwisle, 1964).

The second study that Newman (1970) conducted used the CSA method on 39 children aged 48-60 months to collate a list of associations for 157 words. The experimenter tried to elicit four sentence associations for each of the test words. Some children became tired giving responses to all 157 words in sessions which spanned multiple days, despite the sessions being limited to ten words per session. This highlights the importance of tailoring the experimental procedure to the age and attention-span of the participants.

Considering potential differences between monolingual and bilingual performance in a WA task on children, Sheng et al. (2006) explored paradigmatic and syntagmatic organisation in the two groups. Having reviewed the literature, they observed no consensus that a bilingual advantage exists on lexical-semantic organisation, so sought to investigate this point themselves. They tested Mandarin-English bilingual children aged 5 to 8 years old (*N*= 12) age-matched with monolingual children (*N*= 12) in a repeated WA paradigm (Elbers & van Loon-Vervoorn, 1998). The rationale for a repeated version of the task is that some children may respond first with a paradigmatic response, decreasing this response type with further elicitations, but another child might show a pattern of increasing paradigmatic responses across trials. Cue words included adjectives, nouns, and verbs due to their varying sensitivity to paradigmatic and syntagmatic links and each cue was meant to elicit three responses. Two stimuli lists were created, both 36 words in length. The task was demonstrated by the examiner, who provided feedback and encouraged single-word responses during preparation. The test was in two parts. The first part tested 18 of the words by eliciting 1 response for each word before repeating the list two more times (in a randomised order) to generate a second and third response. In the case of repetitions from previous responses, the experimenter requested a novel response from participants. After a break, the second part tested the remaining 18 words using the same procedure. Bilinguals were tested in both languages using a different list for each language, performed in two separate sessions, 2-7 days apart. Monolinguals were tested once in English.

Sheng et al. (2006) coded responses as paradigmatic or syntagmatic and calculated the proportion of paradigmatic responses as a dependent variable. Initial responses saw greater paradigmatic responding than second or third responses in monolinguals and bilinguals. In bilinguals, the proportion of paradigmatic responses was comparable in both their languages, with a stronger correlation between paradigmatic responses to nouns in the first (L1) and second (L2) languages, compared to verbs. The authors interpret this as indicating that semantic knowledge might transfer between languages easier for nouns than verbs.

There was little difference between monolingual and bilingual paradigmatic response rates and both groups generated paradigmatic responses to adjectives with greater ease than other word forms. The similarity between the two groups stands in opposition to a bilingual advantage in lexical-semantic organisation, which indicates that the development of semantic organisation follows the same rate in bilinguals as it does in monolinguals. The authors note the high mean age of participants (over 7 years) in their study might not have been able to capture developmental differences between monolingual and bilingual performance and that other language pairs ought to be tested to explore variables affecting bilingual language development further.

Through reviewing the literature, it is clear that few studies can be found directly eliciting free associations from infants under the age of five. Most focus on children older than 48 months and while Wojcik and Kandhadai (2020) tested children as young as 3-years-old in a recent study, the full sample included children up to 8 years of age, which encompasses a wide age range (that they needed to cover the period at which the paradigmatic shift is believed to occur). Furthermore, whether the associated responses of adults and children are similar (Koff, 1965) or very different (Woodrow & Lowell, 1916) remains inconclusive.

As with WA studies in the adult literature, there are huge discrepancies in how this task is administered on young children, with some evidence suggesting the mode of delivery (i.e. oral vs. written) might affect response type (Palermo, 1971). The number of prime words varies between studies (14-157), and it is clear from Newman's (1970) study that this factor must be carefully considered to avoid children become fatigued during the task.

A positive feature of these studies, particularly the older ones, is the detailed documentation of the experimental procedure and data handling. The detail included in these studies allows a deeper understanding of the types of responses a young child might produce, such as a child naming something in the immediate environment, which is based on the description of responses excluded from analysis (Palermo, 1971; Palermo & Jenkins, 1964). The elicitation of responses, for example, single or multi-word responses (Newman, 1970), and whether to encourage just one (Wojcik & Kandhadai, 2020) or more (Sheng et al., 2006) responses can all inform the design of a new WA task, specifically for 3-year-olds looking at noun-noun relations suitable for use as experimental stimuli in studies investigating semantic development.

2.1.3 Proposed Research & Rationale

While free association norms are informative for research in the domains of psychology and linguistics, particularly when selecting stimuli for empirical experiments, caution must be taken not to generalise findings from normative studies across different populations such as regional populations, as these will have their own associative norms (Nelson et al., 2004). Equally, the lexical-semantic system has the flexibility to change according to cultural experiences and trends over time. This temporal factor indicates that norms should be periodically updated (Palermo & Jenkins, 1964, 1965) as technological, cultural and global changes might shape new associations or alter old ones. Many of the studies presented here on WAs in children are already very old which gives impetus to conduct a study on children that better reflects the world we are living in today.

Word associations are likely to be modulated by age (Fitzpatrick et al., 2013) and if associations stem from our experience of the world, and our exposure to linguistic input, this will inevitably differ according to the stage of a child's linguistic development. Common relationships between words in young children might be missed if relying on predetermined relations (Wojcik & Kandhadai, 2020) which do not derive from the population of interest. Surprisingly, there are few studies documenting the WAs that children have. Due to the lack of such data, it remains to be seen if infant associations mirror adult associative norms (Arias-Trejo & Plunkett, 2009) or how different types of meanings emerge. **Experiment 1** aims to bridge this gap by compiling a set of common child WAs and comparing their forward associative strength (FSG) to the same word pairs in adult norms.

It would be unwise to explore the primacy of semantic meaning in young children by relying on semantically related word pairs assumed to exist in a child's lexical-semantic system, without first testing them. This is what many studies have had to do to date, relying on adult associative norms for their stimulus selection, despite many of the words tested on adults not being suitable for use in infant studies. Infant studies rely on visual aids to index word recognition since reading and writing are not yet available, making it essential for any words used in studies to be highly imageable. So, developing a task whose focus it is to document common noun-noun WAs in the lexicon at as young an age as possible, could be developed into a stimulus resource for studies investigating the development of lexical-semantic networks in infancy.

To test this would involve testing participants as young as possible. A free assocaition task necessitates speech production, a large enough vocabulary, and the cognitive ability to participate. This makes the youngest possible age for participation around 2-3 years of age, however, previous studies that have collected children's WAs (Entwisle, 1966; Koff, 1965; Palermo & Jenkins, 1964) have tested children older than five years of age. Only one other study has tested children under 5 years (Wojcik & Kandhadai, 2020), though the sample size does not compare in magnitude to early studies testing older children and so there is scope to contribute further WA data to the under-represented age group of under 5-year-olds. Due to a child's limited attention span and need to be stimulated, **Experiment 2** changes the modality of the WA task used in Experiment 1, to an online format to make the task more engaging to this young age group and also to increase the scope of its reach for greater representation and to see if findings can be generalised to other infant groups.

One group of interest is the bilingual infant population. **Experiment 3** continues the online methodology developed in Experiment 2 but recruits 3-year-olds who speak English and one other language to compare the types of WAs bilingual children produce compared to monolingual children. This might reveal similarities or differences between the two groups which can contribute to our understanding of how the lexical-semantic system develops during infancy. One way to do this is to observe the types of words bilinguals use in their related responses. For example, preferential use of translation equivalents and cognates might indicate a more robust lexical-semantic organisation for words that share features between languages in a bilingual speaker.

Documentation of the associations very young children have between words, at different stages of linguistic development would be of particular merit to the research community and Experiment 1-3 is the first step to compiling such a resource. According to Fitzpatrick et al. (2013), WA tasks might be optimally used to track changes in a growing lexicon, as relationships between words are created and strengthened. Thus, a free association task has the potential to explore the development of lexical-semantic connections by collecting descriptive data of the associated words in a young child's productive lexicon. This might help inform our understanding of how meaning evolves from infancy into a mature lexical-semantic system.

2.2 Experiment 1: Word Association Study on Monolingual Children, At Home

2.2.1 Introduction

Experiment 1 aims to document the WAs that three-year-olds have with common nouns (e.g. *car*) to provide a stimulus resource for researchers investigating the development of meaning in infants from a young age. We know that infants are sensitive to semantic meaning, including taxonomic and associative links, by 24-months of age (Arias-Trejo & Plunkett, 2009). This is evidenced by measuring whether infants attend longer to pairs of words that are related in meaning (e.g. *cat-dog*) compared to unrelated word combinations (e.g. *cat-plate*). Even infants as young as 18 months have been shown to listen longer to taxonomically related lists of words than to unrelated lists (Delle Luche et al., 2014), demonstrating a sensitivity to the semantic relationships between words at this young age. One potential confound in this type of study is in their stimuli selection, which is based on the WAs from adults and older children.

There is currently no readily available resource of WAs in young children to inform stimuli selection, hence the reliance on adult word associative norms. This study aims to bridge that gap by collating common noun-noun WAs in 3-year-olds. In order to this, a free association task suited to a younger age group will be developed to test if it can be successfully completed by children as young as 3, and to observe whether the WAs in this younger population resemble those of older children (with regards to the proportion of paradigmatic responses) and adults (in terms of specific cue-response word pairs).

The outcome of this research will either validate the WAs currently being used in infant studies (i.e. those taken from adult associative norms) or provide evidence that younger children have different WAs, while their language is still developing. If the latter is true, the findings from this research may provide a stimulus set of WAs shared by 3-year-olds that is more representative of the associations in young children. Such a stimulus set would be a useful resource for researchers exploring the emergence of semantic meaning in infancy. The WAs of 3-year-olds may indicate some of the first associations that children form and can verbalise, suggesting the primacy of these relationships which in turn would be more likely to be captured in studies that observe the development of semantic meaning in infants as young as 18-months-old.

Research Aims

- To collect free association data from an unrepresented age group (2 to 5 years)
- To compile a resource of noun-noun child-specific WAs for use in infant experiments on semantic development
- To compare infant WAs with adult associative norms

Research Questions

1. Can 3-year-olds understand and complete a repeated word association task, using related responses?
2. Do the associated responses of 3-year-olds demonstrate a prevalence for syntagmatic responses, replicating findings found in slightly older children?
3. How do word associations in 3-year-olds compare to adult word associations?
4. Are the word pairs commonly used in infant semantic priming studies represented in the word associations produced by 3-year-olds in a production task?

2.2.2 Method

Rationale

Newman's (1970) two studies offer the most suitable framework for the current study due to the reduced number of cue words a participant must respond to (*N*= 14), compared to other similar experiments (*N*= 100-200 words). The oral mode of the task is also more suited to younger children who cannot yet write. A further benefit is the detail provided in the methodology employed in the study, which includes the exact script used by the experimenter. The script gives examples to demonstrate the task rather than just explain the task. A child as young as 3 cannot be asked to read or even listen to a description of a task and then be expected to comprehend and perform that task. Instead, demonstrating a task and providing examples is more likey to be understood by a young participant. For Newman's (1970) second experiment, children as young as 48 months were tested using this script, demonstrating its efficacy in explaining the task to young participants. For this reason, the script used for the 'continued associations' version of Newman's WA task will act as a model for the present study, which will be adapted for younger participants.

The fact that Newman's (1970) methodology has been tried and tested on one of the youngest age groups (48 months) documented to have done a free association task like this offers further impetus to use it as a basis on which to develop a task for younger participants. To mitigate for using it on a younger population, participants will have the opportunity to provide more than one response to a cue word. In fact, Newman's (1970) second experiment also attempted to elicit four associations from participants for each of the test words. This will be adopted in the current study as it provides an opportunity for children that are unclear on the task requirements to explore this through the opportunity to offer multiple responses. It could also be that a child first imitates the experimenter through repetition or rhyme (Palermo & Jenkins, 1964) as a tactic to processing the word before being able to offer a related word. Therefore, more than one response will be encouraged but not a mandatory requirement of the task. This approach to the task leads on to the way in which Palermo and Jenkins (1964; 1966) and Palermo (1971) framed the task. It was described as a 'word game' which seems appropriate for young participants: to keep the task fun and engaging, and to reduce any sense of a correct or incorrect response to the task. The idea of the task as a game will be emphasised in the current methodology and if a cue word is proving difficult, the instructions will be to progress on to another word to maintain an element of fun throughout the task, which will hopefully foster completion of the task and reduce attrition rates, commonly found in infant studies.

When considering the analysis of participant responses, Palermo and Jenkins (1964; 1966) and Palermo (1971) rejected responses if a certain threshold (10%) was reached for a certain response types e.g. non-word responses, repetitions and the naming of objects in the room. The current study is one of very few to test children as young as 3 years of age and so the expected response type is unclear and has not yet been replicated. The proposed research is of an exploratory nature and so it seems appropriate to include all responses for analysis to ensure no bias towards what is perceived to be an 'association', in adult terms, is assumed. It might emerge that children of this age find the task cognitively prohibitive and if this is the case, the data to support this i.e. unrelated words, repetitions or non-words, might be indicative of this phenomenon.

Different to all previous studies is the way in which the task is administered. The parent acts as the experimenter for the proposed task, with the rationale being two-fold. First, a young participant will undoubtedly find communicating with a familiar parent or carer preferable to an unfamilair adult who might try to administer the task. This presents a more naturalistic setting in which to elicit responses from the child, thus more likely to reflect the instinctive WAs that come to the child's mind. An unfamilar adult would first need to establish rapport with the child and if successful, it still might affect the child's behaviour or attitude to the task. Second, the closure of the Plymouth BabyLab due to the Covid-19 pandemic prohibited the running of studies in-person and so any experiment during the period of a national lockdown needed to be run remotely. As already outlined, an unfamilair adult running the task might have an adverse effect on the child's ability to do the task or it may influence their behaviour, which could be exacerbated in an online format where there is little possibility to establish rapport, which is ordinarily done through play in a lab setting. Although a parent taking on the role of experimenter has its own inherent drawbacks, such as: the inability to control the precise delivery of the procedure (even if mitigating for this by providing clear instructions and a script to follow); the lack of control in parents influencing their child's responses during the task; and reliance on the parent to report their child's reponses honestly and accurately; considering the age of the children and the current pandemic, this was the best option available and it will be clearly considered when interpreting the findings.

Participants

A total of 140 participants were recruited online using the University of Plymouth (UoP) BabyLab database and its corresponding Facebook page. Recruitment was extended to Essex (*N*= 3), Oxford (*N*= 2), and Lancaster (*N*= 6) Baby Labs through their Facebook pages towards the end of data collection to complete participant numbers in some age bins, resulting in a final sample size of 150.

Participants were divided into seven, two-month age bins: 34-35, 36-37, 38-39, 40-41, 42-43, 44-45, 46-47 to explore word association production across a child's third year of life. Twenty participants were recruited for each age bin.

Exclusion criteria were those speaking or exposed to more than one language according to the Language Exposure Questionnaire (LEQ, Cattani et al., 2014), with an exposure of less than 20% to an additional language considered eligible for participation. Children with a diagnosed developmental or language delay, or with an uncorrected hearing problem were excluded.

Materials

One hundred highly imageable, concrete nouns were selected that feature in the Oxford CDI (Hamilton et al., 2000) and UK CDI (UK-CDI Database, 2016) and that are known by at least 60% of 18-month-olds in both these inventories (see **Appendix A**). The cue words were from nine categories: food and drink, animals, vehicles, body parts, toys, clothes, furniture, and rooms, outside, and household objects. The 100 words were divided into 10 lists of 10 words, ensuring each category was represented in each list.

Two pseudo-randomised orders were created for each of the 10 wordlists to avoid effects of cue order. Care was taken to avoid consecutive words appearing from the same category or linked in some way associatively. Words sharing initial word onset were not presented consecutively. The two orders of the 10 wordlists can be found in **Appendix B**.

Procedure

After receiving ethical approval from the UoP's Ethics Committee, participants meeting the inclusion criteria (e.g. age, language, and developmental profile) were contacted to participate via the UoP BabyLab database or the UoP BabyLab Facebook page. The initial email included the participant information sheet and consent form (see **Appendix C**).

Interested families were then sent an email with instructions on how to administer the WA task at home with their child (**Appendix D**). This second email included a section to consent to the study, a script for parents to use to administer the task and one of the 10 wordlists. On receipt of this, parents were asked to check that the words in the word list were known to their child, and if not, they were instructed to email to request replacement words or a replacement word list.

3. Complete the test using the script below

Follow the script as closely as you can.

Say all 3 examples.

For every word in the list, try to get up to 3 different responses. 1 response per word is absolutely fine though.

Try do all 10 words in one go if possible.

There are no right or wrong answers! Have fun!

Script

"We're going to play a game to see how quickly you can say a word that is connected to a word that I say.

If I say KITCHEN you might say BREAKFAST. (Example 1)

If I say MUMMY you might say DADDY. (Example 2)

If I say DRINK you might say WATER. (Example 3)

Okay, are you ready?

What do you think of if I say …? (Response 1)

And another word when I say…? (Response 2)

And another word?" (Response 3)

Figure 1*: Experiment 1. Word association task instructions and script for parents*

Parents were instructed to follow the script (see **Figure 1**) as closely as possible. The script is based on Newman's (1970) study. The script first demonstrated the task using three examples before the experiment proper. The script prompted three responses to each cue word, but it was made clear that one response per cue word, or even no response, was acceptable as the task was to be framed as a game, thus should be engaging and enjoyable for the participants. Parents were instructed to record their child's responses in the order they were given, in a table provided (see **Table 1**) which was in the instructional email.

Parents were encouraged to record everything the child said. A ‘comments/observations box’ was included with a suggestion to note down if children were naming objects in their immediate environment, or to assist the experimenter in understanding very personal associations arising from the task, based on the child’s unique life experience. This was to enable better coding of the types of responses that children give, to observe for patterns in the data.

List 1	Cue	1st word	2nd word	3rd word
1	aeroplane			
2	bicycle			
3	bread			
4	cereal			
5	door			
6	frog			
7	leg			
8	pasta			
9	slide			
10	toe			

***Table 1**: Experiment 1. Table for parents to record their child’s responses*

Parents returned the completed task by email to the experimenter. The responses were checked for clarity and to avoid making assumptions, parents were contacted to provide further information about ambiguous responses (to assist in later coding of each response). Parents were specifically asked to indicate if seemingly random responses related to something in the immediate environment while performing the task. Previous research on free associations in children (Palermo, 1971; 1964) has shown this to be common behaviour when young participants are unable to produce a response.

A final debriefing email was sent out thanking the family for their participation in the study (see **Appendix E**). A digital certificate was attached for the child and a £5 Amazon voucher code was included for participating.

Families recruited from other Baby Labs were asked to complete two further forms as part of the procedure. The first was a sign-up questionnaire usually given to families wanting to register on the UoP BabyLab database, which documents information about the participant such as age, language background and developmental delays. The second form was a demographic questionnaire for parent/s or caregiver/s of the participant. Both were used for screening purposes. This information was already on file for participants recruited from the UoP BabyLab database.

Pilot Study

The abovementioned procedure was first piloted on children between 2 and 5 years of age (*N*= 14). Very little had to be changed regarding the instructions of the task, since parents commented positively on the clarity of the task. It was also communicated that 10 words was an achievable number of cue words for the target age group and no comment was made regarding the elicitation of three responses per cue word, so it was decided to keep this in the experimental procedure. It did emerge that children at the youngest age of the target group i.e. 24 – 30 months, were not always successful in understanding the task, with some unable to complete it at all. This prompted a change in the minimum age from 24 months to 34 months. Due to availability of resources and a refocusing of the research aims, the upper age limit was set to 47 months to focus on WAs in the third year of life. At this age, most children are able to produce the items being tested in the task, along with a large number of other lexical items. This presents a novel opportunity at which to observe emerging WAs, which has surprisingly received little attention to date.

2.2.3 Data Processing and Analysis

All participant responses were coded by the author and two junior researchers coding a subset of the data (see **Rater-reliability** section). Inferential analyses were performed only on the first response for each cue word, per child, since not all children gave three responses to every cue word. However, related responses from second and third responses have been retained for descriptive analysis, especially when looking at exact cue-target word combinations.

The following steps were taken while coding the responses:

- Spelling errors were corrected unless parents noted that a response had been phonemically transcribed
- When a response could have been classified as more than a single word class, e.g. a noun or verb, and in the absence of contextual clues, nouns were prioritised, or the most common form was selected by looking up its frequency.
- Any contextual information provided by parents/caregivers when asked to clarify certain responses (e.g. if a response was something in the immediate environment) was noted in brackets after the response, to assist coding.
- Missing responses were marked as ‘No Response’

Coding for Response Type

Different response types were identified by analysing the data collected in the pilot study. Seven categories were identified initially:

0 no response given

1 recognised association/category (i.e. from adult norms)

2 association unique to individual (this information was provided by parents)

3 association arising from a previous response given (e.g. PIG- 1^{st}= mud, 2^{nd}= straw, 3^{rd}= moss. The 2^{nd} and 3^{rd} responses relate to the 1^{st} response 'mud' rather than the cue word PIG)

4 related in a general/wider sense (i.e. not an obvious association but a logical connection)

5 repetition of the cue word/ repetition of a response already given

6 naming something in the immediate environment (this information was provided by parents)

7 an unclear association (i.e. cannot be coded 1-6)

Through coding a subset of the test responses, a further four categories were added:

8 rhyme (e.g. CAR - bar)

9 sounding out (e.g. APPLE – 'a' for apple)

10 exemplar listing which includes cue word (e.g. CAR - big car, my car, one car)

11 action/mime or sound to indicate cue word (e.g. LION - roar)

To differentiate between order of response, a code was applied for the 1st, 2nd, and 3rd response. Using:

_a_ = 1st

_b_ = 2nd

_c_ = 3rd

Finally, responses were demarcated by word class i.e. noun, verb, adjective etc.

n_ = noun

v_ = verb

adj_ = adjective

adv_ = adverb

prep_ = preposition

exc_ = exclamation

none_ = no word form

An example of response coding can be found in **Table 2**.

Participant number	Age in months	Gender	bib	bib_coded	bike	bike_coded
IACAN_B6_1_01	45	male	baby	_a_1_n_	NO RESPONSE	_a_0_

Table 2*: Experiment 1. An example of a coded word association response*

After coding all responses, code allocation was reviewed which resulted in category '10' being omitted, thus category '11' will be referred to as '10' from now on. It was found to be an ambiguous category on reflection which could be replaced by other categories. Additionally, some ‘3’ and ‘4’ categorisations were changed to other category types. The former had sometimes been allocated after an ambiguous/random first response to a cue (i.e. category ‘7’) and so not informative for analysis. The latter due to what was first deemed a general association, was in fact an association shared by more than one other participant, thus these were re-categorised as category ‘1’. The revised categories can be found in **Table 3**.

Category	**Description**
0	No response given/ “I don’t know”/ “I don’t want to play”
1	Recognised association (i.e. what an adult might say in response to the word)
2	Association unique to individual (based on parental comments- given in brackets if there are any; or when referencing own life e.g. “my car”)
3	Association arising from a previous response given (e.g. PIG- 1st= mud, 2nd= straw, 3rd= moss. The 2nd and 3rd responses relate to the 1st response ‘mud’ rather than the cue word PIG)
4	Related in a general/wider sense (i.e. not an obvious association but a logical connection e.g. trousers- people)
5	Repetition of the cue word/ repetition of a response already given
6	Naming something in the immediate environment (this will be noted in brackets)
7	An unclear association (i.e. cannot be coded 1-6 or 8-10)

8	Rhyme (e.g. CAR - bar)
9	Sounding out (e.g. APPLE – 'a' for apple)
10	Action/mime or sound to indicate cue word (e.g. LION - roar)

Table 3*: Experiment 1. The revised categories for coding participant responses*

Coding for Paradigmatic and Syntagmatic Associations

After categorising the type of response using the above system, related responses (i.e. category 1, 2, and 4), were tagged as paradigmatic, syntagmatic, or both. Definitions used in previous WA studies with children (Sheng et al., 2006; Wojcik & Kandhadai, 2020), were adopted such that a response was marked:

Paradigmatic if considered a superordinate (e.g. cat-animal), a subordinate (e.g. train-carriage), a synonym (e.g. brush-comb), an antonym (e.g. hot-cold), or a category coordinate (e.g. elephant-dog).

Syntagmatic if able to syntactically follow or precede the cue (e.g. wash-hands, drink-have), or if thematically close (e.g. bed-story)

Both if satisfying the conditions of both a paradigmatic and syntagmatic response.

Rater-reliability

Category Coding

Two junior researchers (*Rater S* and *Rater J*) were given a sub-set of the data (25%) to code. They were briefed on the category types and asked to code 7 participants' responses (5%

of the dataset) together. One participant was randomly selected from each of the seven age bins for this subset. The researchers were given the opportunity to ask questions regarding the category types during the briefing and after the task. Following this, the junior researchers worked independently and were each given a different set of 14 participants' data (10% of the full dataset each) to code, this time with two participants randomly selected from each age bin, ensuring an even spread of male and female participants. This was then cross-checked with the author's category allocation.

Observer agreement for these categorical data were calculated using the Cohen's kappa measure, and the level of agreement interpreted following Landis and Koch (1977).

The inter-rater agreement for *Rater S* was 92%, with a Cohen's k of 0.63, demonstrating substantial agreement (for 10% of the full dataset). The inter-rater agreement for *Rater J* was 90% with a Cohen's k of 0.61, indicating again substantial agreement (for 10% of the full dataset).

Together the two raters had a percentage agreement of 91% and a Cohen's k of 0.62 (for 5% of the full dataset).

Altogether this provides strong evidence that the coding scheme used is objective and likely to elicit the same category codes across raters, rather than by chance.

On closer inspection of the discrepancies, differences were mostly noted between whether an association was a direct association (i.e. Category 1) or a less obvious but still related in a general sense (i.e. Category 4). For analysis, these categories will be collapsed when comparing associated and non-associated responses so this difference in category allocation becomes irrelevant at this point.

Paradigmatic and Syntagmatic Coding

All associated responses were coded as paradigmatic, syntagmatic or both. To test for the reliability of this coding, two different junior researchers collaborated to code a random sample of 10% of the responses. The inter-rater agreement was 93%, with a Cohen's k of 0.92, demonstrating near perfect agreement (for 10% of the full dataset). This provides strong evidence that the coding performed is objective and likely to elicit the same categorisation across raters, rather than by chance.

Associative Strength Analysis

The likelihood of a cue word producing a particular response in a WA task can be indexed using a measure of associative strength (Nelson et al., 2000).

By recording the number of participants producing a particular response to a cue (P), and the total number in the group responding to a given cue (G), forward and backward strength metrics can be calculated. Forward strength (FSG) is the probability that a cue word will produce the 'target' or response word e.g. *cat* -> *dog* whereas backward strength or 'target-to-cue strength' is the probability of the response word producing the cue e.g. *cat* <- *dog* (Nelson et al., 2005). FSG is calculated as P/G resulting in the proportion of participants producing the response out of all of those presented with it. BSG is calculated in the same way.

We performed two FSG calculations: one for all responses (i.e. capturing all 3 attempts for each of the cue words given to a child) and one for first responses only.

To calculate FSG for all responses, we adjusted the value 'G' to be the total number of attempts at producing a response. For example, if the word 'cat' is allocated to 12 participants, 'G' is calculated as 12*3=36. This is a more conservative metric than measuring the number of participants. The number of participants 'N' given a cue is still included for reference. For the FSG calculation for first responses, we used the standard G metric (i.e. number of participants given the cue).

Data Cleaning

First, the total number of responses was calculated per cue word, excluding the number of Category 0 responses i.e. No Response. This provided the 'G' value, or the total number of instances the cue was presented to the group. Due to the subtraction of Category 0 responses from this total, the G value varies per cue word.

Then, the associated response types (i.e. Category 1, 2, 4) were homogenized. Since parents typed in the responses themselves, there was some variability to how this was done e.g. spaces left before the typed response, spelling errors, use of capital letters, multi-word responses. Grouping was first conducted at this level. Subsequently, the following principles were applied when making final changes to grouping responses:

- Plurals were grouped with singular exemplars of the same word following the data processing of related studies (Entwisle, 1966)
- Multi-word utterances containing a noun, gave preference to the noun over other words in the utterance. The rationale for this is due to the nature of this study, which

is to compile a list of noun-noun WAs to be used as stimuli in semantic priming and word meaning studies with infant participants

Each response was preserved in its entirety when grouping with other responses for full disclosure of child responses, and these entries are separated by the forward slash symbol '/' in the results tables.

After grouping responses in this way, the 'P' value could be calculated to reflect the number of participants per group producing a particular response. The strength of the word association (FSG) was calculated for every response that was produced by two or more participants following the procedure used by Nelson et al. (2000). This was done by dividing the P value by the G value: P/G, to generate a proportion which could be compared to other datasets looking at associative strength in WAs (Moss et al., 1996; Nelson et al., 1998).

The backward strength of word combinations was calculated where possible.

An additional metric recorded for reference is the number of idiosyncratic responses for a cue (i.e. those produced only by an individual participant).

2.2.4 Results

A total of 168 participants participated in the WA study altogether. The 14 participants from the pilot study were not included in the final dataset as many fell outside the revised age range. Three participants were excluded from analysis due to exposure to another language for more than 20% of the time (*N*= 2), or an inability to complete the task (*N*= 2).

The final sample size was 150 participants (female=84, male=66), despite an original 140-participant target. This was due to some parents submitting responses much later than anticipated. This was permitted and involved some participants changing between age bins (due to the time elapsed between invite and task completion) which also sometimes resulted in a word list being completed by multiple participants in an age bin. **Table 4** shows the distribution of participants across age bins and the number of male and female participants.

	Male	**Female**	**Total**
34-35	7	15	22
36-37	12	10	22
38-39	8	13	21
40-41	11	10	21
42-43	9	12	21
44-45	10	13	23
46-47	9	11	20
Total	66	84	150

Table 4*: Experiment 1. Number of participants by age bin and gender*

In a few cases, individual target words show greater representation than others. Before completing the task, parents were instructed to check the wordlist allocated to them and to inform the experimenter if their child did not know any of the words in the list. Some words were substituted in lists for this reason. An attempt was made to include omitted words in other lists, but this was not always successful due to the scale of the project and managing a large number of participants at different stages of the procedure simultaneously.

All of a child's responses were included in the analysis and no child was excluded based on a minimum number of responses given or the types of responses given. It was assumed that the target words in a wordlist were known to the child as parents were given the opportunity to request replacement words for their child if some of the words from the wordlist were unfamiliar. The exploratory nature of the study also prompted no exclusion criteria to be placed on number and type of a child's responses.

Responses were input into Excel files and uploaded in R Studio version 1.4.1717 for all further analyses. The R tidyverse and dplyr packages were used.

Descriptive Statistics

A total of 4512 responses were collected from 150 3-year-olds completing the WA task. After subtracting responses categorised as 'No Response' (i.e. Category 0, *N*= 908), a total of 3604 responses remained. This produces an average of 24 responses out of a possible 30, since each child could respond 3 times for each of the 10 cue words.

Considering first responses only, out of a possible 1500 responses (150 participants, each with 10 cue words), 1454 responses remained after subtracting 'No responses' (*N*= 46). The mean response rate was 9.69.

Types of Responses

Considering the specific language used when giving responses, **Figure 2** shows the distribution of word forms used. Nouns were most frequently used, followed by verbs, then adjectives.

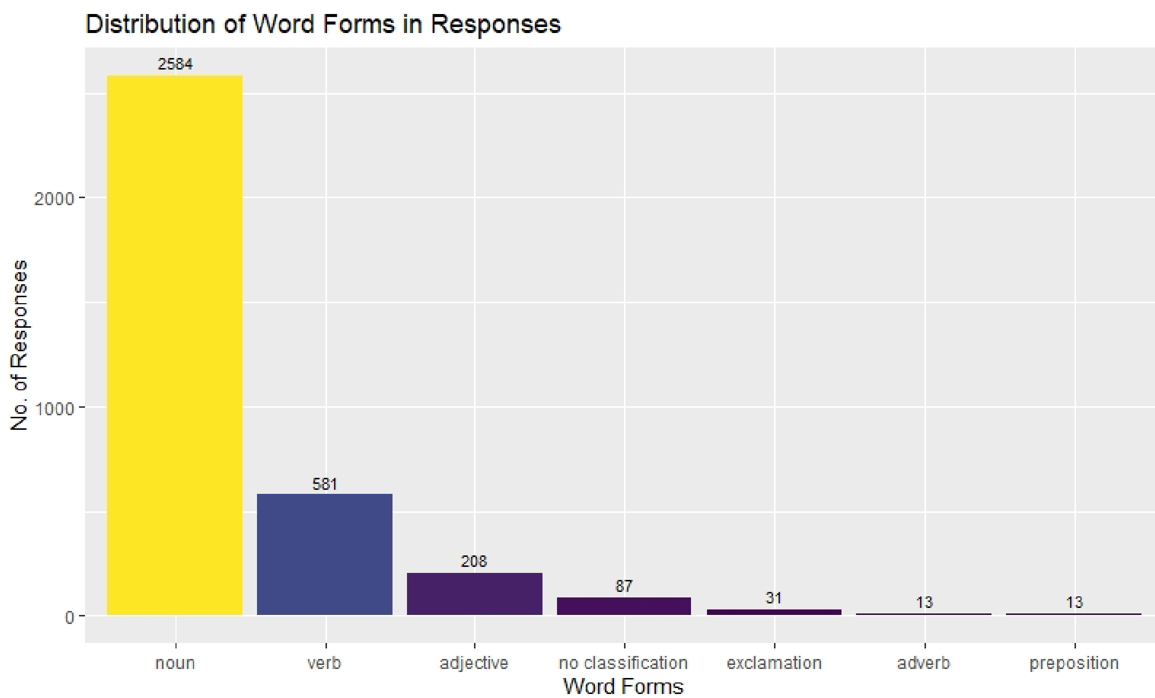

Figure 2*: Experiment 1. The number and type of word forms in participant responses*

The distribution of all responses by response type in **Figure 3** shows that the highest proportion of responses had a direct semantic relationship to the cue word, i.e. Category 1 responses (51%), followed by a high proportion of Category 0 responses (20%). The latter represents a failure to respond or a response such as "I don't know" or "I don't want to play".

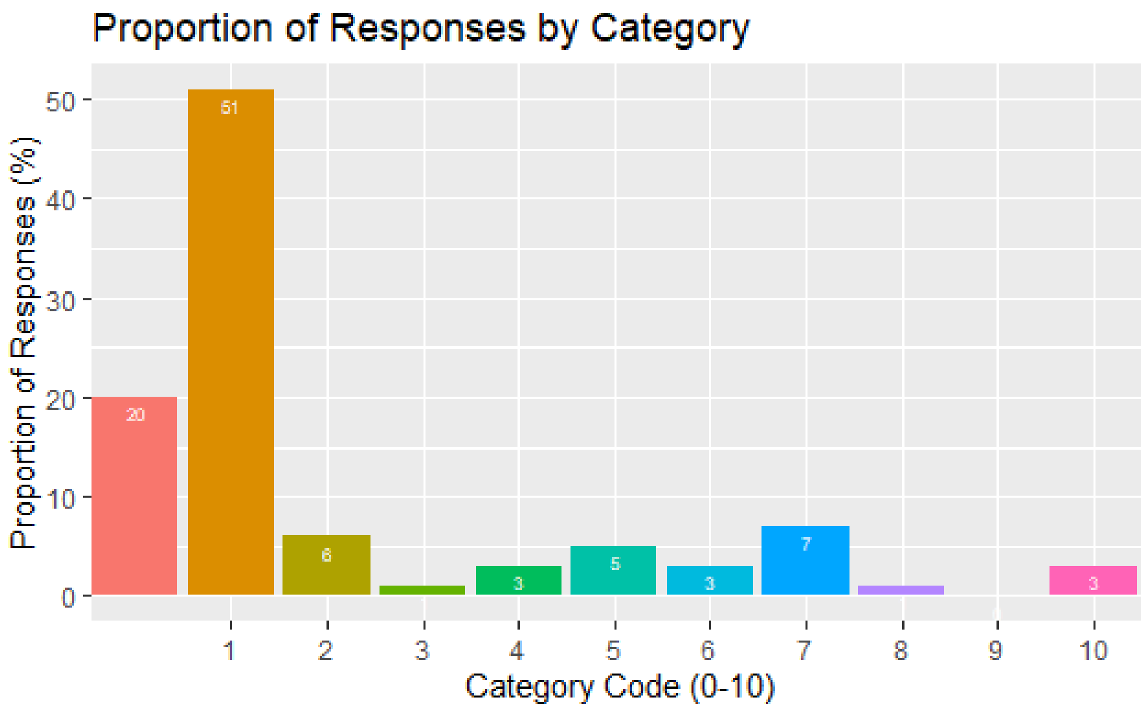

***Figure 3**: Experiment 1. The proportion of WA responses by response category*

By organising responses into no response (Category 0) and collapsing categories representing a related response (Categories 1, 2, 4, 10) and responses which are not related (Categories 3, 5, 6, 7, 8, 9), **Table 5** illustrates the distribution of response types in the whole group, as a percentage and as a raw value.

	No. Responses	Proportion
No Response Given	908	20%
Related Response	2807	62%
Unrelated Response	797	18%

***Table 5**: Experiment 1. Proportion of responses by relatedness of response type*

This indicates that children aged 34 to 47 months can perform a free association task and respond with an appropriately related response over half of the time. A 'No Response' occurred more frequently than an unrelated response.

Related vs. Unrelated Responses

Category 0, or 'No Responses', will now be disregarded to observe patterns between related and unrelated response types by age and gender.

The proportion of related responses by age bin is displayed in **Figure 4** and indicates an increased trend in related responses as a child gets older.

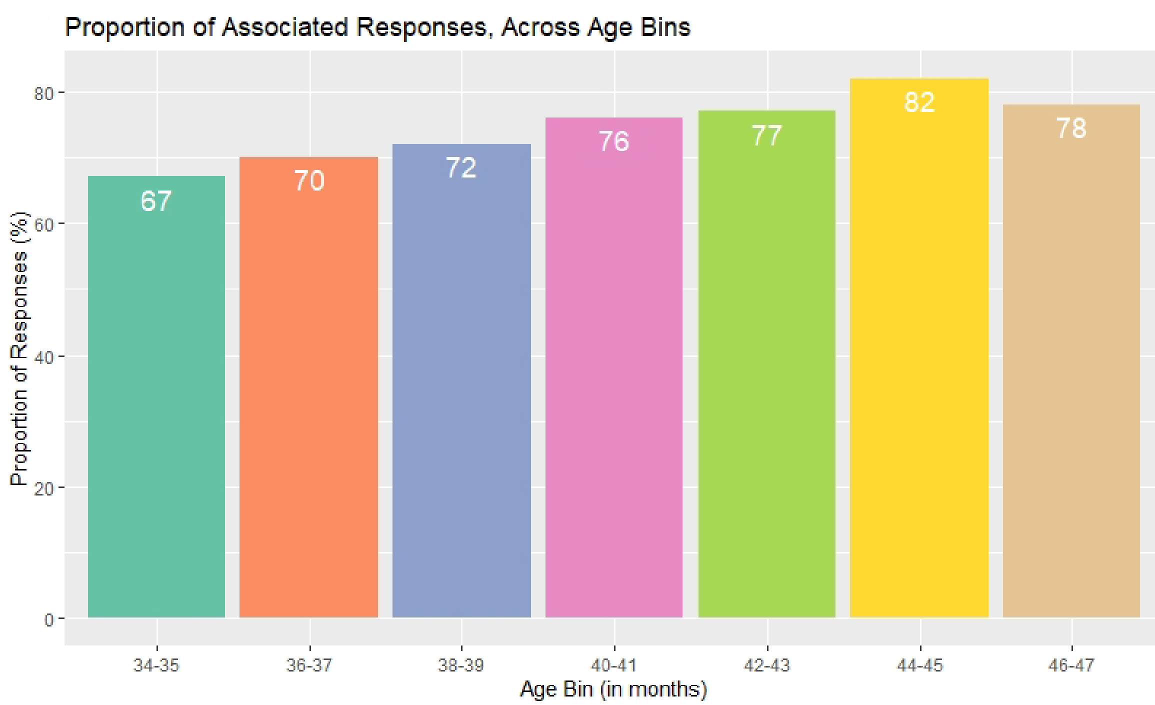

***Figure 4**: Experiment 1. Proportion of related responses by age bin*

Turning to first responses only, and since some participants did not provide a response for each of the 10 cue words, a proportional score of related responses was calculated for each participant. This was the number of related responses divided by the total number of responses (i.e. minus No Responses). The overall mean proportion of related responses to the task was 0.85 (*SD*= 0.21). We ran a Type III ANOVA on participants' proportion of related responses with gender and age bin as fixed factors. There were no significant differences between the proportion of related responses by gender and age, and no interactions between the variables ($ps > 0.1$). Thus these variables were discounted in further analyses.

Paradigmatic vs. Syntagmatic

Semantically-related responses (Categories 1, 2, 4, 10) were categorised as paradigmatic, syntagmatic or both. Following Wojcik and Kandhadai's (2020) method of calculation, responses classified as paradigmatic or both were combined. With this method, 26% of all responses were paradigmatic (or both) and 74% of responses were syntagmatic. In the case of first responses, 25.2% of first responses were paradigmatic (or both) and 74.8% of first responses were syntagmatic.

The proportion of paradigmatic responses was calculated for each individual participant. We ran a Type III ANOVA on participants' proportion of paradigmatic responses for all responses with gender and age bin as fixed factors. There were no significant differences between the proportion of paradigmatic responses by gender and age, and no interactions between the variables ($ps > 0.1$). The same was true when analysing first responses only ($ps > 0.1$).

Taking age as a continuous variable, a Pearson's Correlation Coefficient was calculated to observe for any effect of age maturation on the proportion of paradigmatic responses. When considering all responses, there was a weak negative correlation that was not significant $r(148) = -.01$, $p = .91$. Considering first responses only, there was a weak positive correlation that was also not significant $r(148) = .05$, $p = .55$. The proportion of paradigmatic responses was not affected with increased age between 34 and 47 months.

Associative Strength in Children's Associated Responses

Responses given by two or more children to each of the 100 cue words were processed to calculate their forward word association strength (FSG, Nelson et al., 2000). Looking at all responses, a total of 432 responses had two or more participants producing the same response for a cue word, with all 100 cue words represented in these responses. The full list of cue words with 2 or more of the same response and their associative strengths (*M*= 0.09, *Range*= 0.04 to 0.29) can be found in **Appendix F**.

Looking at first responses only, a total of 188 responses had two or more participants producing the same response for a cue word, with 96 of the cue words represented in these responses. The full list of cue words (organised alphabetically) with 2 or more of the same response as a first response and their associative strengths (*M*= 0.20, *Range*= 0.11 to 0.69) can be found in **Appendix G** .

We then compared the associative strength between the two sets (i.e. all responses vs. first responses). Associative strength in the two sets differed significantly according to Welch's t-test, $t(618) = -19.92$, $p < 0.001$. The associative strength for word pairs in first responses

(*M*= 0.20, *SD*= 0.08) was greater than associative strength in all responses (*M*= 0.07, *SD*= 0.04). Due to this difference, we focus henceforth on first responses.Since one aim of this research was to look at the most common imageable noun-noun associated word pairs in 3-year-olds, we extracted noun-noun word pairs to create a stimulus resource bank. Of the 188 responses shared by two or more children, 115 of these were noun-noun word pairs (see **Appendix H**).

Based on the findings of this research alone and using the metric of associative strength, these 115 word combinations can be considered the most common imageable noun-noun word combinations in this sample of 3-year-olds.

To determine whether the most common WAs in these 3-year-olds are unique to this age group, a comparison will be drawn in the following section using the associative strengths from the adult literature for the same word combinations.

Associative Strength Compared to Adults' Associated Responses

Of the 188 word-word combinations produced as first responses by 2 or more of the 150 3-year-olds in this study, 30 were not characterised in either the Birkbeck or the South Florida norms (though the cue was used); 13 were not used as a cue in the Birkbeck norms, nor documented as an associated response in the South Florida norms; 2 were not documented as an associated response in the Birkbeck norms, nor used as a cue in the South Florida norms; and 4 were not used as a cue in either study, resulting in a total of 49 word pairs found in children's responses, without a value of associated strength in adults[5]. This missing data corresponds to a total of 26%[6] associated responses found in 3-year-olds that is not reflected in adult associative norms.

The resulting 139 word pairs which are represented in the adult data were analysed. Where there was an associative strength available in the two adult studies used for comparison (Moss et al., 1996; Nelson et al., 1998), the mean of the two was taken, but where only one value was available, this was taken to represent the associative strength in adults. The 139 word pairs can be seen in **Appendix I** , ordered alphabetically by cue word.

A *t*-test was run to determine any difference between the associative strength between word pairs in children and adults. There was a significant difference in the associative strength between age groups $t(137) = -4.58$, $p < .001$, indicating stronger associative strength between word pairs in children (M= 0.21, *Range*= 0.11 - 0.69) compared to adults (M= 0.14, *Range*= 0.01 - 0.76). There was a significant, weak positive correlation between the two groups $r(137) = .22$, $p = .01$. This shows a tendency for strongly associated word pairs in adults, to be strongly associated in children too

Concerning the word pairs present and absent in adult norms, there was a significant difference in the associative strengths of word pairs available only in the child data, and those available in the adult and child data $t(186) = -2.56$, $p = .01$. Word pairs present in adult norms showed stronger associative strength in children (M= 0.21, *Range*= 0.11 - 0.69) than those absent in adult norms (M= 0.17, *Range*= 0.11 - 0.36). Word associations not

[5] If considering all responses to a cue word and not just the first response, this increases the total to 181 word pairs not represented in adult norms.
[6] When all responses are considered, this increases the proportion of associated responses in children that are not found in adult norms, to 42%.

represented in the adult data are not as strongly associated in children as those that are represented.

Some of the strongest WAs in the data are not replicated in the adult literature so while no comparison can be made statistically, these may represent different WAs in 3-year-olds that warrant further testing. These word combinations are displayed in Appendix J.

Associative Strength in Word Pairs used in Infant Studies Investigating Semantic Development

Studies cited in the introduction of this thesis, which investigate the development of semantic meaning in infants, were used as a basis for compiling a list of typically selected associatively related word pairings for infant studies, which derive from adult associative norms. These word pairs are compared to the productive word combinations found in this study[7].

The stimuli lists from eight studies, with some studies reporting more than one experiment, were compared. A total of 141 prime-target related word pairs were present in the dataset. Any word pair that was represented more than once in a study or across studies, was counted just once. Sixteen (11%) of the 141 prime-target related word pairs were represented in the associative responses produced by 3-year-olds. These can be found in **Table 6**.

[7] We included all responses, rather than first responses only for this comparison, to increase the likelihood of the word combinations in the studies mentioned being found in our data set.

	Prime	Target	Study	Languages	Experiment type	FSG_3-year-olds	FSG_Moss and Older (1996)	FSG_Nelson et al. (1999)
1	cat	dog	Styles and Plunkett (2009)	English	auditory+visual	0.16	0.667	0.513
	cat	dog	Arias-Trejo and Plunkett (2009)	English	auditory+visual	0.16	0.667	0.513
	cat	dog	Willits et al. (2013)	English	auditory	0.16	0.667	0.513
	cat	dog	Singh (2013)	English+Mandarin	auditory+visual	0.16	0.667	0.513
2	park	swing	Chow et al. (2017)	English	auditory+visual	0.16	0.021	0.061
3	sheep	cow	Styles and Plunkett (2009)	English	auditory+visual	0.14	0.133	0.014
	sheep	cow	Arias-Trejo and Plunkett (2009)	English	auditory+visual	0.14	0.133	0.014
	sheep	cow	Jardak and Byers-Heinlein (2019)	English+French	auditory+visual	0.14	0.133	0.014
	sheep	cow	Singh (2013)	English+Mandarin	auditory+visual	0.14	0.133	0.014
4	dog	kitty	Willits et al. (2013)	English	auditory	0.13	0.511 (cat)	0.667
5	table	chair	Singh (2013)	English+Mandarin	auditory+visual	0.13	not a cue	0.756
6	hair	brush	Chow et al. (2017)	English	auditory+visual	0.11	0.021	0.207
7	cereal	bowl	Chow et al. (2017)	English	auditory+visual	0.1	not found	0.031
8	sofa	cushions	Chow et al. (2017)	English	auditory+visual	0.09	not found	not found
9	window	door	Jardak and Byers-Heinlein (2019)	English+French	auditory+visual	0.09	not a cue	not found
10	rabbit	carrot	Chow et al. (2017)	English	auditory+visual	0.08	0.034	not found
11	hand	foot	Jardak and Byers-Heinlein (2019)	English+French	auditory+visual	0.06	0.095	0.396 (feet)
12	bus	train	Jardak and Byers-Heinlein (2019)	English+French	auditory+visual	0.05	0.041	not found
13	chair	table	Jardak and Byers-Heinlein (2019)	English+French	auditory+visual	0.05	not a cue	0.314
14	leg	hand	Singh (2013)	English+Mandarin	auditory+visual	0.05	0.024	not found
15	shoe	foot	Chow et al. (2017)	English	auditory+visual	0.05	not a cue	0.329 (feet)
16	park	tree	Arias-Trejo and Plunkett (2009)	English	auditory+visual	0.04	0.083	0.115
FSG= forward associative strength								

***Table 6**: Experiment 1. Common prime-target stimuli used in infant studies investigating semantic development*

To compare associative strengths of these word pairs in the two age groups, the FSG was averaged across the two adult studies where possible, otherwise the FSG value was taken from the study it was available in. A t-test was run to determine if there was a difference in the associative strength between words pairs in adults and children. There was a statistically significant difference in the associative strength found between word pairs in adults and children $t(14) = -2.71$, $p = .01$. Associative strength between words pairs is stronger in adults (*M*= 0.25, *Range*= 0.02 - 0.76) than in children (*M*= 0.11, *Range*= 0.04 - 0.16).

Associative strength between specific word pairs in adults and children- commonly used in infant semantic priming studies- was found to be weakly positively correlated $r(12) = .41$, $p = .15$, though this was not statistically significant. The stronger the word pair in adults, the stronger it is in children too.

2.2.5 Discussion

Experiment 1 tested if children as young as 3 years old could successfully complete a WA task and sought to compare any recurring responses in children, to those found in adult norms using forward associative strength as the metric of comparison.

There was strong evidence that children between 34 and 47 months can produce associated responses in a repeated free association task. In fact, 3-year-olds produced related responses for the majority of their responses (62%). After related responses, 'No Responses' (Category 0) were largest in number, accounting for a fifth of all responses.

By looking at first responses only, we see that the overall mean proportion of related responses to the task was 0.85 (*SD*= 0.21) which is higher than when looking at the proportion of related responses from all attempts to respond to a cue word. This suggests that the likelihood of a response being related with subsequent responses to the same cue word, decreases. There was a clear rationale for the three-response-per-cue-word design of the experiment, however, this might not be a reasonable expectation of all children of this age in terms of ability and attention found in children. One way to accommodate for these individual differences might be to elicit a first response and then allow the experimenter to decide whether to pursue further responses, depending on the child's ability and behaviour at test, making it explicit that one response is sufficient. There was no effect of age on children's WA responses, but a tendency for the production of more related responses as age increases. With a larger age spread in participants, it might be easier to see how as age increases, so too does suitability of response in a WA task. The interaction of age and increased vocabulary is likely to drive this trend and future research might collect participant information on vocabulary size to see if related responses correlate with age and vocabulary size.

A large number (*N*= 432) of associated word combinations were produced by two or more 3-year-olds ranging in associative strength from 0.04 to 0.29 (*M*= 0.09), when taking all responses into account. When considering first responses only, 188 word combinations were produced by 2+ children (FSG: *M*= 0.20, *Range*= 0.11 to 0.69). Together this is strong evidence that 3-year-olds can successfully complete a WA task and even produce some of the same responses as their peers, rather than just idiosyncratic responses. This might provide a glimpse into the shared experiences of 3-year-old children, which is represented in their lexical-semantic structure at this age. The fact that some, especially strong connections found between word pairs, seen in children's related responses but not in adult associative norms could reflect this phenomenon too. However, caution must be taken not to generalise, and these findings would need to be replicated to draw any inference about the probability that a particular cue will elicit an expected associated response in a 3-year-old. This will be addressed in Experiment 2.

Seventy-four percent of all the related responses given by 34-47-month-olds were syntagmatic (74.8% of first responses) and there was no effect of age on the rate of paradigmatic responses, in the third year of life. The tendency for 3-year-olds to produce syntagmatic responses in a language production task is in line with the suggestion that associative or thematic semantic meaning shows greater primacy than taxonomic semantic meaning. It also corresponds to the finding that a shift to paradigmatic responses in a WA task occurs much later, at 6 years of age.

Nouns were overwhelmingly the word form of choice in participant responses. This tendency for a noun response corresponds to that found in the wider literature. The fact that all cue words were nouns in this paradigm is likely to have contributed to this outcome. An interesting point to note is that noun responses were not paradigmatic in nature. One assumption relating to word class is that nouns elicit paradigmatic responses at an earlier age (Nelson, 1977). This experiment did not replicate this finding since the majority of responses were syntagmatic, despite all cues being nouns. This might suggest that at 3 years of age, it is too early to notice any signs of a shift to a more adult-like response type i.e. predominantly paradigmatic responses, in a WA task.

Since a feature of this research was to focus on nouns so that a stimuli list might be created for future researchers, it did not provide the opportunity to test WAs to words of other grammatical classes in 3-year-olds. According to Nelson (1977), the syntagmatic-paradigmatic shift is most noticeable in high-frequency adjectives, which has been documented by other research showing a larger proportion of paradigmatic response in children responding to adjectives (Cronin, 2002; Sheng et al., 2006). However, Wojcik and Kandhadai (2020) did not replicate this finding in children as young as 3. Whether 3 years of age is too young an age to observe an effect of adjectives on paradigmatic response, as was the case in the present study with no effect of nouns on paradigmatic responses, remains to be seen.

When comparing the WAs in children to those found in adults, the findings from Experiment 1 indeed suggest that adults and children converge in the likelihood that certain cue words will elicit the same associative responses, however, this is only true for some word pairs. A direct comparison is difficult to make between the associative strengths found in children and adults due to other variables that might be at play. In fact, not much is known about the variables affecting WA behaviour (Fitzpatrick et al., 2013), yet a cautionary comparison can give a sense of any general patterns that might exist.

Analyses suggest a significant difference in the associative strength in adults compared to children, with stronger links in children compared to adults for the word pairs compared in this study. There was evidence for a weak positive correlation between the associative strengths found in adults and children which indicates a pattern of increased strength for word pairs in 3-year-olds, which exist in adults. We might interpret this as a gradual developmental path from the immature to the mature lexical-semantic system. It would be informative to map if this trend strengthens when testing children older than three, at different time intervals, in future research.

Further evidence to corroborate the developmental trajectory of an immature to a mature lexical-semantic system comes from the associative strength of missing WAs from adult norms which are not as strongly associated in children as the WAs found in children and adults. This points to children having unique WAs at the age of three and though these might be associatively weaker when using the metric of associative strength, they nonetheless exist (maybe temporarily) and are shared by others at this age.

The previous point is important to emphasise and explore further. The 100 cue words in this study generated 188 associated responses given as a first response by two or more 3-year-olds, yet only 139 of these associatively-related pairs could be found in adult associative norms. In other words, 26% of responses given by two or more children are not found in adult norms[8], and this includes some of the strongest associated word pairs found in children. This in itself might be indicative of the transitory nature of the immature lexical-semantic system. Some adult associations might not form in infancy and remain into adulthood, instead, these findings suggest that there are unique WAs at 3 years of age which may be replaced by other, more adult-like associations, with increased age and life experience. This could occur in parallel to a subset of word pairs, shown to exist in both children and adults, though the strength of these associations differ.

According to Fitzpatrick et al. (2013), referencing WAs that have not been taken from the target population might not acknowledge the unique characteristics of the population of

[8] This rises to 42% when looking at all responses, not just first responses.

interest. This might be true of the WAs found in the children of this study, but not found in adult norms. Arias-Trejo and Plunkett (2009) proposed that “the absence of a semantic–associative relation in adult norms may be a reliable indicator of its absence in infants”, however, the findings from Experiment 1 would suggest that the absence of a WA in adult associative norms, is not necessarily a reliable indicator of its absence in the developing lexical-semantic system.

This further highlights how caution must be taken when using adult norms to inform stimuli selection for infant studies. The absence of some of the strongest child WAs in the associative norms of adults is of relevance to the wider research field. Studies designed to investigate semantic development in infants rely on the WAs documented in adult norms when selecting appropriate stimuli (i.e. prime and target word pairs). One example is the word pair ‘teddy-bed’/ ‘bed/teddy’ which has both forward and backward strength in 3-year-olds, but is not present in adult norms. This word pair makes intuitive sense to exist as a strong association in the mind of a child, though relying on adult norms would not capture it as a suitable pair for use in an experiment. This example serves to highlight the importance of stepping back from the need to match stimuli on a number of criteria and searching in adult associative norms for the ‘perfect match’ to fulfill these criteria. Rather, we must consider the most child-appropriate word pairs for use in experiments investigating the emergence of semantic meaning in infancy.

Some of the most common prime-target WAs employed in infant semantic priming studies were represented in the productive language of 3-year-olds performing a WA task. However, the actual proportion of the prime-target pairs equates to only 11%, which represents a minority. The sample of 3-year-olds tested in this study clearly share some of the WAs found in adult associative norms, but have their own, more child-specific associations, which can be stronger than word pairs in the adult literature. This suggests a more reliable source of WAs for use in semantic priming studies, needs to come from the WAs documented in children rather than adults, and ideally in children as close in age to the population being tested.

A list of 115 associatively-related word pairs was compiled for this research, which reflects associated first responses to cue words produced by two or more 3-year-olds engaged in a free association task. These word pairs comprise of imageable noun-noun combinations. The list thus provides a resource of stimuli that might be consulted when designing studies investigating semantic development in young children. These word pairs reflect language production and since production succeeds language comprehension, which is what studies investigating semantic development do, it is the closest we might get to knowing the precise WAs children form as their lexical-semantic system undergoes development.

Future Research

An interesting finding to emerge from this study came not from participant responses, but from email correspondence with parents and caregivers administering the task with their children. Some children struggled to understand or engage with the task, despite the parent or caregiver trying multiple times, using a family member to model the task, or leaving a period of time before attempting the task for a second or third time. This raises the question of whether the task itself was unclear or too cognitively demanding for some children, which may have inhibited them from participating, or may have meant they performed the task without fully understanding the type of response to give.

Another observation from email correspondence was the variability in which parents or caregivers administered the task. Despite clear instructions to follow a script for the task, this was not always done. The list of words was also not followed in the order given. This is important as the order of the words in the list had been carefully selected to avoid consecutive words having a semantic or associative meaning, or from sharing the same phonological onset or element of rhyme in any way. This was controlled to avoid any carry-over effect of one word in the list, to the next. Some items in a list were from the same semantic category, due to the limited number of categories available at such a young age, so these may have been given consecutively, based on the knowledge that some parents had changed the order of the words.

There were also instances of the task being divided up and administered over multiple days due to the child's limited attention span or willingness to participate. While the findings do not indicate this confounded performance, in terms of response rate, it could point to the nature of the task needing modification to make it more accessible and engaging to a wider range of children.

With these observations in mind, an alternative format for the experimental procedure will be developed in Experiment 2 to mitigate for any confounding effects of the parent acting as experimenter.

2.3 Experiment 2: Word Association Study on Monolingual Children, Online

2.3.1 Introduction

Findings from Experiment 1 validated the use of a free association task on 3-year-olds when the task is administered by a parent. Having the parent act as 'experimenter' inevitably calls into question the validity of the task's administration, and indeed informal correspondence with participants indicated that there were some deviations from the delivery of the task when performed by different families in their unique home contexts. While this may not directly influence the types of responses a child gives, it warrants a replication study to confirm that when a parent administers the task at home, the types of WAs that a 3-year-old produces in this context are the same types of responses that would be given in a more controlled setting. This potential confound has prompted an adaptation of the original methodology into an online format.

The online WA task did not require the parent to act as the experimenter, but instead uses pre-recorded videos of puppets describing and demonstrating the task. A participant's responses were recorded, and the task was in a more engaging format to optimise the likelihood of the child staying focused on the task and minimising the chance they would become distracted or resort to naming things around them. A further impetus to test online rather than in a lab setting was the inability to test face-to-face currently due to the global pandemic.

In Experiment 2, we asked if the WAs produced by 3-year-olds in the parentally-administered version of the task could be replicated in another modality, that is, in an online format. To what extent the context influenced the responses was addressed, as well as observing if word pairs found in Experiment 1 re-occurred in this online modality, and if their associative strength was replicated.

The task remained very similar in its design through its remote administration, for instance, by using the same cue words and with ten cue words and three responses encouraged for each cue word. However a homogeneous delivery of the task was better achieved by controlling how the task was explained and how responses were recorded.

Piloting and the findings from Experiment 1 indicated that participants under 36 months were not able to consistently complete the task and in general there was only a small effect of age on responses in a WA task. We therefore adjusted the age range for Experiment 2 to 36-39 months. From the ten lists of cue words in Experiment 1, cue words eliciting the most frequent WAs were selected to create two new lists with ten words per list for Experiment 2. Each cue word in List 1 corresponds with its most frequent associate in List 2, thus providing an opportunity to observe forward and backward priming strength between the word pairs.

We predicted that overall, there would be a replication of the most frequent WAs in three-year-olds in the modified online modality. However, due to individual differences, and a high idiosyncratic response rate in young children (Wojcik & Kandhadai, 2020), the strength of the WAs and specific word pairings may differ for Experiment 2. If the parent acting as the ‘experimenter’ was a confounding factor in Experiment 1, then we would expect there to be a marked difference in the types of the responses produced by participants (e.g. fewer related responses). However, we anticipated that some parents may have influenced some responses in Experiment 1 and subjectively chosen the answers they reported, and this may be evident through a larger proportion of unrelated responses, that a parent might have otherwise disregarded. Equally, if the online modality makes the task more engaging, we expect to see a reduction in the naming of objects in the immediate environment and potentially a greater proportion of related responses. The source of the effect might be difficult to tease out, however, yet if the WAs ultimately remain the same, it would validate the delivery of the child-adapted free association task either by a parent or online.

Research Aims

The first aim was to replicate the WA findings from Experiment 1. This was to ensure that any frequent associations found in the data were not generated by chance, nor were they due to subjective parental report. The second aim was to develop an online modality of the task to ensure consistent administration of the task, objective recording of responses, and reduce attrition by boosting engagement with the task. The new procedure adopted the following principles:

- control more stringently for the order of presentation of words in a wordlist
- more clearly demonstrate the task with a ‘model’ participant responding to the example words in the task, rather than the experimenter describing them
- be more engaging by focusing a participant’s attention more effectively, to minimise the chance that a participant stops responding or resorts to naming things in their surrounding environment
- be suitable for online/ home-testing during the continued pandemic

Research Questions

1. Can an online free association task be successfully designed for 3-year-olds?
2. Are the types of responses replicated when the free association task is administered online to 3-year-olds?

2.3.2 Method

Participants

Monolingual toddlers were recruited from the UoP BabyLab database and its social media platform pages (*N*= 24: 13 female, 11 male). The mean age of participants was 1145 days (*SD*= 35.94, *Range*= 1085 to 1196 days) or 37.64 months. Participants were divided into two age bins: 36-37 months and 38-39 months (+/- 15 days) with 12 children in each age bin. CDI III scores (Fenson et al., 2007, lexical component only) were collected from participants, but only approximately a third of parents completed this part of the task (*N*= 7, *M*= 79.43/99, *SD*= 13.62).

Children were excluded from participating if speaking more than one language, if born more than six weeks prematurely, or with a diagnosed language or developmental delay.

Materials

Stimuli

Twenty highly imageable nouns, known by at least 60% of 18-month-olds according to the Oxford Communicative Development Inventory (Hamilton et al., 2000) and UK CDI (UK-CDI Database, 2016) were taken from the 100 words used in Experiment 1. Two new word lists comprising of 10 words each were created. Words were selected if they had generated at least 2 responses for a cue in Experiment 1, and if able to coordinate with another word in the list of 20 cue words, thus having the potential to generate forward and backward strength. Coordinates were divided into two separate lists as shown in **Table 7**.

	List 1	**List 2**
1	chair	table
2	bed	teddy
3	tooth	brush
4	finger	hand
5	key	door
6	sock	foot
7	bowl	cereal
8	head	hair
9	park	swing
10	bath	towel

Table 7*: Experiment 2. Experimental lists of cue words for the online WA task*

Audio and Video Recordings

The script given to parents in Experiment 1 to execute the WA task, was adapted for use online. The task explanation and examples were divided up to be delivered by two puppets, and there was greater exemplification to aid conceptual understanding of the task.

Video recordings were made of the puppets explaining and demonstrating the task by two junior researchers, all directed and overseen by the author. Great effort was taken to make the instructional delivery engaging by using child-directed speech and varying pitch. In addition to the main explanatory video, short motivational clips were recorded of the puppets encouraging participation and praising a participant for attempting the task.

Recordings were also made of the puppets counting down through the number of cue words, so the participant had sight of the duration of the task left to complete.

Cue words were recorded auditorily by the same junior researchers. This was to enable their presentation individually, without the puppets on screen, to minimise distractions while participants were prompted to produce responses to the cue words.

Procedure

After ethical approval had been granted by the UoP ethics panel, parents/ carers with children meeting the eligibility criteria were invited to the study through the UoP database or when expressing an interest through social media adverts. The email invitation included a Participant Information Sheet outlining the procedure, data handling, and a consent form which would later be completed digitally. Parents indicated the day and time they would complete the online experiment and a unique link was generated for the Gorilla Experiment platform, with further instructions on the procedure. Clicking on the link took the participants through a series of tasks, in the following order: study overview screen; participant eligibility questionnaire; consent form; audio and video test screen with equipment eligibility questionnaire; participant and parent/carer demographic questionnaire; word checklist; WA task; pointing word checklist; CDI III (lexis component only); debrief. An experimenter was available for questions and troubleshooting during the time the participant attempted the task.

For the WA task, auditory stimulus words were presented following an explanation of the task. The task explanation took the form of a recording of two puppets doing the task together in which examples of WAs were given (using words not in the stimulus list) and an emphasis was placed on the need to say the first thing that came to mind, as quickly as possible. **Figure 5** shows how the task and experimental interface look from the participant's perspective.

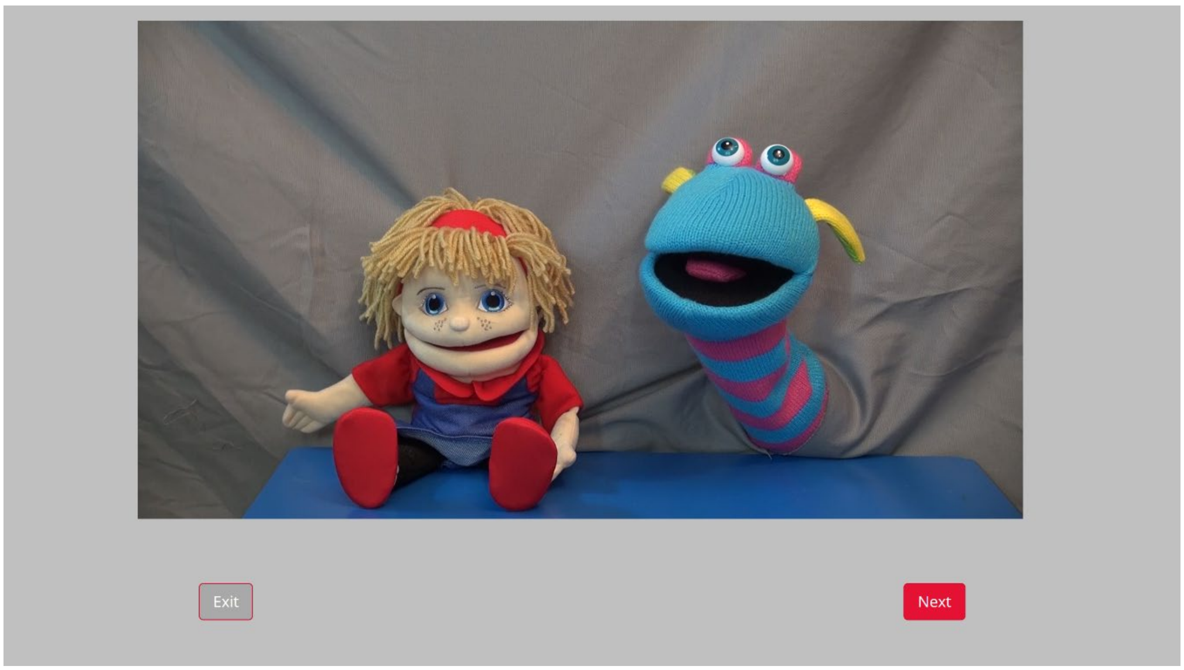

Figure 5*: Experiment 2. Still of the puppets explaining the online WA task*

Following the puppets' instructions, a cue word was played while an abstract, visual attention getter appeared on screen to maintain the child's attention to the task/ on screen. Parents were instructed to encourage their child to produce up to 3 responses per cue word during this time while an audio recording of the child and parent was made through the participant's device.

When clicking on ‘Next’ for a subsequent cue word, a video of the puppets praised the child’s attempt and three text fields appeared for the parent to type the child’s responses in, in the order given (see **Figure 6**). This feature was added in case of an error with the audio recording, or a difficulty understanding the child’s speech, and to analyse how parents record their child’s responses.

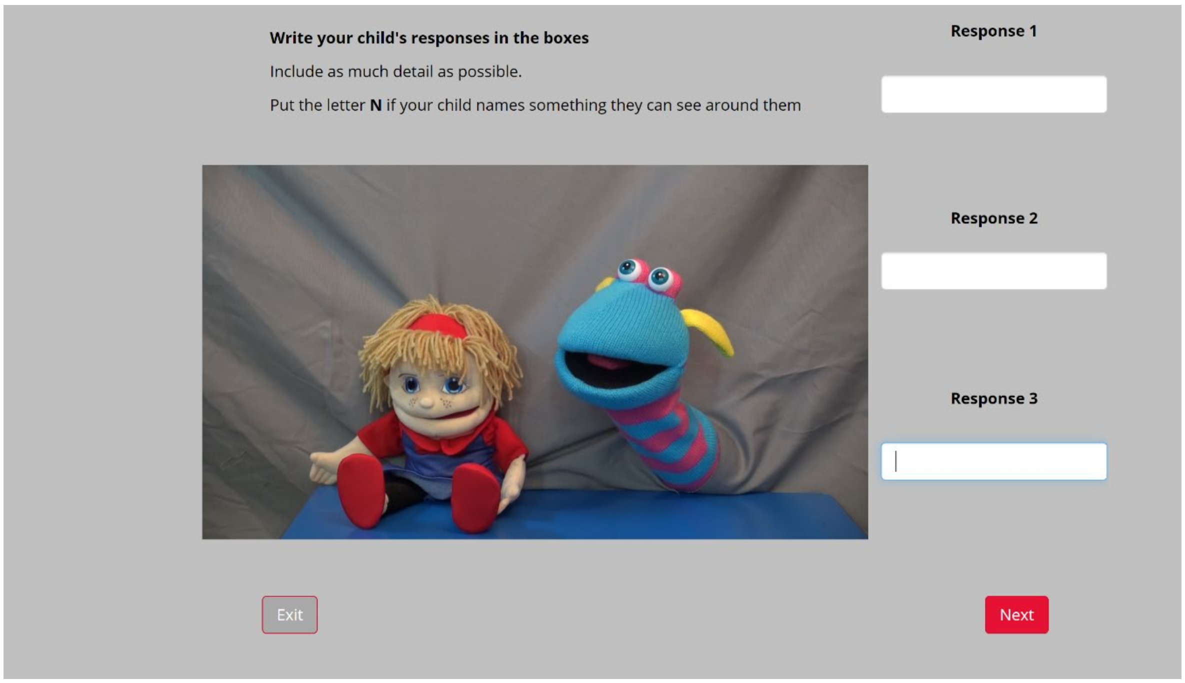

***Figure 6**: Experiment 2. Online text fields for carer to enter child's responses*

On every trial, the parent was able to determine when the child was ready to progress to the next word in the list by clicking on a ‘Next’ button. This allowed for individual differences in the speed at which a participant could produce up to three related words. It was made clear to parents to move on if a child could not think of three responses, or if a child became disengaged. Additionally, an 'Exit' button was present on every screen to end the task if the child did not want to continue.

After 5 words had been presented in this vein, a video of the puppets demonstrated the task again (with a word not included in the stimulus set). The final five words were then tested.

To end, a recording of the puppets demonstrated a final task for the child which directly tested comprehension of the stimulus words. The child was instructed to point at a picture corresponding to a word played over speakers. Two images appeared simultaneously on screen for this task at the same time as audio onset. The parent was instructed to click on a button on the left or right of the screen, depending on the side of the screen (and corresponding picture) that the child pointed at.

Finally, the parent completed a digitalised version of the CDI III (lexis component only)[9] before a final debrief questionnaire asking for any questions or comments relating to the task.

Following completion of the full procedure, parents/ carers were emailed a certificate and Amazon voucher for their participation. Any questions or concerns regarding the task or its data were clarified via email by either party.

[9] This feature did not work correctly online for all participants and follow-up email versions of the CDI were sent, but not consistently completed and returned to the experimenter; thus this data is largely missing for the sample.

Piloting

Various iterations of the Gorilla experiment were trialed on junior researchers and children to ensure the sequence of tasks was optimal, and that the instructions for the parent/ carer were straight-forward and unambiguous. Piloting resulted in the following modifications to the procedure: a hardware eligibility check stage; the optimisation of audio and video for varying bandwidths; restriction of the task for use with the Google Chrome browser; and various modifications to task instructions, to name but a few.

2.3.3 Data Processing and Analysis

Audio responses were transcribed and compared to parental reports of their child's responses. The rate of agreement between the audio transcription and parental report was 92%, providing sufficient evidence to use parental responses for further analysis. The 8% discrepancy in recorded responses was likely due to: the webcam not capturing all responses i.e. a child continued talking when the recording stopped; parents not accurately recording/ not remembering to record all words uttered; parents not acknowledging all responses as valid.

Reponses were cleaned and grouped, then categorised (0-10) by two independent coders (working collaboratively), as previously outlined in Experiment 1. To measure inter-rater reliability, a Cohen's kappa calculation was run. The agreement between raters was 'perfect' with 100% agreement. This high level of agreement indicates that the categories were being applied consistently when different coders categorised responses.

Paradigmatic and syntagmatic coding was compared between raters with a Cohen's kappa calculation. Rater agreement was 'almost perfect' at 96%, $\kappa = 0.82$. This provides strong evidence that the coding performed is objective and likely to elicit the same categorisation across raters, rather than by chance.

2.3.4 Results

Descriptive Statistics

A total of 593 responses were recorded as related or unrelated out of a possible 720 responses. Remaining responses were 'No Responses' (N= 127) which were removed and not included in further analyses. Based on a participant producing up to 3 responses for each of the 10 cue words, an individual participant produced an average of 24.71 responses (SD= 5.80).

Considering first responses only, out of a possible 240 responses (24 participants, each with 10 cue words), 218 responses remained after subtracting 'No responses' (N= 22). The mean response rate was 9.01 (SD= 1.61).

Types of Responses

Figure 7 shows the proportion of all responses by response type. Category 1/ 'Related' responses were most prominent (48%), followed by Category 0/ 'No Responses' (18%), then Category 7/ 'Random' responses (16%).

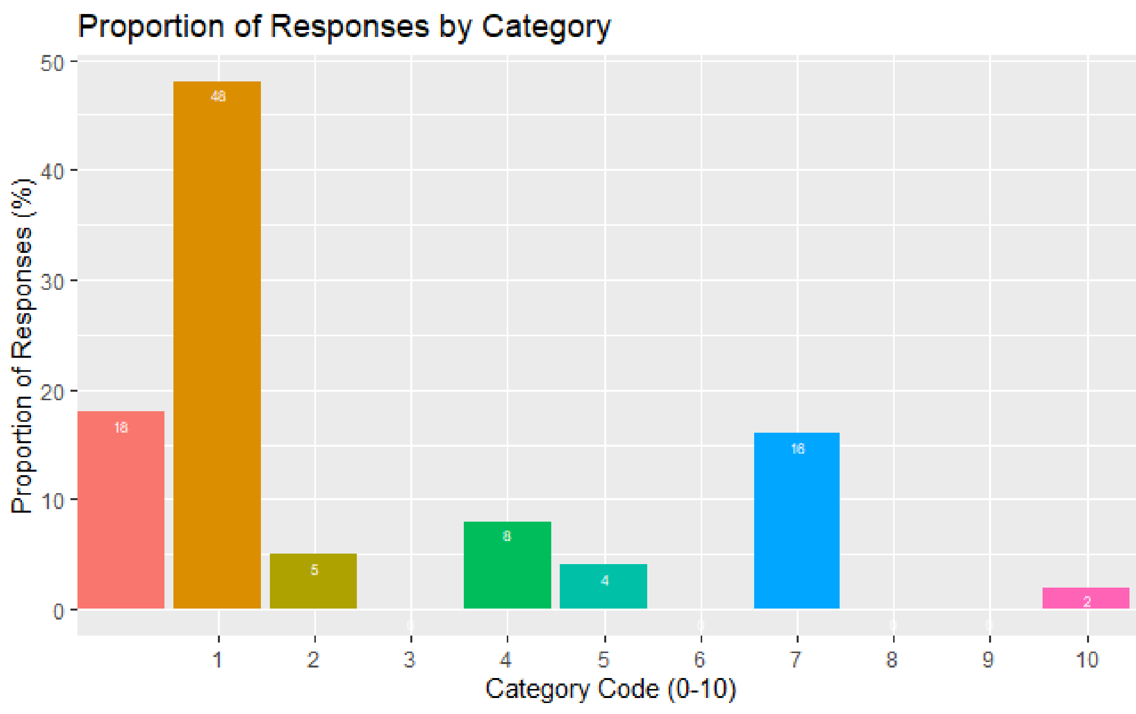

Figure 7*: Experiment 2. Proportion of responses by response category*

Related vs. Unrelated Responses

Organising responses into 'No Responses' (Category 0), a related response (Categories 1, 2, 4, 10) and an unrelated response (Categories 3, 5, 6, 7, 8, 9), **Table 8** illustrates the distribution between the three main response types proportionally and as raw values.

	Num. Responses	**Proportion**
No Response Given	127	18%
Related Response	442	61%
Unrelated Response	151	21%

Table 8*: Experiment 2. Responses by relatedness of response in the online WA task*

This indicates that 36-39-month-olds, when performing a free association task in an online modality, are able to respond appropriately to the task with a high proportion of semantically related responses compared to unrelated, or 'No Response' types.

Turning to first responses only, and since some participants did not provide a response for each of the 10 cue words, a proportional score of related responses was calculated for each participant. This was the number of related responses divided by the total number of responses produced (i.e. minus No Responses). The overall mean proportion of related responses to the task was 0.82 (*SD*= 0.19). We ran a Type III ANOVA on participants' proportion of related responses with gender and age bin as fixed factors. There were no significant differences between the proportion of related responses by gender and age, and no interactions between the variables (ps > .05).

Paradigmatic vs. Syntagmatic Responses

A total of 74% of responses of all the related responses were syntagmatic and 26% were paradigmatic (or both). For first responses, 72% were syntagmatic.

The proportion of paradigmatic responses was calculated for each individual participant.

We ran a Type III ANOVA on participants' proportion of paradigmatic responses for all responses with gender and age bin as fixed factors. There were no significant differences between the proportion of paradigmatic responses by gender and age, and no interactions between the variables (ps > 0.1). The same was true when analysing first responses only (ps > 0.1).

Taking age as a continuous variable, there was a weak negative correlation between the proportion of paradigmatic responses in all responses as age increases, though this was not significant $r(22) = -.35$, $p = .10$. When considering first responses only, the same pattern of findings was true $r(22) = -.20$, $p = .36$. Together this indicates that 3-year-olds predominantly produce related responses that are syntagmatic, and this is not modulated by age (between 36 and 39 months), or gender.

Associative Strength

Responses were cleaned and organised as per Experiment 1. When the same response to a cue word was generated by 2 or more participants, its associative strength was calculated (Nelson et al., 2000). Looking at all responses, a total of 72 responses were given by 2 or more participants with all 20 cue words represented in these word combinations. The full list of cue-response word pairs with their corresponding associative strengths (*M*= 0.08, *Range*= 0.06 to 0.19), organised alphabetically by cue word, can be found in **Appendix K**. Considering first responses only, 25 responses were given by 2 or more participants with 18 of the 20 cue words represented in these word combinations. The list of first response word combinations shared by 2+ children can be found with their corresponding associative strengths (*M*= 0.22, *Range*= 0.17 to 0.42) in **Appendix L**, organised alphabetically by cue word.

We then compared the associative strength between the two sets (i.e. all responses vs. first responses). Associative strength in the two sets differed significantly according to Welch's t-test, $t(95) = -10.01$, $p < 0.001$. The associative strength for word pairs in first responses

(*M*= 0.22, *SD*= 0.07) was greater than associative strength in all responses (*M*= 0.08, *SD*= 0.03). Due to this difference, we focus henceforth on first responses.

The corresponding associative strength for these word pairs was then extracted from adult associative norms (Moss et al., 1996; Nelson et al., 1998) and compared to the child data. Associative strength was averaged across the two adult studies where possible, otherwise an available value from one of the studies was taken to represent the associative strength in adults overall (see **Table 9**).

	Cue	**Response**	**Child FSG**	**Adult FSG**
1	bed	pillow/s	0.17	0.02
2	bed	sleep/ sleeping/ sleep on it	0.17	0.64
3	brush	hair/ we use it for our hair	0.42	0.32
4	cereal	EAT/ eating	0.17	0.03
5	cereal	breakfast	0.25	0.44
6	chair	sit on it/ sit	0.25	0.21
7	door	close the door/ close	0.17	0.03
8	foot	toes	0.25	0.28
9	hair	brush your hair/ brush/ brushing hair in the bath/ brush it/ brush hair	0.25	0.11
10	head	shoulders/ shoulders knees and toes	0.17	0.05
11	key	lock/ lock things with a key/ locking us in	0.25	0.37
12	park	swing/ swings in the park/ we swing/ swings	0.33	0.04
13	sock	wear them on our feet/ pointed to foot/ put on foot/ you put your sock on your feet	0.17	0.17
14	swing	at the park/ park/ park swing	0.17	0.07
15	table	Eat/ eating at the table/ eat/ we eat on it/ we eat pancakes there	0.25	0.03
16	tooth	brush/ brushing your teeth/ brushing/ Brushing teeth/ brush them	0.17	0.12
17	towel	drying off/ dry/ dry hands/ drying/ dry the cat	0.33	0.28

***Table 9**: Experiment 2. Associative strength for word pairs from online child data (first responses) and represented in adult norms*

Seventeen of the 25 associative pairs found in the online free association task were present in the adult associative norms. Eight of the 25 related word-pairs found in children's responses, did not have a value of associated strength in adults, 3 associated word pairs were not characterised in either the Birkbeck or the South Florida norms (though the cue was used); 4 were not used as a cue in the Birkbeck norms, nor documented as an associated response in the South Florida norms; and 1 was not used as a cue in either study. This missing data corresponds to 32%[10] of associated responses found in 3-year-olds that is not reflected in adult associative norms.

The associative strengths of related responses in children from the 8 cue-response pairs not present in adult norms (*M*= 0.21, *Range* = 0.17 – 0.33), were compared to the associative strengths of the 17 cue-response pairs present in children and in adult norms (*M*= 0.23, *Range* = 0.17 – 0.42). There was no significant difference in associative strengths $t(23) = -0.72$, $p = .48$ between cue-response word pairs in children only and for pairs in children and also represented in adult associative norms.

The seventeen word-pairs which were represented in the child and adult data were analysed further.

A *t*-test was run to determine any difference in word associative strength in children and adults. There was no difference in the associative strength between words in the two groups $t(40) = 0.87$, $p = .39$, though the associative strength was slightly higher in children (*M*= 0.22, *Range*= 0.17 - 0.42) compared to adults (*M*= 0.19, *Range*= 0.014 - 0.638). There was no significant correlation between the two groups $r(15) = .23$, $p = .37$, despite a weak positive

[10] This proportion increases to 37% when considering all responses and not just first responses.

tendency. Associative strength seems to be comparable in adults and children and there is some indication that this could correlate positively: word pairs with high associative strength in adults, are also strong in children.

As with Experiment 1, the strongest, imageable noun-noun combinations were identified (*N*= 34) and are displayed in **Table 10**. These represent the strongest, imageable associated word pairs from the online WA task in 36-39-month-olds (first responses in bold, *N*= 9).

	Cue	Response	Associative strength: FSG (P/G)
1	bath	shower/ i have a shower	0.06
2	bath	towel	0.06
3	bath	toys/ Toys (bath toys)/ put toys in the bath/ dinosaur (bath toys)	0.17
4	bed	teddy	0.06
5	**bed**	**pillow/s**	**0.08**
6	bed	toys/ mushroom (soft toy)/ pumpkin (soft toy)	0.14
7	bowl	spoon	0.06
8	brush	combing/ comb	0.06
9	**brush**	**hair/ we use it for our hair**	**0.19**
10	cereal	milk	0.06
11	cereal	porridge	0.06
12	chair	table	0.06
13	door	handle/door handle	0.06
14	finger	hand/ hands/ red ouchie on my hand	0.08
15	**foot**	**toes**	**0.08**
16	foot	socks on it/ socks	0.08
17	foot	shoes/ get you shoes on	0.08
18	**hair**	**brush your hair/ brush/ brushing hair in the bath/ brush it/ brush hair**	**0.14**
19	hand	fingers/ finger	0.06
20	**head**	**shoulders/ shoulders kneesand toes**	**0.06**
21	head	ears/ ears on your head	0.06
22	head	eyes	0.06
23	head	feet	0.06
24	head	hair/ it has hair/ hair on your head/ cradle cap in my hair	0.11
25	key	lock the car/ car	0.06

26	key	lock it up with a door/ lock the door/ door/ a door/	0.11
27	park	slide	0.08
28	**park**	**swing/ swings in the park/ we swing/ swings**	**0.17**
29	**sock**	**wear them on our feet/ pointed to foot/ put on foot/ you put your sock on your feet**	**0.11**
30	**swing**	**at the park/ park/ park swing**	**0.11**
31	teddy	bed/ take them to bed	0.06
32	tooth	toothpaste	0.06
33	**tooth**	**brush/ brushing your teeth/ brushing/ Brushing teeth/ brush them**	**0.14**
34	towel	for my face/ face	0.06

***Table 10**: Experiment 2. Imageable noun-noun, cue-response word pairs in the online WA task (first responses in bold)*

2.3.5 Comparing Experimental Modalities: At Home vs. Online

In the following section, we compare the two experimental modalities: at home (Experiment 1) and online (Experiment 2), whilst acknowledging that Experiment 2 only tests a subset of the stimulus words (*N*= 20) compared to the stimuli used in Experiment 1 (*N*= 100) and therefore differences and similarities might not generalise for all the words tested in Experiment 1.

Descriptive Statistics

There was no difference in response rate between the two experimental modalities $t(172) = .44$, $p = .66$ which indicates that there was no difference in how 3-year-olds approached and responded to the WA task when it was performed by a parent in the home, and when it was demonstrated by a puppet online.

Types of Responses

With regards to response type, the pattern of findings in the online WA task clearly mimics the findings in the at-home version of the task. The online experiment replicates the finding of a large proportion of related responses to a cue word, as found when the WA task was administered in the home. This is especially true for the proportion of Category 1 responses (online- 48%; at-home- 51%), and the high rate of related responses is made more evident when considering the proportion of related responses in first responses only (Experiment 1: *M*= 0.85, *SD*= 0.21; Experiment2: *M*= 0.82, *SD*= 0.19). Category 0 responses (online- 18%; at-home- 20%) were also proportionally comparable.

One difference between modalities is the higher occurrence of Category 7 responses, or seemingly random responses, in the online format (16%)[11] compared to the at-home task (7%).

No effect of gender or age on relatedness of response was found in either modality.

Paradigmatic vs. Syntagmatic Responses

In both modalities, syntagmatic responses occurred more frequently than paradigmatic responses. The rate of paradigmatic responses was not modulated by age or gender.

[11] Three participants did show a very large proportion of Category 7 responses in the online experiment (two responded over half the time, and one almost half of the time with this response type) it is worth noting.

Associative Strength

Considering all related responses in Experiment 1 and 2, 38 word pairs were represented in both experimental modalities as responses given by two or more 3-year-olds for the same cue words. Ten of the word pairs, or 26%, are not represented in adult associative norms. The full list of word pairs found in all responses of both versions of the task can be found in **Appendix M**.

For first responses, 13 word pairs were represented in both experiments (see **Table 11**). One of these word pairs was not represented in adult associative norms (7.69%).

	Cue	**Response**	**Experiment 1 (parent) FSG**	**Experiment 2 (online) FSG**	**FSG_Birkbeck Norms: Moss and Older (1996)**	**FSG_South Florida Norms: Nelson et al. (1998)**
1	bed	sleep	0.27	0.17	not a cue	0.64
2	brush	hair	0.4	0.42	0.20	0.44
3	cereal	eat	0.13	0.17	not documented	0.03
4	chair	sit	0.31	0.25	not a cue	0.21
5	foot	toes	0.14	0.25	0.09	0.47
6	hair	brush	0.13	0.25	0.02	0.21
7	key	lock	0.12	0.25	0.49	0.26
8	park	swing	0.18	0.33	0.02	0.06
9	sock	foot/feet	0.38	0.17	not a cue	0.17
10	swing	park	0.14	0.17	0.07	not a cue
11	table	eat	0.29	0.25	not a cue	0.03
12	teddy	cuddle	0.29	0.25	not a cue	not a cue
13	tooth	brush	0.33	0.17	0.12	0.12

***Table 11**: Word associations replicated in Experiment 1 and 2 as first responses*

The associative strength for related word pairs (in first responses) did not differ in Experiment 1 and 2 $t(11) = 0.02$, $p = .98$, with the average associative strength in the online version (M= 0.24, *Range*= 0.17 – 0.42) equal to that in Experiment 1 (M= 0.24, *Range*= 0.12

– 0.40). Word pairs are associated to an equal degree when the task is administered by a parent/ carer at home, or when done online.

2.3.6 Discussion

Conducting a free association task online with 3-year-olds can be achieved and is a feasible and valid way to deliver this task type with empirical evidence that it generates the same proportion and type of responses as the same task when administered by a parent, in a home setting, for a subset of stimulus words. This points towards a successful version of the task, i.e. online rather than in-person, which has implications for future testing. More diverse populations and age groups might be reached if the task is administered online, including typically and atypically developing children.

Rate of response was comparable in Experiment 1 and 2, and also the type of response. Syntagmatic responses were favoured in both versions of the task.

The almost identical proportion of response types in the online version compared to Experiment 1 suggests that attrition was not reduced by making the task more engaging and visually appealing. There was no evidence of an increased number of related responses, nor a decrease in the rate of Category 5 (i.e. repeating the cue or a previously given response) or Category 6 (i.e. naming something in the immediate environment) responses. Making the task more engaging did not affect response behaviour which suggests the way the task was completed with a parent and online are similar, and reflective of a 3-year-old's true performance at that age. Responses given do not seem to have been by chance.

Parental report of the WAs produced by their children was accurate 92% of the time, suggesting it is an objective and reliable way to record the responses to a free association task in children, making it a comparable modality to the online version of the task.

In terms of the exact associated responses generated to the cue words by two or more children, we saw a replication of thirty-eight word pairs from Experiment 1 (*total*= 432 pairs) and Experiment 2 (*total*= 72 word pairs), when counting all responses given. For first responses only, 13 word pairs appeared in both experiments. There was no difference in the associative strength of these 13 word pairs when the experiment was done at home with a parent or when done online.

The fact that so many word pairs were replicated in the online task suggests these might be more robust and thus more reliable for use in experiments investigating the development of the lexical-semantic system.

While this replication of findings validates the method used, whether the WAs found are representative of other infant populations, is a theoretical question of interest. Thus, Experiment 3 uses the same experimental design as Experiment 2 to pose the question: do bilingual children produce the same WAs as their monolingual counterparts? Exploring patterns of convergence and divergence between monolinguals and bilinguals helps us better understand the interaction of cognition and language development (Bialystok et al., 2003), as well as understand how different infant populations develop.

2.4 Experiment 3: Word Association Study on Bilingual Children, Online

2.4.1 Introduction

Whether the WAs a bilingual child produces show a prevalence of cognates, that is, words sharing semantic meaning and phonology across languages (e.g. *cat* in English and *kat* in Dutch); or translation equivalents, that is, words with the same meaning in both languages, is a question of interest. This might indicate a more robust lexical-semantic organisation for words that share features between languages, whether it is phonological and/or semantic. According to Byers-Heinlein and Werker (2013), a good receptive and productive command of many translation equivalents positively contributes to the richness of the child's semantic system. This is especially true at the initial stages of development and the authors propose that the richness of semantic connections in a bilingual infant surpasses those in a monolingual infant (Byers-Heinlein & Werker, 2013). To investigate this idea further, Experiment 3 explores the rate of translation equivalents and cognates in related responses, using a WA task. If a majority of responses are these types of words, unique to bilinguals, it might support the notion of the special status of these words in the lexical-semantic system.

If bilinguals have a more developed lexical-semantic system compared to monolinguals, that is, more efficient because of more developed connectivity, this might mean bilinguals out-perform monolinguals on a WA task. Evidence for this hypothesis would be taken as a greater proportion of related responses to the cue words, compared to the proportion of random responses. Another outcome supporting a more established semantic network in bilinguals could be responses which more closely resemble mature, adult-like WAs, than those found in Experiment 1 and 2 of this thesis.

We also have to consider the possibility that bilinguals might out-perform monolinguals on a WA task because they have a greater range of words than monolinguals (if counting a bilingual's knowledge of the word in each of their languages, see Águila et al., 2007), rather than because they have a more richly connected lexical-semantic system. It is possible that a bilingual will be able to produce more words when given a cue, compared to a monolingual, however, this might also mean that bilinguals produce greater variability in the responses they give.

If, on the other hand, bilinguals' lexical development closely follows that of monolinguals, we would expect to see a similar pattern of findings in the rate and types of associated responses produced by bilingual 3-year-olds in a WA task.

Research Aims

The aim of Experiment 3 is to observe if bilingual 3-year-olds can produce associated responses to cues in a comparable way to their monolingual counterparts, that is, if they can conceptually navigate a repeated free association task and respond with associated responses to a higher degree of frequency than responding with unrelated words, or not responding at all. A secondary aim is to explore the proportion of associated responses that are known to the participant in both of their languages (i.e. translation equivalents) and, where possible, to see if cognates feature highly in a bilingual's related responses. This might be insightful of lexical-semantic organisation in the bilingual lexicon. A final aim is to compare associated responses in bilingual and monolingual children, to adult associative norms to see if there is convergence or disparity.

Research Questions

1. Do bilingual 3-year-olds perform in the same way to monolingual children in a free association task?
2. Do bilingual participants show a preference for translation equivalents or cognates in their associated responses?
3. Do certain cue-response pairs feature in monolingual, bilingual and adult norms that would be suitable as a bank of stimuli for semantic priming studies?

2.4.2 Method

Participants

Bilingual toddlers were recruited from the UoP BabyLab database and its social media platform pages (*N*= 19: 12 female, 7 male). The mean age of participants was 1155 days (*SD*= 63.03, *Range*= 1082 to 1267 days) or 37.97 months.

Children were excluded from participating if born more than six weeks prematurely; if having an untreated hearing impairment; or if diagnosed with a language or developmental delay.

Children exposed to English and any other language were eligible. On average, a participant was exposed to the English language for 51.73% of the time in an average week (*SD*=19.30;

Range= 27 - 85%[12]). An overview of the languages spoken in addition to English can be seen in **Table 12**.

	Second Language (L2)	**Age in Days**	**LEQ score- Weekly English Exposure (%)**	**Word List**
1	French	1092	45	1
2	French	1127	83	1
3	German	1263	27	1
4	German	1146	-	2
5	German	1267	60	2
6	Greek	1223	72	1
7	Greek	1139	31	1
8	Hungarian	1082	47	2
9	Latin American Spanish	1248	75	2
10	Ndebele	1125	50	1
11	Portugese	1091	85	2
12	Portugese	1086	-	2
13	Russian	1199	28	2
14	Spanish	1177	34	1
15	Spanish	1093	47	1
16	Spanish	1101	50	1
17	Spanish	1198	?	1
18	Spanish	1185	42	2
19	Spanish+Bemba	1111	-	1

***Table 12**: Experiment 3. Participant details in the bilingual WA task*

[12] While this is above the ideal 20% threshold, it does not differ from comparable studies e.g. Sheng (2006): 'averaged 59% of total amount of language use and varied from 30% to 84%'

A participant was deemed sufficiently bilingual if exposed to English and a second language at least 20% of the time in each language, assessed using the LEQ (Cattani et al., 2014). The LEQ is a questionnaire, administered by a researcher, in which a carer of the participant is asked to estimate the number of hours a child is exposed to each of their languages in an average week and the individuals contributing to that exposure. The LEQ calculates a percentage estimate of weekly exposure that a child is exposed to English.

Recruitment of eligible bilingual participants was difficult, and rigor was upheld where possible but there were some participant inclusions not fully adhering to eligibility criteria. For example, not all participants completed the LEQ (*N*= 4) due to technical issues, but email correspondence regarding language background convinced the researcher of the suitability of a participant to participate. Additionally, recruitment of trilinguals was actively avoided, but one participant was included due to knowledge of the L3 amounting to less than 10% of total language exposure, as per Poulin Dubois et al.'s (2018) criteria for bilingual participant inclusion. Finally, since language exposure was measured after participants had completed the task, two participants deemed suitable during email correspondence fell outside the language exposure threshold but were included anyway (see Participant 2 and 11 in **Table 12**).

Participants that were excluded either fell far outside the minimum 20% language exposure threshold (*N*= 2); had a suspected language delay according to a parent (*N*= 1); failed to complete the task fully, despite reminders to complete it (*N*= 3); or did not engage with the task and decided to withdraw (*N*= 3).

Materials

The same materials were used as in Experiment 2.

Procedure

The procedure was the same as Experiment 2, with an additional component after completion of the WA task. This involved a short call over Zoom with the caregiver to i.) complete the LEQ questionnaire and ii.) discuss the associated responses given during the task. The parent was asked all questions outlined in the LEQ and then invited to comment on the accuracy of the percentage of English exposure calculated. Then, parents were asked if the translations of their child's WAs were correct (translated into the second language before the meeting using Google translate). They were next asked to confirm if the translations of the English responses were known to the child. This information was used to calculate the proportion of translation equivalents used by participants in their responses. Finally, parents were asked to determine if the WA in English and its translation bore any phonetic resemblance. This was taken to measure the proportion of a child's responses that were cognates.

In the bilingual version of the WA task, parents were asked to indicate if their child knew the ten target words in English and their second language, or just one of the languages.

2.4.3 Data Processing and Analysis

Reponses were cleaned and grouped as per Experiments 1 and 2, with 10% of responses randomly chosen and categorised by two independent coders working collaboratively.

Agreement of categories between the author and the raters was 'almost perfect' at 95%, $\kappa = 0.90$ according to a Cohen's kappa calculation. The same procedure was followed for coding paradigmatic/ syntagmatic responses with 'substantial' rater agreement at 91%, $\kappa = 0.80$. This provides strong evidence that the coding performed is objective and likely to elicit the same categorisation across raters, rather than by chance.

One difference in the handling of the bilingual data was the response language. Though the experimental language mode was English, some participants responded in their second language, or a combination of English and the second language. All non-English responses were translated into English before categorisation.

2.4.4 Results

Descriptive Statistics

A total of 417 responses were categorised as related or unrelated, out of a possible 570 responses (19*30). The remaining 153 responses were classified as 'No Responses' and were removed and not included in further analyses. An individual participant produced an average of 24.11 responses (*SD*= 6.63; *Range*= 10 – 30)..

Considering first responses only, out of a possible 190 responses (19 participants, each with 10 cue words), 168 responses remained after subtracting 'No responses' (*N*= 22). The mean response rate was 9.33 (*SD*= 1.09).

Types of Responses

Figure 8 shows the proportion of response types in bilingual participants. The highest proportion of responses were related (Category 1: 37%; Category 4: 16%) followed by 'No Responses' (27%).

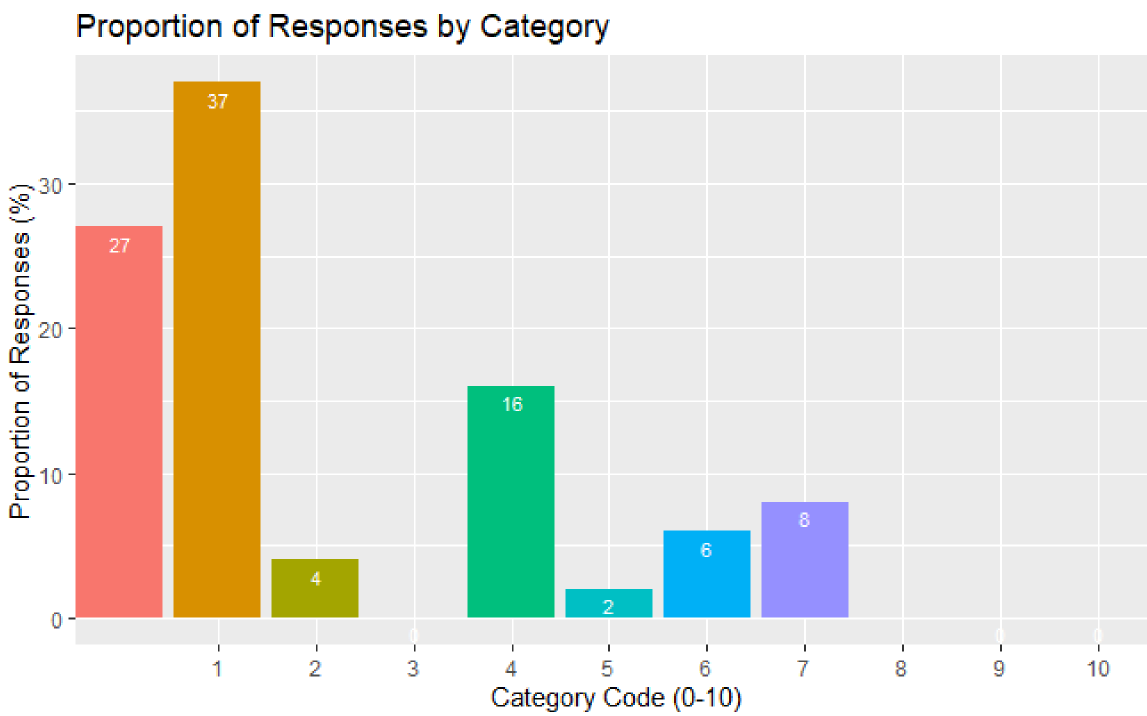

***Figure 8**: Experiment 3. Proportion of responses in bilinguals by response type*

This indicates that bilingual 3-year-olds can perform a free association task and produce related responses (Category 1) more often than not responding (Category 0), or giving another type of response.

Related vs. Unrelated Responses

Related responses were produced by bilingual children more often (57%) compared to unrelated (16%) or 'No Responses' (27%) as outlined in **Table 13**.

	Num. Responses	**Proportion**
No Response	153	27%
Related Response	324	57%
Unrelated Response	93	16%

Table 13*: Experiment 3. Proportion of responses in bilinguals by relatedness of response*

Turning to first responses only, and since some participants did not provide a response for each of the 10 cue words, a proportional score of related responses was calculated for each participant. This was the number of related responses divided by the total number of response (i.e. minus No Responses). The overall mean proportion of related responses to the task was 0.90 (*SD*= 0.19). We ran a Type III ANOVA on participants' proportion of related responses with gender and age as fixed factors. There were no significant differences between the proportion of related responses by gender and age, and no interactions between the variables (*p*s > .50).

Paradigmatic vs. Syntagmatic Responses

A total of 78% of all bilingual 3-year-olds' related responses were syntagmatic and 22% paradigmatic. For first responses, 79.6% were syntagmatic.

The proportion of paradigmatic responses was calculated for each individual participant. We ran a Type III ANOVA on participants’ proportion of paradigmatic responses for all responses with gender and age as fixed factors. There were no significant differences between the proportion of paradigmatic responses by gender and age, and no interactions between the variables (*p*s > 0.1). The same was true when analysing first responses only (*p*s > 0.1).

We next correlated a particpant’s proportion of paradigmatic responses with age (in days) for all responses. There was a weak positive correlation that did not reach significance $r(17) = .04$, $p = .87$. When considering first responses only, there was a positive correlation that was not significant $r(17) = .36$, $p = .14$.

Translation Equivalents and Cognates

The proportion of all responses[13] that were translation equivalents is shown in **Table 14**. A total of 64% of related responses were known to a participant in the first (L1) and second (L2) language, suggesting most bilingual responses are translation equivalents. When a parent did not indicate if their child knew the word in the two languages, it was classified as ‘Incomplete information’.

[13] Due to the small sample size, we looked at all responses given to the cue word for this calculation.

	Known in L1 only	Known in L2 only	Known in L1+L2	Incomplete information
% Responses	8.64	0.62	63.89	26.85

***Table 14**: Experiment 3. Proportion of translation equivalents used by bilinguals in all related responses*

Cognates produced in all responses[14] show a much lower representation in the related responses of bilinguals (see **Table 15**), with only 5% of all responses produced being cognates.

	Cognates	Non-cognates	Incomplete information
% Responses	5.25	92.59	2.16

***Table 15**: Experiment 3. Proportion of cognates used by bilinguals in related responses*

The proportion of cognate responses was calculated per participant and participants were grouped according to whether they spoke a close (German, French) or distant (Hungarian, Russian, Spanish, Portuguese, Greek, Ndbele) language to English. When comparing the proportion of cognate responses between the two groups, there was no significant difference of language distance according to a t-test, $t(17) = .48$, $p = .64$, though children

[14] Due to an already low number of cognates produced in all responses, there would have been even fewer in first responses only and so further analysis was not performed.

who spoke a close language did produce more related responses that were cognates on average (*M*= 0.06, *SD*= 0.04) compared to children speaking a distant L2 (*M*= 0.04, *SD*= 0.06).

Associative Strength

Responses were cleaned and organised as per Experiment 1 and 2. Forward associative strength was calculated (Nelson et al., 2000) for responses produced by 2 or more participants.

Looking at all responses, a total of 51 responses were given by 2 or more participants with all 20 cue words represented in these responses. The full list of cue-response pairs with their corresponding associative strengths (*M*= 0.09, *Range*= 0.07 to 0.20), organised alphabetically by cue word, can be found in **Appendix N**.

Considering first responses only, 16 responses were given by 2 or more participants with 14 of the 20 cue words represented in these word combinations. The list of first response word combinations shared by 2+ children can be found with their corresponding associative strengths (*M*= 0.26, *Range*= 0.20 to 0.38) in **Appendix O**, organised alphabetically by cue word.

We then compared the associative strength between the two sets (i.e. all responses vs. first responses). Associative strength in the two sets differed significantly according to Welch's t-test, $t(65) = -9.16$, $p < 0.001$. The associative strength for word pairs in first responses (*M*= 0.26, *SD*= 0.07) was greater than associative strength in all responses (*M*= 0.09, *SD*= 0.03).

The corresponding associative strength for these word pairs was then extracted from adult associative norms (Moss et al., 1996; Nelson et al., 1998). Associative strength was averaged across the two adult studies where possible, otherwise an available value from one of the studies was taken to represent the associative strength in adults overall. This metric of associative strength in adults was then compared to the child data in Experiment 3 (Bilingual child)[15], Experiment 2 (Monolingual child Online), and Experiment 1 (Monolingual child Parental) (see **Table 16**). Where no value was available 'not documented' has been entered. 'Not a cue' indicates the cue was not used in the study.

[15] We used all responses (*N*= 51) rather than first responses due to the small sample size of first responses (*N*= 16).

	CUE	RESPONSES produced by 2+ participants	Bilingual child_FSG	Monolingual child Online_FSG	Monolingual child Parental_FSG	FSG_Birkbeck Norms: Moss and Older (1996)	FSG_South Florida Norms: Nelson et al. (1998)
1	**bath**	toys/ dinosours (toys in bath)/ happy hippo/ duck/ juguetes (toys)	0.286	0.171	0.130	not documented	not documented
2	**key**	door/ puerta (door)	0.280	0.147	0.210	0.156	0.218
3	**door**	open/ open door	0.267	0.148	0.150	0.146	0.183
4	**brush**	teeth	0.250	not documented	0.160	0.044	0.277
5	**door**	shut/ shut the door/ shutting	0.200	0.111	0.060	0.062	not documented
6	**brush**	hair/ andrew's hair	0.188	0.188	0.190	0.200	0.440
7	**sock**	feet/ foot /pie (foot)	0.185	0.138	0.200	not a cue	0.172
8	**chair**	table/ mesa (table)	0.179	0.063	0.050	not a cue	0.314
9	**tooth**	brush them/ brush your teeth/ brushing your teeth/ cepillarlos (brush	0.172	0.152	0.160	0.115	0.123
12	**hair**	cutting/ cut beard	0.167	not documented	not documented	0.043	0.033
11	**hair**	brush your hair/ to brush	0.167	0.182	0.110	0.021	0.207
14	**towel**	dry	0.167	0.143	0.080	not a cue	0.284
13	**teddy**	cuddle/ cuddeling/ cuddles	0.167	0.100	0.240	not a cue	not a cue
10	**foot**	shoes	0.167	0.083	0.050	0.192	0.108
16	**bed**	toys (soft toys)/ friends (toys)/ dog (Peg, his soft toy)/ teddy	0.148	0.171	not documented	not a cue	not documented
15	**bed**	dormir (sleep)/ sleep	0.148	0.086	0.150	not a cue	0.638
17	**bath**	water/ agua (water)	0.143	not documented	0.100	0.354	0.097
19	**chair**	sit/ sentarse (sit)/ sitting (said in French: assis)	0.143	0.094	0.110	not a cue	0.212
18	**cereal**	milk	0.143	0.061	0.130	not documented	0.204
21	**swing**	outside	0.133	not documented	not documented	not documented	not documented
23	**swing**	swinging	0.133	not documented	not documented	not documented	not documented
22	**swing**	park	0.133	0.174	0.180	0.067	0.101
20	**door**	close/ close door	0.133	0.074	not documented	0.021	0.044
24	**key**	lock	0.120	0.118	0.050	0.489	0.255
25	**hand**	eating	0.118	not documented	not documented	not documented	not documented
26	**hand**	hold hands/ hold hand when crossing the street	0.118	not documented	not documented	0.048	0.012
28	**park**	car/ park your car	0.111	not documented	not documented	0.479	0.108
30	**teddy**	girafa (giraffe)/ it's not a giraffe!	0.111	not documented	not documented	not a cue	not a cue
29	**park**	slide	0.111	0.086	0.160	not documented	not documented
27	**finger**	hand/ hands	0.111	0.074	0.190	0.125	0.268
31	**chair**	eat/ eating	0.107	not documented	0.080	not a cue	not documented
34	**table**	drink/ drinks	0.095	not documented	not documented	not a cue	not documented
35	**table**	eating	0.095	0.143	0.130	not a cue	0.026
33	**table**	breakfast	0.095	0.107	0.070	not a cue	not documented
32	**bath**	play/ jugar (play)	0.095	0.057	not documented	not documented	not documented
37	**key**	going/ going out (to the park)	0.080	not documented	not documented	not documented	not documented
36	**key**	car/ coche (car)	0.080	0.059	0.070	not documented	0.115
39	**head**	hair/ pelo (hair)	0.077	0.114	0.220	0.064	0.186
40	**head**	shoulder/ head, shoulders, knees and toes	0.077	0.057	not documented	0.085	0.021
38	**head**	feet/ foot	0.077	0.057	not documented	0.043 (foot)	0.052
44	**sock**	trousers	0.074	not documented	not documented	not a cue	not documented
41	**park**	play/ jugar (play)	0.074	not documented	0.090	0.021	0.027
43	**sock**	toes	0.074	not documented	0.060	not a cue	not documented
42	**park**	swing/ columpio (swing)	0.074	0.171	0.160	0.021	0.061
45	**bowl**	chocolate/ chocolate (she had just eaten some chcolate from a bowl)	0.071	not documented	not documented	not documented	not documented
47	**bowl**	drink/ special drink	0.071	not documented	not documented	not documented	not documented
46	**bowl**	eat/ eating	0.071	0.067	not documented	not documented	not documented
49	**tooth**	they fall out/ your tooth is falling out	0.069	not documented	not documented	not documented	not documented
48	**tooth**	teeth	0.069	not documented	not documented	0.038	not documented
50	**tooth**	toothbrush	0.069	not documented	0.140	0.019	0.123
51	**tooth**	toothpaste	0.069	0.061	0.140	0.019	0.058

***Table 16**: Experiment 3. Responses to cues by 2+ bilingual children and FSGs compared to monolingual child and adult FSGs*

Thirty-two of the fifty-one associative pairs found in the bilingual free association task were present in adult associative norms. Nineteen cue-response word pairs were not present in the adult data due to 11 associated word pairs not being characterised in either the Birkbeck or the South Florida norms (though the cue was used); 6 were not used as a cue in the Birkbeck norms, nor documented as an associated response in the South Florida norms; and

2 were not used as a cue in either study. This missing data corresponds to 37% of the associated responses given by two or more participants.

Also worth noting is the absence of ten of the cue-response pairs from the data in Experiment 1 and 2, thus a more accurate proportion of missing data from the adult literature might be lower, at 20%. Missing items are displayed in **Table 17**.

	CUE	RESPONSES produced by 2+ participants	Bilingual child_FSG	Monolingual child Online_FSG	Monolingual child Parental_FSG	FSG_Birkbeck Norms: Moss and Older (1996)	FSG_South Florida Norms: Nelson et al. (1998)
1	**bath**	toys/ dinosours (toys in bath)/ happy hippo/ duck/ juguetes (toys)	0.286	0.171	0.130	not documented	not documented
13	**teddy**	cuddle/ cuddeling/ cuddles	0.167	0.100	0.240	not a cue	not a cue
16	**bed**	toys (soft toys)/ friends (toys)/ dog (Peg, his soft toy)/ teddy	0.148	0.171	not documented	not a cue	not documented
29	**park**	slide	0.111	0.086	0.160	not documented	not documented
31	**chair**	eat/ eating	0.107	not documented	0.080	not a cue	not documented
32	**bath**	play/ jugar (play)	0.095	0.057	not documented	not documented	not documented
33	**table**	breakfast	0.095	0.107	0.070	not a cue	not documented
43	**sock**	toes	0.074	not documented	0.060	not a cue	not documented
46	**bowl**	eat/ eating	0.071	0.067	not documented	not documented	not documented

Table 17*: Experiment 3. Bilingual children's related responses, not represented in adult norms*

2.4.5 Comparing Word Associations in Monolingual & Bilingual Children

There was a slightly higher mean response rate in the online, monolingual task (*M*= 24.71, *SD*= 5.80) compared to the online bilingual task (*M*= 24.11, *SD*= 8.15) though the difference was not statistically significant $t(41) = -1.30$, $p = .20$. This indicates that monolingual and bilingual 3-year-olds produce a comparable number of responses in a repeated WA task.

No effect of gender or age was significant in either the monolingual or bilingual group.

There was no difference in the proportion of all related responses in monolingual (*M*= 0.75, *Range* = 0.345 - 1) and bilingual (*M*= 0.79, *Range* = 0.308 - 1) children, $t(41) = .55$, $p = .59$. Therefore, monolingual and bilingual 3-year-olds produce a similar rate of related responses and 'No Responses' in a WA task.

Syntagmatic responses occurred more frequently than paradigmatic responses in monolingual (74%) and bilingual participants (78%). The rate of paradigmatic responses was not modulated by age or gender in either population.

In terms of the exact related responses to cue words that are produced by bilingual and monolingual children, a comparison was done to include word pairs from Experiment 1 as well as Experiment 2. A total of thirty-seven associated word pairs were found across studies. The full list can be found in **Appendix P.** Cue-response pairs were included if present in the bilingual data and at least one of the monolingual studies (Experiment 1 or 2). **Table 18** shows a truncated list of the cue-response pairs in both groups that are highly-imageable, noun-noun combinations most suitable for stimuli selection in infant studies requiring visual stimuli. This list of twenty-three cue-response word pairs best represents the WAs in 3-year-old monolingual and bilingual children that could most reliably be used in studies investigating the development of semantic meaning in the infant lexical-semantic system.

	CUE	RESPONSES produced by 2+ participants	Bilingual child_FSG	Monolingual child Online_FSG	Monolingual child Parental_FSG	FSG_Birkbeck Norms: Moss and Older (1996)	FSG_South Florida Norms: Nelson et al. (1998)
1	**bath**	toys/ dinosours (toys in bath)/ happy hippo/ duck/ juguetes (toys)	0.286	0.171	0.130	not documented	not documented
2	**key**	door/ puerta (door)	0.280	0.147	0.210	0.156	0.218
3	**brush**	teeth	0.250	not documented	0.160	0.044	0.277
4	**brush**	hair/ andrew's hair	0.188	0.188	0.190	0.200	0.440
5	**sock**	feet/ foot /pie (foot)	0.185	0.138	0.200	not a cue	0.172
6	**chair**	table/ mesa (table)	0.179	0.063	0.050	not a cue	0.314
7	**tooth**	brush/ brush your teeth/ brushing your teeth/ cepillarlos (brush your teeth)	0.172	0.152	0.160	0.115	0.123
8	**hair**	brush/ brush your hair	0.167	0.182	0.110	0.021	0.207
9	**foot**	shoes	0.167	0.083	0.050	0.192	0.108
10	**bed**	toys (soft toys)/ friends (toys)/ dog (Peg, his soft toy)/ teddy	0.148	0.171	not documented	not a cue	not documented
11	**bath**	water/ agua (water)	0.143	not documented	0.100	0.354	0.097
12	**cereal**	milk	0.143	0.061	0.130	not documented	0.204
13	**swing**	park	0.133	0.174	0.180	0.067	0.101
14	**park**	slide	0.111	0.086	0.160	not documented	not documented
15	**finger**	hand/ hands	0.111	0.074	0.190	0.125	0.268
16	**key**	car/ coche (car)	0.080	0.059	0.070	not documented	0.115
17	**head**	hair/ pelo (hair)	0.077	0.114	0.220	0.064	0.186
18	**head**	shoulder/ head, shoulders, knees and toes	0.077	0.057	not documented	0.085	0.021
19	**head**	feet/ foot	0.077	0.057	not documented	0.043	0.052
20	**sock**	toes	0.074	not documented	0.060	not a cue	not documented
21	**park**	swing/ columpio (swing)	0.074	0.171	0.160	0.021	0.061
22	**tooth**	toothbrush	0.069	not documented	0.140	0.019	0.123
23	**tooth**	toothpaste	0.069	0.061	0.140	0.019	0.058

***Table 18**: Experiment 3. Imageable, noun-noun cue-response word pairs with high FSG in bilingual and monolingual 3-year-olds*

2.4.6 Discussion

Experiment 3 has demonstrated that a repeated free association task can be administered in bilingual 3-year-olds as successfully as in monolingual 3-year-olds. The responses that bilingual participants produce are highly aligned with their monolingual counterparts in terms of response rate and proportion of related responses. This suggests that both groups perform the task comparably, with no evidence of a bilingual advantage. Though fewer bilingual participants were tested compared to monolingual 3-year-olds, the findings here do not indicate an advanced lexical-semantic system in bilinguals.

Bilinguals do show a higher than chance tendency (62%) to respond with a translation equivalent in a WA task, compared to producing a word known in just one of their languages. This might be due to boosted word retrieval in words sharing meaning across languages. Poulin-Dubois et al. (2018) tested 22-month-old French-English bilingual toddlers in an online lexical retrieval task to test whether, like in the bilingual adult literature, a facilitation effect is present in translation equivalents in the second year of life. The authors found the hypothesised effect of translation equivalents in the dominant and non-dominant languages tested in the bilingual toddlers. The results demonstrated implicit activation of the translation equivalent, when the target word was in the bilingual's other language. Word retrieval was boosted as a consequence. The present study tests productive language in bilinguals and could signal a facilitative retrieval effect of translation equivalents in a WA task. This might indicate stronger networks for translation equivalents in the lexical-semantic system of a bilingual.

Cognates, by contrast were not produced with the same rate of frequency (5%) in the responses of 3-year-olds in a WA task. This might be due to the languages spoken by participants tested in Experiment 3. Many participants spoke L2s which were not typologically similar to English (e.g. Ndeble, Greek, Hungarian) and while speakers of close languages did on average use cognates more than speakers of distant languages, the difference was not significantly different. Future studies might prioritise languages such as Dutch and German, which have a higher number of cognates, compared to languages such as Spanish, which featured heavily in Experiment 3.

Experiment 1, 2 and 3 were compared to generate a list of twenty-three associated noun-noun, imageable word pairs that feature in at least one monolingual study in this chapter and this, bilingual study. Nineteen of the word pairs also feature in adult associative norms. This list is put forward as the most reliable set of stimuli to be used in semantic priming studies in monolingual and bilingual children.

One aim of the next chapter is to test some of the productive WAs found in children in Chapter 1, to validate their use in future studies testing language comprehension.

Limitations

An oversight of the experimental design was the shared language of experimenter (i.e. parent) and the child participant, and how this would influence the response language. Even though the task was designed to be led by videos of puppets, with minimal input required from the parent, this was not the case. Listening to recordings of children completing the task revealed the parent interacted with the child throughout the task, mostly to encourage participation and to give additional responses. However, the language of the parent influenced the child’s performance in the task, such that some children did not perform the task in English. More than one parent said that it created an unnatural environment when the language of instruction was English, as this was not the shared language of child and parent. This serves to highlight the importance of the parent in facilitating the WA task with young children.

As resource availability restricts us from offering the task in another language other than English (as this would involve recording the puppets doing the task in all the languages represented in Experiment 3), it is important that any future iteration of the task addresses this issue in its instructions. Parents should be instructed to complete the task in the most naturally communicative language with their child and not be forced to complete the task in English, as this may affect how the child responds to the task. Translation of the responses into English was done collaboratively with parents, which ensured meaning was accurately captured and this is something that could be continued going forward.

3 Chapter 2: Semantic Priming

3.1 Literature Review

Semantic priming is ubiquitous in a variety of cognitive tasks such as semantic categorisation, naming, and lexical decision which makes it of particular interest to a variety of cognitive disciplines (McNamara, 2005). The effect of semantic priming is thought to occur as part of fundamental mechanisms of retrieval from memory (McNamara, 2005).

Exploring semantic priming helps in understanding aspects of perception and cognition, including but not limited to knowledge representation, discourse comprehension, and word recognition (McNamara, 2005).

The cognitive mechanisms underpinning lexical-semantic processing can be investigated using a semantic priming methodology. The first study of its kind was Meyer and Schvaneveldt's (1971) lexical decision task on twelve teenage participants. The task required participants to decide if two strings of letters, simultaneously presented, constituted two words (e.g. *table-grass*) or not (e.g. *marb-bread*) (McNamara, 2005). Half of the pairs of words were semantically related (e.g. *dog-cat*) and half were unrelated (e.g. *apple-car*). Participants responded 85ms faster on average when the words were semantically related compared to when they were not. This was taken as evidence for the activation of related words during word processing. The term 'semantic priming' originates from this study, and it has been the catalyst for decades' of priming studies in the adult population ever since. While other tasks such as free association and semantic fluency tasks test explicit knowledge of semantic meaning, semantic priming tasks test implicit knowledge, offering a unique opportunity to test hypotheses relating to the primacy and nature of the lexical-semantic system.

The concept of associative relatedness, that is, words which are produced in response to each other (McNamara, 2005), is captured by free association norms (Nelson et al., 1998), while semantic relations are defined as category coordinates (e.g. types of clothing, transport). Some consider words which can be used to define a concept (e.g. apples are red) as having semantic relatedness (e.g. McRae & Boisvert, 1998; Moss et al., 1996), though for words to be 'pure' semantic relations, they cannot be associatively related as well as semantically related (McNamara, 2005). While associatively related words have been well evidenced as priming one another (e.g. Ferrand & New, 2004), a pure semantic priming effect has shown mixed findings. McRae and Boisvert (1998) claim that studies failing to find a semantic priming effect in the absence of associative relatedness, did not use prime-target stimuli with a strong enough relationship.

To compare semantic priming according to relatedness, Lucas (2000) conducted a meta-analysis of twenty-six priming studies. An effect size (Cohen, 1977) of 0.25 was found for pure semantic prime-target relations, compared to 0.49 for associatively related prime-target stimuli. This illustrates that pure semantic priming exists, though to a lesser degree compared to associative priming, and serves to highlight the importance of associative relatedness to boost priming in purely semantically related words, by almost doubling the effect (McNamara, 2005). However, the most reliable effect of semantic priming has been found in words both taxonomically and associatively related (e.g. *chair* - *table*) due to the associative relatedness providing a 'priming boost' (McRae & Boisvert, 1998; Perea & Rosa, 2002).

3.1.1 Evidence of Semantic Priming in Adults using the Visual World Paradigm

Different paradigms have been used to investigate semantic priming in adults, with a common task being the lexical decision task, however, this task requires a participant to make a decision based on their knowledge of language, making it unsuitable for use with very young children. One procedure which is similar to the intermodal preferential looking paradigm used with children, is the Visual World Paradigm (VWP), and so we focus solely on evidence of semantic priming in adults using the VWP, to enable as close a comparison as possible between adult and child findings.

The VWP has long been used for measuring lexical processing including phonological, phonetic, semantic, conceptual processing, and word recognition. The VWP involves a word-to-picture matching task in which participants must identify a picture from a four-picture array that corresponds to an auditory target. Studies have shown that a participant will more readily fixate a visual stimulus that is preceded by a related auditory target, when it is taxonomically or associatively related rather than unrelated (Chow et al., 2017). Exposure to a related word prior to a target word facilitates its processing which is demonstrated by faster and more accurate processing (Nation & Snowling, 1999; Neely, 1991). Cooper (1974) conducted a seminal VWP study in which an auditory target e.g. *Africa* influenced gaze fixations to semantically related objects e.g. *lion* more than unrelated objects did. However, the semantic relationship between words and objects was not systematically controlled. *Africa* and *lion* are associatively related rather than semantically related; therefore, it is unclear from this research alone if associative or semantic features caused the effect. To investigate this line of enquiry further, Huettig and Altmann (2005) tested whether participants would react differently on trials in which the audio target relates to a semantically related or associatively related object in the visual array. Semantically, but not associatively related pictures, attracted overt attention. The semantic overlap was measured using semantic feature norms (Cree & McRae, 2003) with a correlation shown between the amount of overlap and the probability of fixating the semantic distractor picture, thus providing convincing evidence that the overt attention that semantically associated pictures attracted was indeed driven by the degree of semantic overlap from the speech input.

In another experiment controlling for types of semantic relatedness, Mirman and Graziano (2012) tested activation of taxonomically and associatively related words during a word recognition study employing the VWP. Their experimental design was divided into a passive (i.e. participants were instructed to look at the picture corresponding to an auditory target) and active (i.e. to click on the picture corresponding to an auditory target) version. In both instances, fixation patterns were measured. Results showed that both levels of semantic meaning were activated during word comprehension. One salient finding was that this was true even when the task did not require an active response. Interestingly, individual differences were found in the relative activation of taxonomic and thematic/associative relations, which predicted the likelihood of favouring taxonomic relations over thematic/associative relations in a similarity judgement task (Simmons & Estes, 2008). According to the authors, this could indicate that individuals have taxonomic and thematic/associative semantic knowledge to differing strengths which could provide evidence for 'two complementary semantic systems' (Mirman & Graziano, 2012).

Merck et al. (2020) used the VWP procedure on nine adult patients with semantic dementia and fifteen healthy controls to observe the effect of taxonomically and thematically/associatively related distractor pictures on hearing auditory targets. Participants were required to identify the picture that matched the target audio while their eye movements were being tracked. The target picture was presented simultaneously with semantically related and unrelated images. Results indicated that semantic dementia patients and healthy controls showed similar levels of distraction when the semantic distractor pictures were taxonomically related, but semantic dementia patients had a higher sensitivity to thematically/associatively related distractor pictures when trying to fixate the target picture. Their errors predominantly involved taxonomic distractors rather than thematic/associative distractors. The conclusion drawn was that as taxonomic connections deteriorate in patients with semantic dementia, they show a greater reliance on thematic/associative knowledge.

Unlike studies thus far, Huettig and McQueen (2007) demonstrated that the referent of an object is activated even when naming it is not a task requirement and this leads to the subsequent activation of semantically related visual objects. In Experiment 1, Huettig and McQueen (2007) displayed visual objects for three seconds before target onset. Visual objects with the same phonological onset as the auditory target were first fixated, before moving to objects with a semantic or perceptual relationship to the target. Findings indicate that participants accessed the names of objects before hearing the speech input, mirroring the cascaded activation proposed in theories of word recognition (McQueen et al., 2003).

In summary, the VWP is a useful procedure to explore semantic priming in adults. A priming effect is found in all the following cases: when word pairs are not controlled as being associatively or taxonomically related (Cooper, 1974); when words are semantically but not associatively related (Huettig & Altmann, 2005); and when prime-target pairs have taxonomic or thematic/associative relations (Mirman & Graziano, 2012). There is some evidence to suggest that individual differences might modulate the strength of taxonomic versus associative connections (Mirman & Graziano, 2012) and with deteriorating taxonomic knowledge in semantic dementia patients, the lexical-semantic system seems to switch to a reliance on thematic/associative knowledge (Merck et al., 2020).

3.1.2 Evidence of Semantic Priming in Bilingual Adults

As discussed in the introduction, whether a bilingual's two languages are organised in a one- or two-language system remains a question of empirical interest. To investigate this question, visual world studies have tried to uncover whether bilinguals employ language-specific or language-non-specific lexical access, that is, whether competition during lexical access occurs in one language (i.e. the one currently being used for communication or comprehension) or in both languages. In Spivey and Marian's (1999) influential paper, Russian-English bilinguals were played audio sentences in Russian e.g. *Put the stamp below the cross* and presented a four-picture array: a stamp, a marker, and two unrelated images. The English *marker* bears phonological resemblance to the Russian *marku* meaning 'stamp' in English, thus acting as a phonological distractor when translated into the non-target language. When participants heard *marku*, a greater proportion of fixations were made to the picture of a marker than to unrelated distractor pictures. This was taken as evidence of language-non-specific lexical access because the English *marker* influenced eye gaze due to its phonological similarity to the Russian target *marku*. This led the authors to conclude that both languages were activated in parallel (Marian & Spivey, 2003a, 2003b).

However, Weber and Cutler (2004) did not replicate the same finding in their study of Dutch-English participants. A significant difference between the two studies was the types of bilinguals recruited. In Spivey and Marian's (1999) study, bilinguals recruited were living in the US and had relocated from Russia as teenagers. They were using English on a daily basis in an educational context. In comparison, Weber and Cutler (2004) recruited Dutch participants who were living in the Netherlands and who were exposed to English less frequently.

Similarly, Spivey and Marian's (1999) findings were not replicated in Ju and Luce's (2004) study when using authentic Spanish target words on proficient Spanish-English bilingual adults in the US, who had a minimum of 5 years' bilingual language exposure. However, when the targets were manipulated to have English voice onset times, there was evidence of interference. Ju and Luce (2004) propose this might be indicative of parallel activation, but only when the acoustic input and stored representations match.

Canseco-Gonzalez et al.(2010) tested cross-linguistic competitors in the VWP on English-Spanish bilinguals. For example, when instructed to *Click on the beans* in an array in which there was a picture of a moustache, the authors found a weak cross-linguistic effect. The translation of moustache in Spanish is *bigote* which shares the same onset /b/ to the target *beans*. Though, on analysing the effect of age of acquisition, early bilinguals were shown not to exhibit this pattern of cross-linguistic interference.

Taken together, using this paradigm, there is evidence of parallel activation of a bilingual's two languages, but this effect is sensitive to differences in age of acquisition and language exposure.

In addition to addressing the language-specific/non-specific debate, various adult studies have focused on examining lexical processing for two types of words unique in a bilingual's lexicon, namely cognates and translation equivalents (TE). Cognates are words which share the same meaning and have similar phonological features (e.g. *foot* and *voet*- foot in Dutch) whereas TEs are words which represent the same concept, i.e. they are direct translations of one another in any two languages, but they have different phonology (e.g. *car* and *auto*-car in Dutch), therefore sound different.

Concerning the former, a unique pattern of cognate processing compared to non-cognate processing has been found using various experimental procedures (De Groot, 1992; Sanchez-Casas et al., 1992- using a translation task; Lalor & Kirsner, 2001- in a priming paradigm; Lemhöfer & Dijkstra, 2004- in a word recognition task). In adult studies, the effect of cognates on processing is usually documented as facilitatory (see Desmet & Duyck, 2007 for a review) due to the semantic and phonological activation in both languages which is thought to speed up processing. This has been documented for bilingual adult visual (Costa et al., 2000; Dijkstra et al., 1999; 2010, Schwartz et al., 2007) and auditory (Blumenfeld & Marian, 2007; Schelletter, 2002) word recognition studies. Proficiency seems to modulate the effect causing a stronger facilitatory effect for L2 word recognition compared to L1 word recognition (Costa et al., 2000; Kroll & Stewart, 1994) in unbalanced bilinguals.

Lemhöfer et al. (2004) tested double (Dutch and German) and triple (Dutch, German, and English) cognate recognition in a lexical decision task with trilingual participants. Results demonstrated faster reaction times (RTs) in cognates in three languages, over cognates in two languages. From this they highlight that word recognition in a non-native language is affected, not only by a speaker's native language, but equally by another non-native language. Their results were further validated by comparing reaction times (RTs) to a monolingual test group which showed similar RTs for non-cognates, double and triple cognates. They further concluded that this 'multilexical coactivation' of lexical representations in various languages is not limited to a specific number of languages.

Despite these congruent findings, an interference effect has also been documented for recognition of cognates in bilingual adults (Dijkstra et al., 1999; Dijkstra et al., 2010; Schwartz et al., 2007). For example, Dijkstra et al. (1999) found a facilitatory effect of cross-linguistic semantic similarity, but an inhibitory effect of the phonological overlap of cognates. Similarly, processing 'false friends' i.e. words which are similar in form but have different meanings in the two languages (English: *bald*; Spanish: *balde* which means 'bucket' rather than the English meaning of: having no hair) has been shown to impede processing in bilinguals due to an overlap in form but a difference in meaning (Persici et al., 2019). Concerning the investigation of TEs in the bilingual adult literature, facilitation effects have been documented in picture-naming tasks (Christoffels et al., 2006; Costa et al., 2000; Costa et al., 1999; Kroll et al., 2008) and in masked priming tasks (e.g. Basnight-Brown & Altarriba, 2007).

In a VWP experiment testing Spanish-English adult bilinguals, Shook and Marian (2019) found evidence for the activation of the translation of target *duck* (*pato*), because there was increased looking to a picture of a *shovel,* which translates to *pala* in Spanish. Since *pala* is a phonological neighbour of *pato*, the assumption is that the TE of *duck* was activated which in turn activated *pala*/*shovel*.

Priming across languages for TEs (e.g. *dog* – *chien*) is stronger than it is for semantically related words in the two languages (e.g. *dog* - *chat*), but which are not TEs (Schoonbaert et al., 2009). For example, in a bilingual version of the Picture-Word Interference paradigm, a participant is instructed to name a picture in one language which has been super-imposed with a written word in the other language. When the written word is semantically related to the picture, participants take longer to name the picture due to interference caused by semantic activation of the two lexicons (Hermans et al., 1998; Kaushanskaya & Marian, 2007).

In summary, adult bilingual priming studies have revealed parallel activation of both languages in proficient bilinguals (Spivey & Marian, 1999), but not in unbalanced bilinguals (Weber & Cutler, 2004); a sensitivity to acoustic input (Ju & Luce, 2004); and an effect of age (Canseco-Gonzalez et al., 2010) all of which might determine if a bilingual's two languages are activated in parallel or not, and the level of interference resulting from this. Cognates generally facilitate processing (e.g. Lemhöfer & Dijkstra, 2004) with some findings of interference (e.g. Dijkstra et al., 2010; seen also in false friends: Persici et al., 2019). There seems to be a general consensus for a facilitatory effect of TEs (e.g. Basnight-Brown & Altarriba, 2007; Costa et al., 2000; Shook & Marian, 2019), with evidence that priming between TEs is stronger than between cross-linguistic semantically related words (Schoonbaert et al., 2009).

3.1.3 Evidence of Semantic Priming in Monolingual Children

Over the last two decades, developmental psychologists have advanced our understanding of how semantic networks develop, with much of the work focusing on monolingual English-speaking toddlers (Arias-Trejo & Plunkett, 2009, 2013; Mani & Plunkett, 2010).

The Intermodal Preferential Looking Paradigm (IPL, Golinkoff et al., 1987) has been widely used to investigate spoken word recognition in very young children. In this task, a word is spoken or played auditorily while the infant is presented with two visual stimuli on a screen. Preferential looking to the target visual stimulus is thought to index word recognition of the spoken word. The IPL is comparable to the VWP, more commonly used with adults (e.g. Huettig & Altmann 2005; Huettig & McQueen 2007; Kamide et al. 2001), using two rather than four visual stimuli, to reduce the cognitive load for infants. The IPL paradigm captures the tendency a listener has to look at a visual display when a component from the display is referred to, either explicitly or implicitly. It therefore does not necessitate an overt response and results show that a target object is fixated in the absence of any task requirement. Thus, when an infant shows a systematic visual preference for a named target rather than an unnamed distractor, it can be taken as evidence for the mapping of the spoken word to the target visual stimulus (Styles & Plunkett, 2009).

Styles and Plunkett (2009) modified the IPL task to create a primed version of the task to explore infant lexicon organisation. This new method uses the same adult stimulus presentation delivery from priming studies in which two lexical items are presented consecutively. However, instead of a lexical decision task which would be impossible for an infant to complete, it uses looking time so as to track eye movement after auditory presentation of either related or unrelated prime-target word pairs.

Using this adapted method for the first time, Styles and Plunkett (2009) tested seventy-two infants at two age ranges: 18 months and 24 months. A test list comprised twelve trials with half of the word pairs both taxonomically and associatively related according to adult associative norms (Moss et al., 1996), and half unrelated. No stimulus was repeated. When a child's attention was on the screen, this prompted manual initiation of a trial. A blank screen was shown while a carrier phrase (e.g. *Yesterday, I saw a…*) containing the prime word (e.g. *cat*) was played over speakers. There was an inter-stimulus interval (ISI) of 200ms before the target word (e.g. *dog*) followed auditorily. A stimulus onset asynchrony (SOA) of 200ms or 400ms (which were divided equally between the two age groups) began after onset of the target audio, at which point the target and distractor picture appeared on screen and remained for 2500ms. Cameras captured the eye movements of participants with trained coders manually coding a participant's eye movement as left or right every 40ms from when the pictures appear on screen, post-testing. The dependent variable was the proportion of looks to the target (PLT) as a proportion of total time spent looking at the target (T) or distractor (D) (i.e. T/T+D).

Results show that both age groups could correctly identify the named target picture in this primed IPL paradigm. Twenty-four-month-olds systematically preferentially fixated the target on related trials, relative to unrelated trials, showing an effect of priming. The 18-month-old group showed a preference for the target regardless of relatedness, indicating no effect of priming. The authors suggest this could be due to greater variability in the behaviour of 18-month-olds owing to factors such as the loss of data due to younger participants not knowing prime or target words used in the experiment.

In follow-up analyses, CDI scores were correlated with an individual's priming difference score (PLT related – PLT unrelated) to investigate if the size of the lexicon would predict priming magnitude. There was no evidence to suggest vocabulary size, as measured by parental vocabulary reports using the CDI, affects priming. There was no notable effect of SOA on priming. Taken together, findings from this study suggest that a semantic system which encodes relatedness is evident at 24 months but not at 18 months.

Arias-Trejo and Plunkett (2009) used the adapted IPL task (Styles & Plunkett, 2009) to explore if a semantic priming effect can be found in infants as young as 18 and 21 months, thus providing evidence of early connections of word relatedness in the developing lexicon. Word pairs that were both semantically and associatively related were chosen, to increase the probability of finding an effect at such a young age. Borne out of an earlier IPL study that found a priming effect of taxonomically and associatively related word pairs in 24-month-olds but not 18-month-olds (Styles & Plunkett, 2009), this study sought to determine if the locus of the priming effect is the relatedness between words, or if priming is driven by the infant's preference for the target visual stimulus. In Styles and Plunkett (2009), the effect of priming in 24-month-olds was either a result of the word-word relationship, or the word-picture relationship between prime and target as no control condition was included. Arias-Trejo and Plunkett (2009) designed their study such that some trials named the target while others did not. It was hypothesised that a priming effect on trials with unnamed target stimuli would demonstrate an effect of word-picture priming and trials with named targets would show word-word priming and reflect an effect of lexical relatedness.

In Experiment 1, fifty-five 18-month-olds and fifty-six 21-month-olds were tested. A developmental difference between the two groups was found. The 21-month-old group preferentially fixated the target when named and related, but not when preceded by an unrelated prime. The authors propose that the prime word might be interfering with target word processing in the unrelated condition to impede recognition of the target visual stimulus. By contrast, the 18-month-old group preferentially fixated a target when it was named, regardless of whether a trial was related or unrelated. At 18 months, there was no preference for the target when it was unnamed, even when the prime was related to the target.

The authors propose this provides evidence for the existence of semantic-associative links in the lexicons of 21-month-olds and while links between associated words might exist at 18 months, the retrieval of the prime could take longer at this younger age which might be halted at the point of target onset (Arias-Trejo & Plunkett, 2009).

In Experiment 2, Arias-Trejo and Plunkett (2009) simplified the task for 18-month-olds by repeating the prime word as the target word in an attempt to show how relatedness affects target identification (this is known as repetition priming). There were three conditions: Prime-Target (e.g. *boot-boot*), Prime-Look (e.g. *boot-look*), and Neutral-Target (e.g. *juice-boot*). There was an effect of relatedness such that 18-month-olds identified the target on Prime-Target trials through repetition priming, but not on Neutral-Target trials. The authors concluded that this shows a sensitivity to word-word relationships at 18 months.

In a later study, Arias-Trejo and Plunkett (2013) tested associatively related, taxonomically related, or unrelated word pairs on 21- and 24-month-olds in a priming adaptation of the IPL task (Styles & Plunkett, 2009). The authors hypothesised a weaker priming effect for word pairs which are either associatively- or taxonomically- related, rather than both, as tested in previous studies (Arias-Trejo & Plunkett, 2009), due to the absence of the 'priming boost' (Moss et al., 1996; Perea & Rosa, 2002).

Fifty 21-month-olds and fifty-four 24-month-olds were tested. Half of each age group received a word list in which related word pairs had an associative relationship, and half had related word pairs that were taxonomically related. The methodology used closely mirrors Arias-Trejo and Plunkett's (2009) earlier adaptation of the IPL task. There were three instances of each of the three trial types: Prime-Target, Prime-Look, and Neutral-Target, resulting in nine experimental trials. The first trial type consisted of the prime and either an associatively or taxonomically related prime. The second trial type used the same type of word-relatedness but did not name the target. This was to observe for any effect of the prime on the target image in the absence of the target label. The third trial type comprised an unrelated prime and target.

The authors defined associative word pairs as those taken from free association norms (FSG = 39 to 1.1) without categorical relatedness. Taxonomically related word pairs were defined as objects with the same superordinate term (e.g. clothes, *sock-pants*) without associative relatedness, according to adult word association norms (Kiss, 1975; Moss et al., 1996). Targets and distractors shared phonological onset so that infants could not identify the target purely by using the word onset (Fernald et al., 2001). Labels and pictures were used only once in each experimental list. A short SOA of 200ms was used.

There was a significant finding for above chance looking on related (associative and taxonomic), named trials for 21-month-olds and also for unrelated trials. No significant finding for above chance looking was found for related trials that were unnamed in this age group. The 21-month-old group thus did not demonstrate an effect of priming. The 24-

month-old group showed a statistically significant systematic preference to look at the target on related (associative and taxonomic) named trials only, and so demonstrated a priming effect. This was true for associative as well as taxonomic relations. The absence of a priming effect in unnamed related trials indicates that the priming effect documented for named, related trials in this and other studies (Arias-Trejo & Plunkett, 2009) is being driven by the processing of the named target, that is, when the infant hears the prime word followed by the target word, rather than any association between the prime word and the target picture.

Arias-Trejo and Plunkett (2013) interpreted these findings as showing that by 21 months of age, associative and taxonomic relations between words are insufficient to cause a priming effect by themselves through activating the lexical-semantic network. At 21 months, infants show systematic preference for the target when it is preceded by a related or an unrelated prime. However, 24-month-olds do show a sensitivity to the meaning relationships, demonstrated by their preference for the target when preceded by a related prime, compared to an unrelated prime. Thus, by 24 months, infants seem to have developed an interconnected lexical-semantic system, demonstrating links between associatively and taxonomically related words. By contrast, the lexicon at 21 months may not have fully established the same level of structure and so it needs a priming boost in order to reveal the same priming effect found at 24 months. This is supported by priming at 21 months found in a previous study using word pairs that had taxonomic and associative relations (Arias-Trejo & Plunkett, 2009). If 21-month-olds had shown sensitivity to one type of relationship over the other, this might have signalled the primacy of its establishment in the lexical-semantic system. This study in conjunction with others suggests the lexical-semantic system is still emerging between 18-24 months.

While the IPL paradigm is ordinarily used with infants rather than the visual world paradigm (VWP), Chow et al. (2017) recently demonstrated that the VWP can be used with 24-30-month-olds. Like the IPL, the VWP measures eye gaze to visual stimuli, which is taken to reflect language processing. In contrast to the IPL which uses two pictures, the VWP uses four. This offers the chance to simultaneously compare two or more linguistic variables in a single trial, at the same time as including a baseline measure, to explore competing representations (e.g. phonological, semantic) which cannot be achieved with the IPL paradigm.

Chow et al. (2017) conducted two experiments on 24- to 30-month-old toddlers to investigate the time-course of spoken-word recognition in toddlers by way of comparison to an adult-like sensitivity to phonological and semantic information during speech processing (Huettig & McQueen, 2007). Twenty-four critical, target-absent trials (the target word was not visually depicted) and thirteen filler, target-present trials (the target word was visually depicted) constituted an experimental list. Trials involved auditory presentation of the target word (4000ms into a trial) and presentation of four visual stimuli (0-8000ms), which included a phonological distractor, a taxonomic (Experiment 1) or thematic/associative (Experiment 2) distractor, unrelated distractors (2 in target-absent trials, 1 in target-present trials) and the target referent (filler, target-present trials only). Filler trials were to keep the toddlers engaged in the task and not for analysis.

In Experiment 1, twenty-four toddlers aged 24 months and twenty-four toddlers aged 30 months were tested. Results displayed a preference to fixate the taxonomic distractor over the phonological distractor, with the attention to the phonological distractor peaking before the taxonomic distractor. An early phonological preference followed by a later taxonomic preference mirrors findings in the adult literature (Huettig & McQueen, 2007) and shows a quicker extraction of phonological information compared to semantic information.

Experiment 2 tested twenty-four 25-30-month-olds using the same experimental procedure as Experiment 1, though taxonomic distractors were replaced with thematic/associative distractors. The findings were identical in terms of initial phonological then semantic distractor preference, following a similar time-course (i.e. activation of semantic knowledge 750ms after target word onset). Comparing the two experiments, the authors found a quicker and higher peak preference for thematic/associative over taxonomic distractors, which could be interpreted as a preference for this type of semantic relation. However, Chow et al. (2017) note that taxonomic distractors in their experiment did not benefit from the priming boost often found in other experiments combining taxonomic and associative relations (Arias-Trejo & Plunkett, 2009), which may account for the difference.

In addition to behavioural studies, electrophysiological studies using event-related potentials (ERPs) have explored the presence of priming during lexical development. This method enables the observation of language processing without any task demand, making them apt for testing infants. At 24 months, studies show electrical activity similar to that found in adult language processing (Styles & Plunkett, 2009).

In one such electrophysiological study, Torkildsen et al. (2007) presented 24-month-olds with auditory word pairs belonging to the same superordinate category (e.g. *dog* - *horse*) or different categories (e.g. *car* - apple). They found that related pairs showed an early negativity (200-400ms) and unrelated pairs a later negativity (600-800ms). This indicates a lexical priming effect where a prime word semantically related to the target word facilitates lexical-phonological processing (Torkildsen et al., 2007). Furthermore, 24-month-olds displayed the N400 effect for incongruency on unrelated trials compared to related trials. The N400 (Kutas & Hillyard, 1980) is a negative wave of neural activity approximately 400ms after stimulus onset which is sensitive to semantic processing. When the N400 is large in magnitude, it indicates semantic incongruency due to the increased cognitive load required to process meaning between unrelated words, compared to words close in meaning. The N400 effect found at 24 months mirrors findings in the adult literature (Anderson & Holcomb, 1995; Holcomb & Anderson, 1993; Holcomb & Neville, 1990). These findings are in line with behavioural findings which indicate that a semantically organised lexical memory is in place at 24 months. However, Arias-Trejo and Plunkett (2013) highlight that the study did not control for how frequently the taxonomically related words were presented to a participant and Rämä et al. (2013) note that target words were not used in both a related and unrelated condition. This makes it difficult to confidently link the organisation of the lexicon purely to taxonomic principles.

Friedrich and Friederici (2004, 2005) performed EEG studies on 14-month-olds with evidence suggesting a priming effect for related word pairs compared to unrelated word pairs. However, without behavioural support it is unclear if the priming effect is a result of implicit naming (see Mani & Plunkett, 2010) or due to a violation of expectation (Arias-Trejo & Plunkett, 2013).

Rämä et al. (2013) tested 18- and 24-month-old monolingual French-learning infants using an ERP and priming methodology. Taxonomically related and unrelated word pairs were spoken in quick succession on an equal number of trials. The N400-like priming effect was observed in 24-month-olds, and in 18-month-olds who had a high level of word production. This supports previous ERP and behavioural findings and points to taxonomic principles of organisation of the mental lexicon in infants by 24 months and at a younger age, 18 months, but only if scoring highly on a productive vocabulary measure. This last finding indicates that the lexical-semantic system might benefit from advanced productive vocabulary skills to facilitate taxonomic organisation of its structure (Rämä, 2013).

Using the head-turn preference procedure (HPP), Delle Luche et al. (2014) tested twenty-four 18-month-olds to determine if infants at this age are sensitive to the semantic relations between words, without visual stimulus mediation (contrary to the IPL paradigm). Auditory lists of words belonging to the same semantic category (i.e. animals, food) were presented in a first block, followed by a block of lists of words containing mixed semantic categories (i.e. clothes and body parts, or vice-versa). They measured the period of time that a participant attended to the side of the auditory speech stream. The relationship between words was mostly taxonomic though there were some associative relations due to the limited number of words available for testing at this age (however, not between two consecutive items).

The results of the study showed that infants maintained greater attention to lists in which words were semantically related, demonstrated by a pattern of longer listening for related word lists, compared to lists of unrelated words. This provides evidence that at 18 months, meaning was being extracted from the auditory presentation of individual words organised into related or unrelated word lists, and that 18-month-olds were sensitive to the semantic relations between those words. This offers further support for the notion that the emergence of a lexical-semantic structure organised according to some taxonomic principles begins before the second year of life. The findings from this study differ to other studies testing 18-month-olds (Arias-Trejo & Plunkett, 2009, 2013; Styles & Plunkett, 2009), which did not find an effect of priming in this age group. The authors propose the different methodologies between studies might explain the lack of a priming effect in studies in which participants are required to process both auditory and visual information. This might prove to be a more cognitively demanding task for 18-month-olds compared to the relatively simple task of listening during a HPP experiment. What this research does show is that infants are sensitive to the relationships between words in the absence of visual support (e.g. through referents).

To summarise, infant monolingual priming studies show that the lexicon by 24 months seems sufficiently organised to activate words related in meaning to facilitate word recognition, documented in various semantic priming studies at this age (Arias-Trejo & Plunkett, 2013). Younger children show greater variability and while there is some behavioural (Arias-Trejo & Plunkett, 2009; Delle Luche et al., 2014) and electrophysiological (Rämä, 2013) evidence of semantic priming as young as 18 months, the effect is not consistently replicable at 18 or even 21 months. This might be in part due to the nature of the word pairs used e.g. taxonomically or associatively related rather than taxonomically and associatively related, the latter of which provides a 'priming boost', or it may be due to the nature of the methodology. By the end of the second year of life, a lexical-semantic system which encodes words based on semantic and phonological similarity supersedes a system in which words are encoded in isolation (Chow et al., 2017). Only at this point is an adult-like system of connectivity recognisable. What remains unclear is how changes develop to the structure of the semantic system between 18 and 24 months, how infants encode relations between words, the mechanisms involved in these processes, and the degree to which experimental parameters such as word choice and stimulus modality (visual and auditory vs. auditory) can impact the magnitude of a priming effect.

3.1.4 Evidence of Semantic Priming in Bilingual Children

Beyond priming studies testing English-speaking, monolingual infants, few priming studies have been conducted on bilingual infants. In fact, it was as recent as ten years ago that evidence was first documented of the interaction of phonological and lexical access between a bilingual's two languages during development (Von Holzen & Mani, 2012).

Von Holzen and Mani (2012) conducted a phonological priming study testing twenty German-English bilingual toddlers aged 21-43 months and seventeen monolingual toddlers. Prime words were spoken in English and targets in German. Trial types were either phonologically related across languages (e.g. *slide* – *Kleid*, dress), phonologically related through translation (e.g. leg= *Bein*, *Stein*= stone), or unrelated. Results indicated no effect of priming in the monolingual group, but a facilitation effect through phonological priming across languages in the bilingual group, and interference of phonological priming through translation. Together, these effects provide evidence for language non-selective lexical access during bilingual language processing. Evidence from this priming study supports the notion of the interconnectivity of a bilingual's lexica due to the co-activation of a word in a bilingual's two languages.

To further investigate the hypothesis of an interconnected linguistic system in bilinguals, Singh (2013) employed a within-subjects, semantic priming methodology on twenty-one simultaneous, Mandarin Chinese-English 30-month-olds. A key aim was to explore the organisation of a bilingual's lexicon and how words connect within and between languages. To achieve this, within- and cross-language semantic priming was investigated in the dominant and non-dominant language. Four experimental blocks were structured as follows: Mandarin prime–Mandarin target, English prime–English target, Mandarin prime–English target, and English prime–Mandarin target. Three trial types were employed: 1. prime-target (prime word and target word both associatively and semantically related), 2. neutral-target (unrelated prime and target), 3. prime-Ah (prime word followed by the language-neutral exclamation "Ah" to determine if a priming effect can be driven by the visual depiction of a target alone, rather than a lexical representation). The results of the experiment revealed an effect of priming across and within languages, of comparable magnitude. This was true only when prime words were named and in the dominant language. This indicates interconnectivity between languages but also represents a clear effect of language dominance on semantic facilitation, which mirrors findings in the adult literature (Basnight-Brown & Altarriba, 2007). Singh (2013) attributed this to greater word familiarity in the dominant language since more familiar words might have processing privileges compared to less familiar words (Mills et al., 2005). One unexpected finding in this study, replicating a phenomenon in other studies, is the lack of preferential looking to the target on unrelated trials, above chance (e.g. Arias-Trejo & Plunkett, 2009; Mani & Plunkett, 2010). This is highly anomalous as infants have been found to fixate the correct target picture from its auditory label from the age of 14 months (Ballem & Plunkett, 2005).

Singh (2013) suggests it might be modulated by age such that older children have greater word recognition capacities and so might experience interference for target recognition from an unrelated prime which activates its own semantic associates, thus preventing word recognition of the target.

Poulin-Dubois et al. (2018) tested thirty-six 22-month-old French-English simultaneous bilingual children using a touch-screen, word identification task to investigate the effect of cross-language synonyms, known as translation equivalents (TE, e.g., *chien* and *dog*), on lexical processing. Adult studies show a facilitatory effect when processing TEs and so the aim was to observe if this is true in the developing bilingual lexicon, while replicating and extending previous research on young bilinguals (Poulin-Dubois et al., 2013).

Children's knowledge of TEs was assessed using parental report according to the MacArthur Bates Communicative Development Inventory (CDI) in English and French. The task used was the Computerised Comprehension Task (Friend & Keplinger, 2003) which assesses receptive vocabulary in French and English by presenting two images concurrently on a screen and instructing the child which image to select (e.g. *Touch the…*). The child then touches the image they believe refers to the auditory label. Participants were tested in each language on two separate visits. The results clearly indicated that TEs were recognised faster than non-TEs in both the dominant and non-dominant language. This indicates that there is an implicit activation of a target word's translation in this receptive lexical decision task, which facilitates faster word retrieval regardless of language dominance. The finding of a translation facilitation effect for TEs in the developing bilingual lexicon aligns with results in the adult literature. Furthermore, the implicit activation of TEs in young bilinguals provides evidence of an interconnected bilingual lexicon even at 22 months of age and that knowing a word in each language has an impact on lexical access.

To explore how early semantic networks are affected by experience of labels (i.e. monolingual English speaker= *dog*, French-English bilingual= *dog*, *chien*) and concepts (i.e. the concept of a dog is essentially the same in both languages), Jardak and Byers-Heinlein (2019) extended the research begun by Singh (2013). The researchers conducted three priming studies on monolingual (French or English) and French-English bilinguals. Experiment 1 was a semantic priming study testing forty 24-month-old monolingual toddlers (half French-speaking, half English-speaking) in order to replicate a priming effect found in earlier studies (Arias-Trejo & Plunkett, 2009, 2013) at this age. Adapting the method and materials used by Styles and Plunkett (2009), a 12-trial experimental design comprised six related prime-target word pairs, and six unrelated prime-target word pairs. Related trials were both associatively and semantically related to increase the chance of an effect by facilitating a 'priming boost' from both types of relatedness. Prime-target word pairs were counterbalanced to ensure the target followed a related prime and an unrelated prime an equal number of times, between participants. Auditory and visual stimuli were used only once per participant. An ISI of 200ms was used, an SOA of 400ms, and 2500ms free-looking time given per trial. The proportion of target looking for related trials was significantly higher on related trials (d = .32) and above chance for French and English-speaking monolinguals. For unrelated trials, no above chance pattern of looking was reliably found. Vocabulary size did not modulate the priming effect. The authors concluded that the study validated their method while replicating the finding of a priming effect in 24-month-old English-speaking monolingual toddlers.

In Experiment 2, sixteen 24-month-old French-English bilinguals were engaged in a priming study in both of their languages, divided by a short break between language blocks. One aim was to investigate if bilinguals would exhibit an effect of priming at the same age as their monolingual counterparts, and another aim was to observe the effect of language dominance on semantic priming. The procedure replicated Experiment 1, but with four rather than two trial types: dominant-related, dominant-unrelated, non-dominant-related, non-dominant-unrelated. No effect of relatedness was found in bilingual 24-month-olds, that is, the difference in the proportion of looking time to the target on related and unrelated trials was not significant. Due to the unexpected absence of a priming effect and to ascertain if bilinguals and monolinguals might develop word-word relations at a different rate, Experiment 3 sought to find a priming effect in bilinguals at the slightly older age of 30 months, thus exploring the possibility of a different developmental trajectory for monolingual and bilingual toddlers. Experiment 3 tested sixteen 30-month-old French-English bilinguals using a modified version of the materials and method in the previous two experiments. Greater care was taken to ensure no phonological overlap in stimuli or cognates were used during a single trial, with participants completing twenty-four trials presented in two blocks of twelve trials, divided by the carrier phrase language used. Half of the trials used the same language for prime and target (within-language condition), and half used a different language for the prime and target (between-language condition). This generated eight trial types when crossing the conditions of: language of prime, language of target relative to prime, and relatedness. A priming effect was found in bilinguals aged 30 months, demonstrated by a higher proportion of looking to the target on related trials relative to unrelated trials. There was no above chance looking to the target on unrelated trials. This provides evidence that bilinguals have established robust semantic connections within and between languages by 30 months. However, since Jardak and Byers-Heinlein (2019) did not replicate the same findings as the Mandarin-English bilinguals in Singh (2013), the authors propose an effect of language pair as a possible explanation, since the two studies tested different language combinations. Considering all three experiments, Jardak and Byers-Heinlein (2019) present clear evidence for the existence of semantic networks, demonstrated by a priming effect in 24-month-old monolinguals and 30-month-old bilinguals. Whether this is due to a developmental difference between the two populations remains to be seen.

In two experiments comparing TE priming (Experiment 1) and cross-linguistic semantic priming (Experiment 2), Floccia et al. (2020) tested 27-month-old bilinguals, who were speakers of English and one other language. Experiment 1 was concerned with three key foci: 1. excitatory or inhibitory activation between TEs; 2. forward (L1>L2) and backward (L2>L1) priming; 3. language distance (measured according to phonological overlap between TEs, see Floccia et al., 2018 for the metric used). Using the same priming procedure as Arias-Trejo and Plunkett (2009), slight modifications were made to stimuli used per participant based on the language spoken in addition to English (i.e. to avoid phonological overlap within and across languages). Twenty-three simultaneous bilinguals were tested in Experiment 1. Results showed a definite effect of priming, regardless of prime and target language used. An interesting finding diverging from results in the monolingual literature, was above chance looking for unrelated trials (in addition to the commonly reported above chance looking on related trials). No effect of language dominance or language distance was found.

Experiment 2 compared TE priming and cross-language semantic priming with the authors expecting a stronger priming effect in the former based on adult findings. A group of thirty-one 27-month-old simultaneous bilinguals (English and another language) were tested in an experimental design similar to Experiment 1, but with the same stimuli used for all language pairs tested (controlling for phonological overlap after testing). As with Experiment 1, there was an effect of priming in Experiment 2 in both languages, with no difference between TEs and cross-language semantic priming, that is "cat primes chien (French for dog) as much as chat (French for cat)" (Floccia et al., 2020). Participants also looked above chance at the target on unrelated trials. Together, the authors deduced that symmetric priming in 27-month-olds provides evidence for an integrated lexicon early in a bilingual's linguistic development and suggests access to concepts through a parallel system.

To explain the consistent, unexpected finding of target identification on unrelated trials, Floccia et al. (2020) offer two explanations. The first relates to the delayed vocabulary size of a bilingual's languages, compared to a monolingual counterpart (Bialystok et al., 2010) with the suggestion that bilingual 27-month-olds would be better compared to monolinguals of a much younger age, though this does not align with the behaviour of 18-month-old monolinguals, who do not show an effect of priming for unrelated trials (Arias-Trejo & Plunkett, 2009). The alternative explanation points to a distinction being drawn between within-language and between-language interactions. Monolingual 21-month-olds demonstrate interference in unrelated trials, preventing them from looking above chance to the target on unrelated trials (Arias-Trejo & Plunkett, 2009). The same does not seem to be true of between-language activation in bilingual toddlers, which is an area the authors are currently investigating.

Von Holzen et al. (2019) studied the effect of cognates on word recognition in thirty-one German-English bilinguals (18-53 months) and twenty-three German monolinguals (23-47 months). They specifically investigated the amount of overlap in cognates and found that phonological similarity (comparing 0- to 3-feature changes between concrete nouns) facilitated word recognition for the L2 but not L1 in bilingual toddlers. This suggests that learning an L2 does not impact L1 word recognition. The authors believed this to be modulated by the L2 proficiency in this group as the L2 was learned successively to the L1, leading to L2 representations being less established and not robust enough to exert any influence on the L1. Words with 2- and 3-feature changes were not recognised as readily as identical cognates in this study which might indicate interference caused by the simultaneous activation of each representation in the two languages, leading to competition (Von Holzen et al., 2019). On comparing age of acquisition (AoA) of the L2 between participants, those who had acquired the L2 later in life showed a better word recognition performance than those who had acquired the L2 early in life. The authors conclude that the L1 phonological experience affects a bilingual's L2 word recognition.

To summarise, in a bilingual infant population, evidence exists to support language non-selective lexical access during bilingual language processing (Von Holzen & Mani, 2012), including the facilitation of TE word retrieval in the L1 and L2 (Floccia et al., 2020; Poulin-Dubois et al., 2018) with no measurable difference between TE and cross-language semantic priming (Floccia et al., 2020). Cognates, or near-cognates (2-3 feature overlap) facilitate word recognition for the L2 but not L1 in bilingual toddlers, with later AoA improving word recognition of the L2 (Von Holzen et al., 2019). Semantic priming is not found in 24-month-old bilinguals (Jardak & Byers-Heinlein, 2019), but emerges at 30 months with either no (Floccia et al., 2020; Jardak & Byers-Heinlein, 2019) or some (Singh, 2013) effect of language dominance. This effect might be dictated by the language pair tested (Singh, 2013). In both cases, there is ample evidence to suggest an integrated lexicon early in a bilingual's linguistic development with access to concepts through a parallel system. However, contrary to findings in monolinguals at 18, 21 and 24 months, the target in unrelated trials is fixated above chance by 30-month-old (Singh, 2013) and 27-month-old (Floccia et al., 2020) bilinguals, though some paradigms replicated the monolingual finding of no above chance looking (Jardak & Byers-Heinlein, 2019). This might be due to a unique effect of between-language activation in bilingual toddlers and requires further investigation.

Taken together, research to date indicates that an effect of priming occurs later in bilingual toddlers (i.e. 27-30 months) compared to monolingual toddlers (i.e. 24 months), yet the locus of the effect remains unclear as different experimental paradigms restrict direct comparisons to be made, with no difference in the developmental trajectory of these two infant groups possible to determine. Language dominance and AoA do seem to disrupt a replicable priming effect being observed in bilingual toddlers, and the reason underlying a facilitation effect on unrelated trials in bilinguals but not monolinguals remains inconclusive.

3.1.5 Proposed Research and Rationale

Our aim is to investigate emergent representations of word meaning in infants, when different types of relatedness are involved, and when infants' linguistic background varies.

Priming effects of different word pairs are examined using different word combinations (Experiment 5 vs. Experiments 6-9), age groups (Experiments 6, 7, 9 vs. Experiments 5, 8) and languages (Experiment 6 vs. Experiment 7).

The main rationale for this series of experiments was to evaluate the robustness of associative relationships identified in Chapter 1. Our findings in Chapter 1 indicated that a large proportion (Experiment 1= 42%; Experiment 2= 26%) of word pairs strongly associated in the lexicons of 3-year-olds (e.g. *bed - teddy*), were not represented in adult norms. So far researchers have relied on adult norms to select related word pairs for priming studies with infants, which can be problematic (see Chapter 1 for further discussion on this). Therefore, Experiment 5 will use a priming methodology to validate the word associations (WA)s produced by two or more 3-year-olds in both the at-home and online versions of the free association task in Chapter 1. Having compared these with the associations generated from the same cue words in adult WA studies (Moss et al., 1996; Nelson et al., 1998), findings indicate that certain noun-noun associations are clearly more robust than others in both adult and child populations. Testing the WAs found in three-year-olds using a primed IPL task may provide some indication of the primacy of these specific WAs over other, less replicable WAs from adult associative norms.

A second, and necessary, rationale, which was dictated by the UK national lockdown of 2020, was to develop and test online preferential looking data collection, as opposed to face-to-face. Little research has been done to investigate looking behaviour in infants when testing in an online modality, which also means very little literature exists to guide the experimental design of such a study. Consequently, a proof of concept was first required to validate that effects such as increased looking behaviour modulated by linguistic cues, is measurable in children doing the task online. Experiment 4 does this by testing 24-month-olds in an online word recognition task. A key aim is to guide the implementation of an online adaptation of a semantic priming, eye-tracking study for Experiments 6-9.

3.2 Experiment 4: Word Recognition Proof of Concept, Online (at 24 months)

3.2.1 Introduction

Before conducting a semantic priming study in an online modality, it was necessary to design and test a paradigm in which eye movement data could be collected using a participant's webcam in their home context. Two key aspects had to be addressed and validated before collecting data to address key research questions relating to language-mediated looking behaviour in infants. The first was to understand the amount of noise that might be generated by an online procedure, due to factors such as fluctuating infant attention, internet speed and different device types, which may affect timing. The second was to see how much usable data could be collected when trials are presented automatically, that is, not infant-led as would be the case in some lab-settings, such as at the University of Plymouth.

A word recognition task was chosen because of its relatively reliable large effect size and replicability when conducted in a lab setting. In a meta-analysis of typically used methods in language development studies, Bergmann et al. (2018) found an average effect size of d= 1.24 (SE= 0.26) in word recognition studies (N= 6). Thus, choosing this method offered the best chance of developing a proof of concept for an online 'eye-tracking' procedure for paradigms with potentially smaller effect sizes, such as a semantic priming study (e.g. d= .32, Jardak & Byers-Heinlein, 2019) .

In a typical word recognition task, a participant is played an auditory stimulus which is the label of one of two simultaneously presented visual stimuli. In a lab setting, a participant fixates the named visual stimulus for longer than the unnamed visual stimulus, taken as evidence of word recognition. There is evidence of infants able to fixate a target referent as young as 6-9 months (Bergelson & Swingley, 2012) in a look-while-listening procedure[16], with word comprehension and recognition generally available by 12 months (Vihman et al., 2007). Therefore, by testing at the older age of 24 months we had an optimum chance of replicating the same effect in an online modality. If any noise generated by running the experiment online were significant, this might outweigh the effect of a longer proportion of looking time to the target image. Thus, it was hypothesised that the effect of word recognition would be indexed by an increased proportion of looking time to the named visual stimulus, relative to the unnamed stimulus when running the study online. If there is an inherent problem in the methodological design, this effect could be absent.

3.2.2 Method

Pilot studies

Two methods were piloted for their efficacy in collecting eye movement recordings. The first used the platform Zoom and a recording of a visual fixation experiment designed in PsychoPy (adults: *N*=3, infants *N*= 1). The second used Gorilla Experiment Builder (www.gorilla.sc, Anwyl-Irvine et al., 2018) (adults: *N* =2, infants *N* = 4).

In the first pilot study, the experiment was played using the 'Share Screen' functionality in Zoom and the session was recorded. Zoom is limited in its ability to enlarge a specific speaker's video, for example, when trying to capture a closer view of a participant's eye movement. This is because screen space is taken up in this Zoom configuration by three icons: the participant's video, the experimenter's video, and the shared screen, which

[16] An alternative term for an intermodal preferential looking (IPL) procedure.

inevitably reduces the size and resolution of each, resulting in low quality recordings of eye movement. In an attempt to overcome the low-resolution video recording, the experiment was run using the OBS software which allows a video recording to be set as a 'webcam'. When used in combination with video conferencing software such as Zoom, the experimenter's video icon can essentially become the experiment, thus removing the need to 'Share Screen'. This in turn reduces the number of icons on screen from three to two and enables larger, higher resolution video to be captured. However, even with this modification, the resolution from the split screen nature of the recording remained unsatisfactory for some participants, especially those with older devices and slower internet connection speed.

Consequently, this prompted a second approach using the online experimental platform, Gorilla Experiment Builder (www.gorilla.sc, Anwyl-Irvine et al., 2018). Gorilla Experiment Builder can run behavioural studies with the functionality to access a participant's webcam and record, with their consent. However, this option is in Beta and has its limitations. One of which is its inability to simultaneously record a participant and the experiment, or precisely what the participant sees on screen and when. While the timing of stimuli presentation and duration can be precisely programmed into the experiment on Gorilla Experiment Builder, when the experiment is run on a participant's device, some variability may exist because of the differences in devices used, internet browsers, and internet connection speeds, though timing accuracy does seem quite stable (Anwyl-Irvine et al., 2021). Another variable aspect of the webcam recording feature is a potential, yet not guaranteed nor consistent, delay in the command from Gorilla requesting access to a participant's webcam, and the point at which the recording starts. This can be up to 500ms according to one of the developers (personal communication, 23rd May, 2021). Piloting (adults: *N* = 2, infants: *N* = 4) showed only marginal delays (10 - 20ms) so it was decided that in spite of this unpredictable variable, the expected magnitude in difference of looking duration to the target versus the distractor, would not be affected by the noise of the potential delay in recording. Additionally, a design feature was added to the experimental design (see below) to note which trials began recording before visual stimulus onset, and which did not.

Piloting the experiment on adults and infants was crucial to devise satisfactory solutions to these limitations and to decide how to best minimise variability in executing the experiment online. Email correspondence with parents and viewing the data that was successfully generated allowed us to make the following design alterations:

- Participants were restricted to using a laptop or computer. Those without such a device were deemed ineligible. This criterion was set to ensure visual stimulus presentation would be as large and as predictably positioned as possible. Gorilla Experiment Builder's default positioning of two adjacent images is to space them as far apart, to each edge of a device's screen as possible.
- The experiment was programmed to only run on the web browser Google Chrome as there were some upload and display issues with other browsers.
- A calibration phase was added at the start of the experiment to ensure a participant's screen was not working in a 'flipped' mode, and to validate that when an image was presented on the right only, the child looked to the right.
- A short beep of 100ms was added to coincide with the visual stimulus onset. In the absence of seeing when the pictures appeared on screen in a participant's webcam

recording, the beep was a feature to enable the coder to have a reference point when manually coding eye movement. Each trial was checked for the presence of the beep during analysis, to ensure that the webcam recording started ahead of the images being presented on screen.

- Trials were divided into two blocks and separated using a short video to maintain attention. As the experiment could not be driven by the child's attention to the screen on every trial, the short video was a way of re-focusing the child in the event that they had lost interest. Piloting showed inattention to be very infrequent.
- 500ms was added to each trial, resulting in the images remaining on screen for 5500ms (compared to 5000ms in a lab-based experiment). This was to compensate for any delay in the webcam recording or clipping towards the end of recording.

Power Analysis and Sample Size Calculations

Based on the analysis of underpowered findings in infant looking preference studies (Oakes, 2017), a minimum sample size of 20 was chosen.

Participants

Participants were recruited through the UoP BabyLab database and Facebook page. Twenty monolingual, English infants (13 boys, 7 girls) participated. The mean age of participants was 24 months 3 days (*Range* 23 months 3 days - 25 months 28 days). Participants were excluded if they spoke more than one language, were born more than six weeks prematurely, or had a diagnosed language or developmental delay.

Materials

A total of twenty-four target words (e.g. *bed*, *key*) were selected which were familiar, common, highly-imageable nouns known by at least 60% of English monolingual 18-month-olds according to the Oxford Communicative Development Inventory (Hamilton et al., 2000) and the UK CDI (UK-CDI Database, 2016) (see **Table 19** for the list and exact percentages). All words were monosyllabic.

Target	**% known at 18 months OCDI**	**% known at 18 months UKCDI**
bed	85	97
bird	88	88
book	95	98
bowl	58	77
box	48	63
bread	72	77
car	95	97
chair	80	95
cheese	63	78
cot	70	68
dog	98	99
duck	90	86
fish	75	81
foot	70	92
frog	56	68
hair	91	86
key	74	81
pig	77	82
plane	81	72
shoe	99	97
spoon	77	76
swing	64	68
train	66	81
tree	69	78

Table 19: *Experiment 4. Percentage of 18-month-olds with knowledge of the stimuli words used in the Word Recognition Proof of Concept Study*

Auditory stimuli were recorded individually by a female adult with a neutral south-west British accent. The carrier word *Look!* was also recorded separately. Visual stimuli were colour photographs from the internet, cut out from their background and placed centrally on a light grey background to reduce brightness on the screen. Two versions of each image were created: one for presentation on the left of the screen, and one for the right. Animate objects were positioned to face towards the centre of the screen.

Target words were organised into word pairs in which there was no semantic or phonological overlap. The twelve pairs formed one block. In each pair, one word acted as the target and the other as a distractor. The distractor words then became the targets in a second block of trials, and these were paired with a different word that had acted as a target in the first block.

Procedure

Following ethical approval of the experiment by the UoP ethics panel, parents were invited to participate in the study through the Plymouth BabyLab database and through adverts posted to the BabyLab's social media accounts. When a parent expressed interest, further communication moved to email. A participant information sheet was issued and the technical requirements for the online study were reiterated through email communication. A day and time were agreed, on which to complete the study. On the appointed day, an email with instructions for the study were sent to the parent and a unique link to the experiment was activated on the Gorilla Experiment Builder website. By using a unique link, it meant participants could leave the experiment and return to it later, continuing where they left off. The reason behind establishing a day and time to do the online experiment was to ensure a researcher could be available for any questions or support required while participants did the task. Parents were instructed to begin the procedure without their child present, to minimise the time a child would need to stay engaged. It was made clear that the parent would be instructed when to prepare their child for the task.

When clicking on the Gorilla Experiment Builder weblink, an overview of the study was displayed, including the eligibility criteria for participation. The next screen was an eligibility questionnaire, to ensure participants were the right age; were not born more than six weeks prematurely; were exposed only to English; and did not have a language or developmental delay. At this point, a participant could be rejected in which case the parent would see an ineligibility screen and be asked to email the Plymouth BabyLab if they believed this to be incorrect, or if they wanted to find out about other studies running that their child might be eligible for.

If eligible, a participant had to consent to the study by completing an online questionnaire which detailed the procedure, the data collected and the right to withdraw. A copy of the consent form can be found in **Appendix Q**.

Following this, participants progressed to a technical eligibility check so they could test their sound and webcam before the experiment, and to grant Gorilla access to webcam recording. A Gorilla pop-up appeared in the web browser asking for consent to access the webcam, at which point a parent could refuse access if they did not agree to their data being accessed in this way. Furthermore, the recording test established the audio and video recording capabilities of a participant's device and it also allowed parents to playback the recording to fully understand the footage that would be recorded of their child when the experiment began. Throughout the procedure, an 'Exit' button was made available in the bottom left-hand corner of the screen in case a participant chose to withdraw from the study. There was explicit mention in the instructional email that a participant should click on this Exit button if they wanted to withdraw and to request, by email, the withdrawal of any data collected on their child up to that point if they desired, without any explanation for their decision.

Demographic information was collected in a series of short online questionnaires before the experiment proper.

The experimental procedure began by instructing the parents to place their child on their lap, with their device's webcam focused on their child's eyes. This was facilitated with set-up instructions and test recordings. When a parent deemed the position of their child satisfactory, they could begin the task.

The experiment was preceded by four calibration trials in which the word *Look* was followed by the word *biscuit* and an image of a biscuit appeared on the left-hand side of the screen. This process was repeated on the left side with the word *monkey* and a corresponding image. The two words were then repeated with the same images now appearing on the right-hand side of the screen. Neither of the words were used as primes, targets, or distractors on critical trials. The calibration phase established a baseline for the participant's individual looking pattern and to validate the image was presented on the correct side and not in a 'flipped screen' mode.

The parent controlled the start of the priming task by clicking on a button. This began with a short animation lasting 5000ms which displayed abstract black and white images to maintain the child's interest, as it was unclear how attentive a child would be using an online paradigm. Then, the automatic presentation of trials began and did not stop in their delivery until all trials had been presented, which lasted for about three minutes.

Each experimental trial began with a smiley fixation point in the centre of the screen for 1000ms to focus the child's attention to the middle of the screen. This was replaced by two visual stimuli, positioned on the left and right sides of the screen for 5500ms. In an equivalent lab-based study, a trial would last 5000ms but an additional 500ms was added in case of clipping at the end of the recording. The auditory stimuli began with a beep for 100ms to coincide with visual stimulus onset, necessary for analysis. This preceded a silence and the carrier *Look* before target word onset at 2500ms. Each trial was thus divided into a 2500ms pre-naming and 2500ms (+500ms) post-naming window.

After 12 trials, a 5000ms attention-getting video played to maintain the child's attention before a second block of 12 trials resumed. The video also separated the two blocks in which visual stimuli acted as targets in one block and distractor pictures in the other. The order of blocks was counterbalanced. The side of the target visual stimulus was counterbalanced. The experiment ended with a 5000ms 'reward' video.

To complete the procedure, the parent marked a list of target words as known or unknown to the child, before a final debrief screen, inviting any questions or comments and a chance to mark whether any technical difficulties had been experienced during the tasks.

After completion of the full procedure, a participant's data were downloaded, and the calibration trials checked to confirm audio and video recording was satisfactory. Questionnaires were reviewed to see if the participant had experienced any technical issues or if further information relating to their responses in the questionnaires was required. A final email was sent, requesting clarification pertaining to comments in the questionnaires (where necessary) and issuing a certificate and £5 Amazon voucher to acknowledge participation. The final email also included a short debrief of the study's aims and application and invited the participant to ask questions if necessary.

3.2.3 Results

Technical Specifications

Devices were restricted to laptops or computers, yet this can still mean a range of screen sizes. Gorilla records the device type used by a participant, including its screen size. The average viewpoint size on screens used was 1432x742 with parents classifying the mean quality of audio as 5 (Very good, on a scale of 1 to 5). Most participants were using the latest operating systems for their devices, and the latest version of Chrome. The full range of devices and technical specifications can be seen in **Table 20**.

Particpant	Participant Device Type	Participant Device	Participant OS	Participant Browser	Participant Monitor Size	Participant Viewport width	Participant Viewport height	Audio Quality (1-5)
0hl65s6s	computer	Desktop or Laptop	Windows 10	Chrome 87.0.4280.88	1536x864	1536	754	Clear enough- 4
26o5cpdq	computer	Desktop or Laptop	Mac OS 10.14.5	Chrome 86.0.4240.193	1440x900	1440	821	Very clear- 5
4k5u5xs8	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.183	1366x768	1349	625	Very clear
4uoxz04w	computer	Desktop or Laptop	Windows 10	Chrome 85.0.4183.121	1536x864	1438	704	NA
ksry8lfl	computer	Desktop or Laptop	Mac OS 10.13.6	Chrome 86.0.4240.80	1680x1050	1680	971	NA
reuryabw	computer	Desktop or Laptop	Mac OS 10.14.0	Chrome 87.0.4280.67	1440x900	1050	752	Clear enough
22vg0z4l	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.75	1536x864	1519	722	Clear enough
5ym93g5p	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.75	1366x768	1366	625	Very clear
cpqtjso9	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.193	1366x768	1349	625	Very clear
qiamsumn	computer	Desktop or Laptop	Windows 10	Chrome 67.0.3396.99	1366x768	1349	662	Very clear
uibpbg89	computer	Desktop or Laptop	Windows 7	Chrome 86.0.4240.111	1920x1080	1920	1009	Very clear
ye42nool	computer	Desktop or Laptop	Windows 10	Chrome 87.0.4280.66	1280x800	1280	689	Very clear
lfioaben	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.198	1280x720	1280	610	Very clear
odcoevc8	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.198	1920x1080	1920	969	Very clear
plrudr83	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.183	1368x912	1368	783	Very clear
pmwyldgf	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.198	1680x1050	1680	939	Clear enough
xtu5nbo8	computer	Desktop or Laptop	Windows 7	Chrome 86.0.4240.193	1536x864	1198	630	Clear enough
heojqujc	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.111	1366x768	1349	657	Very clear
iiahp11j	computer	Desktop or Laptop	Mac OS 10.15.7	Chrome 86.0.4240.193	1440x900	1200	667	Very clear
s5xh3nt0	computer	Desktop or Laptop	Windows 10	Chrome 86.0.4240.111	1366x768	1366	625	Very clear

Table 20: *Experiment 4. Overview of Device Types Used in the Online Word Recognition study on 24-month-olds*

Data Processing and Analysis

Using a bespoke online encoder developed by the UoP School of Psychology technical team, videos of individual trials were uploaded and automatically split into 50ms frames. For each frame, the primary coder, blind to the visual and auditory stimuli presented, assessed the digital videos off-line frame by frame, manually marking the position of the participant's eye position as left, right, away, or indeterminate by using four corresponding keys on the keyboard. This information was saved in .csv format and later downloaded for analysis.

A second, skilled coder manually coded ten per cent of the full dataset. Inter-rater reliability agreement between coders was 87% and according to a Cohen's Kappa calculation, was moderately reliable, $\kappa=0.47$. On further inspection of the discrepancy between the two coders, out of the total 13% disagreement, 6% was specific to whether a gaze was indeterminate or not, meaning the gaze was still on screen, but unclear where exactly. This might explain the lower-than-expected reliability measure, which will feed into the coding of data for the semantic priming experiments (i.e. in the form of training on how to determine indeterminate looks).

Trials were excluded from analysis if a child did not fixate for a minimum of 750ms, somewhere on the screen (left, right or indeterminate), or if the child did not know the target word based on a parent's report of their child's word knowledge. The latter ensured that an infant was evaluated only on their understanding of known words.

The raw .csv files, generated by coding eye movements using the UoP Encoder, were uploaded in R Studio version 1.4.1717 for all further analyses. The R tidyverse and dplyr packages were used.

Descriptive Statistics

When aggregating all participants' looking time by condition, on average, participants spent 82% of the time looking at either the left or right side of the screen, with an additional 16% of the time looking at the screen but at an indeterminate point on the screen (i.e. neither clearly left nor right). This time also accounts for saccades between the left and right side of the screen. Finally, 2% of looks per participant were looks away from the screen.

Out of a possible 480 trials (a maximum of 24 trials for each of the 20 participants), a total of 459 trials were included for analysis. Reasons for exclusion were entirely due to the target word not being known to the child (21 trials or 4.38% of trials), which was measured by parental report. No trials were excluded due to inattentiveness, measured as <750ms spent looking at the screen per trial. The average number of valid trials per participant was 22.95 (SD= 1.4). In summary, 24-month-old infants were very engaged in an online looking task when administered in the home. By way of comparison, in a meta-analysis looking at looking while listening studies, among other methods, Bergmann et al. (2018) used a linear mixed effects model to predict an exclusion rate of 30% of data for this task type, including minimum looking time criteria. In a more recent study, Byers-Heinlein et al. (2021) saw an exclusion rate of 5.07% for equipment failure, parental interference, and fussiness, in addition to 23.03% data loss due to infants not attending to objects during the specified window of analysis.

There was no effect of gender on response rate $t(18) = .44$, $p = .66$ though girls gave slightly more responses on average (*M*= 23.14, *SD*= 1.21) compared to boys (*M*= 22.85, *SD*= 1.52).

Proportion of Looking Time to the Target

A participant's looks were aggregated by condition (i.e. target, distractor, away, indeterminate) and the proportion of time spent looking at the target compared to the distractor was calculated for the pre- and post- naming windows.

The pre-naming window of analysis was set at 200ms – 2500ms which allows for an initial 200ms shift eye gaze (Fernald et al., 1998, 2001) from an attention-getter to one of the pictures, followed by 2300ms of free-looking. The post-naming window was set at 2700ms - 5000ms to allow for initial processing of the onset of the audio, followed by the same amount of free-looking time (equivalent to 46 frames of 50ms per trial, per participant).

The proportion of looking time (PLT) towards the target visual stimulus, relative to the distractor stimulus, was calculated as the dependent variable for the pre-naming and post-naming windows, per trial:

PLT to target/ (PLT to target + PLT to distractor)

A two-tailed, paired t-test was run on the PLT in the pre-naming and post-naming windows of analyses. Twenty-four-month-olds looked at the target longer in the post-naming window (*M*= 0.62, *SD*= 0.07) compared to the pre-naming window (*M*= 0.50, *SD*= 0.07) (see

Figure 9, with the white square indicating the mean). The difference between looking behaviour in these two periods was significant with a very large effect size, $t(19) = 17.22$, $p < .0001$, $d= 1.61$. This indicates that participants looked longer at the target picture after it had been named, indexing word recognition.

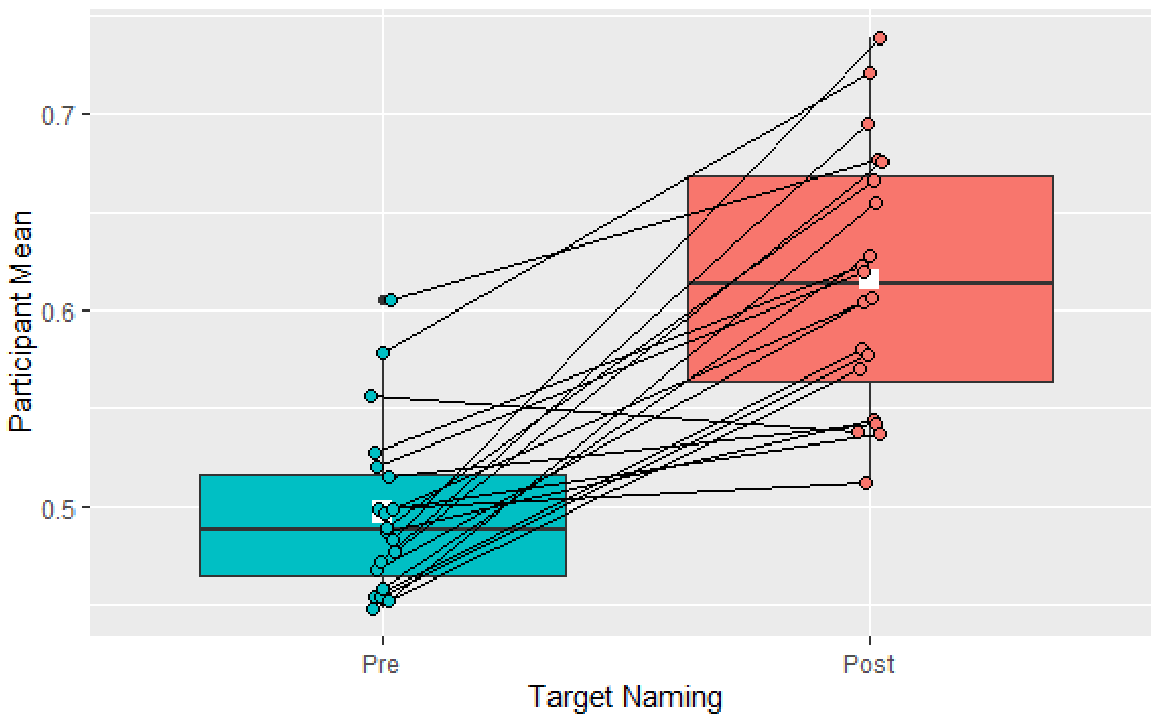

***Figure 9**: Experiment 4. Proportion of Looking Time Pre- and Post-naming during an Online Word Recognition study with 24-month-olds*

3.2.4 Discussion

A simple word recognition experiment was run using the online experimental platform Gorilla Experiment Builder (www.gorilla.sc, Anwyl-Irvine et al., 2018) as a proof of concept to test the feasibility of running online preferential looking experiments with infants. The results from Experiment 4 indicate that with some modifications to lab-based procedures, an online version of an infant methodology can indeed be run successfully. Experiment 4 adapts the Inter-modal Preferential Looking (IPL) task into an online modality, providing a validation of the general testing paradigm and procedure for the experiments that will follow. As far as we are aware, this procedure is the first of its kind to be conducted completely online with young children using Gorilla Experiment Builder.

The results clearly showed that infants aged 24 months looked at a picture on-screen longer when the picture was named, compared to a picture that was unnamed. This is an expected outcome which indexes word recognition in children and replicates previous lab-based findings (Vihman et al., 2007). The novelty lies in the fact that the 24-month-olds were performing the task online, in their own homes and using their own devices. Participants were not overly distracted by their surroundings, nor were there significant issues with differing device types and internet speeds. Noise inevitably exists due to these factors, but when running experiments with a large, expected effect size, such as the word recognition task reported here (d= 1.61), the effect should be robust enough to withstand this inherent noise. Compared to lab-based studies, the effect size found in Experiment 4 is in fact larger in magnitude (e.g. Bergmann et al., 2018, found an average effect size of d= 1.24 in a meta-analysis) which is a promising finding for the studies that follow.

Interestingly, participants remained engaged throughout the procedure despite the fact that trials were not infant-led, that is, they ran automatically without pause. This is a very different approach to many lab-based studies in which the start of every trial is initiated by the experimenter when the infant's attention is focused on the computer screen (e.g. Arias-Trejo & Plunkett, 2009; Chow et al., 2016; Floccia et al., 2020; Singh, 2013; Styles & Plunkett, 2009). Automatic presentation of trials was borne out of necessity while using Gorilla Experiment Builder to administer the task online. According to the findings of this study, running the experiment without pause does not seem to have had a negative impact on a child's ability to perform the task. This may be thanks to the features integrated into the design of the experiment such as fixation points and video rewards at the start, middle and end of the procedure.

Participants also remained engaged in the face of a twenty-four-trial experimental design, which is double the number of trials commonly used in infant studies at this age (Arias-Trejo & Plunkett, 2009; 2013; Jardak & Byers-Heinlein, 2019). This is encouraging support for future studies as using this number of trials will help with the power of future studies in the case of potential data loss occurring, as mentioned above (i.e. distraction, technical issues etc.).

With regards to this particular study, there was very little attrition or data loss (<5%) compared to some lab-based studies, which can lose up to 30% according to a meta-analysis performed by Bergmann et al. (2018). This might be due to a participant feeling more relaxed in their home environment compared to a lab environment. By informally looking at the experimental videos, children did not seek out contact as frequently with a parent by turning around, as they do in the lab. Similarly, the child might have felt more at ease on a parent's lap, rather than in an unfamiliar car seat/ booth in a lab. These hypotheses are supported by the data; there was a high proportion of looks on-screen to the left or right (82%) versus off-screen (2%). This amount is likely to be larger considering looks on-screen but to an indeterminate location (16%) may have actually been looks left or right. One explanation might be the manual coding of eye movement which minimised data loss, compared to lab-based studies in which the eye-tracker losing signal leads to data loss.

Taken together, these findings provide encouraging support that other infant paradigms might be suited for adaptation to online testing. What remains to be seen is whether paradigms with smaller effect sizes, such as the effect sizes found in priming studies (e.g. *d*= .32, Jardak & Byers-Heinlein, 2019), can be evidenced using the same online procedure. Findings from this study indicate that infants can complete twice as many trials as other, comparable priming studies specify, while still maintaining attention. Using an increased number of trials will help increase power for testing such hypotheses.

3.3 Experiment 5: Semantic Priming Validation of Child WAs, Online (at 3 years)

3.3.1 Introduction

Having validated that a word recognition experiment can be run online for participants as young as 24 months in Experiment 4, Experiment 5 addresses the main research question of whether the unique child word associations found in Experiment 1 and 2 will demonstrate a measurable difference in a receptive task such as a semantic priming study. This is achieved by comparing the magnitude of any semantic priming effect in child-specific associations to adult-specific associations and to associations found in both adults and children. The PLT (proportion of looking time) to target vs. distractor images on related trials is measured to test our predictions. Based on the findings in the WA (word association) task, it is hypothesised that adult associations not represented in the child WA data may not show any semantic priming effect, or the effect may be smaller in magnitude compared to the word pairs found in children's associations. In contrast, child-specific associations and those represented in both child and adult WA data are expected to show a consistent priming effect. If the associative forward strength (FSG) values from Experiment 1 and 2 are a reliable metric for how strongly words are connected in the lexical-semantic system of 3-year-olds, we might expect the prime-target pairs found in both adults and children to demonstrate a stronger effect of semantic priming as these had a higher FSG than the child-specific word pairs. A stronger effect of priming in child-specific word pairs might indicate stronger receptive knowledge of these than productive knowledge (as measured in the WA task) or simply that a child's attention will be maintained for longer for the unique child WAs since their experience of the world at the age of three is represented in these word pairings.

In Experiment 4, we demonstrated that an adapted version of the IPL paradigm could be administered remotely using an online platform and with an automatic, rather than child-led, delivery of trials. The IPL task in Experiment 4 replicated a significant finding of word recognition that has previously been found in lab-based studies (Vihman et al., 2007) however, other experimental paradigms have not yet been tested online, which warrants further investigation.

Therefore, Experiment 5 aims to test the validity of a semantic priming IPL task in an online modality in the first instance. Second, we test whether a measurable difference exists between three types of prime-target word pairings (i.e. child-specific associations, child and adult associations, adult-specific associations) due to the small semantic priming effect size anticipated (e.g. *d*= .32, Jardak & Byers-Heinlein, 2019) in this experimental paradigm, and whether testing online introduces noise which conceals any effect of semantic priming.

3.3.2 Method

Power Analysis and Sample Size Calculations

A power analysis calculation was performed using an effect size extrapolated from Jardak and Byers-Heinlein (2019). This was achieved by using the t value from the reported t-test (*t*= 2.04) along with the sample size (*N*= 40) to extract a Cohen's d value using the equation d=t*(√ (2/n)) resulting in: 2.04*(√(2/40)) and a Cohen's d value of .32.

The resulting power analysis using this effect size showed a sample size of 39 participants would be sufficient with 80% power. Since the study design has four list orders, it was decided that 40 participants should be tested to enable a more elegant experimental design.

Participants

Forty 3-year-old healthy, English monolingual toddlers were tested (21 boys, 19 girls). The average age of participants was 37 months 3 days (*Range* = 35 months 3 days - 39 months 6 days). Productive vocabulary size was measured using the word list component of the MacArthur-Bates Communicative Development Inventory III (Dale et al., 1998) . These scores can be found in **Table 21**. A further four participants were tested but excluded due to technical issues with the experimental platform, Gorilla Experiment Builder, during testing.

	Participant number	Gender	CDI	Age in days
1	SPV_001	female	NA	1102
2	SPV_002	male	80	1150
3	SPV_003	male	74	1136
4	SPV_004	female	97	1183
5	SPV_005	female	95	1176
6	SPV_006	female	96	1153
7	SPV_007	female	93	1085
8	SPV_007a	female	96	1085
9	SPV_008	male	83	1168
10	SPV_009	male	58	1132
11	SPV_010	male	86	1201
12	SPV_011	female	85	1206
13	SPV_012	female	73	1176
14	SPV_013	female	83	1155
15	SPV_014	male	70	1148
16	SPV_015	male	92	1104
17	SPV_017	female	88	1183
18	SPV_018	female	94	1149
19	SPV_020	male	88	1188
20	SPV_021	male	98	1195
21	SPV_022	female	88	1092
22	SPV_026	male	70	1098
23	SPV_027	male	82	1088
24	SPV_028	male	65	1120
25	SPV_029	female	84	1182
26	SPV_031	female	85	1149
27	SPV_032	male	79	1154
28	SPV_034	male	84	1125
29	SPV_036	male	79	1130
30	SPV_037	female	94	1115
31	SPV_038	male	93	1101
32	SPV_039	female	92	1117
33	SPV_041	male	67	1074
34	SPV_042	female	86	1183
35	SPV_043	male	98	1114
36	SPV_044	male	90	1118
37	SPV_045	female	79	1091
38	SPV_047	male	87	1098
39	SPV_048	male	76	1085
40	SPV_049	female	95	1084

Table 21: *Experiment 5. Participant details for 3-year-olds performing an online Semantic Priming Study (CDI score= total number of words known out of 99)*

Materials

Forty-eight common, highly imageable nouns were selected from the productive language of 3-year-olds (see Chapter 1 of this thesis). It can be assumed that these words will also be receptively known by children aged 3 years.

Prime-target pairs were formed based on the WAs found in 3-year-olds in Chapter 1, but also from WAs documented as having a high associative forward strength in adults (Moss et al., 1996; Nelson et al., 2004). This resulted in three prime-target conditions: (i) unique child associations documented in the WAs of 3-year-olds (see Chapter 1), (ii) validated adult associations (i.e. word pairs documented in both the adults' WAs and the WAs of 3-year-olds), (iii) unvalidated adult associations (i.e. only found in the adult data, not in 3-year-olds' associated responses). There were four trials per condition and 12 control/ unrelated trials. Word pairs in unrelated trials had no attested associative or taxonomic relation, nor did distractor/target pairings in all trial types. Word pairs did not share phonological onset or rhyme. The full list of test stimuli can be found in **Table 22**.

	Association Type	Prime	Target	FSG_child Expriment 1	FSG_child Experiment 2	FSG_adult Birkbeck Norms: Moss and Older (1996)	FSG_adult South Florida Norms: Nelson et al. (1998)	Distractor
1	Adult+Child	chair	table	0.05	0.06	n/c	0.31	hand
2	Adult+Child	key	door	0.21	0.14	0.16	0.22	toys
3	Adult+Child	finger	hand	0.19	0.06	0.13	0.27	bib
4	Adult+Child	sock	foot	0.20	0.08	n/c	0.17	mouse
5	UniqueAdult	nappy	bib	n/a	n/a	n/c*	n/c*	puddle
6	UniqueAdult	elephant	mouse	n/a	n/a	0.07	0.09	foot
7	UniqueAdult	plate	cup	n/a	n/a	0.24	0.05	teddy
8	UniqueAdult	apple	banana	n/a	n/a	0.02	0.02	swing
9	UniqueChild	park	swing	0.16	0.17	0.02	0.06	banana
10	UniqueChild	bed	teddy	0.13	0.06	n/c	n/a	cup
11	UniqueChild	boots	puddle	0.16	n/c	n/a	n/a	table
12	UniqueChild	bath	toys	0.13	0.17	n/a	n/a	door
13	Unrelated	box	mouth	n/a	n/a	n/a	n/a	train
14	Unrelated	duck	hair	n/a	n/a	n/a	n/a	cheese
15	Unrelated	fish	car	n/a	n/a	n/a	n/a	bread
16	Unrelated	bus	pig	n/a	n/a	n/a	n/a	house
17	Unrelated	frog	plane	n/a	n/a	n/a	n/a	slide
18	Unrelated	bike	shoe	n/a	n/a	n/a	n/a	peas
19	Unrelated	cat	bread	n/a	n/a	n/a	n/a	plane
20	Unrelated	cake	house	n/a	n/a	n/a	n/a	pig
21	Unrelated	boat	peas	n/a	n/a	n/a	n/a	hair
22	Unrelated	cot	slide	n/a	n/a	n/a	n/a	mouth
23	Unrelated	pen	train	n/a	n/a	n/a	n/a	shoe
24	Unrelated	hat	cheese	n/a	n/a	n/a	n/a	car
Key	n/a	no associate noted						
	n/c	not used as a cue						
	*	n/c but included as commonly used in infant studies						

Table 22: *Experiment 5. Stimuli List for an Online Semantic Priming Study on 3-year-olds*

Twenty-four photographs of real objects were chosen to act as visual stimuli. Each visual stimulus was cut out of its background and presented centrally on a 50% grey background. The twenty-four images were seen twice by each participant: once as the target, and once as a distractor, appearing in different blocks to avoid an effect of repetition. The presentation side of the target was counterbalanced across participants and images were flipped accordingly, to always face inwards.

Each prime/ target word was individually recorded as auditory stimuli by a female speaker with a neutral British south-west accent. Words were spoken in a child-directed manner.

Each word was recorded three times and the single best token selected for its clarity and typicality, which was usually the middle of the three. Three neutral carrier phrases: *I want a/ an…*, *I have a/ an…*, *I saw a/ an...* were recorded in the same manner. The carrier phrase and prime word were concatenated into a single audio file for each trial. The target words were presented in isolation. Auditory and visual stimuli were presented using the experimental platform, Gorilla Experiment Builder (www.gorilla.sc, Anwyl-Irvine et al., 2018).

Four list orders were created to counterbalance presentation side of the target image. Block order was also counterbalanced (see **Table 23** for an example of a list order). No infant saw more than two consecutive trials from the same relatedness condition.

Trial	Association Type	Prime	Target	Distractor
1	UniqueChild	park	swing	banana
2	Unrelated	box	mouth	train
3	Unrelated	duck	hair	cheese
4	UniqueAdult	nappy	bib	puddle
5	Unrelated	fish	car	bread
6	UniqueAdult	elephant	mouse	foot
7	Unrelated	bus	pig	house
8	Unrelated	frog	plane	slide
9	Adult+Child	chair	table	hand
10	Adult+Child	key	door	toys
11	Unrelated	bike	shoe	peas
12	UniqueAdult	plate	cup	teddy
13	Adult+Child	finger	hand	bib
14	Unrelated	cat	bread	plane
15	UniqueChild	bed	teddy	cup
16	UniqueChild	boots	puddle	table
17	Unrelated	cake	house	pig
18	Unrelated	boat	peas	hair
19	Adult+Child	sock	foot	mouse
20	Unrelated	cot	slide	mouth
21	UniqueChild	bath	toys	door
22	Unrelated	pen	train	shoe
23	Unrelated	hat	cheese	car
24	UniqueAdult	apple	banana	swing

***Table 23**: Experiment 5. Example List Order of Trials*

Each list contained twenty-four trials, which is double the number of trials used in other studies (Jardak & Byers-Heinlein, 2019), but the number of trials constituted an appropriate length of experiment for the age group, as validated in Experiment 4.

Procedure

Parents were invited to participate via the UoP BabyLab database or corresponding social media pages. An information sheet about the study was emailed along with instructions for the study and a unique link to the Gorilla Experiment Builder website. A time was arranged for the parent to access the link when a researcher was available by email for questions or assistance.

The procedure replicated Experiment 4 in terms of pre-testing components, which included: eligibility checks, consent, collection of participant and demographic information, and instructions on how to position the child and how to run the experiment. The testing itself was procedurally different and is explained below.

Each trial began with a smiley fixation point in the centre of the screen for 1000ms to focus the child's attention to the middle of the screen. This was replaced by a blank screen and the carrier phrase embedded with a prime word (e.g. *I saw a… cat*) played auditorily. An ISI of 200ms was then followed by the target word (e.g. *dog*) and an SOA of 400ms at which point two images appeared: one on the left-hand side of the screen, and one on the right. One of the images was a referent to the target word, and one was a distractor image. Both images remained on screen for a further 2600ms. After twelve trials, a short animation was played to maintain the child's interest. The second block of twelve trials then followed automatically. The experiment ended with a short animation. The parent could exit the task at any point by clicking on the 'Exit' button.

Parents completed a word checklist for the words featured in the test as well as the vocabulary component of the CDI III at the end of the procedure. A final debrief screen allowed the parent to leave comments or ask questions. Parents were contacted by email after the data was downloaded and checked. The email enclosed a certificate and £5 Amazon voucher code as a thank you for participation and, where necessary, the parent was asked questions relating to comments they had left in the questionnaires and debrief sections of the procedure. A short debrief was provided.

3.3.3 Results

Data Processing and Analysis

Using UoP-developed bespoke software, webcam recordings of individual calibration and experimental trials were uploaded by participant and automatically split into 50ms frames. Calibration recordings were checked first to understand the looking behaviour of an individual (e.g. subtle/ obvious saccades, the orientation of the screen in relation to the child's position etc.), and to validate that looks were being made to the side on which the target image was being displayed.

Each video of a trial was played in full, with audio, before analysis began. Since there was no recording of the visual stimuli in the video, hearing the audio did not influence manual coding of the eye gaze as the target location was unknown. This pre-analysis step was two-fold. First, it enabled us to check the target word had recorded, indicating there was not a significant delay in the Gorilla command to begin webcam recording (600ms into a recording, the images were presented on screen with an SOA of 400ms after target word onset, thus a reliable marker that the trials were running to the time programmed and not affected by internet speed). A second reason was to understand a participant's looking pattern and head movement, to help when coding for left/right looks. This was verified by observing the parent's eye movement too i.e. if the screen was low down, the parent's gaze would indicate this which helped determine if the child was looking on or off screen.

For experimental trials, the primary coder marked if a child was looking left, right, on-screen but at an indeterminate location (which also accounts for saccades across the screen), or off-screen, using four keys on the keyboard. This was done manually for each 50ms frame for every trial. The coding was automatically saved in a .csv file which was later imported into R for analysis. A second coder coded a 10% subset of the data to test for rater reliability. Inter-rater reliability agreement between coders was 91% with a Cohen's Kappa k of 0.80, indicating substantial agreement.

Trials were excluded if i.) a participant failed to look at the screen for a minimum time of 750ms (or 15 frames, each measuring 50ms) as per Jardak and Byers-Heinlein (2019) on each trial, ii.) the length of a given trial was under 2500ms as this signified that a technical error must have occurred, iii.) if a parent had marked either the prime word or target word as unknown to the child. Trials with webcam recordings without audio were excluded if the parent could not verify that sound had been played during the experiment.

A participant was excluded if fewer than 50% of related and unrelated trials were available for analysis after excluding individual trials based on the above criteria.

Analyses were completed in R Studio version 1.4.1717. The R tidyverse and dplyr packages were used.

Descriptive Statistics

Out of a possible 960 trials (a maximum of 24 trials for each of the 40 participants), a total of 920 trials were included for analysis. Reasons for exclusion were due to: insufficient trial length, taken as <2500ms (11 trials or 1% of trials); inattentiveness, measured as <750ms spent looking at the screen per trial (11 trials or 1% of trials); prime or target word unknown to child (8 trials or 1% of trials); technical error (10 trials or 1% of all trials). Zero participants had to be replaced due to not meeting the minimum threshold number of trials, per condition.

The average number of valid trials per participant was 23 (SD= 1.99). This high number indicates children were very engaged in an online looking task when administered in the home.

There was no effect of gender on response rate, $t(38) = .96$, $p = .35$, though girls responded on average slightly more (M= 23.32, SD= 1.92) than boys (M= 22.71, SD= 2.05) to the task.

Out of the four trial types, participants completed an average of 3.85/4 (SD= 0.59) trials for unique child word pairs, 3.8/4 (SD= 0.69) trials for validated adult word pairs, 3.75/4 (SD= 0.59) trials for unvalidated adult associations, and 11.65/12 (SD= 0.86) trials for unrelated word pairs.

Proportion of Looking Time to the Target

The window of analysis was set at 200-2000ms which coincides with visual stimulus onset, an allowance of 200ms for an initial saccade, and a free-looking period of 1800ms[17].

The proportion of looking time (PLT) towards the target visual stimulus, relative to the distractor stimulus, was calculated as the dependent variable for each trial as:

PLT to target/ (PLT to target + PLT to distractor)

A two-tailed, paired t-test was run on related and unrelated trials to compare if a significant difference exists between looking behaviour on the two trial types. The difference of pairs follows a normal distribution according to a Shapiro-Wilk's test (p= .19). Three-year-olds looked longer on related trials (M= 0.51, SD= 0.07) compared to unrelated trials (M= 0.48, SD= 0.07) and the difference was significant $t(39) = 2.39$, $p = .02$, d= .38. See **Figure 10** (the white square indicates the mean).

[17] Floccia et al. (2020) identified the first 1700ms as the time period in which differences between conditions are seen in 27-month-olds, using the same task type on bilinguals.

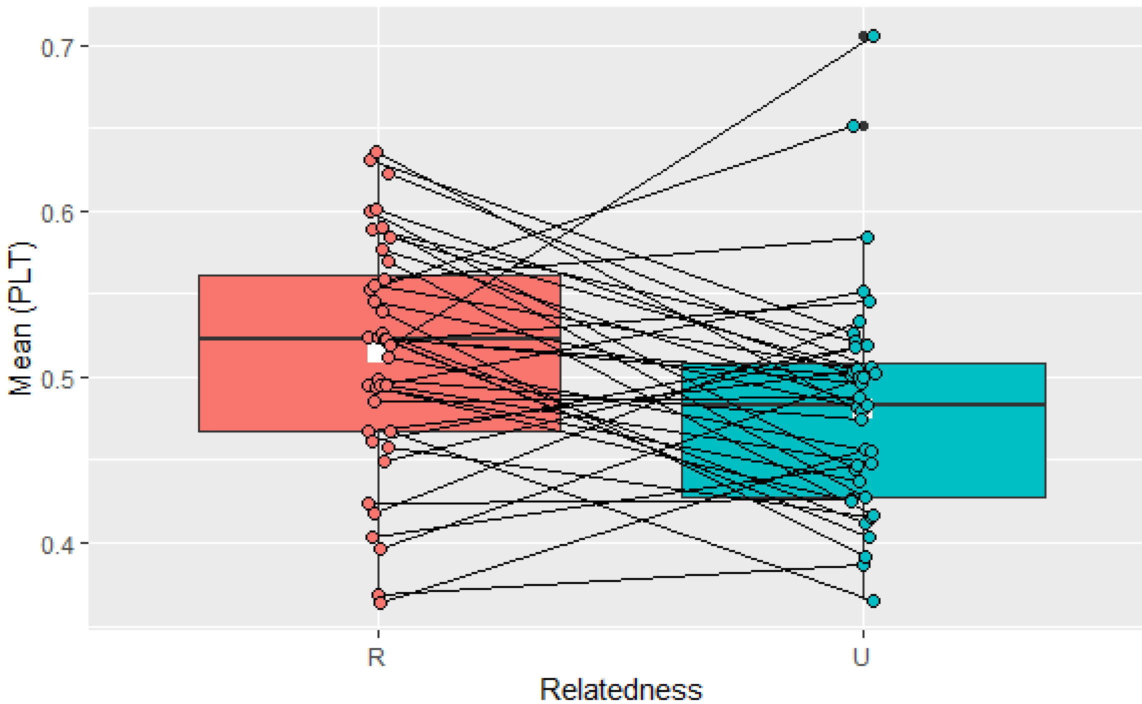

Figure 10*: Experiment 5. Proportion of looking to a target visual stimulus on semantically related and unrelated trials in an online semantic priming study on 3-year-olds*

A follow-up, one sample t-test was performed to investigate if looking was above chance (0.5) on related and unrelated trials. Comparisons to chance (0.5) with PLT indicated that 3-year-olds did not look significantly above chance in related ($t(39) = 1.28$, $p = .10$) or unrelated trials ($t(39) = -1.76$, $p = .96$).

In sum, the looking patterns of 3-year-olds indicate some sensitivity to the different relationship between words, demonstrated by a target preference when trials were related. However, there is no evidence of target recognition which is usually interpreted as above chance looking. The target not being recognised in unrelated trials replicates previous lab-based studies (e.g. Arias-Trejo & Plunkett, 2009; Styles & Plunkett, 2009), but the lack of target recognition on related trials is unexpected.

To examine the effect of association type (unique child, unique adult, adult and child, and unrelated), a one-way, repeated measures ANOVA was run on PLT with association type as a fixed factor. The PLT was statistically different for association type,

$F(2.57, 100.1) = 13.13$, $p < .0001$, generalized $\eta 2 = .19$.

Planned pairwise comparisons were performed with a Bonferroni adjustment to identify the locus of the difference. Post-hoc analyses revealed that the PLT to the target for child-specific associations ($M = 0.59$, $SD = 0.12$) differed significantly to adult-specific associations ($M = 0.45$, $SD = 0.12$; $p < 0.0001$), to adult-child associations ($M = 0.50$, $SD = 0.13$; $p = .003$), and to control trials ($M = 0.48$, $SD = 0.07$; $p < .0001$). Other pairwise comparisons were not statistically significant. This data is visualised in **Figure 11**.

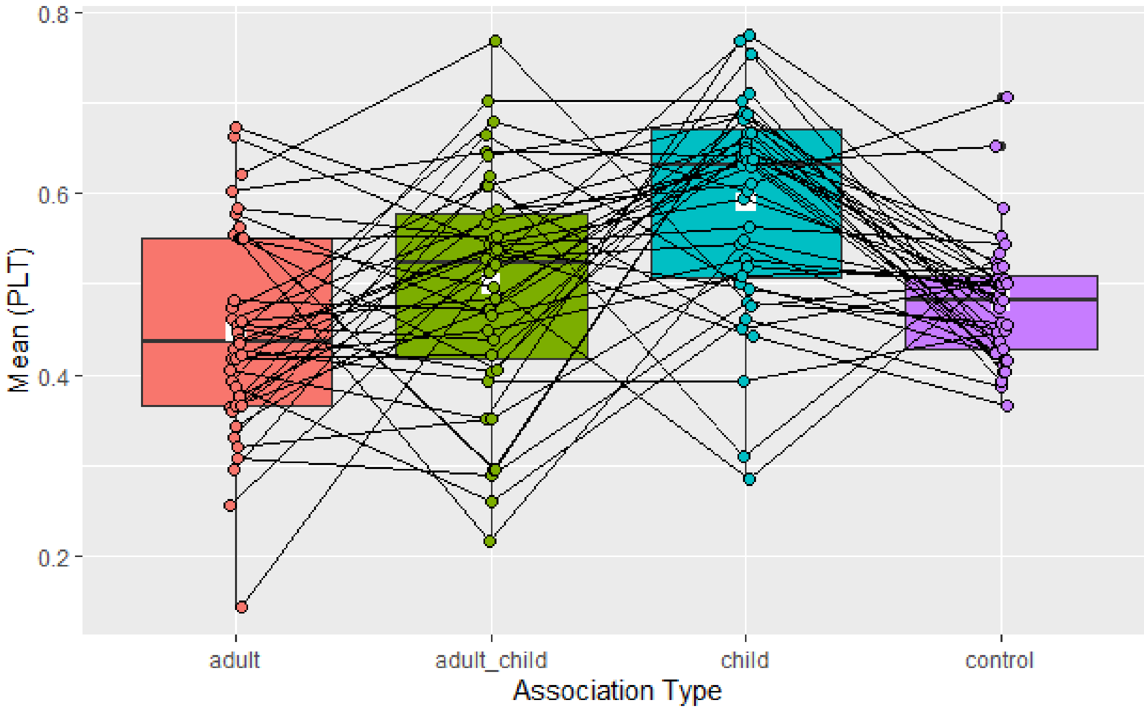

Figure 11*: Experiment 5. Proportion of looking time to the target by word association type in 36-39-month-olds doing an online semantic priming task*

Comparisons to chance (0.5) with PLT indicated that 3-year-olds looked significantly above chance in trials with child-specific associations ($t(39) = 4.82$, $p < .0001$), but not in trials with adult-child associations ($t(39) = -0.01$, $p = .5$). In trials with adult-specific associations, PLT above chance was not significant ($t(39) = -2.96$, $p = .1$) as was the case in unrelated trials ($t(39) = -1.76$, $p = .96$)

Together this shows that children looked longer at the target when the prime-target word pair had been generated in the WA task (see Experiment 1-3), compared to other WA types tested here. Only in this condition was looking above chance. The lack of above chance looking for adult or adult-child associations, and unrelated word pairs usually suggests no target recognition was indexed.

We ran a correlation between CDI scores and priming difference scores, which were calculated by subtracting the PLT on unrelated trials from the PLT on related trials per child, as per Jardak and Byers-Heinlein (2019). There was no effect of productive vocabulary size on priming $r(37)= .06$, $p= .7$.

Time-course Analysis

The PLT to the target for related and unrelated trials was averaged across participants for each 50ms time bin and plotted (see **Figure 12**). Visual inspection suggests that the curves start to diverge at approximately 125ms.

Figure 12: *Experiment 5. Time-course of looking behaviour in 36-39-month-olds for semantically related and unrelated trials with the significant divergence in behaviour indicated by a boxed area*

To determine where any difference in looking behaviour occurred on related and unrelated trials during the time-course of word recognition, a non-parametric statistical cluster analysis was performed (see Maris & Oostenveld, 2007), which has been successfully employed by various studies investigating preferential looking (Floccia et al., 2020; Von Holzen et al., 2019; Von Holzen & Mani, 2012). Paired t-tests were run for each time bin, followed by identifying clusters with significant *t* vales and comparing these to a Monte Carlo distribution. Comparisons using the time-course analysis revealed a significant difference in looking behaviour between 450 and 850ms post visual stimulus onset (cluster *t* statistics = 27.99, Monte Carlo $p = 0$) between related and unrelated trials, with the unrelated condition showing reduced looking in this period compared to related trials. This area is marked by a box in **Figure 12**. This analysis suggests that the priming effect, as indexed by the difference in PLT in the related and unrelated conditions, occurs at around 450ms after target onset.

3.3.4 Discussion

A semantic priming study was run on 3-year-olds in their home environment with a very high proportion of valid trials and low levels of data loss (3%). This indicates that running an adapted infant, primed IPL experiment in a participant's home is a valid technique, which is worth developing and extending to test other infant populations and research questions.

The results of Experiment 5 replicate the findings of lab-based studies, that is, a semantic priming effect resulted from a higher proportion of a child's looking time to a target picture when the prime word was associatively related to the target word (according to free association responses), compared to trials on which the prime and target words were unrelated.

In unrelated trials, 3-year-olds did not look above chance at the target visual stimulus, suggesting that 3-year-olds did not recognise the target word. This replicates in-lab study findings (e.g. Arias-Trejo & Plunkett, 2009; Styles & Plunkett, 2009). Unexpectedly, children also did not look above chance on related trials despite a higher proportion of PLT relative to unrelated trials (discussed in more detail below). We only discovered above chance looking in related trials when we analysed the effect of word association (WA) type on PLT.

The main aim of Experiment 5 was to ascertain if the unique child WAs found in Experiments 1-3 would demonstrate a measurable difference in a receptive task such as a semantic priming study. To explore this, we compared PLT for each WA type (unique child, unique adult, child and adult, unrelated). The results clearly demonstrated that the priming effect was modulated by WA type. Related word pairs with the highest PLT were those taken from the productive vocabularies of 3-year-olds, tested in a WA task (see Chapter 1, Experiment 1-3). This WA type was the only of the four types tested with an above chance probability of looks towards the target image. The absence of above chance looking when all three WA types were combined to calculate PLT on related trials might suggest an online modality was not sensitive enough to capture this effect, particularly for WAs not robust in a child's lexical-semantic system (i.e. those stemming from adult associative norms). This is in spite of the semantic priming effect indexed. The finding that PLT for child-specific WAs differed significantly to the two other WA types (adult specific, child and adult) provides some evidence that an effect of semantic priming only occurred in the combined related data due to the associative boost provided by the child specific WAs.

As hypothesised, WAs not found in the productive vocabularies of 3-year-olds, but prominent in the associated responses of adults performing a WA task, did not show a strong effect of priming in this experiment. This deserves attention as many studies exploring the primacy of connections in the lexical-semantic system of infants have relied on associative norms from the adult literature to drive decisions regarding experimental stimuli for their studies. Studies which might not have seen a priming effect could be a result of stimuli selected, and the assumption that a WA in the adult lexical system is equivalently robust in the infant system. In experiments that did find a priming effect, further analysis on the stimuli selected could help inform other researchers on the best word pairs to select for infant studies.

A finding that we did not expect to see was the lack of a priming effect in child-adult associations, that is, word pairs documented in our own findings of Chapter 1 (for 3-year-olds) and in adult associative norms (Moss et al., 1996; Nelson et al., 2004). We expected an effect of priming based on the assumption that forward associative strength (FSG) values were a reliable metric for the strength of a relationship between words (remember that FSG= the likelihood that a cue will elicit a certain target). The FSG for child-specific associations was weaker than the FSG for adult-child associations in Experiment 1-3, thus we had expected to see a stronger effect of priming in child and adult WAs compared to child specific WAs. However, there was no semantic priming effect in child and adult WAs. This finding could indicate that the FSG is not a reliable metric for WAs in young children, or more specifically, not a reliable metric for receptive knowledge (which is what this study tests, compared to the productive knowledge tested in Experiment 1-3). Additionally, due to the noise inherent in infant data, it might not be appropriate to compare absolute values between the two groups. This line of enquiry would require further investigation. Another explanation for no semantic priming in child and adult WAs may be explained by the associative nature of the child-specific WAs compared to the more taxonomic child and adult WAs (see Experiment 9 for a discussion on this).

Taken together, Experiment 5 replicates in-lab findings in as far as a semantic priming effect was measured, but the lack of above chance looking on (combined) related trials requires further investigation to determine if the finding was unique to this experiment, or if it more broadly represents an issue with the sensitivity of an online testing procedure. Notwithstanding this concern, time-course analysis indicated that priming occurred at 450-850ms, which is not that dissimilar to in-lab studies (e.g. 630ms in Chow et al., 2017; 600ms in Floccia et al, 2020). This could suggest that noise from factors such as device and internet speed variability were not large enough to conceal the effect of semantic priming, at least not when testing 3-year-old children.

3.4 Experiment 6: Semantic Priming in Monolingual Children, Online (30 months)

3.4.1 Introduction

Experiment 5 demonstrated that a semantic priming effect was present in 3-year-olds when the experiment was an adapted version of a primed IPL task run in the home of a participant, using their own device. Experiment 6 aims to test this same phenomenon, but at the younger age of 30 months. The first aim is to replicate a semantic priming effect when testing online, in an even younger population. Semantic priming is a robust finding in monolingual infants as young as 24 months (Arias-Trejo & Plunkett, 2013; Jardak & Byers-Heinlein, 2019) so we would expect to find a clear effect at 30 months, even when the task is administered in a home setting. Thus, it is expected that 30-month-olds will sustain attention for a target referent for a longer period when the preceding word is related, compared to when the preceding word is unrelated, which we expect to disrupt attention.

The second aim for re-running a semantic priming study online is to act as a monolingual control study from which to compare an English-Dutch/German[18] bilingual sample of 30-month-old toddlers (see Experiment 7). To do so, we will employ the methodology used by Jardak and Byers-Heinlein (2019) which tested 30-month-old French-English bilinguals relative to French and English monolingual groups. Experiment 7 will investigate the effect of English-Dutch/German cognates on semantic priming, extending previous research on language non-selective lexical access during infant bilingual language processing. Consequently, this has dictated the stimuli selection for this monolingual, control study. Since Experiment 6 tests English monolinguals, trials in which there is an English-

[18] The original study was designed for English-Dutch bilinguals, but this was extended to include English-German bilinguals after commencing data collection, due to recruitment difficulty (see Experiment 7: Discussion section for further information).

Dutch/German cognate present should be attended to in the same way as trials in which cognates are not present and validating this finding in Experiment 6 is a necessary step to validate any effect of cognate status in the bilingual sample tested in Experiment 7.

3.4.2 Method

Participants

Forty 30-month-old healthy, English monolingual toddlers were tested (22 boys, 18 girls). The average age of participants was 29 months, 27 days (*Range* = 29 months 21 days - 30 months 18 days). Five additional participants were excluded and replaced for technical reasons.

Materials

Forty-eight common, imageable nouns were chosen as prime and target stimuli. Selection of words was based on the record of 60% of 18-month-olds understanding the words according to the UK CDI (UK-CDI Database, 2016) and Oxford CDI norms of vocabulary development (Hamilton et al., 2000). Words pairs were created for a 24-trial-long experimental design. Twelve of the target words were preceded by related prime words embedded in a carrier phrase, and twelve target words were preceded by unrelated primes.

Related prime and target words either had an attested forward association according to adult associative norms (Moss et al., 1996) or were category coordinates, with no associative forward strength for unrelated prime-target word pairs or between the prime and the distractor. There was no phonological overlap (i.e. onset or rhyme) between the word pairs.

Common word pairings from previous infant priming studies were prioritised, for a replicable study design, but the stimuli set from the English-Dutch/German bilingual cognate semantic priming study (Experiment 7) took precedence so that this study could act as a control. For this reason, the flexibility of stimuli selection was restricted, and decisions heavily influencing stimuli selection will be discussed in more detail in Experiment 7 (e.g. cognate status, phonological overlap between translation equivalents, number of syllables etc.). The stimuli set can be found in **Table 24**.

	Type of Relatedness	Prime	Target	Prime-Target (FSG) using Nelson et al. (1998)	Prime-Target (FSG) using Moss and Older (1996)	Distractor	Prime-Distractor (FSG) using Nelson et al. (1998)	Prime-Distractor (FSG) using Moss and Older (1996)	Infant Priming Study using the word pair
1	Associative	park	tree	0.115	0.083	coat	No association noted	No association noted	Arias-Trejo & Plunkett (2009)
2	Associative	glass	milk	No association noted	No association noted	book	No association noted	No association noted	Floccia et al. (2020) cup-milk
3	Associative	garden	house	No association noted	Not used as a cue	bus	No association noted	Not used as a cue	
4	Taxonomic	sock	trousers	No association noted	Not used as a cue	eye	No association noted	Not used as a cue	
5	Taxonomic	cat	bird	No association noted	0.021	doll	No association noted	No association noted	
6	Taxonomic	foot	hand	(feet) 0.095	0.021	door	No association noted	No association noted	Jardak & Byers-Heinlein (2019) hand-foot
7	Taxonomic	train	bus	0.03	0.062	pen	No association noted	No association noted	Jardak & Byers-Heinlein (2019)
8	Taxonomic	boots	coat	No association noted	No association noted	frog	No association noted	No association noted	
9	Taxonomic	spoon	bowl	No association noted	No association noted	window	No association noted	No association noted	Singh (2014)
10	Taxonomic	car	lorry	(truck) 0.111	(truck) 0.024	tree	No association noted	No association noted	Styles & Plunkett (2009)
11	Taxonomic	chair	bed	0.013	Not used as a cue	hand	No association noted	Not used as a cue	
12	Taxonomic	chicken	mouse	No association noted	Not used as a cue	ball	No association noted	Not used as a cue	
13	Unrelated	sofa	flower	No association noted	No association noted	dog	No association noted	No association noted	
14	Unrelated	box	frog	No association noted	No association noted	lorry	No association noted	No association noted	
15	Unrelated	nose	window	No association noted	Not used as a cue	bowl	No association noted	No association noted	
16	Unrelated	clock	ball	No association noted	Not used as a cue	house	No association noted	Not used as a cue	
17	Unrelated	tiger	star	No association noted	Not used as a cue	bed	No association noted	Not used as a cue	
18	Unrelated	banana	heart	No association noted	Not used as a cue	mouse	No association noted	Not used as a cue	
19	Unrelated	cheese	dog	No association noted	Not used as a cue	trousers	No association noted	Not used as a cue	
20	Unrelated	nappy	eye	Not used as a cue	Not used as a cue	flower	Not used as a cue	Not used as a cue	
21	Unrelated	cake	doll	No association noted	No association noted	bird	No association noted	No association noted	
22	Unrelated	bottle	door	No association noted	Not used as a cue	heart	No association noted	Not used as a cue	
23	Unrelated	key	pen	No association noted	No association noted	milk	No association noted	No association noted	
24	Unrelated	horse	book	No association noted	Not used as a cue	star	No association noted	Not used as a cue	

***Table 24**: Stimuli used in Experiments 6-9 with FSG calculated for the Prime-Target and Prime-Distractor pair combinations. FSG= forward associative strength*

Trials were organised into two blocks so that the twelve target visual stimuli in Block 1 became the distractor visual stimuli in Block 2 and the twelve distractor visual stimuli in Block 1 became the target visual stimuli in Block 2. Different word pairs were created in Block 1 compared to Block 2, so that when a target became a distractor, this was in a new word pairing.

Four list orders were created (see **Appendix R - U**), counterbalancing for side of presentation of target image and for block order. The order of trials was organised so that no more than three consecutive trials were critical trials (i.e. related word pairs, compared to unrelated word pairs). Care was taken to avoid consecutive trials having semantic or phonological overlap.

Auditory targets and primes were recorded in one session by an adult female with a neutral British, south-west accent in a child-directed tone. Three tokens of each word and carrier phrase were recorded and the best was manually selected in each case. Auditory primes were concatenated with one of three carrier phrases: *I saw a…*, *I want a …*, *I have a…*.

Visual stimuli were photographs from the internet, used in previous studies in the Plymouth BabyLab. The background was removed from each image and set on a pale grey background. Each of the twenty-four visual stimuli were presented twice: once as the target, and once as a distractor.

Procedure

The experiment was run online on the Gorilla Experiment Builder platform. The procedure was identical to that described in Experiment 5, but using the stimuli set outlined in the previous section.

3.4.3 Results

Data Processing and Analysis

Webcam recordings were prepared and coded as per Experiment 5. One coder coded the eye-movement (left, right, indeterminate, away) for every participant and a second coder independently coded a 10% randomly selected sample of the same data. Inter-rater reliability agreement between coders was 90% with a Cohen's Kappa k of 0.64, indicating substantial agreement.

The same exclusion criteria were applied as for Experiment 5. Trials were excluded which were: attended to for less than 750ms; under 2500ms in duration; contained a prime or target word unknown to the child.

Descriptive Statistics

Out of a maximum 960 trials (24 trials for each of the 40 participants), a total of 864 trials were included for analysis. Excluded trials were a result of insufficient trial length, taken as <2500ms (30 trials: 3% of trials); inattentiveness, measured as <750ms spent looking at the screen per trial (12 trials: 1% of trials); prime or target word unknown to child (42 trials: 4% of trials); technical error (12 trials: 1% of all trials).

The mean number of valid trials per participant was 21.6 (*SD*= 2.67). This high number indicates children were engaged in an online looking task when administered in the home.

There was no effect of gender on response rate, $t(38) = .73$, $p = .47$, though girls responded on average slightly more (*M*= 21.94, *SD*= 2.24) than boys (*M*= 21.32, *SD*= 3.00) to the task.

Proportion of Looking Time to the Target

The window of analysis was set at 200ms – 2000ms which coincides with visual stimulus onset, an allowance of 200ms for an initial saccade, and a free-looking period of 1800ms.

The proportion of looking time (PLT) towards the target visual stimulus, relative to the distractor stimulus, was calculated as the dependent variable for each trial as:

PLT to target/ (PLT to target + PLT to distractor)

A two-tailed, paired t-test was run on PLT for related and unrelated trials to determine if an effect of priming was present. Thirty-month-olds looked longer on unrelated trials (*M*= 0.61, *SD*= 0.08) compared to related trials (*M*= 0.54, *SD*= 0.08) and the difference was significant, $t(39) = 4.78$, $p < .0001$, $d = -.76$. This data is visualised in **Figure 13** (the white square indicates the mean).

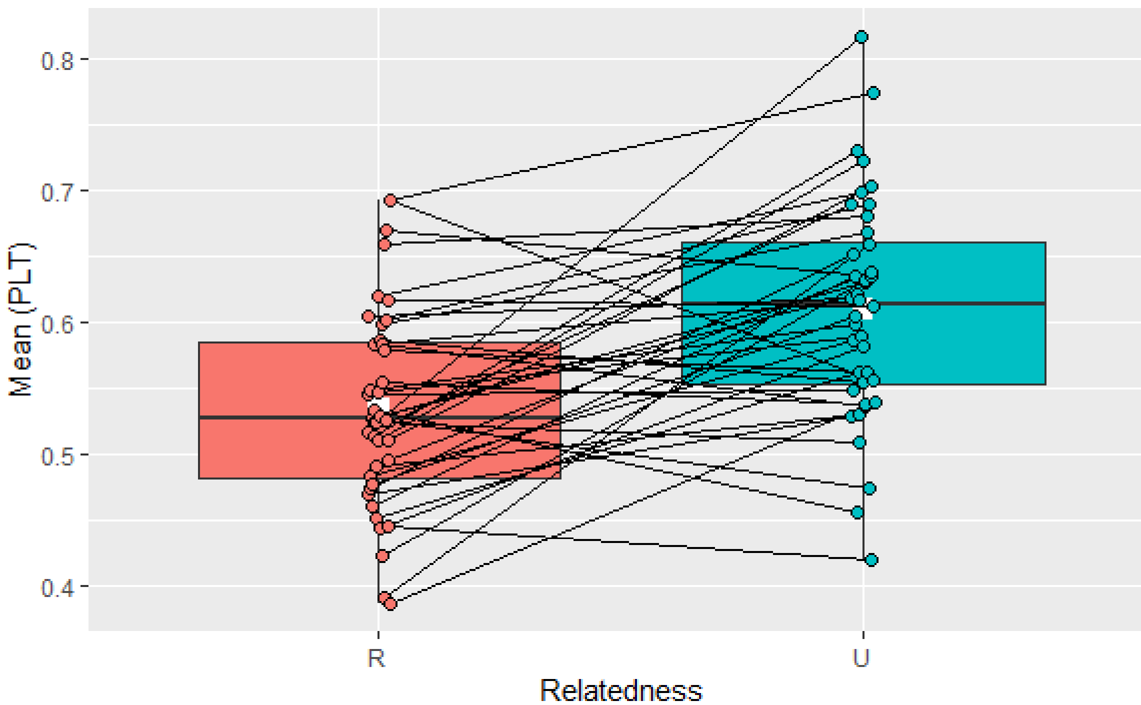

Figure 13: *Experiment 6. Proportion of looking time to a target visual stimulus on semantically related and unrelated trials in an online semantic priming study on 30-month-old monolinguals*

Follow-up, one sample t-tests were performed to investigate if looking was above chance on related and unrelated trials. Comparisons to chance (0.5) with PLT indicated that 30-month-olds did look significantly above chance in related ($t(39) = .2.93$, $p = .003$), and unrelated trials ($t(39) = 8.31$, $p <.0001$).

In summary this shows correct target identification on related and unrelated trials in 30-month-olds, but with increased looking on unrelated trials which would indicate an inverse priming effect occurred, contrary to what was expected.

The correlation between CDI scores and the priming difference scores (PLT related trials-PLT unrelated trials) was not significant $r(35) = .11$, $p = .5$.

Time-course Analysis

The PLT to the target for related and unrelated trials was averaged across participants for each 50ms time bin and plotted (see **Figure 14**). Visual inspection suggests that the two curves diverge at approximately 250ms.

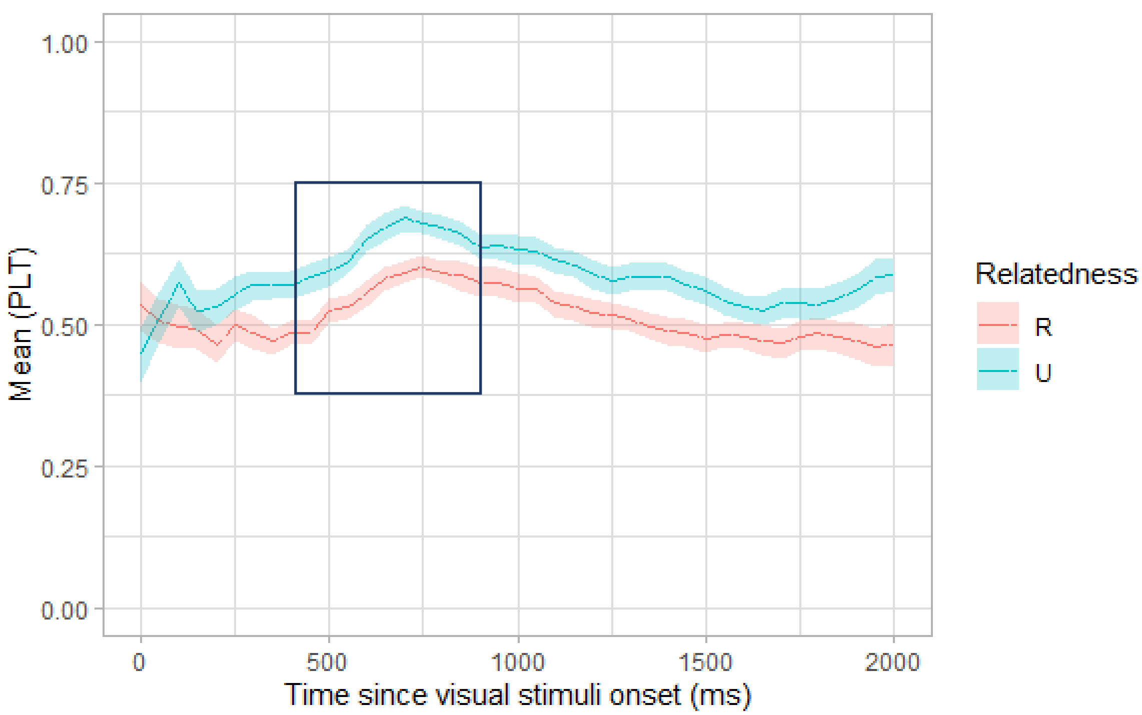

Figure 14: *Experiment 6. Time-course of looking behaviour in 30-month-olds for semantically related and unrelated trials with the significant divergence in behaviour indicated by the boxed area*

As with Experiment 5, a cluster analysis was performed (Maris & Oostenveld, 2007) to determine where significant differences occur between looking behaviour on related and unrelated trials in the time-course of word recognition. Paired t-tests were run for each time bin, followed by identifying clusters with significant *t* vales and comparing these to a Monte Carlo distribution. The difference in looking behaviour for related and unrelated trials was identified between 400 and 900ms post visual target onset (cluster *t* statistics = -30.42, Monte Carlo p = .0002), with the related condition showing reduced looking in this period compared to unrelated trials. This area is marked by a box in **Figure 14**. This analysis suggests the difference in PLT in the related and unrelated conditions, occurs at around 400ms post visual stimuli onset.

3.4.4 Discussion

A semantic priming study was conducted on 30-month-olds with unexpected results. It was hypothesised that a related prime before a target word would increase the proportion of looking time to the target picture, relative to unrelated trials. The opposite was found: 30-month-olds looked longer at the target on unrelated trials, compared to related trials and this difference in looking time was significant. This finding does not replicate the findings of Experiment 5 or those from comparable lab studies (e.g. Jardak & Byers-Heinlein, 2019).

Interestingly, target word recognition was indexed in both related and unrelated trials. The former is a common finding in semantic priming studies on infants, though the latter is not (e.g. Arias-Trejo & Plunkett, 2009; Styles & Plunkett, 2009). This suggests that 30-month-olds recognised the target on both trial types and were not influenced by the relatedness of the prime. Participants also seem to have been more interested in the target on unrelated trials, compared to related trials. This is in contrast to 36-month-olds (Experiment 5) who did not demonstrate word recognition through above chance looking, but who were affected by the relatedness of the prime since related trials were attended to significantly more than unrelated trials.

The age of the children we tested restricted the semantic categories we could select words from for the cue words in the WA task. Because of this, it is possible that semantic inhibition may have caused the unexpected inverse priming effect. According to Chow et al. (2018), 18-month-olds, inhibit a previously activated semantic category when switching to a different semantic category. The previously activated category becomes irrelevant so to efficiently direct attention to the new category, the irrelevant information is inhibited. However, if 18-month-olds then must activate the previous category, access to that category is impaired. This behaviour is in line with adult inhibitory processes. As Chow et al. (2018) note, the effect of inhibitory processes take time to become apparent and so this may be one possible explanation for our unexpected findings. Determining if this is causing the inverse priming effect would take careful consideration and a re-design of the current stimuli lists, so we will first turn to alternative explanations which can be more simply tested.

Since a priming effect was evidenced in Experiment 5, there are three plausible factors which might explain the unexpected findings in Experiment 6. Two relate to differences between Experiment 5 and 6, namely, the age of participants in the studies, and the stimuli used. The third relates to the online modality of the experimental design and the little evidence available to confidently rule out any interference from running an infant study in this way.

Regarding the latter, one concern we had before running an infant study online was the potential introduction of noise to such a degree that a priming effect would not be apparent. Although this is not completely true of Experiment 5, it cannot be ruled out fully since we did find an effect of priming in Experiment 5, but not above chance looking. Furthermore, the effect of priming was stronger when the word pairs were strongly associated in a child's lexicon, as evidenced by the related responses given by 3-year-olds in Experiments 1-3. Together this suggests that there could be an interaction between relatedness (i.e. how associated the word pairs are to a child) and modality, which is causing the unexpected findings in Experiment 6.

At this point, it is worth noting some of the similarities and differences between this experiment and Experiment 5. Firstly, this study on 30-month-olds showed a lower response rate (21.6/24) and higher data loss (9%) compared to Experiment 5 (23/24, 3% respectively). This might be due to the age of the participants or relate to the experimental design itself. In terms of similarity, turning to time-course analyses, while the divergence in looking behaviour for related and unrelated trials is similar in Experiment 5 (450-850ms) and Experiment 6 (400-900ms), PLT to the target is higher on related trials in Experiment 5, but unrelated trials in Experiment 6. This might be due to the different stimulus sets used in each experiment.

Experiment 5 used mainly associatively related words, but Experiment 6 used mainly taxonomically related words. The reason for this was due to the difficulty in finding word pairs that would also be suitable for use when translated into Dutch and German (for Experiment 7). That is, words which did not share phonological overlap through translation and words which were cognates for the cognate condition only, but not otherwise. This resulted in certain word pairs being selected that have not featured in the stimuli lists of previous infant semantic priming studies. It could be that words pairs were selected that are not robust in the lexical-semantic systems of 30-month-olds, because they are not prime-target stimuli with a strong enough relationship (McRae & Boisvert, 1998), in addition to the associative/ taxonomic bias in the different stimulus lists, as mentioned previously. The lack of relatedness through association is problematic for the stimuli used in Experiment 6 as related word pairs which have both a semantic and an associative relationship have been shown to provide a 'bootstrapping' effect to ensure priming is optimised in participants with limited vocabularies (Arias-Trejo & Plunkett, 2009). This is supported by a meta-analysis (Lucas, 2000) evidencing a larger effect size for associatively related prime-target stimuli (d= .49) compared to pure semantic prime-target relations (d= .25).

Finally, age should not be ruled out as having contributed to the unexpected finding as participants in this experiment are 6-9 months younger than those tested in Experiment 5. While this does not seem a likely explanation because semantic priming has been observed from 24 months in lab-based studies (Arias-Trejo & Plunkett, 2013; Jardak & Byers-Heinlein, 2019), there might have been an interaction between the age of participants and the stimuli chosen.

To understand why 30-month-olds were more attracted to a target visual stimulus on unrelated trials compared to related trials, a replication study is necessary and will be the focus of Experiment 8. This study should identify if age contributed to the outcome, or if testing online played a role. Both these variables need testing before other semantic priming studies can be confidently run in an online modality, at this age or younger. Before exploring these variables, the preliminary findings from Experiment 7 will be presented, which is the bilingual study for which Experiment 6 was a control.

3.5 Experiment 7: Semantic Priming in Bilingual Children, Online (30 months)

3.5.1 Introduction

Experiment 7 was run in parallel to Experiment 6, which tested 30-month-old English monolinguals in an online IPL semantic priming study. The aim of the current study was to replicate the effect of semantic priming in 30-month-old French-English bilinguals found by Jardak and Byers-Heinlein (2019). Three aspects of Experiment 7 are novel compared to the study by Jardak and Byers-Heinlein: 1. The semantic priming task is run in an online modality wherein trials run automatically and are not child-led; 2. The language pairing is English-Dutch/ German[19]; 3. Cognates are specifically tested for their effect on semantic priming.

It was hypothesised that an effect of semantic priming would be evident in English-Dutch/ German bilinguals, despite the task being run online. Since the experiment was run only in English, trials which include a cognate as a prime or target should facilitate processing if the participant's L2 is English, but if English is the L1 then the effect may not be facilitatory (see Von Holzen et al., 2019). Due to the inconclusive findings in bilingual infant priming studies, language dominance may (Singh, 2013) or may not (Floccia et al., 2020; Jardak & Byers-Heinlein, 2019) modulate the priming effect and the language pair tested might contribute to the disparate findings (Singh, 2013).

This experiment is presented in this thesis despite its very low participant number (*N*= 4) because it justifies Experiment 6. The author of this thesis had organised to spend three

[19] The original study was designed for English-Dutch bilinguals, but this was extended to include English-German bilinguals after commencing data collection, due to recruitment difficulty (see Discussion section for further information).

months at the MPI in Nijmegen to collect data, but due to the pandemic, this had to be postponed many times, with important consequences for the recruitment.

3.5.2 Method

Participants

Four 30-month-old healthy, English-Dutch/English-German bilingual toddlers were tested (1 boy, 3 girls). A participant qualified as bilingual if exposed to one of the two languages for at least 20% of the time during an average week. This was assessed using questions from a Language Exposure Questionnaire (LEQ, Cattani et al., 2014). More detailed language history information was collected through questionnaires (see Procedure) for each participant for analyses.

Materials

The materials for English-Dutch/German participants were identical to those in Experiment 6 (see **Table 25**). Twelve trials used related prime and target words, and twelve were unrelated. With regards to the cognate status of words in related trials, three trials comprised cognate words as the prime, target and distractor word; three trials used cognates for the prime and target word, but a non-cognate as the distractor; three trials used non-cognates as prime, target and distractor words; and three trials used non-cognate as prime words, but cognates as target and distractor words. The same distribution of conditions was applied to unrelated trials.

Prime, target and distractor words in a single trial did not overlap phonologically within or between languages. Care was taken to not have consecutive trials with a semantic or phonological overlap. No more than three consecutive trials were from the same condition. The four list orders used in Experiment 6 were used in this study and they counterbalanced for target location and block order.

					Prime			Target			Distractor		
	Trial	**Prime Type**	**Target Type**	**Distractor Type**	**English**	**Dutch**	**German**	**English**	**Dutch**	**German**	**English**	**Dutch**	**German**
1	related	cognate	non-cognate	non-cognate	**sock** /sɒk/	**sok** [sɔk]	**Socke** /zɔkə/	**trousers** /ˈtraʊ.zəz/	**broek** [bruk]	**Hose** /ho:zə/	**eye** /aɪ/	**oog** [ox]	**Auge** /aʊ̯gə/
2	related	cognate	non-cognate	non-cognate	**cat** /kæt/	**kat** [kɑt]	**Katze** /kat͡sə/	**bird** /bɜːd/	**vogel** ['voxəl]	**Vogel** /fo:gəl/	**doll** /dɒl/	**pop** [pɔp]	**Puppe** /pʊpə/
3	related	cognate	non-cognate	non-cognate	**park** /pɑːk/	**park** [pɑrk]	**Park** /paːɐ̯k/	**tree** /triː/	**boom** [bom]	**Baum** /baʊ̯m/	**coat** /kəʊt/	**jas** [jɑs]	**Mantel** /mantəl/
4	related	cognate	cognate	cognate	**foot** /fʊt/	**voet** [vut]	**Fuß** /fu:s/	**hand**	**hand** [hɑnt]	**Hand**	**door** /dɔːr/	**deur** [dør]	**Tür** /tyːɐ̯/*
5	related	cognate	cognate	cognate	**train**	**trein** [trɛin]	**Zug**	**bus** /bʌs/	**bus** [bʏs]	**Bus** /bʊs/	**pen** /pen/	**pen** [pɛn]	**Stift** /ʃtɪft/*
6	related	cognate	cognate	cognate	**glass**	**glas** [xlɑs]	**Glas** /gla:s/	**milk** /mɪlk/	**melk** [mɛlk]	**Milch**	**book** /bʊk/	**boek** [buk]	Buch /bu:x/
7	unrelated	cognate	non-cognate	non-cognate	**sofa** /ˈsəʊ.fə/	**sofa** [so·fa]	**Sofa** /zo:fa:/	**flower** /flaʊər/	**bloem** [blum]	**Blume** /blu:mə/	**dog** /dɒg/	**hond** [hɔnt]	**Hund** /hʊnt/
8	unrelated	cognate	non-cognate	non-cognate	**box** /bɒks/	**box** [bɔks]	**Kasten** /kastən/*	**frog** /frɒg/	**kikker** [ˈkɪkər]	**Frosch** /fʁɔʃ/*	**lorry** /ˈlɒr.i/	**vrachtwagen** ['vrɑxtwaxə(n)]	**LKW** /lkv/
9	unrelated	cognate	non-cognate	non-cognate	**nose** /nəʊz/	**neus** [nøs]	**Nase** /na:zə/	**window** /ˈwɪn.dəʊ/	**venster** [ˈvɛnstər]	**Fenster** /fɛnstɐ/	**bowl** /bəʊl/	**kom** [kɔm]	**Schüssel** /ʃʏsəl/
10	unrelated	cognate	cognate	cognate	**clock**	**klok** [klɔk]	**Uhr** /uːɐ̯/*	**ball** /bɔːl/	**bal** [bɑl]	**Ball** /bal/	**house**	**huis** [hœys]	**Haus**
11	unrelated	cognate	cognate	cognate	**tiger** /ˈtaɪ.gər/	**tijger** [ˈtɛixər]	**Tiger** /tɪgɐ/	**star** /stɑːr/	**star** [stɑr]	**Stern** /ʃtɛn/	**bed** /bed/	**bed** [bɛt]	**Bett** /bɛt/
12	unrelated	cognate	cognate	cognate	**banana** /bəˈnɑː.nə/	**banaan** [baˈnan]	**Banane** /bana:nə/	**heart** /hɑːt/	**hart** [hɑrt]	**Herz** /hɛɐ̯t͡s/	**mouse** /maʊs/	**muis** [mœys]	**Maus** /maʊ̯s/
13	related	non-cognate	non-cognate	non-cognate	**boots** /buːts/	**laarzen** [lars]	**Stiefel** /ʃti:fəl/	**coat** /kəʊt/	**jas** [jɑs]	**Mantel** /mantəl/	**frog** /frɒg/	**kikker** [ˈkɪkər]	**Frosch** /fʁɔʃ/*
14	related	non-cognate	non-cognate	non-cognate	**spoon** /spuːn/	**lepel** [ˈlepəl]	**Löffel** /lœfəl/	**bowl** /bəʊl/	**kom** [kɔm]	**Schüssel** /ʃʏsəl/	**window** /ˈwɪn.dəʊ/	**venster** [ˈvɛnstər]	**Fenster** /fɛnstɐ/
15	related	non-cognate	non-cognate	non-cognate	**car** /kɑːr/	**auto** [ˈauto]	**Wagen** /va:gən/	**lorry** /ˈlɒr.i/	**vrachtwagen** ['vrɑxtwaxə(n)]	**LKW** /lkv/	**tree** /triː/	**boom** [bom]	**Baum** /baʊ̯m/
16	related	non-cognate	cognate	cognate	**chair** /tʃeər/	**stoel** [stul]	**Stuhl** /ʃtʊl/	**bed** /bed/	**bed** [bɛt]	**Bett** /bɛt/	**hand** /hænd/	**hand** [hɑnt]	**Hand** /hant/
17	related	non-cognate	cognate	cognate	**garden** /ˈgɑː.dən/	**tuin** [tœyn]	**Garten** /gaːɐ̯tən/*	**house** /haʊs/	**huis** [hœys]	**Haus** /haʊ̯s/	**bus** /bʌs/	**bus** [bʏs]	**Bus** /bʊs/
18	related	non-cognate	cognate	cognate	**chicken** /ˈtʃɪk.ɪn/	**kip** [kɪp]	**Hähnchen** /hɛ:nçən/	**mouse** /maʊs/	**muis** [mœys]	**Maus** /maʊ̯s/	**ball** /bɔːl/	**bal** [bɑl]	**Ball** /bal/
19	unrelated	non-cognate	non-cognate	non-cognate	**cheese** /tʃiːz/	**kaas** [kas]	**Käse** /kɛ:zə/	**dog** /dɒg/	**hond** [hɔnt]	**Hund** /hʊnt/	**trousers** /ˈtraʊ.zəz/	**broek** [bruk]	**Hose** /ho:zə/
20	unrelated	non-cognate	non-cognate	non-cognate	**nappy** /ˈnæp.i/	**luier** ['lœyjər]	**Windel** /vɪndəl/	**eye** /aɪ/	**oog** [ox]	**Auge** /aʊ̯gə/	**flower** /flaʊər/	**bloem** [blum]	**Blume** /blu:mə/
21	unrelated	non-cognate	non-cognate	non-cognate	**cake** /keɪk/	**taart** [tart]	**Kuchen** /ku:xən/	**doll** /dɒl/	**pop** [pɔp]	**Puppe** /pʊpə/	**bird** /bɜːd/	**vogel** ['voxəl]	**Vogel** /fo:gəl/
22	unrelated	non-cognate	cognate	cognate	**bottle** /ˈbɒt.əl/	**fles** [flɛs]	**Flasche** /flaʃə/	**door** /dɔːr/	**deur** [dør]	**Tür** /tyːɐ̯/*	**heart** /hɑːt/	**hart** [hɑrt]	**Herz** /hɛɐ̯t͡s/
23	unrelated	non-cognate	cognate	cognate	**key** /kiː/	**sleutel** [ˈsløtəl]	**Schlüssel** /ʃlʏsəl/	**pen** /pen/	**pen** [pɛn]*	**Stift** /ʃtɪft/*	**milk** /mɪlk/	**melk** [mɛlk]	**Milch** /mɪlç/
24	unrelated	non-cognate	cognate	cognate	**horse** /hɔːs/	**paard** [part]	**Pferd** /p͡fe:ɐ̯t/	**book** /bʊk/	**boek** [buk]	**Buch** /bu:x/	**star** /stɑːr/	**star** [stɑr]	**Stern** /ʃtɛn/
* trials were re-ordered for English-German participants and 2 trials added to ensure an equal number of trials per condition													

***Table 25**: Experiment 7. Stimuli for English-Dutch and English-German 30-month-old bilinguals in a semantic priming study*

For English-German participants, the German translations for English words resulted in some trials changing condition from the English-Dutch stimuli set due to a change in cognate/non-cognate status. This rendered two trials redundant and required two additional trials to be added to ensure there were an equal number of trials for all conditions (see **Table 26**). This took the trial total to 26 since we did not want to alter the original experiment but did want to include the two additional trials for the elegance of the experimental design.

				Prime		Target		Distractor	
	Trial	**Prime Type**	**Target Type**	**English**	**German**	**English**	**German**	**English**	**German**
25	unrelated	cognate	non-cognate	**fish** /fɪʃ/	**Fisch** [fɪʃ]	**bike** /baɪk/	**Fahrrad** [ˈfa:ɐ̯ra:t]	**arm** /ɑːm/	**Arm** [arm]
26	unrelated	cognate	cognate	**boat** /bəʊt/	**Boot** [bo:t]	**arm** /ɑːm/	**Arm** [arm]	**bike** /baɪk/	**Fahrrad** [ˈfa:ɐ̯ra:t]

***Table 26**: Experiment 7. Additional trials for the adapted English-German bilingual semantic priming study on 30-month-olds*

Procedure

The experiment was run online on the Gorilla Experiment Builder platform. The procedure was almost identical to that in Experiment 6. In addition to the monolingual procedure, bilingual participant and demographic questionnaires included additional questions relating to language exposure which were taken from the LEQ (Cattani et al., 2014). This was to check eligibility for the study regarding the amount of language exposure a child had in each language[20].

A further modification was to the word list. In addition to a parent marking whether prime and target words were known to the child in English, they were also asked if these words were known to the child in their second language.

All correspondence and instructions, including the consent form, were translated into Dutch (with the help of the MPI lab assistants) so that parents could either follow the study with English or Dutch instruction. The experiment itself was only ever run in English.

The Dutch (Zink & Lejaegere, 2002) /German (Szagun et al., 2009) CDI (lexical component only) was sent to participants by email after task completion, to answer in addition to the English CDI III (word component only), which was already integrated into the online Gorilla Experiment Builder procedure.

3.5.3 Results

Data collection is ongoing and providing analyses on four participants is not reliable. However, descriptive data at this point show that the mean number of valid trials per participant was 21.6 (*SD*= 2.67) and PLT was higher for unrelated trials (*M*= 0.46, *SD*= 0.30) compared to related trials (*M*= 0.45, *SD*= 0.11) in bilingual 30-month-olds.

[20] Language exposure was also informally assessed during email correspondence with interested participants by asking parents for an approximation of exposure to each language to ensure only eligible participants took part.

3.5.4 Discussion

This key study was due to be conducted during a research placement at the Max Planck Institute in Nijmegen, The Netherlands, during the period March 2020 – June 2020. Arrangements with the institution had been made and the study approved, however, due to the global pandemic, this placement was cancelled and though every effort was made to delay the placement to another time, it had to be cancelled completely. Recruitment of Dutch-English bilingual children would have been through this institution, and this was the motivation for designing the stimuli for a Dutch-English population.

When other research studies moved to online testing, this study also moved online, and great effort was made to find eligible bilingual participants for the study. This began by recruiting through the Plymouth BabyLab database and Facebook page. This was not very successful, mostly due to participants speaking an additional language, falling out of the age range, or not having enough exposure to English and Dutch in a typical week. Specialist Facebook Groups were contacted next, to promote the study on their pages, as well as contacting other UK BabyLabs to ask for help with recruitment, and Dutch organisations such as the Dutch Embassy in London. When this amounted to only little uptake, the study was adapted (i.e. in relation to translation equivalents and cognates) so that German-English bilingual families could participate. Again, specific groups were made aware of the study on social media platforms (e.g. the German Saturday school in Leicester), as well as advertising through the BabyLab database and our own social media pages. The final recruitment effort was to pay for targeted advertising on Facebook so that English-Dutch and English-German bilingual families with young children would encounter the advert on their news feed, or it would appear as an advert in relevant Facebook groups.

Only four participants had been successfully tested before the unexpected findings from Experiment 6 emerged during analysis. This caused us to stop actively recruiting for Experiment 7 until we could determine what was causing the looking behaviour in Experiment 6.

The intention is to continue this study beyond the PhD thesis, adapting the stimuli set if necessary. Studying bilingual language development was the key focus of my PhD before the pandemic prohibited me from going abroad to collect data. This unexpected turn of events forced me to adapt my research questions and study design to a predominantly monolingual focus and I plan to return to my initial area of interest when possible.

3.6 Experiment 8: Semantic Priming in Monolingual Children, Online (3 years)

3.6.1 Introduction

To understand why the unexpected finding of increased looking to the target on unrelated trials occurred in Experiment 6 (testing 30-month-old monolinguals), Experiment 8 aims to test if age was a factor in this outcome by running the same experimental design and stimuli set from Experiment 6, on 3-year-olds. There was a significant finding of semantic priming in Experiment 5, testing 3-year-olds, so by testing the same age group but with a different stimuli list (i.e. that used in Experiment 6, on 30-month-olds), we might understand if age contributed to the asymmetric priming effect, or if it was due to another variable. If a semantic priming effect is found in this age group, using the stimuli list from Experiment 6, then age might be taken as a contributing factor of the outcome. If this study replicates the same asymmetric finding as Experiment 6 then the root of the cause must either be the modality/ experimental procedure, or the stimuli list.

3.6.2 Method

Participants

Eighteen 30-month-old healthy, English monolingual toddlers were tested (7 boys, 11 girls). The average age of participants was 37 months, 6 days (*Range* = 35 months 21 days - 39 months 24 days). Two additional participants were excluded for technical reasons.

Materials

The materials were identical to those used in Experiment 6.

Procedure

The experiment was run online on the Gorilla Experiment Builder platform. The procedure was identical to that in Experiment 6.

3.6.3 Results

Data Processing and Analysis

Webcam recordings were prepared and coded as per Experiment 5 and 6. The main coder coded eye movement for all participants. A second coder independently coded a 10% randomly selected sample of the same data. Inter-rater reliability agreement between coders was 90% with a Cohen's Kappa k of 0.64, indicating substantial agreement.

The same exclusion criteria were applied as for Experiment 5 and 6. Trials were excluded which were: attended to for less than 750ms; under 2500ms in duration; contained a prime or target word unknown to the child.

Descriptive Statistics

Out of a possible 432 trials (a maximum of 24 trials for each of the 18 participants), a total of 385 trials were included for analysis. Reasons for exclusion were due to: insufficient trial length, taken as <2500ms (39 trials or 10.13% of trials); inattentiveness, measured as <750ms spent looking at the screen per trial (6 trials or 1.56% of trials); prime or target word unknown to child (1 trial or 0.26% of trials); technical error (1 trial or 0.26% of all trials).

Zero participants had to be replaced due to not meeting the minimum threshold number of trials, per condition.

The average number of valid trials per participant was 21.39 (*SD*= 1.50). This high number indicates children were engaged in an online looking task when administered in the home.

There was no effect of gender on rate of response $t(16) = -2.02$, $p = .06$, though boys responded on average more (*M*= 22.13, *SD*= 1.55) than girls (*M*= 20.8, *SD*= 1.23).

Proportion of Looking Time to the Target

Proportion of looking time was calculated as per Experiment 5 - 7 (i.e. PLT to target/ (PLT to target + PLT to distractor)) between 200ms – 2000ms (i.e. 200ms after visual stimuli onset and 1800ms free-looking time)

A two-tailed, paired t-test was run on PLT for related and unrelated trials. Three-year-olds looked longer on unrelated trials (*M*= 0.53, *SD*= 0.1) compared to related trials (*M*= 0.49, *SD*= 0.09) but the difference was not significant, $t(17) = -2.05$, $p = .06$, $d = -.48$. This result is visualised in **Figure 15** (the white square indicates the mean).

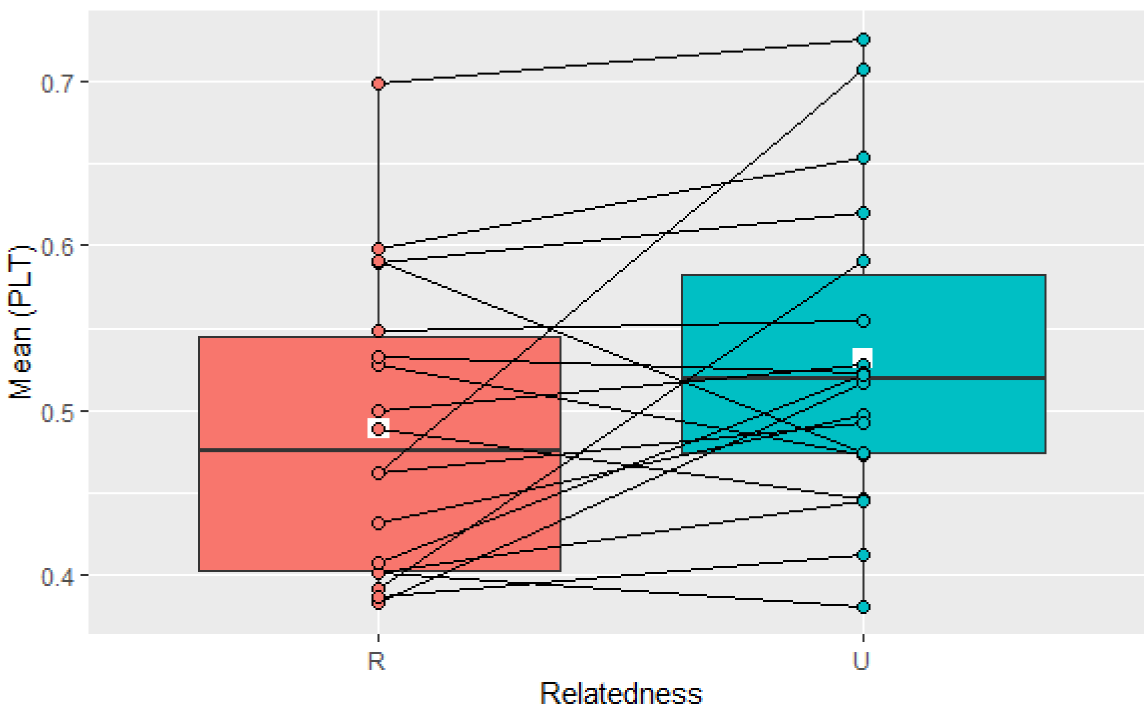

Figure 15: *Experiment 8. Proportion of looking time to the target on semantically related and unrelated trials using the stimuli list from Experiment 6 on 3-year-olds*

Follow-up, one sample t-tests were performed to investigate if looking was above chance on related and unrelated trials. Comparisons to chance (0.5) with PLT indicated that 3-year-olds did not look significantly above chance in related (t(17) =0-.51, p = .07) or unrelated trials (t(17) = 1.384, p = .09).

In summary this shows no significant difference between unrelated trials and related trials, and target looking above chance, which might indicate a lack of target recognition.

No correlation was found between CDI scores and PLT scores r(16) = -0.11, p = .67.

Time-course Analysis

The PLT to the target for related and unrelated trials was averaged across participants for each 50ms time bin between 0ms and 2000ms and plotted (see **Figure 16**). Visual inspection suggests that the two curves diverge at approximately 200ms and again at 1700ms.

To determine differences between looking behaviour on related and unrelated trials, a cluster analysis was performed (Maris & Oostenveld, 2007), as per Experiments 4 - 7.

A difference in looking behaviour for related and unrelated trials was identified between 1600ms and 1800ms (cluster *t* statistics = -14.02, Monte Carlo *p* = .01), though this period is likely to be after initial word recognition has occurred. This area is marked by a box in **Figure 16**. Since no difference in looking behaviour was found at the early point of word recognition, this suggests a similar looking behaviour on related and unrelated trials for participants in this experiment.

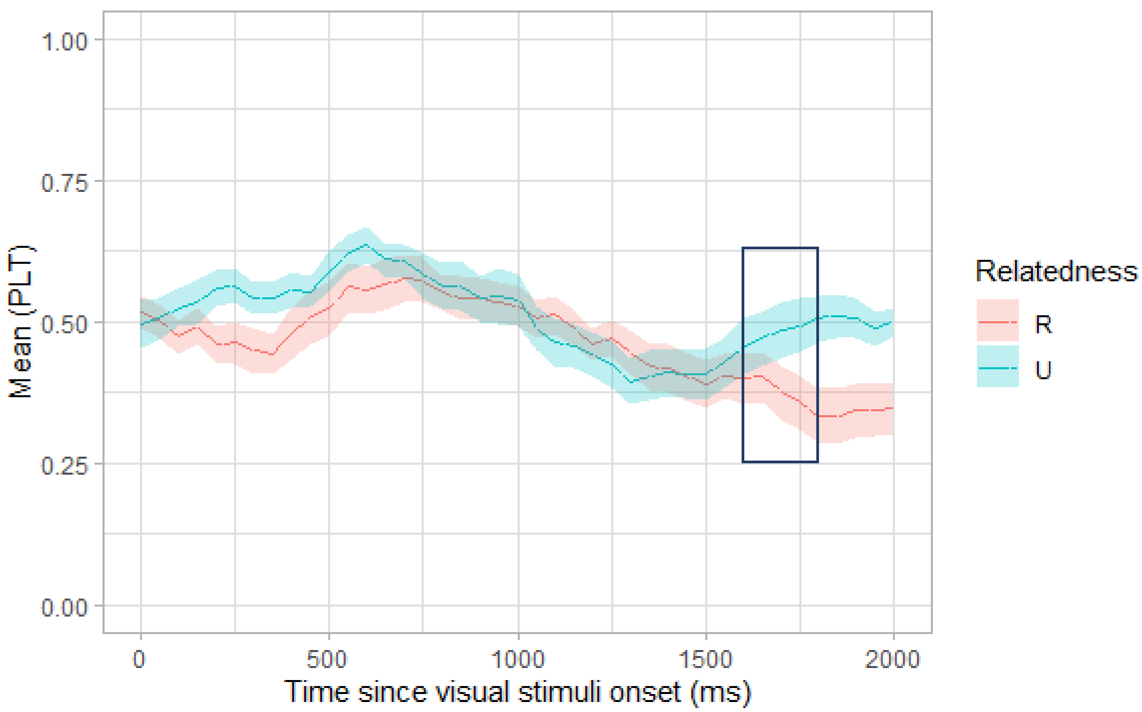

Figure 16: *Experiment 8. Time-course of looking behaviour in 3-year-olds for semantically related and unrelated trials using the stimuli set from Experiment 6*

3.6.4 Discussion

This online semantic priming study used the stimuli list from Experiment 6 (Age= 30-month-olds; Results= high PLT to target on unrelated trials) on the same age group tested in Experiment 5 (Age= 3-years-old; Results= high PLT to target on related trials) to determine if age was the locus of the 'asymmetric priming' effect in Experiment 6. The findings from this study replicated those in Experiment 6 in as far as unrelated trials were attended to for longer than related trials, however, unlike Experiment 6, there was no significant difference between PLT for related and unrelated trials and no above chance looking on either trial type. The similar pattern of increased looking on unrelated trials suggests that age was not the locus of the effect. The time-course analysis showed no qualitative difference in looking behaviour at the point of word recognition (i.e. no significant clusters found) when prime and target were related or unrelated. This differs to the finding in Experiment 6. In Experiment 6, there was increased looking behaviour toward the target image on unrelated trials compared to related trials and this looking behaviour was sustained for the duration of a trial. This is supported by a statistically significant *p* value in the t-test, and significant *t* values in a cluster analysis. By contrast, visual inspection of the time-course of looking behaviour in this experiment converges for some of the trial, although the initial looking behaviour at the point of word recognition does differ. The time-course analysis suggests that 3-year-olds initially looked more towards the target image on unrelated trials compared to related trials, but the looking behavior for both trial types becomes very similar from about 750ms.

We now consider why 30-month- and 36-month-olds had increased PLT on unrelated trials. The reason for this could be due to the target images on unrelated trials being more visually appealing than target images on related trials, or that distractor images on related trials are more visually attractive than the targets on related trials. Although care was taken when selecting stimuli, this is a potential confound that might have contributed to the effect found in Experiment 6. Cluster analyses did identify the end of the trial (1600-1800ms) as a point of divergence in looking behaviour on related and unrelated trials. This could point to target images in unrelated trials, or distractor images in related trials being more engaging to 3-year-olds.

Before exploring whether some images were more visually appealing to participants, a replication study testing 30-month-olds will be run to help clarify the role that age plays in determining the magnitude of the 'asymmetric priming' effect. The effect of experimental modality (online vs. in-lab) has not yet been explored, so these two features will drive the next experiment.

3.7 Experiment 9: Semantic Priming in Monolingual Children, In-lab (30 months)

3.7.1 Introduction

Experiment 9 coincides with the re-opening of the Plymouth BabyLab after the Covid-19 pandemic, making in-lab testing possible again. The aim of this study was to validate the online experimental design of a semantic priming study, by running it again on 30-month-olds in the lab. This enables evaluation of the effect of running the study online with this age group, and to compare the findings found in two older age groups doing the study online, when the stimuli were different.

If a semantic priming effect is found in this age group, then the modality of the experiment may contribute strongly to looking behaviour when the task is administered online. If there is an 'asymmetric effect' that is pronounced and sustained for the duration of a trial, like in Experiment 6, we might conclude that age contributes to the looking behaviour, but modality contributes less. If there is an 'asymmetric effect' at the point of word recognition but not sustained for the full trial, like in Experiment 8, then we can conclude that 30-month-olds are behaving in the same way as 3-year-olds and so age is not a confounding factor, but that modality could be. The final factor not yet explored in depth is the difference in stimuli sets, which could be confounding all experimental outcomes. To explore this as a possibility, the findings from Experiment 6, 8, and 9 will then be compared by running an item-level analysis.

3.7.2 Method

Participants

Eighteen 30-month-old healthy, English monolingual toddlers were tested (9 boys, 9 girls). The average age of participants was 29 months, 21 days (*Range* = 29 months 10 days - 30 months 9 days). One additional participant was excluded due to significant data loss.

Materials

The materials were identical to those used in Experiment 6.

Procedure

The procedure was identical to that used in Experiment 6, but the experiment was run in the Plymouth University BabyLab. On arrival, the parent and child were given time to settle into the BabyLab. The child played while the parent completed a consent form, word list and demographic information. When ready, the parent and toddler were taken to the eye-tracking booth where the child was placed in a car seat approximately 65cm away from the computer monitor which was fitted with the Tobii TX300 eye tracker with a sampling rate of 120 Hz. The parent stood behind the child, out of detection from the eye-tracker. A five-point calibration was performed before starting the experiment. The in-lab delivery of experimental trials, including timing, was identical to the Gorilla Experiment Builder online procedure apart from one feature: trials were child-led so that a trial was only initiated by the experimenter (seated outside the eye-tracking booth) when a child was actively looking at the screen. The experimenter was blind to the stimuli conditions. After testing, the parent and child were given a certificate and small gift for participating, during the debrief.

3.7.3 Results

Descriptive Statistics

Out of a possible 432 trials (a maximum of 24 trials for each of the 18 participants), a total of 366 trials were included for analysis. Reasons for exclusion were due to inattentiveness, measured as <750ms spent looking at the screen per trial (53 trials or 12.27% of trials); prime or target word unknown to child (14 trials or 3.24% of trials); technical error (7 trials or 1.62% of all trials). Zero participants had to be replaced due to not meeting the minimum threshold number of trials, per condition.

The average number of valid trials per participant was 20.33 (*SD*= 3.46). This indicates children were engaged in the task.

There was no effect of gender on rate of response $t(16) = 0$, $p = 1$ and boys responded on average (*M*= 20.33, *SD*= 3.24) at the same rate as girls (*M*= 20.33, *SD*= 3.87).

Proportion of Looking Time to the Target

Proportion of looking time was calculated as per Experiment 5-8 (i.e. PLT to target/ (PLT to target + PLT to distractor)).

A two-tailed, paired t-test was run on PLT for related and unrelated trials. Thirty-month-olds looked longer on unrelated trials (*M*= 0.55, *SD*= 0.08) compared to related trials (*M*=

0.49, *SD*= 0.08) and the difference was significant, $t(17) = -3.84$, $p = .001$, $d = -.91$. This data is visualized in **Figure 17** (the white square indicates the mean).

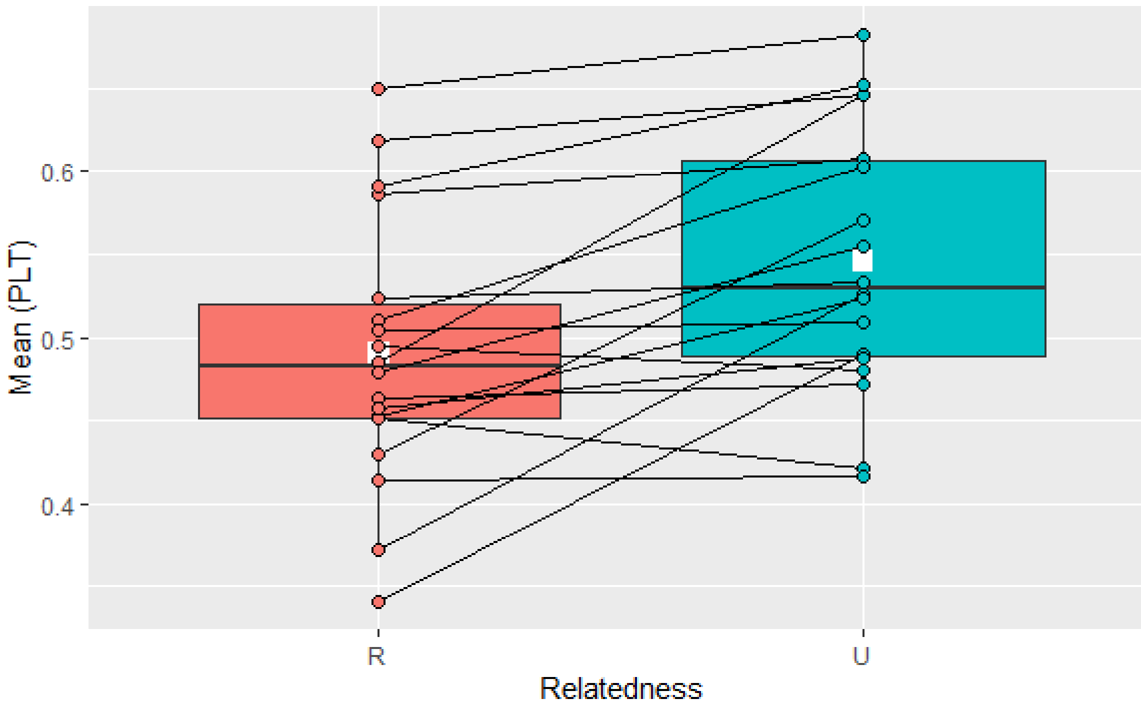

Figure 17: *Experiment 9. Proportion of looking time to the target on semantically related and unrelated trials using the stimuli list from Experiment 6 on 30-month-olds tested in the lab*

Follow-up, one sample t-tests were performed to investigate if looking was above chance on related and unrelated trials. Comparisons to chance (0.5) with PLT indicated that 30-month-olds did not look significantly above chance in related trials ($t(17) = -0.51$, $p = .69$) but they did on unrelated trials ($t(17) = 2.42$, $p = .01$).

In summary this shows a significant difference between unrelated and related trials, such that participants looked above chance at the target on unrelated trials, but not related trials. This suggests word recognition for unrelated trials but not related trials.

Time-course Analysis

The full trial duration post image onset was observed. The PLT to the target for related and unrelated trials was averaged across participants for each 50ms[21] time bin between 2800ms and 5300ms and plotted (see **Figure 18**). The prime and carrier lasted 2200ms, followed by a 200ms ISI and the target word. There was an SOA of 400ms at which point the two images were on screen by 2800ms. Looking time was set to 2500ms after this.

Visual inspection suggests that the two curves diverge at approximately 800ms.

[21] This was automated using the eyetrackingR package in R Studio.

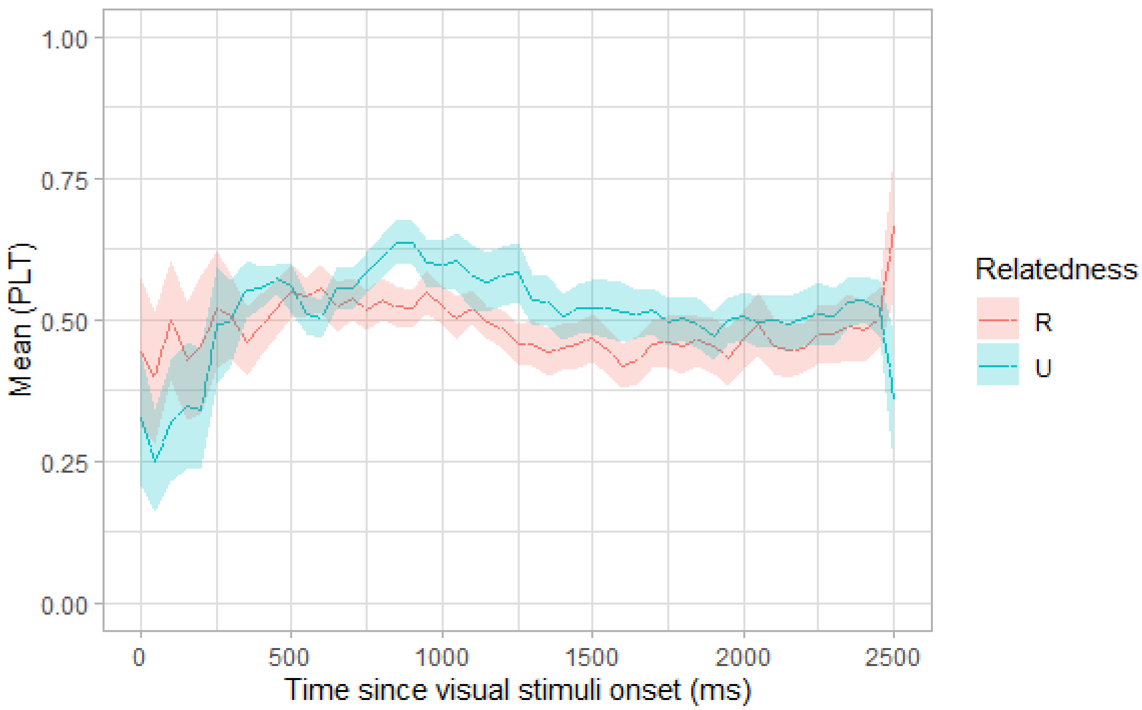

Figure 18: *Experiment 9. Time-course of looking behaviour in 30-month-olds for semantically related and unrelated trials using the stimuli set from Experiment 6, tested in-lab*

To determine differences between looking behaviour on related and unrelated trials, a cluster analysis was performed (Maris & Oostenveld, 2007), as per Experiments 5 - 8. No clusters were found, indicating no difference in looking behaviour at any particular time period between related and unrelated trials.

3.7.4 Discussion

This in-lab semantic priming study on 30-month-olds used the same stimuli list as Experiment 6 and 8, with a similar set of findings: increased looking to the target visual stimulus on unrelated trials relative to related trials. This confirms that age did not overtly drive behaviour, since Experiment 9 tested 30-month-olds while Experiment 8 tested 3-year-olds, with comparable looking behaviour in the two studies. Findings from this experiment also suggest that experiment modality did not cause increased PLT on unrelated trials since Experiment 9 was run in the lab, while Experiment 6 and 8 were done online.

However, the saliency of certain visual stimuli may have driven looking behaviour across ages and modalities on unrelated trials, which necessitates further analyses and will be dealt with in the next section.

3.7.5 Comparing Semantic Priming Experiments 6, 8,and 9

Results from Experiments 6, 8, and 9 showed a similar pattern of findings, especially when contrasted to Experiment 5 (see **Table 27** for an overview).

	Experiment 6: 30 months, online	Experiment 8: 36 months, online	Experiment 9: 30 months, in-lab	Experiment 5: 36 months, online, different stimulus set
Response rate	*M* = 21.6, *SD* = 2.67	*M* = 21.39, *SD* = 1.50	*M* = 20.33, *SD* = 3.46	*M* = 23.00, *SD* = 1.99
PLT Related Trials	*M* = 0.54, *SD* = 0.08	*M* = 0.49, *SD* = 0.09	*M* = 0.49, *SD* = 0.08	*M* = 0.51, *SD* = 0.07
PLT Unrelated Trials	*M* = 0.61, *SD* = 0.08	*M* = 0.53, *SD* = 0.01	*M* = 0.55, *SD* = 0.08	*M* = 0.48, *SD* = 0.07
T-test significance	*t* (39) = 4.78, ***p* < .0001**, *d* = -.76	*t* (17) = -2.05, *p* = .06, *d* = -.48	*t* (17) = -3.84, ***p* = .001**, *d* = -.91	*t* (39) = 2.39, ***p* = .02**, *d* = .38
Above chance looking Related Trials	(*t* (39) = .2.93, ***p* = .003**)	(*t* (17) =0-.51, *p* = .07)	(*t* (17) = -0.51, *p* = .69)	(*t* (39) = 2.93, ***p* = .003**)
Above chance looking Unrelated Trials	(*t* (39) = 8.31, ***p* <.0001**)	(*t(17)* = 1.384, *p* = .09)	(*t* (17) = 2.42, ***p* = .01**).	(*t* (39) = -1.76, *p* = .096)
Timecourse & Cluster Analysis	400 and 900ms, cluster t statistics = -30.42, Monte Carlo ***p* = .0002**	1600ms and 1800ms, cluster *t* statistics = 14.42, Monte Carlo ***p* = .01**	No clusters	450 and 850ms, cluster t statistics = 27.99, Monte Carlo ***p* = 0**

***Table 27**: Key findings from Experiments 5, 6, 8, and 9 compared. (Response rate: number of valid trials out of 24.)*

One key difference is the stimuli chosen in Experiment 5 compared to Experiments 6, 8, and 9. For this reason, the data from Experiments 6, 8, and 9 were combined and the mean PLT calculated per target stimulus, per Experiment and Relatedness condition.

Two ANOVAs were run on PLT for the related and unrelated conditions separately. Experiment was taken as a repeated measure. On related trials the effect of experiment on PLT was statistically significant, $F(2, 36)= 5.03$, $p = .01$, $\eta 2= .06$. Post-hoc pairwise comparisons were performed with a Bonferroni adjustment to identify the locus of the difference. The PLT on related trials differed significantly between Experiment 6 and Experiment 8 ($t(18)= 3.22$, $p = .01$), such that PLT was on average higher in Experiment 6 ($M= 0.51$, $SD= 0.09$) compared to Experiment 8 ($M= 0.45$, $SD= 0.09$), but not different compared to Experiment 9 ($M= 0.48$, $SD= 0.08$).

On unrelated trials the effect of experiment on PLT was statistically significant, $F(2, 36)= 27.14$, $p < .001$, $\eta 2= .1$. Post-hoc pairwise comparisons with a Bonferroni adjustment revealed an effect of experiment such that PLT differed significantly between Experiment 6 and Experiment 8 ($t(18)= 6.09$, $p < .001$), and between Experiment 6 and Experiment 9 ($t(18)= 5.84$, $p < .001$). PLT was significantly higher in Experiment 6 ($M= 0.59$, $SD= 0.12$), compared to Experiment 8 ($M= 0.51$, $SD= 0.08$) and compared to Experiment 9 ($M= 0.54$, $SD= 0.11$).

Taken together, this indicates a significant difference between Experiment 6 compared to the other two experiments (see **Figure 19** for mean PLT by experiment and relatedness).

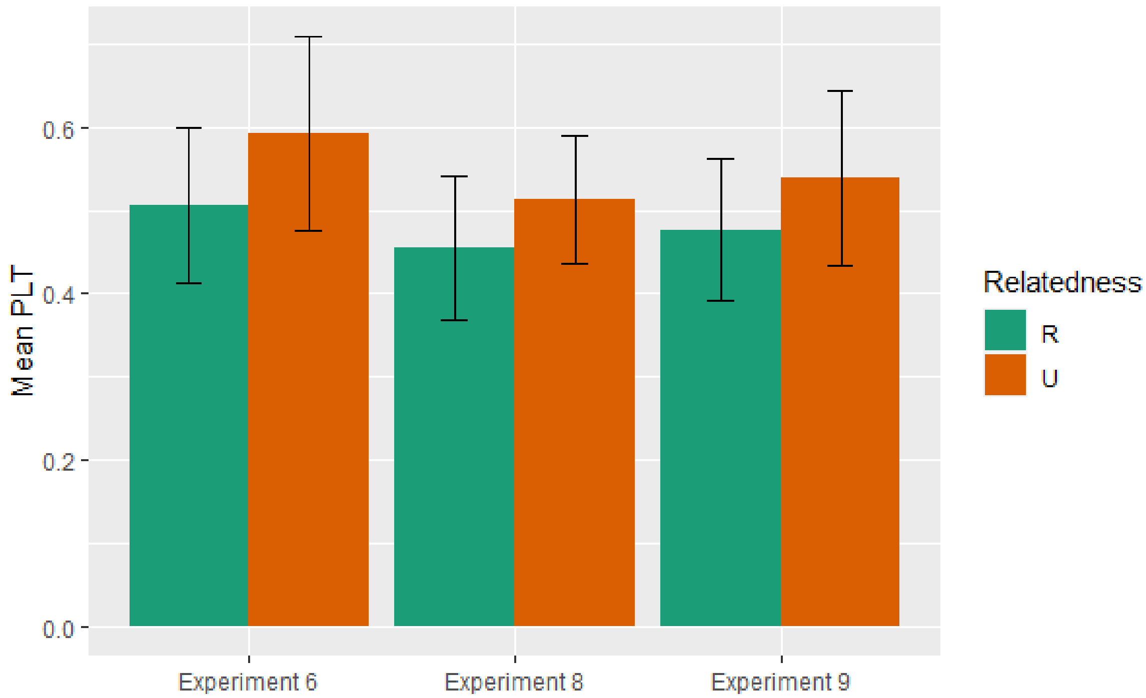

Figure 19*: Experiment 6, 8, & 9. Mean PLT, grouped by target stimulus, per Experiment and Relatedness*

To follow up on this finding, the mean PLT per target stimulus[22] was calculated (see **Figure 20** for the related condition and **Figure 21** for the unrelated condition). Visual inspection indicates that some stimuli attracted a high PLT in the related (e.g. *bed* and *pen* show PLT > 0.6) and unrelated (e.g. *book, doll, milk,* and *window* show PLT > 0.7) conditions and a low PLT in related (e.g. *bowl*, *doll*, *frog* have PLT < 0.3) and unrelated conditions (e.g. *house, bed* and *star* have PLT < 0.4).

[22] Not all items acted as a target in both the related and unrelated conditions. The missing items from the related condition were: ball, dog, flower, heart, window. The missing items from the unrelated condition were: bird, bus, coat, hand, tree.

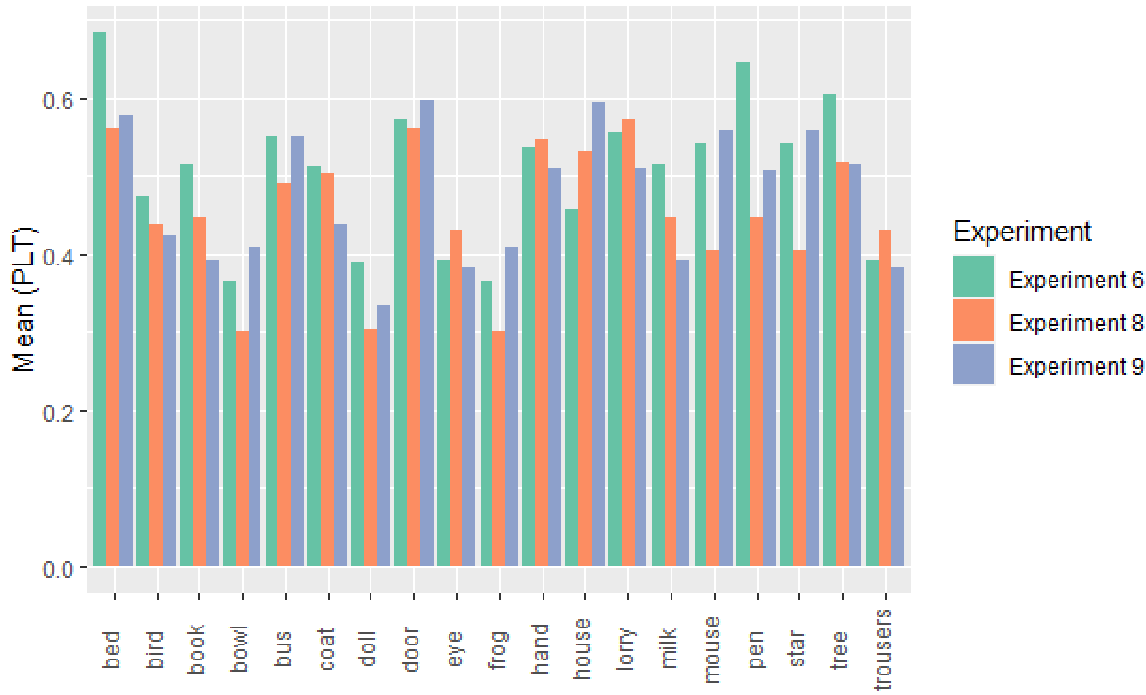

Figure 20: *Experiment 6, 8, & 9. Mean PLT per item on <u>related</u> trials*

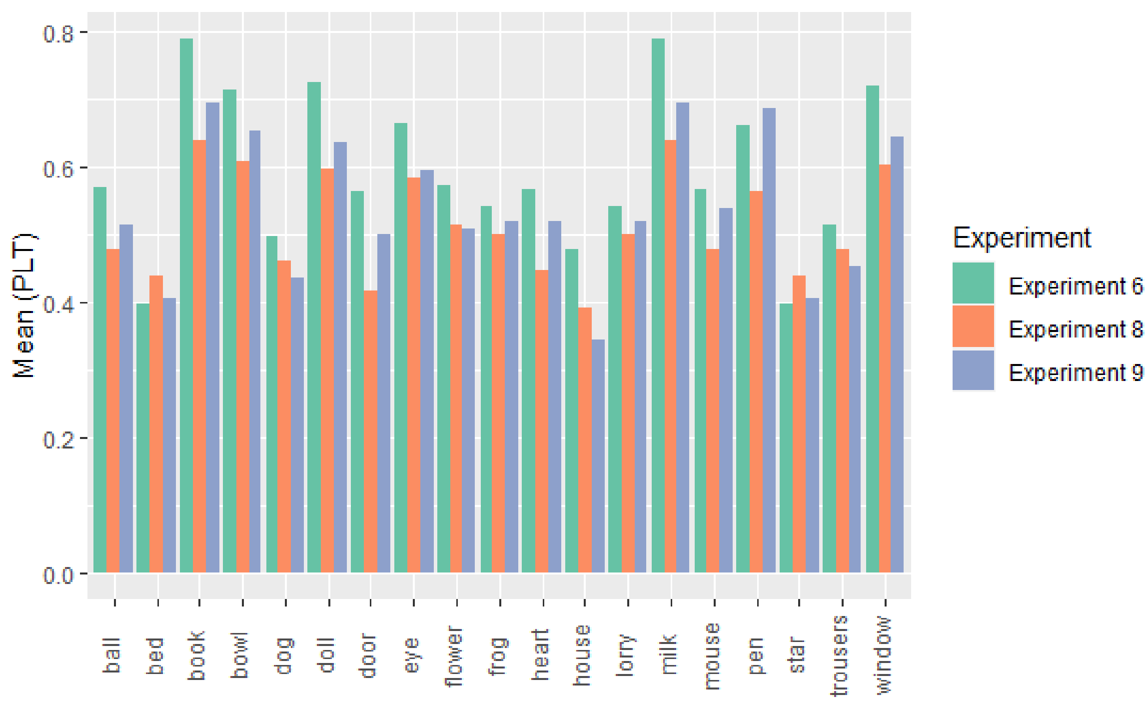

Figure 21*: Experiment 6, 8, & 9. Mean PLT per item on <u>unrelated</u> trials*

To confirm that certain stimuli attracted more attention than others, outliers were calculated as the overall mean PLT +/- 2 standard deviations. This resulted in two outliers: *book* (PLT= 0.71) and *milk* (PLT= 0.71) in the grouped dataset, both in the unrelated condition.

These items were removed from the combined dataset and the two one-way ANOVAs were re-run on PLT for the related and unrelated conditions separately. There was no effect of Experiment in the related condition, $F(2,48) = 1.28$, $p = .29$. This indicates that the PLT on related trials is the same for Experiment 6 (*M*= 0.51, *SD*= 0.1), Experiment 8 (*M*= 0.46, *SD*= 0.09), and Experiment 9 (*M*= 0.49, *SD*= 0.08). See **Figure 22** for a visualisation of this data.

There was no effect of Experiment in the unrelated condition, $F(2,48) = 2.74$, $p = .07$, also indicating no difference in PLT on unrelated trials for Experiment 6 (*M*= 0.57, *SD*= 0.10), Experiment 8 (*M*= 0.50, *SD*= 0.07), and Experiment 9 (*M*= 0.52, *SD*= 0.10). See **Figure 22** for a visualisation of this data.

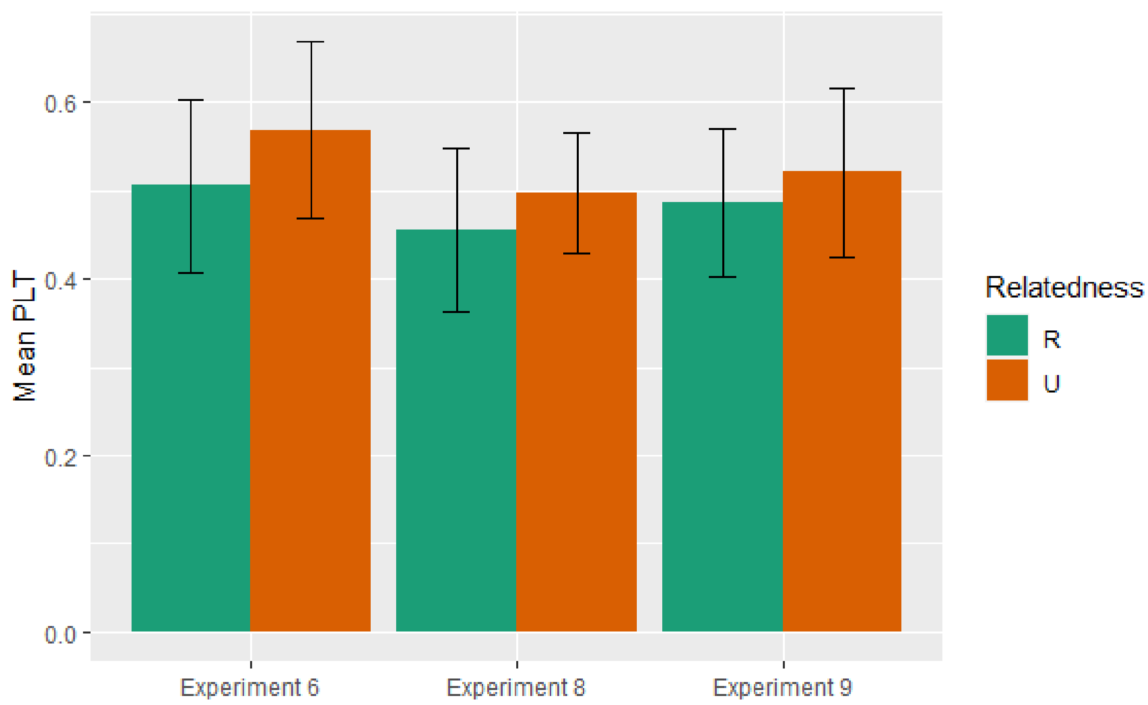

***Figure 22**: Experiment 6, 8, & 9. Mean PLT, grouped by target stimulus, per Experiment and Relatedness after excluding outliers*

Having statistically established no difference between Experiment 6, 8, and 9, after excluding outliers, we now look at the relatedness of word pairs in more detail (see **Table 28**).

	Trial	**Prime**	**Target**	**Prime-Target (FSG) using Nelson et al. (1998)**	**Prime-Target (FSG) using Moss and Older (1996)**	**Distractor**	**Prime-Distractor (FSG) using Nelson et al. (1998)**	**Prime-Distractor (FSG) using Moss and Older (1996)**	**Infant Priming Study using the word pair**
1	Associative	park	tree	0.115	0.083	coat	No assoc	No assoc	Arias-Trejo & Plunkett (2009)

2	Associative	glass	milk	No assoc	No assoc	book	No assoc	No assoc	Floccia et al. (2020) cup-milk
3	Associative	garden	house	No assoc	Not a cue	bus	No assoc	Not a cue	
4	Taxonomic	sock	trousers	No assoc	Not a cue	eye	No assoc	Not a cue	
5	Taxonomic	cat	bird	No assoc	0.021	doll	No assoc	No assoc	
6	Taxonomic	foot	hand	(feet) 0.095	0.021	door	No assoc	No assoc	Jardak & Byers-Heinlein (2019) hand-foot
7	Taxonomic	train	bus	0.03	0.062	pen	No assoc	No assoc	Jardak & Byers-Heinlein (2019)
8	Taxonomic	boots	coat	No assoc	No assoc	frog	No assoc	No assoc	
9	Taxonomic	spoon	bowl	No assoc	No assoc	window	No assoc	No assoc	Singh (2014)
10	Taxonomic	car	lorry	(truck) 0.111	(truck) 0.024	tree	No assoc	No assoc	Styles & Plunkett (2009)
11	Taxonomic	chair	bed	0.013	Not a cue	hand	No assoc	Not a cue	
12	Taxonomic	chicken	mouse	No assoc	Not a cue	ball	No assoc	Not a cue	

Table 28*: Experiment 6, 8, & 9. Related Stimuli Pairs (No assoc= No association noted)*

To compare associatively related word pairs (1-3 in **Table 28**) to taxonomically related word pairs (4-12 in **Table 28**), we took the combined PLT data of Experiment 6, 8, and 9 and ran a t-test to compare the PLT for the two types of relatedness. There was no significant effect of relatedness on PLT, $t(10) = -0.38$, $p = .72$, despite associatively related word pairs (M= 0.50, SD= 0.07) receiving slightly more attention compared to taxonomically related word pairs (*M*= 0.49, *SD*= 0.07). It must be stressed that there were more taxonomically related trials (*N* = 684) compared to associatively related trials (*N* = 228) which makes the analysis too under-powered to be reliable due to too few instances of the latter, but this exploratory analysis suggests no significant difference between type of relatedness on PLT, despite this finding in other studies (McNamara, 2005).

We will now consider the individual visual stimuli, to explore why unrelated trials attracted more attention than related trials. In the individual experiment datasets, there was no significant difference in PLT for related and unrelated trials in Experiment 8, but there was in Experiment 6 and Experiment 9. However, after excluding outliers and re-calculating PLT for related and unrelated trials for each experiment, only Experiment 6 showed a significant difference in PLT on related and unrelated trials (see **Table 29**). This suggests that Experiment 6 may have been an outlier.

	Experiment 6. 30 months, online	**Experiment 8. 36 months, online**	**Experiment 9. 30 months, in-lab**
PLT Related Trials	*M* = 0.54, *SD* = 0.07	*M* = 0.49, *SD* = 0.09	*M* = 0.50, *SD* = 0.1
PLT Unrelated Trials	*M* = 0.59, *SD* = 0.09	*M* = 0.52, *SD* = 0.10	*M* = 0.53, *SD* = 0.08
T-test significance	*t* (39) = -3.29, *p* = .002	*t* (17) = -1.67, *p* = .11	*t* (15) = -1.98, *p* = .07

Table 29*: Experiment 6, 8, & 9 summary PLT statistics after outliers excluded*

For example, returning to **Figure 21** it is clear that Experiment 6 had some of the highest PLT per target stimulus in the unrelated condition (e.g. *book*, *doll*, *milk*, *window* had PLT > 0.7). We took the mean PLT from the combined data and calculated outliers as mean +/- 2SDs in the Experiment 6 dataset and found five outliers. This includes *milk* and *book* as per the group outliers (though with higher PLT in Experiment 6), and three additional outliers: *bowl*, *doll*, and *window* (see **Table 30**).

Visual Stimulus	Relatedness	Mean (PLT)
book	unrelated	0.79
bowl	unrelated	0.71
doll	unrelated	0.72
milk	unrelated	0.79
window	unrelated	0.72

Table 30*: Experiment 6. Outliers measured as mean PLT +/- 2SDs*

Re-calculating the mean PLT for related and unrelated trials in Experiment 6, after excluding the five outliers from the dataset, shows that there was no difference in PLT, $t(39) = 0.16$, $p = .87$. Children looked at the target on related trials (M= 0.56, SD= 0.09) at the same rate as on unrelated trials (M= 0.56, SD= 0.09).

3.7.6 General Discussion of all Priming Experiments

In summary, no effect of semantic priming was found in Experiment 6, 8, or 9, even after excluding outliers from our analysis. We did not find a difference between trials which were associatively related compared to trials which were taxonomically related, when trying to explain the lack of a semantic priming effect.

By removing outliers in follow-up analyses, we demonstrated no effect of Experiment on our combined dataset, which combined data from Experiment 6, 8, and 9. Removing outliers from the combined dataset also showed no difference in looking behaviour on related and unrelated trials for Experiment 8 and Experiment 9. In contrast, Experiment 6 did still show a significant finding of PLT for related and unrelated trials after removing group outliers (i.e. *book* and *milk*), but when excluding additional outliers which were specific to Experiment 6 (i.e. *bowl*, *doll*, *window*), the difference in PLT by relatedness was no longer significant. Together this suggests that Experiment 6 likely had outlying findings when compared to Experiment 8 and 9, which may be due to certain visual stimuli which attracted a large proportion of attention, particularly on unrelated trials.

For example, *book* attracted a high PLT when used as a target in unrelated trials. A reason for the saliency of *book* is likely due to its bright, colourful format, especially when paired with the neutral target image of *milk*. This is coupled with the fact that *book* acting as a target was paired with the distractor *milk*, and *milk* acting as a target was again paired with *book* as a distractor. This was not intentional and likely due to an oversight while developing experimental materials that could be used for English-Dutch/German bilinguals in Experiment 7. A long list of criteria regarding phonology, syllables, translation equivalents and cognate status for the bilingual study seems to have resulted in a lack of rigor given to the stimuli selection for the baseline monolingual study. As such, *book* paired with *milk* in block 1 and block 2 of the experiment might have meant a participant was primed through repetition by the two images in block 1 and so drawn to the more visually appealing *book* in the pair, in block 2.

What is clear through analysis, is that the design of stimuli list orders was imperfect through the repetition of word pairs and even the repetition of visual stimuli (i.e. once as a target, and once as a distractor). It must be noted, however, that the same method of experimental list design was used in Experiment 5 (i.e. a visual stimulus was seen once as the target and once as the distractor, in two different blocks) and follows the experimental design of other studies (e.g. Floccia et al., 2020). However, these studies still showed an effect of semantic priming in spite of this.

While some attempt has been made to explain the unexpected finding of a high proportion of looking behavior on unrelated trials, by looking at the visual stimuli used in individual trials, one must be careful to not draw conclusions on this basis alone as there could be an interaction between stimuli and modality. For example, our follow-up analyses in Experiment 6, 8, and 9 more closely resemble our findings from Experiment 5. That is, there was no evidence of semantic priming when word pairs were taken from adult norms or when word pairs were intuitively deemed related. This might suggest an online modality is not sensitive enough to show an effect of priming for all WAs (word associations), since we only found an effect of semantic priming (coupled with word recognition, demonstrated by above chance looking) when the word pairs originated from child specific WAs, documented in Chapter 1 of this thesis. However, modality alone cannot account for our unexpected findings. We should have observed an effect of semantic priming in Experiment 9 since it was administered in the lab, but we did not. The lack of an effect in this case cannot be attributed to the experiment's modality but might be explained by the stimuli and experimental design. For example, the word pairs were predominantly taxonomic (9 out of 12 trials). Evidence from the adult literature indicates that associatively related word pairs have the potential to double the semantic priming effect compared to purely taxonomic relations between words (McNamara, 2005). It is likely that this, combined with issues relating to the saliency of certain stimuli and errors in the experimental design might be why a lab-based study at 30 months did not evidence an effect of semantic priming.

For future studies, a better experimental design is one which counterbalances a word as a target and a distractor between participants rather than within participants (but note Torkildsen et al., 2007 did not do this). Equally, word pairs which are associatively related should be prioritised, especially when administered in an online modality in which it remains to be seen whether it is a sensitive enough modality to capture the same effects found in infants, in a lab setting. Replication studies testing both age groups on a new, fully counterbalanced stimuli list is a necessary step to fully assert that a semantic priming study can be replicated online, at 30 months and 3 years.

4 General Discussion

Investigating how infants first establish relationships between words is a necessary step towards understanding the qualitative shift children make from an immature lexical-semantic system in which either 1. words are 'islands' of concept-to-word mappings (Anglin, 1970; Arias-Trejo & Plunkett, 2009; McNeill, 1970), which develop into an organised and complex interconnected network of semantic relationships (Friedrich & Friederici, 2011, 2017; Gaskell & Marslen-Wilson, 2002) which characterises a mature, adult system, or 2.

children are born with the aforementioned mature system in place (Rämä et al., 2013) and must organise newly-learned words into it. Since little is known about the word-word associations in infants that establish the network of meanings in the mature semantic system (Arias-Trejo & Plunkett, 2009), this thesis sought to contribute to our understanding by documenting the word associations (WA)s monolingual English children can produce, at as young an age as practicably possible. These WAs were then compared to monolingual English adults, and bilingual infant groups to explore similarities and differences. It was hypothesised that this might highlight the primacy of certain relationships between words and elucidate the difference between connections in a mature lexical-semantic system, compared to an immature one. Similarly, observing how words unique to a bilingual speaker (i.e. translation equivalents and cognates) are organised in the lexical-semantic system might help us understand how a network of meaning establishes that can accommodate more than one language.

An overview of the key findings from each experiment is presented in the section that follows, before moving on to discuss the implications of these findings and the future directions this research could take.

4.1 Overview of Main Findings

Chapter 1 examined the relationships 3-year-olds have between words using a language production, free association task adapted for at-home and online testing. A key aim was to explore if such a task has the potential to reveal insight pertaining to the development of lexical-semantic connections. This was approached by collecting descriptive data of the associated words in a young child's productive lexicon and comparing these to the prime-target relationships commonly found in adult associative norms, searching for similarities and differences. Another aim was to create a resource of WAs in young children to inform stimuli selection in infant studies and move away from a reliance on adult word associative norms.

To that purpose, Experiment 1 collected the related responses of 150 34-47-month-olds using a free association experimental design from previous studies on older children (Newman, 1970) and adapted it to make it more suitable for the target age group, since no other studies had tested at this age previously. As a direct consequence of the Covid-19 pandemic, the task was developed for at-home testing, relying on the parent to assume the role of experimenter. Since very few studies have run studies using a free association task on young children, the first aim was to test if 34-47-month-olds could successfully complete the task. A further aim was to see how the responses given would compare to older children and adults. The findings clearly demonstrated that children between 34 and 47 months could successfully complete a free association task, with the majority of responses (62%) having a relationship to the cue word. Not only were many responses related, but a large proportion were also shared by more than one child (*N*= 432) and not simply idiosyncratic. Most related responses were syntagmatic (74%), replicating previous research for the prevalence of this association type at the early stages of language development. The associative strength for related word pairs featuring in Experiment 1 and adult associative norms (Moss et al., 1996; Nelson et al., 2004) showed a weak positive correlation. However, some of the strongest associations in children were not represented in the adult norms and when looking at the word pairs selected as stimuli in other infant studies (Jardak & Byers-

Heinlein, 2019), only a minority (11%) were found in the WAs produced by 34-47-month-olds in Experiment 1. It was concluded that children do have unique WAs, and they are what one might expect children of this age to hold as associates (e.g. *bed-teddy*), however, these are not captured in adult norms and so would go overlooked when selecting experimental stimuli when accessing adult norms as a resource.

Experiment 2 was a replication study of Experiment 1 using an online methodology aimed at testing the reliability of parental report when testing outside the lab, at the same time as attempting to increase instruction clarity, improve testing homogeneity, and promote task engagement. In the online procedure, pre-recorded videos of puppets demonstrated the WA task and the participant's responses were audio recorded. We explored if the experimental context influences WA responses by observing the exact associations produced and their associative strengths. The age group was adjusted to 36-39 months since Experiment 1 demonstrated only a small effect of age on responses. The cue words eliciting the most frequent WAs in Experiment 1 were selected to create two new lists with ten words per list for Experiment 2. The results indicated that the same proportion and type of responses were produced as per Experiment 1, with thirty-eight exact word pairs replicated (out of 63), with comparable associative strength. The online format used in Experiment 2 did not seem to increase engagement, but parental report when compared to audio recordings indicated objective recording. As with Experiment 1, some of the WAs produced were not represented in adult associative norms (26%).

Using the same procedure as Experiment 2, Experiment 3 tested whether bilingual children would produce the same WAs as their monolingual counterparts. We were interested in observing if bilinguals would show a preference for cognates or translation equivalents to explore if shared features between languages might indicate a more robust lexical-semantic organisation of these word types. Another aim was to compare monolingual to bilingual performance to test whether bilinguals have a more established semantic network (Byers-Heinlein & Werker, 2013). The findings validated that the task could successfully be completed by bilingual 3-year-olds, replicating the same response rate and proportion of related responses found in monolinguals. Translation equivalents were produced over half the time (62%) suggesting that they might facilitate lexical retrieval during language production, with stronger networks for translation equivalents in the lexical-semantic system of a bilingual (Bilson et al., 2015). Cognates were not produced as frequently (5%), likely due to the typologically distant languages spoken by participants.

On comparing responses in Experiment 1, 2, and 3, nineteen imageable, associated noun-noun word pairs were identified as featuring in at least one monolingual study, the bilingual study, and in adult associative pairs. This is presented as the most reliable set of stimuli for future infant priming studies.

Having established the feasibility of WA tasks in toddlers and isolated a set of child-specific word pairs, Chapter 2 sought to develop an online version of a primed IPL methodology from which to explore semantic priming in monolingual and bilingual infants. Predominantly borne out of a necessity due to the Covid-19 global pandemic, the first aim was to design and validate a procedure reliable enough to capture the replicable effect of semantic priming in infants, before observing for any qualitative difference of lexical representations in the developing lexicon of monolingual and bilingual infants.

Experiment 4 was a proof of concept to test the feasibility of running an online preferential looking experiment with infants in their home environments. At the time of testing, there were no other studies attempting to run a time-sensitive online testing procedure with young children using the Gorilla Experiment Builder platform. In Experiment 4, twenty 24-month-olds completed a word recognition task which was automatically administered (compared to an infant-led procedure commonly used in Baby Labs). Participants saw two pictures on screen in a pre-naming phase before an auditorily played word identified the target picture and allowed free looking in a post-naming phase. The aim was to determine if a procedure using an automatic trial delivery, while testing in the home environment on different devices with variable internet speeds, would impede finding the expected effect of target recognition. The intention was to design and validate a simple online procedure that could be adapted for other paradigms, such as the primed IPL paradigm. The results demonstrated that infants looked longer at the target picture relative to the distractor picture after the onset of the auditory label. This indicates that with some modifications to lab-based procedures, an online version of an infant methodology can indeed be run successfully, replicating in-lab findings. This is supported by the strong effect size (*d*= 1.61) which is larger than comparable lab-based experiments (e.g. Bergmann et al., 2018, found an average effect size of *d*= 1.24 in a meta-analysis), indicating noise from an online procedure does not cancel out evidencing an effect, at least not one as large as this. Similarly, data loss of an online procedure (Experiment 4 <5%) may be less than in-lab procedures (e.g. up to 30% according to a meta-analysis performed by Bergmann et al., 2018).

Experiment 5 used the procedural findings from Experiment 4 to adapt the primed IPL task (Styles & Plunkett, 2009) into an online semantic priming procedure. Forty, three-year-olds participated in the online study. There was again little data loss (3%) and a high proportion of valid trials when testing online. A prime word in a carrier phrase preceded the target label of one of two visually presented pictures. The relationship between the prime and target word was either associatively related or unrelated. Associative relatedness was based on the WAs documented in 3-year-olds' responses to a free association task in Chapter 1 of this thesis, and on adult associative norms (Moss et al., 1996; Nelson et al., 2004). The aim was to determine if unique child WAs found in Experiments 1-3 would demonstrate a measurable priming difference compared to adult-specific associations and associations found in both adults and children. Results did indeed reveal an effect of WA type such that related word pairs with the highest PLT (proportion of looking time calculated as: time spent looking at the target/ time looking at target or distractor) were those taken from the productive vocabularies of 3-year-olds. Semantic priming was not observed in the other two association types and there was no difference between the two conditions. It was surprising to find no priming effect in the WAs that are documented in both adult associative norms and Experiments 1-3, and this might be due to the online modality's sensitivity to evidence such an effect. Similarly, while there was evidence of semantic priming overall, this was not supported by above chance looking in related trials which might be due to the insensitivity of the online procedure. Above chance looking was only found in child-specific related word pairs, when analysing PLT by WA type. This might indicate that an online semantic priming procedure is only sensitive enough to capture a priming effect when WAs are robust in a child's lexical-semantic system (i.e. the child-specific WAs in this experiment) and in fact could suggest that the finding of a semantic priming effect in the combined dataset (i.e. all WA types together) benefited from an associative boost (McRae & Boisvert, 1998; Moss et al., 1996; Perea & Rosa, 2002) from the child-specific word pairs.

Another semantic priming study, Experiment 6, ran in parallel to Experiment 5. A key difference to Experiment 5 was the way in which related word pairs were chosen as stimuli. Experiment 6 used the more common approach of selecting word pairs based on associative strength from adult norms, those taken from other infant IPL methodologies, or selected based on taxonomic relatedness. Experiment 6 tested the online priming IPL procedure on monolingual 30-month-olds (*N*= 40). As well as being a replication study, a key aim for Experiment 6 was to act as a monolingual control study from which to compare the effect of English-Dutch/German cognates on semantic priming in the bilinguals tested in Experiment 7. Surprisingly, monolingual 30-month-olds looked significantly longer at the target on unrelated trials, compared to related trials. This does not replicate the findings of Experiment 5 or other lab studies (e.g. Jardak & Byers-Heinlein, 2019). Age, experiment modality and stimuli selection were all discussed as possible reasons for the unexpected finding and it was concluded that a replication study was necessary to help determine the locus of the effect.

Running in parallel to Experiment 6, Experiment 7 aimed to extend previous research on language non-selective lexical access during infant bilingual language processing. The effect of cognates, language dominance, and language pair were areas of interest in this online bilingual semantic priming study. Only five participants were successfully tested due to the

Covid-19 pandemic, which resulted in the cancellation of a data collection placement in The Netherlands. Consequently, this remains an ongoing study.

To investigate the possible causes of the inverted priming effect found in Experiment 6, the design of Experiment 8 was largely similar to Experiment 6, with only age differentiating the two. While Experiment 6 tested 30-month-olds, Experiment 8 tested 3-year-olds (*N*= 18), to explore whether age contributed to the unexpected target preference finding in unrelated trials compared to related trials. Recall that a priming effect had been found in Experiment 5 on 3-year-olds, with a different set of stimuli. However, the results in Experiment 8 showed that 3-year-olds looked longer at the target on unrelated trials, just like the 30-month-olds in Experiment 6. It was decided that a final replication study considering the effect of experiment modality (i.e. online vs. in-lab) should be explored before drawing any conclusions.

To that purpose, Experiment 9 replicated the online procedure used in Experiment 6 and 8, but in a lab setting on 30-month-old toddlers (*N*= 18). Again, participants spent longer looking at the target on unrelated trials compared to related trials, replicating the same finding as Experiment 6 and 8. Stimuli selection was discussed as a key factor driving this effect and follow up analyses showed that not only was Experiment 6 likely an outlier, but that certain word pairs had strongly contributed to the inverted priming effect. Omitting these word pairs and re-running the analyses showed no effect of semantic priming for Experiment 6, 8, and 9. The sensitivity of the online modality was discussed as a possible reason for the lack of a semantic priming effect, though this ought to have been diminished in Experiment 9 which replicated the procedure in the BabyLab. The lack of a semantic priming effect in the lab, using the same stimuli, suggests the prime and target did not have a robust enough relationship in the lexical-semantic systems of the young participants to evidence a semantic priming effect. Although analyses indicated no statistical difference between associatively related and taxonomic word pairs when combining data from Experiment 6, 8, and 9, it is our suspicion that this nonetheless contributed to the outcome. Experiment 5 did show an effect of semantic priming with associatively related word pairs, which was at its strongest when separating out the WAs produced by 3-year-olds in a WA task (Experiments 1-3). For the WAs taken from adult norms, there was no effect of semantic priming, which is the same result replicated in Experiment 6, 8, and 9.

4.2 Limitations

One limitation to Chapter 1 of this thesis is the way in which we determined whether a child's response to the WA task was related or not. Since this research wanted to move away from a reliance on adult associative norms to a more accurate set of child-specific associations when selecting stimuli, a more objective measure of relatedness should have been selected. We determined relatedness according to our own perception of what is considered associated, as determined by adult norms. While we did this to enable comparison of child specific WAs to adult WAs and due to not knowing how successful the task could be completed at such a young age (36-39 months), it does stand in opposition to an attempt to move away from adult norms as we still used them to define relatedness in children. While we did try to establish a method of categorisation which captures different types of relatedness (e.g. 'Category 2' referred to an association specific to the child's own experience) this is an area to develop further, though it might lead to increased numbers of idiosyncratic responses. To more objectively define relatedness, we could use semantic feature norms (McRae et al., 2005) and also corpora such as CHILDES (MacWhinney, 2000) to drive decisions on relatedness rather than relying on instinct and adult norms when categorising responses. After this, the same comparisons and analyses performed in Chapter 1 could be made on WAs from adult norms.

A limitation to Chapter 2 of this thesis was trying to test hypotheses using an online modality, not sufficiently validated yet. While Experiment 4 did try to first validate the online modality by replicating in-lab findings, online, there was always a concern that testing online would not be sensitive enough to capture a small effect size, despite us testing at an age (30-36 months) by which children should demonstrate a sensitivity to semantic priming (e.g. see Arias-Trejo & Plunkett 2009, 2013 for evidence at 24 months). Consequently, further methodological development is needed to reliably replicate expected findings from the lab, to ensure reliable testing of new hypotheses. As we will discuss in the **Future Methodological Research** section, stimuli selection is an area that could be improved to eliminate it as potentially interfering with looking behaviour. By doing this we are more likely to establish if testing online affects infants' looking behaviour.

4.3 Implications

4.3.1 Development of lexical-semantic connections

Replicating WAs from adult associative norms in children as young as three years is a clear indication of the maturing lexicon in this age group. A weak positive correlation between the associative strengths found in the same word pairs in adults and children could be indicative of a progression towards a mature system. Coupled with generally weaker associative strength in the unique child word pairs, if associative strength is a reliable metric to measure this phenomenon, then the adult-like associations may strengthen with age as the child-specific associations reduce, never reaching maturation and so not represented in adult norms. Styles and Plunkett (2009) recommend further research to determine if different relationship types are "weighted differently across development". This would need further investigation to support such a notion and we discuss this future line of enquiry in the **Future Directions** section of this General Discussion.

The fact that a semantic priming effect was only found in Experiment 5 that used related word pairs that had associative relations compared to all other experiments which had related word pairs composed of mostly taxonomic relations might teach us two things. First, as we will argue below, it could be that associative meaning is more robust in the lexical-semantic system than taxonomic meaning. Second, that the word pairs used in Experiment 5, which derive from the productive vocabularies of 3-year-olds, have more established connections in the lexicon, than other word pairs used in the other stimulus list (used in Experiments 6 to 9). If true, this might point to these word-to-word relationships being established before other adult-like relationships. Further studies are underway to replicate the semantic priming effect found at 36-39 months in Experiment 5, in the same age group and at the younger ages of 30 months and 24 months. These studies use a slightly different stimuli list which employs better counterbalancing to ensure visual stimuli do not interfere with any effect of priming that might exist.

Findings from Chapter 2 indicate that lexical-semantic links might be more robust in the lexical-semantic system of a 3-year-old when they capture associative meaning compared to taxonomic meaning. The rationale for this lies in the lack of a semantic priming effect in

Experiment 6, 8, and 9, but a clear effect in Experiment 5. For example, taxonomic relationships were more prevalent in the stimulus word pairs of Experiment 8 (result= no semantic priming) compared to Experiment 5 (result= evidence of semantic priming), when testing the same age group (36-39 months). Other semantic priming studies have selected prime-target pairs which had both an associative and a taxonomic relationship (Arias-Trejo & Plunkett, 2009) to capitalise on the associative boost effect (Moss et al., 1996; Shelton & Martin, 1992) which increases the strength of a semantic priming effect. Considering the prime-target word pairs in Experiments 6-9, only three of the twelve pairs in the related condition were associative, with all others taxonomic, which might be why no semantic priming effect was found. In contrast, Experiment 5 found a strong priming effect in child-specific WAs (*N*= 4), which were all associatively related in children, but not in adults (except for *park-swing*). No priming effect was found in the shared adult and child WAs, nor in the adult-specific WAs. For related prime-target pairs in these two conditions (*N*= 8), most were associatively and taxonomically related (5/8), or associatively related (2/8), rather than taxonomically related (1/8). As such, we would have expected the associatively related pairs to have boosted priming sufficiently to evidence a semantic priming effect, but this was not the case. This provides strong evidence for the robustness of our child-specific WAs in the lexical-semantic systems of 3-year-olds, above other types of WAs that are commonly used in infant semantic priming studies.

Associative links may arise due to early experience to a conjunction of events: experience of the real world (e.g. playing with toys in the bath) and exposure to recurring words during those moments. Therefore the links between toys and bath for example might be of two kinds: links between visual representation and lexical forms. In contrast, taxonomic links may emerge from a re-representation of meaning within an existing lexicon, based solely on abstract knowledge.

The adult (see Lucas, 2000, for a review) and child literature (e.g. Delle Luche et al., 2014) shows that pure taxonomic relationships can evidence a priming effect, even if the effect size is much smaller than with associative semantic priming. In infant studies, there is behavioural (see Delle Luche et al., 2014, using the head-turn preference procedure at 18 months) and brain imaging (Rämä et al., 2013; Torkildsen et al., 2007 at 18 months; and Sirri & Rämä, 2015 at 24 months) evidence that taxonomic relationships do exist some 12 months before the age we tested at, demonstrating that taxonomic links already exist in the lexical-semantic network from a young age. Therefore, we can only surmise that the online modality prevented us finding this same effect. The only anomaly to this conclusion is Experiment 9, in which an effect of priming should have been found, since Experiment 9 was run in the lab. In this case, the lack of an associative boost in prime-target relatedness might explain why we found no semantic priming and that purely taxonomic relationships were insufficient to elicit a priming effect because the links between taxonomic connections alone were not robust enough in the lexical-semantic systems of young children to facilitate lexical retrieval.

In terms of our research questions regarding the lexical organisation of words in bilinguals compared to monolinguals, we were interested in comparing performance in a WA task and also the effect of English-Dutch/German cognates on semantic priming. Our key aim was to explore the organisation of the lexical-semantic structure during bilingual language development and whether we could find evidence for more robust connections in this emergent system for translation equivalents and/or cognates. In comparing the performance of monolingual and bilingual 3-year-olds in a WA task, we have been able to provide evidence to support the view that bilinguals develop similarly to their monolingual counterparts (Bilson et al., 2015), specifically when completing a WA task. Response rate and type were comparable in both groups, indicating that neither group found the task any more difficult than the other. Furthermore, some WAs were shared across groups, signifying comparable productive ability and more importantly, a similar organisation of related words in the lexicon.

At the same time, we have shown that translation equivalents (TEs) are used in the productive vocabularies of bilinguals with greater frequency than words known in just one of their languages. This supports the facilitative effect of TEs found in previous studies investigating language reception (e.g. Poulin-Dubois et al., 2018). It might also be that TEs have stronger connections in the lexical-semantic system of bilinguals due to their shared features between languages. Byers-Heinlein and Werker (2013) propose that a bilingual infant with a high number of TEs in their receptive and productive vocabularies have a richer network of semantic connections compared to monolingual infants. This is supported by evidence from TE priming studies with simultaneous bilingual 27-month-old toddlers (English and one other language) which showed a priming effect for TEs (e.g. *cat* and *chat*) which was equal in magnitude to the priming effect in a cross-language semantic condition (e.g. *cat* and *chien*- dog in French; Floccia et al., 2020). Further evidence comes from a word recognition task which indicates TEs have a facilitative effect on word retrieval such that TEs are recognised faster than non-TEs in both of a simultaneous bilingual's languages (see Poulin-Dubois et al., 2018 and their experiments on 22-month-old French-English infants). This suggests there is implicit activation of the target and its translation during a receptive task, for TEs that exist in an interconnected lexicon already by 22 months, and that knowing a word in each language has an impact on lexical access. When considering the two models of developmental bilingual word processing discussed in the introduction (BLINCS: Shook & Marian, 2013; DevLex-II: Zhao & Li, 2010, 2013), our findings most closely align with the predictions of DevLex-II. The model suggests that additional learning creates stronger connections between translation equivalents and so translation equivalents could contribute to stronger lexical connections in the developing semantic system. If true, this would imply that translation equivalents are processed faster for language production, which explains why we found a high proportion of translation equivalents in the related responses given by bilingual toddlers.

While TEs have evidenced facilitation, cognates seem to either facilitate or inhibit. For example, the effect of cognates on word recognition in sequential German-English 23-47-month-olds facilitated word recognition in the weaker L2 when phonological feature overlap was almost identical, but did not facilitate word retrieval with increasing feature changes (Von Holzen et al., 2019). The authors suggest that words with more phonological feature changes might cause interference due to the simultaneous activation of each representation. This suggests that the cognate has to be almost identical to have a facilitative effect. We discuss the impact of this on our continued testing of Experiment 7 in the **Future Directions** section.

4.3.2 Adult Associative Norms as a Resource for Designing Infant Experiments

The findings of this thesis demonstrate a concordance between the organisational structures of adults and children in terms of associative relations. Experiments 1-3 clearly show that children share some of the associations that adults exhibit in a mature lexical-semantic system. This aligns with theoretical analyses of vocabulary development using CDI data by Hills et al. (2010) who hypothesise that toddlers preferentially acquire words based on adult associative norms. Despite this, over a quarter of the associated responses given by children (Experiment 1= 42%; Experiment 2= 26%) do not feature in adult associative norms, including some of 3-year-olds' more common associations (e.g. *bed-teddy*). This has implications for the experimental design of infant studies, which to date have had to rely on adult associative norms to select associatively related word pairs as experimental stimuli. Various researchers have identified this as problematic (Arias-Trejo & Plunkett, 2013) and while some primed IPL studies have captured an effect of semantic priming using adult associations as a resource for stimuli selection, choosing word pairs taken directly from the productive vocabularies of young children is more representative of the word knowledge in the developing lexicon. In fact, Experiment 5 statistically demonstrated this in an online primed IPL study where the only statistically significant PLT to the target referent in related trials was for the word pairs taken directly from the productive associations in Experiments 1-3.

As Arias-Trejo and Plunkett (2013) note, young children might over- or under-extend taxonomic categories, with relationships that do not conform to adult norms or that exist in spite of their absence in the mature semantic system. The findings of this thesis would support this idea rather than supporting the earlier suggestion from Arias-Trejo and

Plunkett (2009) that "… the absence of a semantic–associative relation in adult norms may be a reliable indicator of its absence in infants…" This research goes a step further to propose that the associations in adult norms might under-represent the associations of children, as they might not capture the unique developmental stage and life experience of 3-year-olds. This is mirrored in Fitzpatrick et al.'s (2013) assertion that WAs stemming from a population other than the population of interest might not acknowledge the unique characteristics of that group.

When designing stimuli for infant studies, it is common practice to approach adult associative norms with a 'shopping list' of carefully controlled criteria such as age of acquisition, syllable length, phonology etc. and selecting the word pairs found in adult norms that can fulfil all these criteria, however, a more natural language approach could be to select word pairs originating from the productive vocabulary of the age group in question and when language production has not yet been attained, the closest age group to the one being tested might be used for reference. This research presents a resource of such word pairs and is a stepping-stone towards a more robust resource for researchers looking for infant-specific associated word pair stimuli.

Using associative strength as a metric to determine stimuli selection from adult norms might not be reliable. In Experiment 5, we found no effect of semantic priming for the WAs taken from Experiments 1-3 and adult associative norms (Moss et al., 1996; Nelson et al., 2004). This was surprising and raises the question of how reliable a metric it is for the strength of a relationship between words, or at least, whether the expected priming effect can be observed in a possibly less sensitive implementation of the IPL, such as when used online. A series of further online studies are underway to address this question by attempting to demonstrate a semantic priming effect in the same age groups tested in this thesis, when using a more rigorously chosen stimulus list. The lack of a semantic priming effect in some of the studies of this thesis (Experiments 6, 8, and 9) led us to believe that some prime-target stimuli may have attracted unexpectedly high or low attention (supported through identification of outliers). To ensure the replication studies do not have this problem, we modified the stimulus set to balance animacy across related and unrelated trials, and ensured target images feature in related and unrelated conditioned, balanced between participants.

4.3.3 Testing Infants in their Home Environment

One of the significant implications of the findings from this thesis is the validation of a testing protocol in a home environment using a participant's own device. While the findings from Chapter 2 did not always mirror lab findings, this is not wholly due to the experiment modality, but possibly due to stimuli selection too (Experiments 6-9). In support of online testing, Experiment 4 clearly evidenced word recognition in 24-month-olds with an effect size which rivals equivalent lab-based studies. Furthermore, using time-course analyses we were able to replicate similar timings for word recognition in-lab and online modalities (see Experiments 5 and 6). This was found despite connection speed differences and device variability. Furthermore, data loss as discussed previously, was lower in magnitude compared to lab-based studies, for example, when considering the attentiveness of participants (i.e. minimum looking time on-screen). Finally, Experiment 2 and 3 introduced a new procedure for an online WA task, which produces the same results as an equivalent at-home procedure.

Thus, testing online still requires some fine-tuning and replicable results, especially in a semantic priming paradigm, but it does have the potential to reach a more diverse and representative sample of participants in infant research by offering those unable to make lab visits, an alternative way of participating in research. Testing in the home environment compared to testing in a lab on a university campus might be less intimidating and enable us to reach different demographics more easily too. All this is in addition to testing in a more naturalistic environment, in which young participants are certain to be more familiar and comfortable with, thus more likely to naturally behave as they would in everyday situations.

4.4 Future Directions

4.4.1 Future Methodological Research

There were some minor methodological oversights in the experiments presented in Chapter 2 of this thesis, notably regarding the word pairs chosen for use in Experiments 6-9. This may have contributed to the null effect of semantic priming and it would be beneficial to replicate the semantic priming effect we did find in Experiment 5, to determine the extent to which the online modality may have inhibited this finding in Experiments 6, 8, and 9. Three new online experiments are currently underway which rectify some of the errors made in stimuli selection previously e.g. targets feature in both a related and unrelated condition between participants; word pairs/images appear only once for a given participant; and most importantly, the taxonomic/associative relationship is more carefully controlled to aid analysis of any associative boost that exists. This should minimise the chance that the absence of a semantic priming effect is attributable to the stimuli selection, rather it should indicate if a primed IPL task, administered online, can replicate in-lab findings, even when no associative relationships exists between prime-target pairs (i.e. no associative boost), as has been evidenced in lab studies (e.g. Arias-Trejo & Plunkett, 2013).

The time-course of word recognition is another avenue which could be explored online further. From the experiments in this thesis, there is some evidence (e.g. Experiments 5 and 6) that online testing might be sensitive enough to capture the time-course of lexical processing, however, our findings span a wider range (a couple of hundred milliseconds) than lab studies might be able to achieve. With the advancement of online testing capabilities, this could certainly improve. Note, the procedure we used on Gorilla was still in Beta (i.e. recording via webcam) and not fully validated for widespread testing. Thus, advancements on Gorilla, coupled with the findings from infant testing using LookIt (Scott et al., 2017; Scott & Schulz, 2017), in addition to the work being done on automatic coding of eye movement (e.g. see the Labvanced platform: https://www.labvanced.com/), could open up other paradigms for testing young participants online, such as the visual world paradigm (VWP), which could be used to test the existence of any competition between different types of semantic relatedness (see Chow et al., 2017 for successful in-lab testing of 24- and 30-month-olds using the VWP).

4.4.2 Future Theoretical Research

It would be informative to track a child's WA production as they age to see if the words produced do start to more closely resemble adult WAs, in terms of exact words produced and their associative strength. This could be approached in one of two ways. First, corpora such as CHILDES (MacWhinney, 2000) could be analysed to determine whether the child-specific WAs we found in Chapter 1 of this thesis stem from the speech input of caregivers or from a child's environment. One way to achieve this is to extract the nouns that occur in the speech stream 3-5 words before the cue word and determine the frequency of these occurrences, running a correlation between these and the WAs we found in the word association task, taking the associative strength (FSG) as the metric for the latter. It would also be informative to document the frequently-occurring noun-noun combinations in general as these may differ to the child-specific WAs we found in the productive vocabularies of three-year-olds. Additionally, this analysis could be done longitudinally to see if age modulates the frequency and type of WAs children are exposed to, as documented in CHILDES. The same could be repeated with common WAs from adult norms to identify when these types of WAs become more frequent in the speech stream and if this correlates with the age at which children start to produce more adult-like associations.

Second, it could form the basis of a longitudinal study which tracks the WAs of the same individuals periodically, at different milestones. For example, school is described as a period in which responses might standardise due to the same routine and language exposure shared by a class, compared to adolescence when different interests emerge (Wojcik & Kandhadai, 2020). Similarly, responses to a WA task have been shown to become more paradigmatic in nature in older children (Newman, 1970) compared to younger children (Brown & Berko, 1960; Entwisle et al., 1964; Ervin, 1961), though the proposition of a thematic/associative to taxonomic shift in the organisation of the lexicon is not supported in all testing modalities (e.g. in a cued-recall task Blewitt & Toppino, 1991), yet adult-like responses to the WA task do seem to be predominantly paradigmatic (e.g. Wojcik & Kandhadai, 2020). The findings from such a longitudinal study would have practical and theoretical value. A practical application would be as a resource to inform stimuli selection for infant studies which more closely reflects the relationships between words for the age group that is being tested, rather than using adult norms. The theoretical value would lie in mapping the connections between words and the strength of these. This could inform us about how the development of the lexical-semantic system emerges and matures based on age and experience, for example, and the point at which the "paradigmatic shift" (Nelson, 1977) occurs, signaling the transition to a more mature, adult-like lexical-semantic system.

The opposite effect, that is, what happens to our early WAs as our lexical-semantic system matures, is also of theoretical interest. For example, the early, child-specific WAs documented in Experiments 1-3 do not feature as strong associates in adulthood, reflected in their absence in adult norms, yet that is not to say these associates disappear from the adult lexicon completely. It could be that traces of these early connections do exist in the adult lexical-semantic system and one way to investigate this could be to use the online primed IPL procedure from Experiment 5, on an adult population. Even if the early WAs do not elicit a strong priming effect in adults as they did in children, a priming effect may still exist, which would point to the existence of these associates into adulthood, albeit lower in associative strength. If the early associates do not show an effect of priming in adults, this might suggest that the lexical-semantic system is a dynamic system which is able to rewrite WAs constantly between infancy and adulthood, as well as strengthening some WAs more than others. In fact, according to Wojcik and Kandhadai (2020) peaks and troughs might be found in the proportion of idiosyncratic and associated responses given across development, which could be documented in the first (see above) proposed extension to the research in this thesis.

So far, we have focused on future directions of research with monolingual participants, but the findings from Experiment 7, when complete, might enable us to explore research questions related to the development of the lexical-semantic system in bilingual speakers. It is hypothesised that Experiment 7 will demonstrate a stronger effect of semantic priming for cognates, compared to non-cognates, with the effect amplified when both prime and target are cognates. However, cognates may also inhibit word recognition if used as distractors and depending on the level of phonological feature overlap. On collecting our full sample of data, we could consider feature overlap in the analysis to account for any effect of facilitation or inhibition we might find. Furthermore, using parental report, we could determine if primes and targets in the experimental stimuli are known to participants in one or both of their languages to explore the interplay of TE knowledge and cognate status. Together this will help determine if a bilingual advantage exists for lexical-semantic organisation, for these specific word types.

4.5 Conclusion

The findings from this thesis contribute to our understanding of lexical-semantic development in toddlers. Chapter 1 provides evidence for the shared adult-like and unique child-like word associations that young children have in their lexicon, whether monolingual or bilingual. Further, it presents a resource of the most commonly produced word associations in 3-year-olds, which might be beneficial to developmental researchers looking for more representative associations when selecting prime-target word pairs in their experiments. Chapter 2 validated the strength of the child -like associations in a semantic priming study and supports the argument for stronger links between associatively related words in the developing lexicon, compared to taxonomically related words.

5 Appendices

Appendix A: Experiment 1. The Percentage of 18-Month-Olds Knowing the Words Used as Cues in the Word Association Task

Word	OCDI % known at 18 months	UKCDI % known at 18 months
aeroplane/plane	81	72
apple	75	82
arm	56	75
ball	98	99
balloon	84	83
banana	91	94
bath/bathtub	94	98
bed	85	97
bee	60	69
bib	75	66
bicycle/bike	69	72
bin	70	83
bird	88	88
biscuit	88	86
boat	62	69
book	95	98
boots	54	65
bottle	65	80
bowl	58	77
box	48	63
bread	72	77
brush	72	77
bubbles	61	85
rabbit	77	77
bus	69	81
butterfly	54	63
cake	54	74
car	95	97
carrots	48	74
cat	94	94
cereal	26	67
chair	80	95

cheese	63	78
chicken	58	72
coat	77	90
cot	70	68
cow	83	82
cup	79	83
dog	98	99
doll	60	73
door	87	96
duck	90	86
ear	84	83
elephant	54	70
eye	86	96
finger	82	79
fish	75	81
flower	77	68
feet	70	92
fork	46	65
frog	56	68
garden	73	72
hair	91	86
hand	77	85
hat	87	89
head	75	89
high chair	68	78
horse	76	78
house	57	78
key	74	81
leg	59	81
lion	65	79
lorry/truck	61	58
monkey	57	90
mouse	54	67
mouth	76	91
nappy	92	98
nose	94	94
orange	37	63
park	38	72
pasta	35	60

peas	47	70
pen	53	70
pig	77	82
plate	52	66
pushchair/buggy/stroller	77	82
pyjamas/Pjs/jim jams/	54	80
settee/sofa/couch	48	74
sheep	69	76
shoes	99	97
slide	59	72
sock	92	91
spoon	77	76
stairs	81	86
swing	64	68
table	64	78
teddy/teddy bear	85	91
phone/telephone/mobile	87	91
tiger	50	72
toast	70	84
toe	71	76
tooth/teeth	75	85
toothbrush	86	94
towel	57	68
toy	60	82
train	66	81
tree	69	78
trousers/pants/britches	55	76
television/telly/TV	77	89
window	63	78

Appendix B: Experiment 1. The Wordlists Used in the Word Association Task

Order 1

List 1	
1	plane/ aeroplane
2	bread
3	frog
4	cereal
5	door
6	bicycle/bike
7	leg
8	pasta
9	slide
10	toe

List 2	
1	apple
2	bin
3	duck
4	chair
5	sock
6	garden
7	lion
8	peas
9	high chair
10	tooth/teeth

List 3	
1	arm
2	bird
3	cheese
4	pig
5	hair
6	spoon
7	lorry/truck
8	pen
9	bubbles
10	toothbrush

List 4	
1	ball
2	rabbit
3	stairs
4	boat
5	settee/sofa/couch
6	hand
7	monkey
8	ear
9	biscuit
10	towel

List 5	
1	balloon
2	chicken
3	bus
4	coat
5	eye
6	swing
7	mouse
8	plate
9	hat
10	toy

List 6	
1	banana
2	train
3	butterfly
4	cot
5	finger
6	book
7	pushchair/buggy
8	boots
9	table
10	head

List 7	
1	bath/bathtub
2	mouth
3	teddy/teddy bear
4	cow
5	nappy
6	brush
7	cake
8	pyjamas/Pjs/jim jams
9	fish
10	tree

List 8	
1	bed
2	car
3	horse
4	cup
5	flower
6	nose
7	bottle
8	phone/telephone/mobile
9	elephant
10	trousers

List 9	
1	bee
2	shoes
3	bowl
4	tiger
5	feet
6	house
7	orange
8	dog
9	carrots
10	television/telly/TV

List 10	
1	bib
2	cat
3	fork
4	doll
5	key
6	park
7	box
8	sheep
9	toast
10	window

Order 2:

List 1	
1	bread
2	plane/ aeroplane
3	slide
4	door
5	bicycle/bike
6	frog
7	leg
8	cereal
9	toe
10	pasta

List 2	
1	apple
2	high chair
3	garden
4	lion
5	sock
6	chair
7	peas
8	duck
9	bin
10	tooth/teeth

List 3	
1	toothbrush
2	pig
3	spoon
4	arm
5	pen
6	bubbles
7	lorry/truck
8	cheese
9	hair
10	bird

List 4	
1	settee/sofa/couch
2	biscuit
3	monkey
4	ear
5	towel
6	boat
7	stairs
8	hand
9	rabbit
10	ball

List 5	
1	coat
2	balloon
3	swing
4	mouse
5	hat
6	eye
7	chicken
8	bus
9	toy
10	plate

List 6	
1	boots
2	finger
3	book
4	head
5	butterfly
6	table
7	cot
8	train
9	banana
10	pushchair/buggy

List 7	

List 8	

List 9	

1	bath/bathtub
2	teddy/teddy bear
3	nappy
4	cake
5	tree
6	cow
7	brush
8	fish
9	pyjamas/Pjs/jim jams
10	mouth

1	trousers
2	cup
3	horse
4	bed
5	elephant
6	car
7	nose
8	phone/telephone/mobile
9	flower
10	bottle

1	house
2	carrots
3	dog
4	television/telly/TV
5	bowl
6	shoes
7	tiger
8	orange
9	feet
10	bee

List 10	
1	bib
2	sheep
3	key
4	park
5	doll
6	toast
7	box
8	window
9	fork
10	cat

Appendix C: Experiment 1. Participant Information Sheet/ Consent Form

UNIVERSITY OF PLYMOUTH

FACULTY OF HEALTH: MEDICINE, DENTRISTRY AND HUMAN SCIENCES

RESEARCH INFORMATION and CONSENT FORM

Name of Principal Investigators

Nadine Fitzpatrick & Caroline Floccia

Title of Research

Infant and child associative norms (IACAN)

Aim of research

This research aims to collate a database of the associations that infants and children have for common words (e.g. hair - brush) as their language develops. This database will then be made available as a reference resource for other researchers when designing infant and child experiments testing the development of meaning in language.

Description of procedure

To document word associations in your child, you will be given a word list of ten words. Begin by checking the wordlist to make sure your child understands all the words. If your child does not understand one or more of the words, please email to let us know so we can give you alternative words.

Following a simple script, you will give your child three examples of the task before beginning with the word list. During the test, your child can give up to three associations for each word on the wordlist. You will need to record what your child says in a table (sent in the email you will receive with the word list) for each word, in the order that they say them. When you have finished all ten words, return the completed table as an email. Try to complete the task in one go if possible. It should not take more than 15-30 minutes.

There are no wrong answers for this task, so you are encouraged to write down everything your child says. If you want to include any comments or observations, please feel free to include these in your email.

Description of risks

None.

Benefits of proposed research

There is no direct short-term benefit, however; long-term we hope that this research can contribute to a better understanding of language development in typically developing children, which can benefit by helping children who are not developing typically.

Right to withdraw

At any moment during this study, you can choose to withdraw without providing any justification. Your data would then be removed from the study. If you are dissatisfied with the way the research is conducted, please contact the principal investigator in the first instance: telephone number 01752584822. If you feel the problem has not been resolved please contact the secretary to the Faculty of Health: Medicine, Dentistry and Human Sciences Research Ethics Committee: Mr Maurice Bottomley hhsethics@plymouth.ac.uk +44 1752 586992.

I am the *parent /legal guardian of ____

The objectives of this research have been explained to me.

I understand that *she/he is free to withdraw from the research at any stage, and ask for *his/her data to be destroyed if I wish. I understand that *his/her anonymity is guaranteed, unless I expressly state otherwise.

I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations). I understand that video footage of my child during the procedure will be used for scoring purposes. It will be kept safely on the investigators' hard drive at the University, and destroyed 5 years after publication.

Under these circumstances, I agree for him/her to participate in the research.
** delete as appropriate*

Name: ……………………………………….

Signature: ……………………………………………. Date:………………………..

Appendix D: Experiment 1. Word Association Task Instructions (sent by email)

1. Consent to take part in the study	
Who are you giving consent for?	*Full name*
Have you read the consent form/ information sheet?	*Yes/No*
Do you consent to take part in this study?	*Yes/No*

2. Check your child knows the words in the word list below
If there are two options (e.g. tummy/belly), choose the most familiar to your child and indicate the one you choose.

If your child does not know 1 or more words, email us and we can give you different words to use. This is not a problem.

Word list		**1st response**	**2nd response**	**3rd response**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10		Email us for your	unique word list	

Comments/ observations	*e.g. Make a note if your child is naming things in their immediate environment or making connections to something happening in their life right now.*

3. Complete the test using the script below
Follow the script as close as you can.
Say all 3 examples.
For every word in the list, try to get up to 3 different responses. 1 response per word is absolutely fine though.
Try do all 10 words in one go if possible.
There are no right or wrong answers! Have fun!

Script "We're going to play a game to see how quickly you can say a word that is connected to a word that I say. If I say KITCHEN you might say BREAKFAST. (Example 1) If I say MUMMY you might say DADDY. (Example 2) If I say DRINK you might say WATER. (Example 3) Okay, are you ready? What do you think of if I say …? (Response 1) And another word when I say…? (Response 2) And another word?" (Response 3)
4. Enter the word/s your child says in the table in the correct order (1st, 2nd, 3rd response). Write down everything they say. If there is no response, enter NO RESPONSE. Leave comments/ observations in the box provided. **Email the completed table back and receive a £5 thank you Amazon voucher and Infant Scientist certificate for taking part!**

Appendix E: Experiment 1. Debrief Email

Hello **PARENT**,

That's absolutely fantastic! Thank you for returning the completed task. **CHILD** did so well!

Could you indicate if any of the following responses were items from the immediate surroundings/ an association that might not be clear to someone who does not know your child:

- CUE WORD- Child’s response
- CUE WORD- Child’s response
- CUE WORD- Child’s response

This will help me when I try to categorise the types of responses children give.

Please find a certificate attached for **CHILD**. There is a star (and rocket!) for every study s/he has participated in!

Here is your £5 Amazon voucher code:

CODE

Thank you both for supporting our research. Please get in touch if you have any questions or would like to withdraw from this study.

Hopefully we’ll see you in the BabyLab in the near future!

Kind regards,
Nadine

Appendix F: Experiment 1. All related responses given by 2+ children in the parentally-administered WA task (i.e. not limited to first responses)

	Cue	Response	No. participants receiving cue (G)	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	**apple**	Eat/ eat it/ you can eat it	38	7	0.18	19
		Juice in it/ apple juice	38	2	0.05	19
		Pear	38	2	0.05	19
		Red	38	2	0.05	19
		yummy/ they are yummy	38	2	0.05	19
2	**arm**	Leg/ legs	43	4	0.09	19
		finger	43	3	0.07	19
		hand/ hands/ DEF: It's something that you make your hand grab something. Hand	43	3	0.07	19
		Body	43	2	0.05	19
		Elbow	43	2	0.05	19
		Head	43	2	0.05	19
3	**ball**	kick/ kick kick/ kicking	30	8	0.27	13
		football	30	4	0.13	13
		Throw/ throwing/ throw up high	30	3	0.10	13
4	**balloon**	pop	33	4	0.12	20
		holding/ Holding a balloon/ We hold them	33	3	0.09	20
		Party	33	2	0.06	20
		red	33	2	0.06	20
5	**banana**	Eat/ eat it/ eat the banana/ eating	32	5	0.16	21
		apple	32	2	0.06	21
		Yellow	32	2	0.06	21
		Fruit	32	2	0.06	21
6	**bath**	toy/ toys/ bathy toys/ put the toys in	40	5	0.13	22
		water	40	4	0.10	22
		to wash ourselves/ wash/ wash hair	40	3	0.08	22
		bubbles	40	2	0.05	22
		duck/ duckies	40	2	0.05	22
		splashing/ splash	40	2	0.05	22
7	**bed**	sleep/ to sleep	39	6	0.15	17

		teddy/ teddy bear/ Lambie (Teddy)/ cuddle up with teddies	39	5	0.13	17
		blanket	39	3	0.08	17
		cushion	39	2	0.05	17
8	**bee**	honey	35	5	0.14	16
		bumblebee/ Bumble Bee	35	3	0.09	16
		Flower/ flowers	35	3	0.09	16
		Fly	35	2	0.06	16
9	**bib**	Baby/ A Baby	36	6	0.17	27
		Food	36	2	0.06	27
		No bib/ No (She doesn't wear a bib anymore, her decision. This 'No' is her saying no to wearing a bib.)	36	2	0.06	27
10	**bicycle/ bike**	bell	36	4	0.11	24
		ride/ riding	36	2	0.06	24
		scooter	36	2	0.06	24
		wheels	36	2	0.06	24
11	**bin**	Rubbish/ Rubbish in the bin/ put rubbish in it	35	7	0.20	13
		Smelly Bin/ smelly	35	2	0.06	13
		Lid	35	2	0.06	13
12	**bird**	fly	38	3	0.08	23
		Nest	38	3	0.08	23
		outside/ Bird outside	38	2	0.05	23
		feather/ feathers	38	2	0.05	23
13	**biscuit**	chocolate	33	4	0.12	15
		eat/ Eat! (shouts excitedly)/ eating	33	3	0.09	15
		Kitchen	33	2	0.06	15
		yummy	33	2	0.06	15
14	**boat**	Water/ In the water/ We was on a boat on water	35	5	0.14	16
		Sail/ sailing	35	3	0.09	16
		sea	35	3	0.09	16
15	**book**	Bedtime	32	4	0.13	12
		Read/ reading	32	4	0.13	12
		Story/ read story	32	4	0.13	12
		pictures	32	3	0.09	12
		pages	32	2	0.06	12
16	**boots**	Puddle/ puddles/ muddy puddles/ Splashing in muddy puddles	32	5	0.16	19
		Walk/ walking	32	3	0.09	19
		Feet	32	2	0.06	19

17	**bottle**	Water/ Water bottle	32	6	0.19	10
		Milk	32	4	0.13	10
		Cup	32	2	0.06	10
		Juice	32	2	0.06	10
		drink	32	2	0.06	10
		lid	32	2	0.06	10
18	**bowl**	breakfast	42	4	0.10	20
		food/ Tasty food	42	3	0.07	20
		dinner	42	2	0.05	20
		Shredded Wheat/ Shreddies	42	2	0.05	20
19	**box**	toys/ toys in it (obsessing over toys that morning!)	37	4	0.11	23
		Make (makes models from boxes)/ making/ make something	37	3	0.08	23
		Stuff/ Put stuff in it	37	2	0.05	23
20	**bread**	eat/ I eat it	39	3	0.08	23
		Toast	39	3	0.08	23
		butter	39	2	0.05	23
		honey	39	2	0.05	23
		kitchen	39	2	0.05	23
		Egg/ eggy	39	2	0.05	23
21	**brush**	hair/ Sophie’s long hair/ brush everyone hair/ Brush hair	32	6	0.19	19
		Teeth/ brush your teeth	32	5	0.16	19
		floor	32	2	0.06	19
22	**bubbles**	Pop/ DEF: They're something that pop. Pop the bubbles	46	8	0.17	19
		blow/ blowing/ Blow bubbles	46	4	0.09	19
		bath	46	3	0.07	19
		float/ float in the sky	46	2	0.04	19
		Water	46	2	0.04	19
23	**bus**	car/ cars	37	3	0.08	25
		train	37	2	0.05	25
		wheel/ wheels	37	2	0.05	25
		big/ big bus	37	2	0.05	25
		red	37	2	0.05	25
24	**butterfly**	wings/ yellow wings/ yellow and blue wings	30	4	0.13	19
		flying/ fly away	30	3	0.10	19
25	**cake**	birthday	35	4	0.11	20
		eat/ eat it	35	3	0.09	20

		chocolate/ choc	35	2	0.06	20
		sprinkles	35	2	0.06	20
26	**car**	Drive/ drive somewhere/ driving	36	3	0.08	18
		wheels/ Wheels to bump	36	3	0.08	18
		Beep/ beep beep	36	2	0.06	18
		Seat/ car seat	36	2	0.06	18
27	**carrots**	eat/ eat them/ eating	34	5	0.15	18
		Crunch crunch/ crunchy	34	2	0.06	18
		rabbit	34	2	0.06	18
28	**cat**	Dog	49	8	0.16	19
		Cat food/ food	49	2	0.04	19
		Elephant	49	2	0.04	19
		Meow	49	2	0.04	19
29	**cereal**	Milk/ Blue milk	39	5	0.13	23
		bowl	39	4	0.10	23
		eat/ eat cereal	39	3	0.08	23
		spoon	39	2	0.05	23
		Weetabix	39	2	0.05	23
		breakfast	39	2	0.05	23
30	**chair**	sit/ sit down	38	4	0.11	17
		breakfast	38	3	0.08	17
		eat/ you can eat	38	3	0.08	17
		Table	38	2	0.05	17
31	**cheese**	Eat/ eating/ eat it/ DEF: Easy- it's something that you eat and it's so squeezey. Apple	35	7	0.20	19
		Doggy/ dogs/ doggie	35	3	0.09	19
32	**chicken**	eat/ eating/ we eat it	38	4	0.11	27
		egg/ eggs	38	4	0.11	27
		Cock a doodle doo	38	2	0.05	27
33	**coat**	rain/ When it's just raining got to put your coat on	32	4	0.13	18
		red	32	3	0.09	18
		out/ going out	32	2	0.06	18
		cold	32	2	0.06	18
		hood/ hood on	32	2	0.06	18
		jacket	32	2	0.06	18
		Sleeve/ sleeves	32	2	0.06	18
34	**cot**	baby/ babies	28	7	0.25	14
		Sleep/ go to sleep	28	2	0.07	14
35	**cow**	moo	36	6	0.17	14
		milk/ ae some milk	36	5	0.14	14
		Pig	36	3	0.08	14

		farm	36	2	0.06	14
		fields/ In the field	36	2	0.06	14
		Goats (loves Pennywell)/ Daddy goat	36	2	0.06	14
36	**cup**	drink/ drink it/ Daddy drinking	34	8	0.24	7
		milk	34	4	0.12	7
		Water	34	4	0.12	7
37	**dog**	woof/ woof woof	39	6	0.15	22
		cat/ Kitty cat	39	5	0.13	22
		walk	39	2	0.05	22
38	**doll**	Chair	37	2	0.05	26
39	**door**	Open/ open it	34	5	0.15	21
		Shut	34	2	0.06	21
40	**duck**	Quack/ quack quack/ they go quack quack	38	11	0.29	18
		water	38	3	0.08	18
		swim/ swimming	38	3	0.08	18
41	**ear**	Earrings	33	3	0.09	17
		Listen/ listening ears	33	3	0.09	17
		Mummy/ on mummy	33	3	0.09	17
		eye/ eyes	33	3	0.09	17
42	**elephant**	Big	38	4	0.11	13
		trunk	38	4	0.11	13
		Ears	38	3	0.08	13
		Stomp stomp/ stomp	38	2	0.05	13
43	**eye**	I spy/ spy	37	4	0.11	20
		Nose	37	3	0.08	20
		eye lash/ eyelash	37	2	0.05	20
		ball/ balls	37	2	0.05	20
		head/ On my head	37	2	0.05	20
44	**foot/ feet**	toes	38	4	0.11	23
		shoes	38	2	0.05	23
		sock/ socks	38	2	0.05	23
		hands	38	2	0.05	23
		walk	38	2	0.05	23
45	**finger**	Hand/ hands	37	7	0.19	17
		nail/ nails	37	2	0.05	17
		Point/ pointing	37	2	0.05	17
		thumb	37	2	0.05	17
		Touch/ touch nose	37	2	0.05	17
46	**fish**	water/ lives in water	36	5	0.14	22
		eat/ eating	36	3	0.08	22
		fish finger/ fingers/ Eat fish fingers	36	3	0.08	22
		tank/ In the tank	36	3	0.08	22

		Swim/ swimming	36	3	0.08	22
		Sharks	36	2	0.06	22
47	**flower**	bee/ bees/ buzzy bee	36	4	0.11	18
		Grow	36	2	0.06	18
		petals	36	2	0.06	18
		Water	36	2	0.06	18
		daisy	36	2	0.06	18
		Pretty	36	2	0.06	18
48	**fork**	Knife	36	7	0.19	13
		Spoon	36	6	0.17	13
		Eat/ to eat	36	4	0.11	13
49	**frog**	ribbit/ Frog says ribbit	35	4	0.11	23
		Water	35	3	0.09	23
		Green	35	2	0.06	23
		jump/ Jump (and she jumps)/ jumps in	35	2	0.06	23
50	**garden**	trees/ apple tree	41	3	0.07	30
		Chair	41	2	0.05	30
		grass	41	2	0.05	30
		Pea	41	2	0.05	30
		trampoline/ Trampoline (has one in the garden)	41	2	0.05	30
		bee/ Bees in the garden	41	2	0.05	30
		Play	41	2	0.05	30
51	**hair**	brush	35	4	0.11	22
		Head	35	3	0.09	22
52	**hand**	fingers/ fingers	34	5	0.15	13
		foot	34	2	0.06	13
		Hair	34	2	0.06	13
53	**hat**	head/ It goes on your head	33	3	0.09	21
		Wear a hat/ wear it/ we can wear a hat	33	3	0.09	21
54	**head**	Hair/ Hair (pointing to his hair)/ hair on	36	8	0.22	19
		Ears	36	2	0.06	19
		Eyes	36	2	0.06	19
		Mummy head/ mummy	36	2	0.06	19
		brain	36	2	0.06	19
55	**high chair**	breakfast	35	2	0.06	18
		eat	35	2	0.06	18
		Drink	35	2	0.06	18
		food/ can eat food	35	2	0.06	18
56	**horse**	Clip clop	40	2	0.05	22
		riding/ ride on them	40	2	0.05	22
		Tail	40	2	0.05	22

		neigh	40	2	0.05	22
57	**house**	Tree	39	3	0.08	25
		windows	39	3	0.08	25
		Light. On and off/ lights	39	2	0.05	25
		Toy/ "My got toys in my house"	39	2	0.05	25
58	**key**	door/ open the door	42	9	0.21	13
		car/ daddy's car/ mummy's car	42	3	0.07	13
		open/ can I open the door	42	3	0.07	13
		lock/ Lock the door	42	2	0.05	13
		Unlock keys/ unlocking	42	2	0.05	13
59	**leg**	Feet/ foot	41	7	0.17	16
		toe/ toes	41	4	0.10	16
		hands	41	2	0.05	16
		Arm	41	2	0.05	16
		Head	41	2	0.05	16
		knee	41	2	0.05	16
		walk	41	2	0.05	16
60	**lion**	roar/ they roar	38	9	0.24	17
		claws/ got big claws	38	2	0.05	17
		Baby lion	38	2	0.05	17
		Tail	38	2	0.05	17
		Zoo/ see them in the zoo	38	2	0.05	17
61	**lorry/ truck**	Wheel/ wheels	38	4	0.11	22
		Digger	38	3	0.08	22
		drive/ DEF: It's something that drives	38	2	0.05	22
62	**monkey**	banana	35	5	0.14	13
		elephant	35	3	0.09	13
		Swing/ swing in branches	35	3	0.09	13
		cheeky/ cheeky monkey	35	3	0.09	13
		tree/ trees	35	3	0.09	13
		Oo oo aa/ ooo ooo ooo	35	2	0.06	13
63	**mouse**	cheese/ Eats cheese	36	2	0.06	25
		Run/ running away	36	2	0.06	25
		It squeaks/ goes squeak	36	2	0.06	25
		Squeak	36	2	0.06	25
		tree/ trees	36	2	0.06	25
64	**mouth**	teeth	30	7	0.23	17
		eat/ eating	30	2	0.07	17
		Gum/ gums	30	2	0.07	17
		Tongue	30	2	0.07	17
		Talking	30	2	0.07	17
65	**nappy**	Bayb/ Babies	32	5	0.16	18

		Poo/ poop/ We don't poop	32	4	0.13	18
		bum/ nappies go on your bum	32	2	0.06	18
		Night time/ Nighttime when you wear a nappy	32	2	0.06	18
66	**nose**	Bogies/ Boogeys	37	3	0.08	18
		Glasses/ glasses(Glasses were on nanny's nose whilst doing the task)	37	2	0.05	18
		Nostril	37	2	0.05	18
		Tongue	37	2	0.05	18
67	**orange**	Red	43	3	0.07	22
		orange juice	43	2	0.05	22
		Apple	43	2	0.05	22
		fruit	43	2	0.05	22
		Yellow	43	2	0.05	22
68	**park**	swing/ swings/ go on the swings	45	7	0.16	23
		slide/ Slide everyday/ Slide on the slide/ slides/ go on the slide	45	7	0.16	23
		Play/ play at the park/ playing	45	4	0.09	23
		roundabout/ go on the roundabout	45	2	0.04	23
		tree/ trees	45	2	0.04	23
69	**pasta**	Eat/ eat up/ eating	37	7	0.19	23
		sauce/ pasta sauce/ saucy sauce	37	4	0.11	23
		cheese	37	3	0.08	23
		tomato/ tomatoes	37	2	0.05	23
70	**peas**	eat/ Eating/ We eat them	37	8	0.22	24
71	**pen**	Draw/ drawing/ DEF: It's something that you draw with	41	6	0.15	16
		Pencil	41	4	0.10	16
		Paper/ Colour on paper	41	3	0.07	16
		Chickens/ Chicken (she said the chicken is in a pen)	41	2	0.05	16
		colouring	41	2	0.05	16
		Crayon	41	2	0.05	16
		write/ writing	41	2	0.05	16
72	**phone/ telephone/ mobile**	Hello/ say hello	32	3	0.09	20
		Ring/ ring ring	32	3	0.09	20
		Watching/ watch	32	2	0.06	20

73	**pig**	'Oink'/ Oink Oink	42	6	0.14	22
		Pink	42	3	0.07	22
		Peppa/ Peppa pig	42	2	0.05	22
		Farm	42	2	0.05	22
		House/ houses	42	2	0.05	22
74	**plane/ aeroplane**	Fly	35	4	0.11	19
		Sky	35	4	0.11	19
		people	35	2	0.06	19
		Sit down/ People sit down	35	2	0.06	19
75	**plate**	Eat/ eating/ We eat food off the plate	34	6	0.18	20
		Food	34	5	0.15	20
		lunch	34	2	0.06	20
		spoon	34	2	0.06	20
		washing up	34	2	0.06	20
76	**pushchair/ buggy**	push/ People push/ pushing	31	4	0.13	16
		Baby/ babies	31	3	0.10	16
		pram	31	2	0.06	16
		wheels	31	2	0.06	16
		Chair/ big chair	31	2	0.06	16
		raincover	31	2	0.06	16
77	**pyjamas/ Pjs/ jim jams**	Bed/ Sleep in bed	38	8	0.21	19
		Sleep	38	2	0.05	19
		Bath/ bath (bedtime routine)	38	2	0.05	19
		Bedtime/ At bed time	38	2	0.05	19
		nice and warm/ warm	38	2	0.05	19
78	**rabbit**	Peter Rabbit/ Peter/ Peter (loves Peter Rabbit)	39	4	0.10	21
		Benjamin (loves Peter Rabbit)/ Benjamin bunny	39	3	0.08	21
		carrot/ carrots	39	3	0.08	21
		Hop	39	3	0.08	21
		Rabbit ears/ big ears	39	2	0.05	21
		tail	39	2	0.05	21
79	**settee/ sofa/ couch**	pillow/ pillows	35	4	0.11	17
		cushion	35	3	0.09	17
		Tellie/ TV/ watching tv	35	3	0.09	17
		Blanket	35	2	0.06	17
		cuddles	35	2	0.06	17
		sit	35	2	0.06	17
80	**sheep**	Cow/ cows	36	5	0.14	17
		Lamb/ lambs	36	5	0.14	17
		Grass	36	4	0.11	17

		horse/ horseys	36	2	0.06	17
81	**shoes**	walk/ Go for a walk	40	4	0.10	20
		Feet	40	2	0.05	20
		Put on/ Shoes on	40	2	0.05	20
82	**slide**	"weeeeeee"/ 'weeee'	33	4	0.12	16
		Swing	33	3	0.09	16
		down	33	2	0.06	16
		Ladder	33	2	0.06	16
		Park/ In the park	33	2	0.06	16
83	**sock**	feet/ foot/ put them on your feet	35	7	0.20	16
		smelly/ smelly sock	35	5	0.14	16
		Dressed/ get dressed	35	3	0.09	16
		on/ put on/ on to play	35	3	0.09	16
		Toes	35	2	0.06	16
84	**spoon**	Fork	34	3	0.09	21
		Yoghurt	34	3	0.09	21
		Breakfast	34	2	0.06	21
		Eating	34	2	0.06	21
		Knife	34	2	0.06	21
		Bowl	34	2	0.06	21
85	**stairs**	Upstairs/ Daddy do work upstairs	30	5	0.17	20
		climb/ climbing	30	3	0.10	20
		Shoes (I usually put our shoes on the stairs to go upstairs)/ shoes. Muddy shoes up the stairs	30	2	0.07	20
86	**swing**	park/ play park/ They are at the park but we can't go to the park because of the germs.	34	6	0.18	20
		sit on/ sitting/ We Sit on them	34	4	0.12	20
		fun	34	2	0.06	20
		hand/ Hand in the air	34	2	0.06	20
87	**table**	Chair/ chairs	30	4	0.13	19
		eat/ eating	30	4	0.13	19
		breakfast	30	2	0.07	19
		Food	30	2	0.07	19
88	**teddy/ teddy bear**	cuddle/ cuddling/ cuddly/ cudddles	29	7	0.24	15
		Bed/ going to bed/ Into bed	29	5	0.17	15
		Sleep/ sleeping	29	2	0.07	15
89	**television/ telly/ TV**	watch	36	4	0.11	20

90	**tiger**	Lion/ Um.. lion!/ Yes…lion! (makes lion noises and pretends to be a lion)	39	5	0.13	23
		"Roar"/ raaarrgh/ rahhhhh/ rawr	39	4	0.10	23
		Stripes/ stripy/ stripey	39	3	0.08	23
		dinosaur	39	2	0.05	23
		orange/ orange lines	39	2	0.05	23
		Sharp teeth	39	2	0.05	23
91	**toast**	jam	50	6	0.12	27
		Eat	50	4	0.08	27
		Bread	50	4	0.08	27
		butter	50	3	0.06	27
		Toaster	50	3	0.06	27
		breakfast	50	2	0.04	27
		honey/ And honey	50	2	0.04	27
		peanut butter	50	2	0.04	27
92	**toe**	feet/ foot	41	4	0.10	21
		nail/ nails	41	2	0.05	21
		Shoe/ shoes	41	2	0.05	21
		sock	41	2	0.05	21
93	**tooth/ teeth**	brush/ brushing/ You brush you teeth very slowly	43	7	0.16	15
		toothbrush/ use a toothbrush/ Unicorn rainbow brush	43	6	0.14	15
		Toothpaste/ use toothpaste/ Pink toothpaste	43	6	0.14	15
		mouth	43	3	0.07	15
		Water	43	2	0.05	15
94	**toothbrush**	Toothpaste/ paste	35	9	0.26	14
		Clean (When I clean her teeth we talk about teeth being shiny and clean)/ Teeth clean/ DEF: It's something I clean my teeth with (action).	35	3	0.09	14
		teeth/ Brush teeth	35	3	0.09	14
95	**towel**	bathroom	36	2	0.06	16
		green	36	2	0.06	16
		swimming	36	2	0.06	16
		Dry	36	3	0.08	16
		bath	36	5	0.14	16

96	**toy**	play/ Play with toys/ To play with/ We play with the toys	37	5	0.14	25
		dinosaur	37	2	0.05	25
		Game/ play game	37	2	0.05	25
		Teddy bear	37	2	0.05	25
		train/ trains	37	2	0.05	25
97	**train**	Choo choo	31	4	0.13	19
		Thomas	31	3	0.10	19
		track/ tracks/ train track	31	3	0.10	19
		Santa	31	2	0.06	19
		wheels	31	2	0.06	19
98	**tree**	leaf/ leaves	32	6	0.19	18
		bird/ birds	32	4	0.13	18
		apples/ Picking apples	32	2	0.06	18
		squirrels	32	2	0.06	18
99	**trousers**	leg/ legs	34	4	0.12	15
		wear/ wear some	34	2	0.06	15
		Jeans	34	2	0.06	15
		pants	34	2	0.06	15
		put it on/ Put them on when we get dressed	34	2	0.06	15
		socks	34	2	0.06	15
100	**window**	Door	44	4	0.09	24
		Clean/ cleaning	44	3	0.07	24
		Flowers/ Flowers too	44	3	0.07	24
		Glass	44	3	0.07	24
		Open	44	2	0.05	24
		raining/ rain	44	2	0.05	24
		Shut	44	2	0.05	24
		curtain/ Curtains and blinds	44	2	0.05	24

Appendix G: Experiment 1. Related responses produced by 2+ children (as first responses) in the parentally-administered WA task

	Cue	Response	No. participants receiving cue (G)	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	apple	pear	13	2	0.15	5
		eat/eat it/you can eat it	13	6	0.46	5
2	arm	hand/hands	17	2	0.12	8
		leg/legs	17	4	0.24	8
3	ball	kick/kicking	14	5	0.36	8
		football	14	2	0.14	8
4	balloon	holding/we hold them/holding a balloon	15	3	0.2	9
		pop	15	2	0.13	9
5	banana	eat it/eat/eating	14	5	0.36	9
6	bath/ bathtub	wash/to wash ourselves/wash hair	15	3	0.2	10
7	bed	teddy/teddy bear/lambie (teddy)/cuddle up with teddies	15	3	0.2	7
		sleep/to sleep	15	4	0.27	7
8	bee	flower/flowers	14	2	0.14	9
		honey	14	3	0.21	9
9	bib	a baby/baby	17	3	0.18	11
		food	17	3	0.18	11
		no (she doesn't wear a bib anymore, her decision. this 'no' is her saying no to wearing a bib.)	17	2	0.12	11
10	bicycle	riding/ride	16	2	0.13	13
11	bin	rubbish/rubbish in the bin/put rubbish in it	14	6	0.43	6
12	bird	feathers/feather	15	2	0.13	10
		nest	15	3	0.2	10
13	biscuit	eat/eat! (shouts excitedly)/eating	14	3	0.21	9
		chocolate/chocolate biscuit	14	2	0.14	9

14	boat	water/in the water/we was on a boat on water/swim in water	15	4	0.27	7
		sailing/sail	15	2	0.13	7
15	book	read/reading/read story	14	5	0.36	5
		pages	14	2	0.14	5
16	boots	puddle/puddles/jumping in muddy puddles/muddy puddles/splashing in muddy puddles	14	2	0.14	7
		walking/walk	14	2	0.14	7
		wellies	14	2	0.14	7
17	bottle	water/water bottle	15	3	0.2	8
		milk	15	2	0.13	8
		lid	15	2	0.13	8
18	bowl	breakfast	14	2	0.14	9
19	box	make (makes models from boxes)/make something/making	17	3	0.18	10
		put stuff in it/stuff	17	2	0.12	10
20	bread	eat/i eat it	16	2	0.13	10
		toast	16	2	0.13	10
21	brush	hair/brush everyone hair/brush hair/sophie's long hair/hairbrush	15	6	0.4	8
		teeth/brush your teeth	15	3	0.2	8
22	bubbles	pop/def: they're something that pop. pop the bubbles/	18	5	0.28	10
		blowing/blow/blow bubbles	18	2	0.11	10
23	bus	big/big bus	14	2	0.14	11
		we go on the bus/take us somewhere we like to go	14	2	0.14	11
24	butterfly	flying/fly away/flies	14	3	0.21	8
		wings/yellow wings/yellow and blue wings	14	3	0.21	8
25	cake	birthday	15	2	0.13	10
		eat/eat it/after you make it you eat the cake	15	3	0.2	10
26	car	wheels/wheels to bump	16	3	0.19	8
27	carrots	eat/eat them/eating	14	3	0.21	8
		crunchy/crunch crunch	14	2	0.14	8
28	cat	dog	19	6	0.32	10
29	cereal	eat cereal/eat	16	2	0.13	14

30	chair	sitting/sit/sit down	13	4	0.31	8
31	cheese	eat it/eat/eating/def: easy- it's something that you eat and it's so squeezey	15	5	0.33	8
32	chicken	eggs/egg	15	3	0.2	11
		eat/eating/we eat it	15	2	0.13	11
33	coat	cold	14	2	0.14	10
		jacket	14	2	0.14	10
34	cot	baby/babies/baby sleeps	13	3	0.23	7
35	cow	milk/ae some milk	14	3	0.21	7
		moo	14	2	0.14	7
		pig	14	2	0.14	7
36	cup	drink/drink it/ daddy drinking	15	2	0.13	5
		tea	15	2	0.13	5
		milk	15	2	0.13	5
		water	15	3	0.2	5
37	dog	woof/woof woof	15	3	0.2	10
		cat/kitty cat	15	5	0.33	10
38	doll	boy	17	2	0.12	11
39	door	open/open it	16	3	0.19	12
40	duck	water	13	3	0.23	6
		quack	13	5	0.38	6
41	ear	listen/listening ears	14	2	0.14	8
42	elephant	trunk	15	4	0.27	5
43	eye	spy/I spy	15	2	0.13	10
44	feet	toes	14	2	0.14	11
45	finger	hand/hands	16	5	0.31	7
46	fish	swimming/swim	14	2	0.14	8
		water/lives in water	14	4	0.29	8
		fingers/fish finger/eat fish fingers	14	2	0.14	8
47	flower	bee/bees/buzzy bee	15	4	0.27	7
		petals	15	2	0.13	7
48	fork	spoon	18	5	0.28	5
		knife	18	3	0.17	5
		to eat/eat	18	4	0.22	5
49	frog	frog says ribbit/ribbit	16	2	0.13	12
		water	16	3	0.19	12
50	hair	cut your hair/cutting	15	2	0.13	7
		brush	15	2	0.13	7
		head	15	3	0.2	7
51	hand	foot	14	2	0.14	6

		fingers/finger	14	3	0.21	6
52	hat	head/it goes on your head	14	2	0.14	9
		wear it/we can wear hat	14	2	0.14	9
53	head	hair/hair on	15	6	0.4	7
54	key	lock/lock the door	17	2	0.12	6
		door	17	7	0.41	6
55	leg	foot/feet	18	5	0.28	10
56	lion	roar/they roar/raah	13	9	0.69	3
57	lorry/ truck	drive/def: it's something that drives	15	2	0.13	10
		wheel/wheels	15	2	0.13	10
58	monkey	elephant	14	3	0.21	7
		swing/swing in branches	14	2	0.14	7
		banana	14	3	0.21	7
59	mouse	squeak/goes squeak/it squeaks	15	3	0.2	12
60	mouth	hair off/hair	15	2	0.13	9
		teeth	15	2	0.13	9
		tongue/points to tongue	15	3	0.2	9
61	nappy	baby/babies	16	3	0.19	12
		put nappy on/put on people	16	2	0.13	12
62	nose	bogies/boogeys	15	2	0.13	11
63	orange	fruit	16	2	0.13	10
64	park	swings/swing/go on the swings	17	3	0.18	12
		play/play at a park/playing	17	3	0.18	12
65	pasta	eat/eating/eat up	16	5	0.31	9
		cheese/cream cheese	16	2	0.13	9
		dinner	16	2	0.13	9
66	peas	eat/we eat them/eating	13	3	0.23	
67	pen	draw/drawing/ it's something that you draw with	16	4	0.25	10
		write/writing	16	2	0.13	10
68	phone/ telephone/ mobile	ring/ring ring	14	2	0.14	8
69	pig	oink/oink oink	16	3	0.19	10
		pink	16	2	0.13	10
70	plane	people	16	2	0.13	10
		sky	16	3	0.19	10
71	plate	lunch	14	2	0.14	7

		food/we eat food off the plate	14	5	0.36	7
		eating/eat	14	3	0.21	7
72	pushchair/ buggy	push/pushing/people push	14	2	0.14	6
		baby/babies	14	2	0.14	6
		pram/maia goes in pram	14	3	0.21	6
		wheels	14	2	0.14	6
73	pyjamas/ Pjs/jim jams	bed	16	3	0.19	10
		bedtime/at bed time	16	2	0.13	10
74	rabbit	peter rabbit/peter (loves peter rabbit)	15	4	0.27	8
		rabbit ears/big ears	15	2	0.13	8
		hop	15	2	0.13	8
75	settee/ sofa/ couch	cushion	14	3	0.21	8
		pillows/pillow	14	3	0.21	8
76	sheep	cow/cows	18	5	0.28	6
		lambs/lamb	18	4	0.22	6
		grass	18	2	0.11	6
77	shoes	feet	14	2	0.14	10
78	slide	park/in the park	16	2	0.13	10
		ladder	16	2	0.13	10
		swing	16	2	0.13	10
		down	16	2	0.13	10
79	sock	foot/feet/put them on your feet	13	5	0.38	6
		smelly/smelly sock	13	4	0.31	6
80	spoon	breakfast	15	2	0.13	8
		fork	15	3	0.2	8
81	stairs	upstairs	14	3	0.21	9
		climb/climbing	14	3	0.21	9
82	swing	fun	14	2	0.14	8
		we sit on them/sit on/sit/sitting	14	3	0.21	8
		park/play park/they are at the park but we can't go to the park because of the germs.	14	2	0.14	8
83	table	eat/eating/eat dinner	14	4	0.29	7
		chair/chairs	14	3	0.21	7
84	teddy/ teddy bear	cuddle/cuddling/cuddles/cuddly	14	4	0.29	7
		bed/going to bed/into bed/bedtime	14	4	0.29	7

85	television/ telly/TV	watch	14	3	0.21	8
86	tiger	rawr/"roar" (nb. he leapt up with sound affects)/rahhhhh/raaarrgh	14	3	0.21	8
		stripes/stripy/stripey	14	2	0.14	8
87	toast	breakfast/mommy breakfast	19	2	0.11	10
		eat	19	3	0.16	10
		bread	19	4	0.21	10
		jam	19	2	0.11	10
88	toe	foot/feet	17	4	0.24	10
89	tooth/ teeth	toothbrush/use a toothbrush	18	4	0.22	8
		brush/brushing/you brush you teeth very slowly/tiny little brushs (she puts two fingers together)/unicorn rainbow brush	18	6	0.33	8
90	toothbrush	mouth/brushing my mouth	12	2	0.17	5
		teeth/teeth clean/brush teeth	12	3	0.25	5
		toothpaste/paste	12	4	0.33	5
91	towel	bath/bath (he'd just got out the bath)/bath time	14	5	0.36	6
		bathroom	14	2	0.14	6
92	toy	play/play time/to play with/we play with the toys	14	4	0.29	10
93	train	choo choo	14	3	0.21	8
		tracks/track/train track	14	3	0.21	8
94	tree	leaf/leaves	15	4	0.27	8
		birds/bird	15	2	0.13	8
		apples/picking apples	15	2	0.13	8
95	trousers	put it on/put them on when we get dressed	14	2	0.14	6
		leg/legs	14	3	0.21	6
		wear some/wear	14	2	0.14	6
		jeans/red jeans	14	2	0.14	6
96	window	raining/rain	18	2	0.11	12
		door	18	2	0.11	12

Appendix H: Experiment 1. Related, imageable noun responses produced by 2+ children (as first responses) in the parentally-administered WA task

	Cue	Response	No. participants receiving cue (G)	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	bin	rubbish/rubbish in the bin/put rubbish in it	14	6	0.43	6
2	key	door	17	7	0.41	6
3	brush	hair/brush everyone hair/brush hair/sophie's long hair/hairbrush	15	6	0.4	8
4	head	hair/hair on	15	6	0.4	7
5	sock	foot/feet/put them on your feet	13	5	0.38	6
6	plate	food/we eat food off the plate	14	5	0.36	7
7	towel	bath/bath (he'd just got out the bath)/bath time	14	5	0.36	6
8	dog	cat/kitty cat	15	5	0.33	10
9	toothbrush	toothpaste/paste	12	4	0.33	5
10	cat	dog	19	6	0.32	10
11	finger	hand/hands	16	5	0.31	7
12	fish	water/lives in water	14	4	0.29	8
13	teddy/ teddy bear	bed/going to bed/into bed/bedtime	14	4	0.29	7
14	fork	spoon	18	5	0.28	5
15	leg	foot/feet	18	5	0.28	10
16	sheep	cow/cows	18	5	0.28	6
17	boat	water/in the water/we was on a boat on water/swim in water	15	4	0.27	7
18	elephant	trunk	15	4	0.27	5
19	flower	bee/bees/buzzy bee	15	4	0.27	7
20	rabbit	peter rabbit/peter (loves peter rabbit)	15	4	0.27	8
21	tree	leaf/leaves	15	4	0.27	8
22	arm	leg/legs	17	4	0.24	8
23	toe	foot/feet	17	4	0.24	10
24	cot	baby/babies/baby sleeps	13	3	0.23	7
25	duck	water	13	3	0.23	6
26	sheep	lambs/lamb	18	4	0.22	6
27	tooth/ teeth	toothbrush/use a toothbrush	18	4	0.22	8
28	bee	honey	14	3	0.21	9

29	butterfly	wings/yellow wings/yellow and blue wings	14	3	0.21	8
30	cow	milk/ae some milk	14	3	0.21	7
31	hand	fingers/finger	14	3	0.21	6
32	monkey	elephant	14	3	0.21	7
33	monkey	banana	14	3	0.21	7
34	pushchair/ buggy	pram/maia goes in pram	14	3	0.21	6
35	settee/ sofa/couch	cushion	14	3	0.21	8
36	settee/ sofa/couch	pillows/pillow	14	3	0.21	8
37	stairs	upstairs	14	3	0.21	9
38	table	chair/chairs	14	3	0.21	7
39	toast	bread	19	4	0.21	10
40	train	tracks/track/train track	14	3	0.21	8
41	trousers	leg/legs	14	3	0.21	6
42	bed	teddy/teddy bear/lambie (teddy)/cuddle up with teddies	15	3	0.2	7
43	bird	nest	15	3	0.2	10
44	bottle	water/water bottle	15	3	0.2	8
45	brush	teeth/brush your teeth	15	3	0.2	8
46	chicken	eggs/egg	15	3	0.2	11
47	cup	water	15	3	0.2	5
48	hair	head	15	3	0.2	7
49	spoon	fork	15	3	0.2	8
50	car	wheels/wheels to bump	16	3	0.19	8
51	frog	water	16	3	0.19	12
52	nappy	baby/babies	16	3	0.19	12
53	plane	sky	16	3	0.19	10
54	pyjamas/ Pjs/jim jams	bed	16	3	0.19	10
55	bib	a baby/baby	17	3	0.18	11
56	bib	food	17	3	0.18	11
57	park	swings/swing/go on the swings	17	3	0.18	12
58	park	play/play at a park/playing	17	3	0.18	12
59	fork	knife	18	3	0.17	5
60	toothbrush	mouth/brushing my mouth	12	2	0.17	5
61	apple	pear	13	2	0.15	5
62	ball	football	14	2	0.14	8

63	bee	flower/flowers	14	2	0.14	9
64	biscuit	chocolate/chocolate biscuit	14	2	0.14	9
65	book	pages	14	2	0.14	5
66	boots	puddle/puddles/jumping in muddy puddles/muddy puddles/splashing in muddy puddles	14	2	0.14	7
67	boots	wellies	14	2	0.14	7
68	bowl	breakfast	14	2	0.14	9
69	coat	jacket	14	2	0.14	10
70	cow	pig	14	2	0.14	7
71	feet	toes	14	2	0.14	11
72	fish	fingers/fish finger/eat fish fingers	14	2	0.14	8
73	hand	foot	14	2	0.14	6
74	hat	head/it goes on your head	14	2	0.14	9
75	plate	lunch	14	2	0.14	7
76	pushchair/ buggy	baby/babies	14	2	0.14	6
77	pushchair/ buggy	wheels	14	2	0.14	6
78	shoes	feet	14	2	0.14	10
79	swing	park/play park/they are at the park but we can’t go to the park because of the germs.	14	2	0.14	8
80	tiger	stripes/stripy/stripey	14	2	0.14	8
81	towel	bathroom	14	2	0.14	6
82	trousers	jeans/red jeans	14	2	0.14	6
83	bird	feathers/feather	15	2	0.13	10
84	boat	sailing/sail	15	2	0.13	7
85	bottle	milk	15	2	0.13	8
86	bottle	lid	15	2	0.13	8
87	bread	toast	16	2	0.13	10
88	cake	birthday	15	2	0.13	10
89	cup	drink/drink it/ daddy drinking	15	2	0.13	5
90	cup	tea	15	2	0.13	5
91	cup	milk	15	2	0.13	5
92	flower	petals	15	2	0.13	7
93	lorry/ truck	wheel/wheels	15	2	0.13	10
94	mouth	hair off/hair	15	2	0.13	9
95	mouth	teeth	15	2	0.13	9
96	nose	bogies/boogeys	15	2	0.13	11

97	orange	fruit	16	2	0.13	10
98	pasta	cheese/cream cheese	16	2	0.13	9
99	pasta	dinner	16	2	0.13	9
100	plane	people	16	2	0.13	10
101	pyjamas/Pjs/jim jams	bedtime/at bed time	16	2	0.13	10
102	rabbit	rabbit ears/big ears	15	2	0.13	8
103	slide	park/in the park	16	2	0.13	10
104	slide	ladder	16	2	0.13	10
105	slide	swing	16	2	0.13	10
106	spoon	breakfast	15	2	0.13	8
107	tree	birds/bird	15	2	0.13	8
108	tree	apples/picking apples	15	2	0.13	8
109	arm	hand/hands	17	2	0.12	8
110	doll	boy	17	2	0.12	11
111	sheep	grass	18	2	0.11	6
112	toast	breakfast/mommy breakfast	19	2	0.11	10
113	toast	jam	19	2	0.11	10
114	window	raining/rain	18	2	0.11	12
115	window	door	18	2	0.11	12

Appendix I: Experiment 1. Related responses given by 2+ children (as first responses) in the parentally-administered WA task and represented in adult associative norms

Cue	Response	Child FSG	Adult FSG
apple	pear	0.15	0.15
	eat/eat it/you can eat it	0.46	0.01
arm	hand/hands	0.12	0.08
	leg/legs	0.24	0.54
ball	kick/kicking	0.36	0.07
	football	0.14	0.04
balloon	pop	0.13	0.15
bath/bathtub	wash/to wash ourselves/wash hair	0.20	0.02
bed	sleep/to sleep	0.27	0.64
bee	honey	0.21	0.22
bib	a baby/baby	0.18	0.63
bicycle	riding/ride	0.13	0.19
bin	rubbish/rubbish in the bin/put rubbish in it	0.43	0.33
bird	feathers/feather	0.13	0.06
	nest	0.20	0.04
biscuit	eat/eat! (shouts excitedly)/eating	0.21	0.02
	chocolate/chocolate biscuit	0.14	0.08
boat	water/in the water/we was on a boat on water/swim in water	0.27	0.24
	sailing/sail	0.13	0.14
book	read/reading/read story	0.36	0.33
	pages	0.14	0.06
boots	walking/walk	0.14	0.02
bottle	water/water bottle	0.20	0.03
bowl	breakfast	0.14	0.02
bread	eat/i eat it	0.13	0.03
brush	hair/brush everyone hair/brush hair/sophie’s long hair/hairbrush	0.40	0.32
	teeth/brush your teeth	0.20	0.16
bubbles	pop/def: they're something that pop. pop the bubbles/	0.28	0.02
	blowing/blow/blow bubbles	0.11	0.03
butterfly	flying/fly away/flies	0.21	0.08
	wings/yellow wings/yellow and blue wings	0.21	0.09
cake	birthday	0.13	0.07

	eat/eat it/after you make it you eat the cake	0.20	0.09
car	wheels/wheels to bump	0.19	0.04
carrots	eat/eat them/eating	0.21	0.01
	crunchy/crunch crunch	0.14	0.01
cat	dog	0.32	0.59
cereal	eat cereal/eat	0.13	0.03
chair	sitting/sit/sit down	0.31	0.21
cheese	eat it/eat/eating/def: easy- it's something that you eat and it's so squeezey	0.33	0.02
chicken	eggs/egg	0.20	0.02
	eat/eating/we eat it	0.13	0.02
coat	cold	0.14	0.07
	jacket	0.14	0.15
cot	baby/babies/baby sleeps	0.23	0.64
cow	milk/ae some milk	0.21	0.35
	moo	0.14	0.06
	pig	0.14	0.02
cup	drink/drink it/ daddy drinking	0.13	0.03
	tea	0.13	0.07
	water	0.20	0.06
dog	cat/kitty cat	0.33	0.59
door	open/open it	0.19	0.16
duck	water	0.23	0.02
	quack	0.38	0.11
ear	listen/listening ears	0.14	0.03
elephant	trunk	0.27	0.21
feet	toes	0.14	0.28
finger	hand/hands	0.31	0.24
fish	swimming/swim	0.14	0.08
	water/lives in water	0.29	0.09
	fingers/fish finger/eat fish fingers	0.14	0.04
flower	petals	0.13	0.17
fork	spoon	0.28	0.33
	knife	0.17	0.41
	to eat/eat	0.22	0.05
fish	water	0.19	0.02
hair	cut your hair/cutting	0.13	0.04
	brush	0.13	0.11
	head	0.20	0.03
hand	foot	0.14	0.13

	fingers/finger	0.21	0.23
hat	head/it goes on your head	0.14	0.22
head	hair/hair on	0.40	0.13
key	lock/lock the door	0.12	0.37
	door	0.41	0.19
leg	foot/feet	0.28	0.13
lion	roar/they roar/raah	0.69	0.03
lorry/ truck	drive/def: it's something that drives	0.13	0.02
	wheel/wheels	0.13	0.02
monkey	swing/swing in branches	0.14	0.01
	banana	0.21	0.05
mouth	teeth	0.13	0.18
	tongue/points to tongue	0.20	0.08
orange	fruit	0.13	0.15
park	swings/swing/go on the swings	0.18	0.04
	play/play at a park/playing	0.18	0.02
peas	eat/we eat them/eating	0.23	0.01
pen	write/writing	0.13	0.07
phone/ telephone/ mobile	ring/ring ring	0.14	0.27
pig	oink/oink oink	0.19	0.04
	pink	0.13	0.02
plane	sky	0.19	0.09
plate	food/we eat food off the plate	0.36	0.19
	eating/eat	0.21	0.05
pushchair	baby/babies	0.14	0.07
pyjamas/Pjs/ jim jams	bed	0.19	0.15
	bedtime/at bed time	0.13	0.02
rabbit	peter rabbit/peter (loves peter rabbit)	0.27	0.01
	rabbit ears/big ears	0.13	0.05
	hop	0.13	0.02
settee/sofa/ couch	cushion	0.21	0.05
sheep	cow/cows	0.28	0.07
	lambs/lamb	0.22	0.09
shoes	feet	0.14	0.33
slide	park/in the park	0.13	0.02
	swing	0.13	0.11
	down	0.13	0.09
sock	foot/feet/put them on your feet	0.38	0.17
spoon	fork	0.20	0.50

stairs	upstairs	0.21	0.14
	climb/climbing	0.21	0.23
swing	fun	0.14	0.01
	park/play park/they are at the park but we can't go to the park because of the germs.	0.14	0.08
table	eat/eating/eat dinner	0.29	0.03
	chair/chairs	0.21	0.76
television/ telly/TV	watch	0.21	0.09
tiger	rawr/"roar" (nb. he leapt up with sound affects)/rahhhhh/raaarrgh	0.21	0.02
	stripes/stripy/stripey	0.14	0.08
toast	breakfast/mommy breakfast	0.11	0.07
	eat	0.16	0.02
	bread	0.21	0.36
	jam	0.11	0.01
toe	foot/feet	0.24	0.58
tooth/ teeth	toothbrush/use a toothbrush	0.22	0.02
	brush/brushing/you brush you teeth very slowly/tiny little brushs (she puts two fingers together)/unicorn rainbow brush	0.33	0.10
toothbrush	mouth/brushing my mouth	0.17	0.02
	teeth/teeth clean/brush teeth	0.25	0.16
	toothpaste/paste	0.33	0.32
towel	bath/bath (he'd just got out the bath)/bath time	0.36	0.05
	bathroom	0.14	0.02
toy	play/play time/to play with/we play with the toys	0.29	0.10
train	choo choo	0.21	0.03
	tracks/track/train track	0.21	0.18
tree	leaf/leaves	0.27	0.18
trousers	leg/legs	0.21	0.01
	wear some/wear	0.14	0.02
	jeans/red jeans	0.14	0.04
window	door	0.11	0.15

Appendix J: Experiment 1. Related responses given by 2+ children (as first responses) in the parentally-administered WA task, and not represented in adult associative norms

Cue	Response	Child FSG	South Florida Norms: Nelson et al. (1998)	Birkbeck Norms: Moss and Older (1996)
banana	eat it/eat/eating	0.36	n/d	n/c
pasta	eat/eating/eat up	0.31	n/d	n/c
sock	smelly/smelly sock	0.31	n/d	n/c
teddy/ teddy bear	cuddle/cuddling/cuddles/cuddly	0.29	n/c	n/c
teddy/ teddy bear	bed/going to bed/into bed/bedtime	0.29	n/c	n/c
flower	bee/bees/buzzy bee	0.27	n/d	n/d
pen	draw/drawing/ it's something that you draw with	0.25	n/d	n/c
monkey	elephant	0.21	n/d	n/d
pushchair/buggy	pram/maia goes in pram	0.21	n/d	n/c
settee/ sofa/couch	pillows/pillow	0.21	n/d	n/d
swing	we sit on them/sit on/sit/sitting	0.21	n/d	n/d
balloon	holding/we hold them/holding a balloon	0.2	n/d	n/d
bed	teddy/teddy bear/lambie (teddy)/cuddle up with teddies	0.2	n/d	n/c
dog	woof/woof woof	0.2	n/d	n/d
mouse	squeak/goes squeak/it squeaks	0.2	n/d	n/d
nappy	baby/babies	0.19	n/c	n/c
bib	food	0.18	n/c	n/d
box	make (makes models from boxes)/make something/making	0.18	n/d	n/d
bee	flower/flowers	0.14	n/d	n/d
boots	puddle/puddles/jumping in muddy puddles/muddy puddles/splashing in muddy puddles	0.14	n/d	n/d
boots	wellies	0.14	n/d	n/d
bus	big/big bus	0.14	n/d	n/d
bus	we go on the bus/take us somewhere we like to go	0.14	n/d	n/d
hat	wear it/we can wear hat	0.14	n/d	n/d
plate	lunch	0.14	n/d	n/d
pushchair/ buggy	push/pushing/people push	0.14	n/d	n/c
pushchair/ buggy	wheels	0.14	n/d	n/c
trousers	put it on/put them on when we get dressed	0.14	n/d	n/d

bottle	milk	0.13	n/d	n/c
bottle	lid	0.13	n/d	n/c
bread	toast	0.13	n/d	n/d
cup	milk	0.13	n/d	n/d
eye	spy/I spy	0.13	n/d	n/d
frog	frog says ribbit/ribbit	0.13	n/d	n/d
mouth	hair off/hair	0.13	n/d	n/d
nappy	put nappy on/put on people	0.13	n/c	n/c
nose	bogies/boogeys	0.13	n/d	n/d
pasta	cheese/cream cheese	0.13	n/d	n/c
pasta	dinner	0.13	n/d	n/c
plane	people	0.13	n/d	n/d
slide	ladder	0.13	n/d	n/d
spoon	breakfast	0.13	n/d	n/d
tree	birds/bird	0.13	n/d	n/d
tree	apples/picking apples	0.13	n/d	n/d
bib	no (she doesn’t wear a bib anymore, her decision. this ‘no’ is her saying no to wearing a bib.)	0.12	n/c	n/d
box	put stuff in it/stuff	0.12	n/d	n/d
doll	boy	0.12	n/d	n/d
sheep	grass	0.11	n/d	n/d
window	raining/rain	0.11	n/d	n/c

Appendix K: Experiment 2. All related responses given by 2+ children in the online WA task

	Cue	Response	No. participants receiving cue (G)*3 attempts	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	bath	we play/ play	36	2	0.06	19
		shower/ i have a shower	36	2	0.06	19
		wash time/ wash yourself in the bath	36	2	0.06	19
		towel	36	2	0.06	19
		not wash my hair/ scrub your hair	36	2	0.06	19
		mummy/ Mummys bath	36	3	0.08	19
		toys/ Toys (bath toys)/ put toys in the bath/ dinosaur (bath toys)	36	6	0.17	19
2	bed	lay on it/ lay in bed	36	2	0.06	15
		teddy	36	2	0.06	15
		pillow/s	36	3	0.08	15
		sleep/ sleeping/ sleep on it	36	3	0.08	15
		toys/ mushroom (soft toy)/ pumpkin (soft toy)	36	5	0.14	15
3	bowl	Poppy - our cat has bowls and he puts the food in them/ cat food	36	2	0.06	13
		spoon	36	2	0.06	13
		you eat food out of your bowl/ eating	36	3	0.08	13
4	brush	combing/ comb	36	2	0.06	20
		pink	36	2	0.06	20
		hair/ we use it for our hair	36	7	0.19	20
5	cereal	milk	36	2	0.06	17
		EAT/ eating	36	2	0.06	17
		porridge	36	2	0.06	17
		breakfast	36	4	0.11	17
6	chair	mummy/ mum	36	2	0.06	11
		table	36	2	0.06	11
		sit on it/ sit	36	3	0.08	11
7	door	handle/door handle	36	2	0.06	17
		close the door/ close	36	2	0.06	17

		outside	36	2	0.06	17
		shut/ shutting	36	3	0.08	17
		opening/ open door/ open/ open and shut them	36	4	0.11	17
8	finger	wiggle your finger/ wiggly worms	36	2	0.06	14
		hand/ hands/ red ouchie on my hand	36	3	0.08	14
9	foot	toes	36	3	0.08	14
		socks on it/ socks	36	3	0.08	14
		shoes/ get you shoes on	36	3	0.08	14
10	hair	wash it with soap/ wash hair and go to bed/ wash your hair	36	3	0.08	13
		brush your hair/ brush/ brushing hair in the bath/ brush it/ brush hair	36	5	0.14	13
11	hand	fingers/ finger	36	2	0.06	14
		Wash/ wash your hands	36	2	0.06	14
12	head	shoulders/ shoulders kneesand toes	36	2	0.06	21
		ears/ ears on your head	36	2	0.06	21
		eyes	36	2	0.06	21
		feet	36	2	0.06	21
		hair/ it has hair/ hair on your head/ cradle cap in my hair	36	4	0.11	21
13	key	lock the car/ car	36	2	0.06	19
		lock/ lock things with a key/ locking us in	36	4	0.11	19
		lock it up with a door/ lock the door/ door/ a door/	36	4	0.11	19
14	park	we play/ play in the mud	36	2	0.06	20
		slide	36	3	0.08	20
		swing/ swings in the park/ we swing/ swings	36	6	0.17	20
15	sock	smelly/ smelly welly	36	2	0.06	15
		pointed to foot	36	3	0.08	15

		wear them on our feet/ pointed to foot/ put on foot/ you put your sock on your feet	36	4	0.11	15
16	swing	wee-weeeee/ weee	36	2	0.06	11
		sit on it	36	2	0.06	11
		up	36	2	0.06	11
		Down	36	2	0.06	11
		push/ push up high/ push it	36	3	0.08	11
		at the park/ park/ park swing	36	4	0.11	11
17	table	dinner and tea/ tea	36	2	0.06	13
		breakfast	36	3	0.08	13
		food/ dinner with food	36	3	0.08	13
		Eat/ eating at the table/ eat/ we eat on it/ we eat pancakes there	36	5	0.14	13
18	teddy	bed/ take them to bed	36	2	0.06	17
		Cuddle/ cuddly unicorn/ cuddling/ cuddle them	36	4	0.11	17
19	tooth	toothpaste	36	2	0.06	17
		bite with them/ indicated biting	36	2	0.06	17
		mummy/ mum	36	2	0.06	17
		brush/ brushing your teeth/ brushing/ Brushing teeth/ brush them	36	5	0.14	17
20	towel	Washing	36	2	0.06	10
		for my face/ face	36	2	0.06	10
		drying off/ dry/ dry hands/ drying/ dry the cat	36	6	0.17	10

Appendix L: Experiment 2. Related responses given by 2+ children (as first responses) in the online WA task

	Cue	Response	No. participants receiving cue (G)	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	bath	toys/ Toys (bath toys)/ put toys in the bath/ dinosaur (bath toys)	12	4	0.33	8
2	bed	toys/ mushroom (soft toy)/ pumpkin (soft toy)	12	2	0.17	6
		pillow/s	12	2	0.17	6
		sleep/ sleeping/ sleep on it	12	2	0.17	6
		lay on it/ lay in bed	12	2	0.17	6
3	bowl	you eat food out of your bowl/ eating	12	3	0.25	6
4	brush	hair/ we use it for our hair	12	5	0.42	8
5	cereal	EAT/ eating	12	2	0.17	7
		breakfast	12	3	0.25	7
6	chair	sit on it/ sit	12	3	0.25	7
7	door	close the door/ close	12	2	0.17	8
8	foot	toes	12	3	0.25	7
9	hair	brush your hair/ brush/ brushing hair in the bath/ brush it/ brush hair	12	3	0.25	5
10	head	shoulders/ shoulders kneesand toes	12	2	0.17	9
11	key	lock/ lock things with a key/ locking us in	12	3	0.25	8
12	park	slide	12	2	0.17	6
		swing/ swings in the park/ we swing/ swings	12	4	0.33	6
13	sock	wear them on our feet/ pointed to foot/ put on foot/ you put your sock on your feet	12	2	0.17	7
14	swing	at the park/ park/ park swing	12	2	0.17	6
15	table	breakfast	12	2	0.17	5
		food/ dinner with food	12	2	0.17	5
		Eat/ eating at the table/ eat/ we eat on it/ we eat pancakes there	12	3	0.25	5
16	teddy	Cuddle/ cuddly unicorn/ cuddling/ cuddle them	12	3	0.25	6
17	tooth	brush/ brushing your teeth/ brushing/ Brushing teeth/ brush them	12	2	0.17	9
18	towel	drying off/ dry/ dry hands/ drying/ dry the cat	12	4	0.33	4

Appendix M: Experiment 2. Word associations replicated in Experiment 1 and 2

	CUE	RESPONSES repeated by 2+ participants	FSG_Online	FSG_Parental	FSG_Birkbeck Norms: Moss and Older (1996)	FSG_South Florida Norms: Nelson et al. (1998)
1	**brush**	hair/ we use it for our hair	0.188	0.190	0.200	0.440
2	**hair**	Brush/ brush it/ brush your hair/ brushing hair in the bath	0.182	0.110	0.021	0.207
3	**swing**	park/ at the park/ park swing	0.174	0.180	0.067	0.101
4	**park**	swing/ swings/ swings in the park/ we swing	0.171	0.160	0.021	0.061
5	**bath**	toys/ put toys in the bath/ dinosaur (bath toys)/ Toys (bath toys)	0.171	0.130	not documented	not documented
6	**tooth**	brush	0.152	0.160	0.115	0.123
7	**door**	Open/ open door/ open and shut them/ opening	0.148	0.150	0.146	0.183
8	**key**	door/ lock it up with a door/ lock the door	0.147	0.210	0.156	0.218
9	**table**	Eat/ eating at the table/ we eat on it	0.143	0.130	not a cue	0.026
10	**towel**	dry/ drying/ drying off	0.143	0.080	not a cue	0.284
11	**sock**	foot/feet: put on foot/ you put your sock on your feet/ wear them	0.138	0.200	not a cue	0.172
12	**foot**	Toes	0.125	0.110	0.085	0.466
13	**foot**	socks/ socks on it	0.125	0.050	0.043	not documented
14	**cereal**	breakfast	0.121	0.050	0.548	0.333
15	**key**	lock/ lock things with a key/ locking us in	0.118	0.050	0.489	0.255
16	**head**	hair/ hair on your head/ it has hair/ cradle cap in my hair	0.114	0.220	0.064	0.186
17	**door**	Shut/ shutting	0.111	0.060	0.062	not documented
18	**table**	breakfast	0.107	0.070	not a cue	not documented
19	**teddy**	Cuddle/ cuddle them/ cuddling	0.100	0.240	not a cue	not a cue
20	**hand**	finger/ fingers	0.095	0.150	0.095	0.358
21	**chair**	sit/ sit on it	0.094	0.110	not a cue	0.212
22	**swing**	sit on it	0.087	0.120	not documented	not documented
23	**park**	slide	0.086	0.160	not documented	not documented
24	**bed**	sleep	0.086	0.150	not a cue	0.638
25	**foot**	shoes	0.083	0.050	0.192	0.108
26	**finger**	hand	0.074	0.190	0.125	0.268
27	**sock**	smelly/ smelly welly	0.069	0.140	not a cue	not documented
28	**teddy**	bed/ take them to bed	0.067	0.170	not a cue	not a cue
29	**bowl**	food/ you eat food out of your bowl	0.067	0.070	not documented	0.017
30	**chair**	table	0.063	0.050	not a cue	0.314
31	**tooth**	toothpaste	0.061	0.140	0.019	0.058
32	**cereal**	milk	0.061	0.130	not documented	0.204
33	**cereal**	EAT/ eating	0.061	0.080	not documented	0.031
34	**key**	car/ lock the car	0.059	0.070	not documented	0.115
35	**bed**	teddy	0.057	0.130	not a cue	not documented
36	**bath**	wash time/ wash yourself in the bath	0.057	0.080	not documented	0.024
37	**head**	ears/ ears on your head	0.057	0.060	not documented	not documented
38	**head**	eyes	0.057	0.060	not documented	not documented

Appendix N: Experiment 3. All related responses given by 2+ bilingual children in the online WA task

	Cue	Response	No. participants receiving cue (G)*3 attempts	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	bath	toys/ juguetes/dinosours (toys in bath)	30	4	0.13	12
		water	30	3	0.10	12
		jugar(play)/play	30	2	0.07	12
2	bed	sleep/ domir	30	4	0.13	13
		toys (soft toys)/friends (toys)	30	2	0.07	13
3	bowl	eat/ eating	30	2	0.07	11
		chocolate/ chocolate /she had just eaten some chcolate from a bowl	30	2	0.07	11
4	brush	teeth	24	3	0.13	7
		hair	24	2	0.08	7
5	cereal	eating	24	2	0.08	12
		milk	24	2	0.08	12
		drink/drinking (the milk)	24	2	0.08	12
6	chair	table	30	4	0.13	9
		sit/ sitting (said in French: assis)/sentarse	30	4	0.13	9
		eat/eating	30	3	0.10	9
7	door	open (she said it in english)/open door	24	4	0.17	8
		shut the door/shut (she said it in english)/shutting	24	3	0.13	8
		close door/close	24	2	0.08	8
8	finger	hand/hands	30	3	0.10	13
9	foot	walking	24	2	0.08	13
		shoes	24	2	0.08	13
10	hair	to brush/brush your hair	24	2	0.08	12
		shaking/shake	24	2	0.08	12
11	hand	eating	24	2	0.08	12
		hold hands/hold hand when crossing the street	24	2	0.08	12
		finger/s	24	2	0.08	12
12	head	hair /pelo/wet hair	30	3	0.10	14
		shoulder/head, shoulders, knees and toes	30	2	0.07	14
		foot/ feet	30	2	0.07	14
13	key	door/ puerta	30	6	0.20	10
		lock	30	3	0.10	10
		the car (coche)	30	2	0.07	10
		going/going out (to the park)	30	2	0.07	10

14	park	car/park your car	30	3	0.10	13
		slide	30	3	0.10	13
		play/ jugar	30	2	0.07	13
		swing/columpio	30	2	0.07	13
15	sock	foot/feet/pie	30	5	0.17	12
		trousers	30	2	0.07	12
		toes	30	2	0.07	12
16	swing	park	24	3	0.13	11
		outside	24	2	0.08	11
		swinging	24	2	0.08	11
17	table	breakfast	24	2	0.08	15
		Eating	24	2	0.08	15
		drink/s	24	2	0.08	15
18	teddy	Cuddeling/cuddle/s	24	3	0.13	12
19	tooth	brush them/brush your teeth (cepillarlos)/brushing your teeth	30	5	0.17	17
		tooth brush	30	2	0.07	17
		they fall out/your tooth is falling out	30	2	0.07	17
		tooth paste	30	2	0.07	17
20	towel	dry	24	2	0.08	10

Appendix O: Experiment 3. Related responses given by 2+ bilingual children (as first responses) in the online WA task

	Cue	Response	No. participants receiving cue (G)	No. participants producing response (P)	Associative strength: FSG (P/G)	Idiosyncratic responses
1	bath	water	10	2	0.20	6
2	bed	sleep/ domir	10	3	0.30	4
3	brush	teeth	8	3	0.38	4
4	cereal	eating	8	2	0.25	5
5	chair	table	10	2	0.20	4
		sit/ sitting (said in French: assis)/sentarse	10	2	0.20	4
6	door	open (she said it in english)/open door	8	3	0.38	4
7	finger	hand/hands	10	2	0.20	5
8	foot	walking	8	2	0.25	6
9	hair	to brush/brush your hair	8	2	0.25	5
10	head	hair /pelo/wet hair	10	2	0.20	7
11	key	the car (coche)	10	2	0.20	5
		door/ puerta	10	2	0.20	5
12	park	car/park your car	10	3	0.30	7
13	teddy	Cuddeling/cuddle/s	8	3	0.38	5
14	towel	dry	8	2	0.25	4

Appendix P: Experiment 3. Word associations replicated in Experiment 1, 2 and 3

	CUE	RESPONSES produced by 2+ participants	Bilingual child_FSG	Monolingual child Online_FSG	Monolingual child Parental_FSG	FSG_Birkbeck Norms: Moss and Older (1996)	FSG_South Florida Norms: Nelson et al. (1998)
1	**bath**	toys/ dinosours (toys in bath)/ happy hippo/ duck/ juguetes (toys)	0.286	0.171	0.130	not documented	not documented
2	**key**	door/ puerta (door)	0.280	0.147	0.210	0.156	0.218
3	**door**	open/ open door	0.267	0.148	0.150	0.146	0.183
4	**brush**	teeth	0.250	not documented	0.160	0.044	0.277
5	**door**	shut/ shut the door/ shutting	0.200	0.111	0.060	0.062	not documented
6	**brush**	hair/ andrew's hair	0.188	0.188	0.190	0.200	0.440
7	**sock**	feet/ foot /pie (foot)	0.185	0.138	0.200	not a cue	0.172
8	**chair**	table/ mesa (table)	0.179	0.063	0.050	not a cue	0.314
9	**tooth**	brush/ brush your teeth/ brushing your teeth/ cepillarlos (brush your teeth)	0.172	0.152	0.160	0.115	0.123
10	**hair**	brush/ brush your hair	0.167	0.182	0.110	0.021	0.207
11	**towel**	dry	0.167	0.143	0.080	not a cue	0.284
12	**teddy**	cuddle/ cuddeling/ cuddles	0.167	0.100	0.240	not a cue	not a cue
13	**foot**	shoes	0.167	0.083	0.050	0.192	0.108
14	**bed**	toys (soft toys)/ friends (toys)/ dog (Peg, his soft toy)/ teddy	0.148	0.171	not documented	not a cue	not documented
15	**bed**	dormir (sleep)/ sleep	0.148	0.086	0.150	not a cue	0.638
16	**bath**	water/ agua (water)	0.143	not documented	0.100	0.354	0.097
17	**chair**	sit/ sentarse (sit)/ sitting (said in French: assis)	0.143	0.094	0.110	not a cue	0.212
18	**cereal**	milk	0.143	0.061	0.130	not documented	0.204
19	**swing**	park	0.133	0.174	0.180	0.067	0.101
20	**door**	close/ close door	0.133	0.074	not documented	0.021	0.044
21	**key**	lock	0.120	0.118	0.050	0.489	0.255
22	**park**	slide	0.111	0.086	0.160	not documented	not documented
23	**finger**	hand/ hands	0.111	0.074	0.190	0.125	0.268
24	**chair**	eat/ eating	0.107	not documented	0.080	not a cue	not documented
25	**table**	eating	0.095	0.143	0.130	not a cue	0.026
26	**table**	breakfast	0.095	0.107	0.070	not a cue	not documented
27	**bath**	play/ jugar (play)	0.095	0.057	not documented	not documented	not documented
28	**key**	car/ coche (car)	0.080	0.059	0.070	not documented	0.115
29	**head**	hair/ pelo (hair)	0.077	0.114	0.220	0.064	0.186
30	**head**	shoulder/ head, shoulders, knees and toes	0.077	0.057	not documented	0.085	0.021
31	**head**	feet/ foot	0.077	0.057	not documented	0.043 (foot)	0.052
32	**park**	play/ jugar (play)	0.074	not documented	0.090	0.021	0.027
33	**sock**	toes	0.074	not documented	0.060	not a cue	not documented
34	**park**	swing/ columpio (swing)	0.074	0.171	0.160	0.021	0.061
35	**bowl**	eat/ eating	0.071	0.067	not documented	not documented	not documented
36	**tooth**	toothbrush	0.069	not documented	0.140	0.019	0.123
37	**tooth**	toothpaste	0.069	0.061	0.140	0.019	0.058

Appendix Q: Experiment 4-8. Online Consent Form

Aim of research

This project is investigating whether a word recognition experiment we usually run in our Plymouth BabyLab can be run as an online study.

We are interested in how much time babies spend looking at two pictures on a screen when one of the pictures is named, but the other isn't. We are trying to find out if eye movements indicate word recognition.

Description of procedure

You will run this experiment on your own device at home. You will be taken through a series of steps to set up your: device, lighting, seating etc. and to test your video and audio.

After reading this page, you will need to consent to the study by clicking the boxes below. By consenting, you agree to the recording of audio and video for analysis.

For the experiment, you will seat your child on your lap facing the webcam. The experiment will begin automatically. For every trial, your child will see two pictures on the screen and hear a word. There are 24 trials. We will record your child's face and eye movements to see where he/she is looking during this time. There will be short animations at the beginning, middle, and end of the experiment to keep your child interested and focused. The task should take about 2-3 minutes.

Description of risks

None.

Benefits of proposed research

There is no direct short-term benefit, however; long-term we hope that this research can contribute to a better understanding of language development in typically developing children, which can benefit by helping children who are not developing typically.

Right to withdraw

At any moment during this study, you can choose to withdraw without providing any justification. Your data would then be removed from the study.

If you are dissatisfied with the way the research is conducted, please contact the principal investigator in the first instance: 01752584822, plymouthbabylab@plymouth.ac.uk. If you feel the problem has not been resolved please contact the secretary to the Faculty of Health: Medicine, Dentistry and Human Sciences Research Ethics Committee: Mr Maurice Bottomley hhsethics@plymouth.ac.uk or +44 1752 586992.

I am the parent/legal guardian of... (please write child's full name)

☐ The objectives of this research have been explained to me.

☐ I understand that my child is free to withdraw from the research at any stage, and ask for his/her data to be destroyed if I wish. I understand that his/her anonymity is guaranteed, unless I expressly state otherwise.

☐ I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations).

☐ I understand that video footage of my child during the procedure will be used for scoring purposes. It will be kept safely on the investigators' hard drive at the University, and destroyed 5 years after publication.

Next

Appendix R: Experiment 6-9. List Order 1

Trial	Relatedness	Prime	Target	Target_Location	Distractor
1	R	sock	trousers	L	eye
2	R	cat	bird	R	doll
3	U	sofa	flower	L	dog
4	R	train	bus	R	pen
5	U	nose	window	R	bowl
6	R	foot	hand	L	door
7	U	clock	ball	R	house
8	R	glass	milk	R	book
9	U	tiger	star	L	bed
10	U	box	frog	R	lorry
11	R	park	tree	L	coat
12	U	banana	heart	L	mouse
13	U	nappy	eye	L	flower
14	R	car	lorry	R	bird
15	R	chair	bed	L	tree
16	U	cake	doll	R	window
17	R	garden	house	R	bus
18	R	boots	coat	L	hand
19	U	key	pen	R	ball
20	R	spoon	bowl	R	frog
21	U	cheese	dog	L	trousers
22	U	horse	book	R	milk
23	U	bottle	door	L	heart
24	R	chicken	mouse	L	star

Appendix S: Experiment 6-9. List Order 2

Trial	Relatedness	Prime	Target	Target_Location	Distractor
1	U	nappy	eye	L	flower
2	R	car	lorry	R	bird
3	R	chair	bed	L	tree
4	U	cake	doll	R	window
5	R	garden	house	R	bus
6	R	boots	coat	L	hand
7	U	key	pen	R	ball
8	R	spoon	bowl	R	frog
9	U	cheese	dog	L	trousers
10	U	horse	book	R	milk
11	U	bottle	door	L	heart
12	R	chicken	mouse	L	start
13	R	sock	trousers	L	eye
14	R	cat	bird	R	doll
15	U	sofa	flower	L	dog
16	R	train	bus	R	pen
17	U	nose	window	R	bowl
18	R	foot	hand	L	door
19	U	clock	ball	R	house
20	R	glass	milk	R	book
21	U	tiger	star	L	bed
22	U	box	frog	R	lorry
23	R	park	tree	L	coat
24	U	banana	heart	L	mouse

Appendix T: Experiment 6-9. List Order 3

Trial	Relatedness	Prime	Target	Target_Location	Distractor
1	R	sock	trousers	R	eye
2	R	cat	bird	L	doll
3	U	sofa	flower	R	dog
4	R	train	bus	L	pen
5	U	nose	window	L	bowl
6	R	foot	hand	R	door
7	U	clock	ball	L	house
8	R	glass	milk	L	book
9	U	tiger	star	R	bed
10	U	box	frog	L	lorry
11	R	park	tree	R	coat
12	U	banana	heart	R	mouse
13	U	nappy	eye	R	flower
14	R	car	lorry	L	bird
15	R	chair	bed	R	tree
16	U	cake	doll	L	window
17	R	garden	house	L	bus
18	R	boots	coat	R	hand
19	U	key	pen	L	ball
20	R	spoon	bowl	L	frog
21	U	cheese	dog	R	trousers
22	U	horse	book	L	milk
23	U	bottle	door	R	heart
24	R	chicken	mouse	R	star

Appendix U: Experiment 6-9. List Order 4

Trial	Relatedness	Prime	Target	Target_Location	Distractor
1	U	nappy	eye	R	flower
2	R	car	lorry	L	bird
3	R	chair	bed	R	tree
4	U	cake	doll	L	window
5	R	garden	house	L	bus
6	R	boots	coat	R	hand
7	U	key	pen	L	ball
8	R	spoon	bowl	L	frog
9	U	cheese	dog	R	trousers
10	U	horse	book	L	milk
11	U	bottle	door	R	heart
12	R	chicken	mouse	R	star
13	R	sock	trousers	R	eye
14	R	cat	bird	L	doll
15	U	sofa	flower	R	dog
16	R	train	bus	L	pen
17	U	nose	window	L	bowl
18	R	foot	hand	R	door
19	U	clock	ball	L	house
20	R	glass	milk	L	book
21	U	tiger	star	R	bed
22	U	box	frog	L	lorry
23	R	park	tree	R	coat
24	U	banana	heart	R	mouse

Bibliography

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242-275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Águila, E., Ramon-Casas, M., Pons, F., & Bosch, L. (2007). La medida del léxico productivo inicial: Aplicación de un cuestionario para la población bilingüe. In Diez-Itza, E. (Ed.), *Estudios de desarrollo del lenguaje y educación* (Vol. 32). Oviedo, Spain: ICE Monografias de Aula Abierta

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502-518. https://doi.org/10.1016/j.jml.2006.12.004

Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, *32*(2), 177-190. https://doi.org/10.1111/j.1469-8986.1995.tb03310.x

Anglin, J. M. (1970). *The growth of word meaning*. M.I.T. Press.

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*(4), 1407-1425. https://doi.org/10.3758/s13428-020-01501-5

Anwyl-Irvine, A.L., Massonié, J., Flitton, A., Kirkham, N.Z., & Evershed, J.K. (2019). Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods, 52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1536), 3633-3647. https://doi.org/10.1098/rstb.2009.0146

Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: associative and taxonomic priming effects in the infant lexicon. *Cognition*, *128*(2), 214-227. https://doi.org/10.1016/j.cognition.2013.03.008

Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language*, *32*(1), 159-173. https://doi.org/10.1017/S0305000904006567

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition*, *35*(5), 953-965. https://doi.org/10.3758/BF03193468

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343-355. https://doi.org/https://doi.org/10.1016/0013-4694(85)90008-2

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences of the United States of America.*, *114*(49), 12916–12921. https://doi.org/10.1073/pnas.1712966114

Bergelson, E., & Aslin, R. (2017). Semantic specificity in one-year-olds' word comprehension. *Language Learning and Development*, *13*(4), 481-501. https://doi.org/10.1080/15475441.2017.1324308

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America.*, *109*(9), 3253-3258. https://doi.org/10.1073/pnas.1113380109

Berglund, E., Eriksson, M., & Westerlund, M. (2005). Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. *Scandinavian Journal of Psychology*, *46*(6), 485-491. https://doi.org/10.1111/j.1467-9450.2005.00480.x

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/https://doi.org/10.1111/cdev.13079

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition* , *13*(4), 525-531. https://doi.org/10.1017/s1366728909990423

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, *24*(1), 27-44. https://doi.org/10.1017/S014271640300002X

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition*, *140*, 122-134. https://doi.org/10.1016/j.cognition.2015.03.013

Blewitt, P., & Toppino, T. C. (1991). The development of taxonomic structure in lexical memory. *Journal of Experimental Child Psychology*, *51*(2), 296–319. https://doi.org/10.1016/0022-0965(91)90038-T

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development*, *87*(6), 1893-1908. https://doi.org/10.1111/cdev.12554

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, *116*(2), 289-296. https://doi.org/https://doi.org/10.1016/j.cognition.2010.05.004

Brown, R., & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development*, *31* (1), 1-14. https://doi.org/10.2307/1126377

Byers-Heinlein, K., Tsui, R. K., van Renswoude, D., Black, A. K., Barr, R., Brown, A., Colomer, M., Durrant, S., Gampe, A., Gonzalez-Gomez, N., Hay, J. F., Hernik, M., Jartó, M., Kovács, Á. M., Laoun-Rubenstein, A., Lew-Williams, C., Liszkowski, U., Liu, L., Noble, C., Potter, C. E., … Singh, L. (2021). The development of gaze following in monolingual and bilingual infants: A multi-laboratory study. *Infancy*, *26*(1), 4-38. https://doi.org/https://doi.org/10.1111/infa.12360

Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: evidence from bilingual infants. *Cognition*, *128*(3), 407-416. https://doi.org/10.1016/j.cognition.2013.05.010

Canseco-Gonzalez, E., Brehm, L., Brick, C. A., Brown-Schmidt, S., Fischer, K., & Wagner, K. (2010). Carpet or Cárcel: The effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes*, *25*(5), 669-705. https://doi.org/10.1080/01690960903474912

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177-208. https://doi.org/10.1080/026432997381664

Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreckx, F., Dennis, I., & Floccia, C. (2014). How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests? *International Journal of Language & Communication Disorders*, *49*(6), 649-671. https://doi.org/10.1111/1460-6984.12082

Chow, J., Aimola Davies, A., & Plunkett, K. (2017). Spoken-word recognition in 2-year-olds: The tug of war between phonological and semantic activation. *Journal of Memory and Language*, *93*, 104-134. https://doi.org/https://doi.org/10.1016/j.jml.2016.08.004

Chow, J., Aimola Davies, A., Fuentes, L., & Plunkett, K. (2018) The vocabulary spurt predicts the emergence of backward semantic inhibition in 18-month-old toddlers. *Developmental Science, 22* (2). https://doi.org/10.1111/desc.12574

Christoffels, I. K., de Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, *54*(3), 324-345. https://doi.org/https://doi.org/10.1016/j.jml.2005.12.004

Chwilla, D. J., Kolk, H. H. J., & Mulder, G. (2000). Mediated priming in the lexical decision task: Evidence from event-related potentials and reaction time. *Journal of Memory and Language*, *42*(3), 314-341. https://doi.org/10.1006/jmla.1999.2680

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Lawrence Erlbaum Associates, Inc.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428. https://doi.org/10.1037/0033-295X.82.6.407

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, *8*(2), 240-247. https://doi.org/10.1016/S0022-5371(69)80069-1

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, *45*(4), 721-736. https://doi.org/10.1006/jmla.2001.2793

Comesaña, M., Fraga, I., Moreira, A. J., Frade, C. S., & Soares, A. P. (2014). Free associate norms for 139 European Portuguese words for children from different age groups. *Behavior Research Methods*, *46*(2), 564-574. https://doi.org/10.3758/s13428-013-0388-0

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology* *6*(1), 84-107. https://doi.org/10.1016/0010-0285(74)90005-X.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283. https://doi.org/10.1037//0278-7393.26.5.1283

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, *41*(3), 365-397. https://doi.org/10.1006/jmla.1999.2651

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491-511. https://doi.org/10.1016/j.jml.2004.02.002

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of xEperimental Psychology: General*, *132*(2), 163. https://doi.org/10.1037/0096-3445.132.2.163

Cronin, V. S. (2002). The syntagmatic-paradigmatic shift and reading development. *Journal of Child Language*, *29*(1), 189-204.

Curtin, S., Byers-Heinlein, K., & Werker, J. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*(4), 492-504. https://doi.org/10.1016/j.wocn.2010.12.002.

Dale, P. S., Reznick, J. S., & Thal, D. J. (1998). A parent report measure of language development for three-year-olds. *Infant Behavior & Development*, *21*, 370.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987-1006. https://doi.org/10.3758/s13428-018-1115-7

de Groot, A. M. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 1001–1018. https://doi.org/10.1037/0278-7393.18.5.1001

de Groot, A. M. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 824-845. https://doi.org/10.1037/0278-7393.15.5.824

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, *47*(1), 91-124. https://doi.org/https://doi.org/10.1006/jmla.2001.2840

DeAnda, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Lexical processing and organization in bilingual first language acquisition: Guiding future research. *Psychological Bulletin Journal*, *142*(6), 655-667. https://doi.org/10.1037/bul0000042

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283.

Delle Luche, C., Durrant, S., Floccia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science*, *17*(6), 948-955. https://doi.org/10.1111/desc.12164

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496-518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284-301. https://doi.org/https://doi.org/10.1016/j.jml.2009.12.003

Elbers, L., & van Loon-Vervoorn, A. (1998). Acquiring the lexicon: Evidence from Dutch research. In S. Gillis & A. De Houwer (Eds.), *The acquisition of Dutch* (pp. 301 – 378). John Benjamins.

Entwisle, D. R. (1966). Form class and children's word associations. *Journal of Verbal Learning and Verbal Behavior*, *5*(6), 558-565. https://doi.org/https://doi.org/10.1016/S0022-5371(66)80091-9

Entwisle, D. R., Forsyth, D. F., & Muuss, R. E. H. (1964). The syntactic-paradigmatic shift in children's word associations. *Journal of Verbal Learning and Verbal Behavior*, *3*, 19-29.

Ervin, S. M. (1961). Semantic shift in bilingualism. *The American Journal of Psychology*, *74*, 233-241. https://doi.org/10.2307/1419408

Fay, D., & Cutler, A. (1977). Malapropisms and the Structure of the Mental Lexicon. *Linguistic Inquiry*, *8*(3), 505-520.

Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, *41*(4), 469-495. https://doi.org/https://doi.org/10.1006/jmla.1999.2660

Fenson, L., Marchman, V., Thal, D., P., Reznick, J., & Bates, E. (2007). *MacArthur- Bates Communicative inventories (CDI): Words and Gestures; Words and Sentences*. Baltimore, MD: Paul Brookes.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228-231. https://doi.org/10.1111/1467-9280.00044

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*(4), 1003-1015. https://doi.org/10.1111/1467-8624.00331

Ferrand, L., & New, B. (2004). Semantic and Associative Priming in the Mental Lexicon. In *Mental lexicon: "Some words to talk about words"* (pp. 25-43). Nova Science Publishers.

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: a unified JavaScript framework for online studies. In *International Conference on Computational Social Science (Cologne)*.

Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. J. (2013). Establishing the reliability of word association data for investigating individual and group differences. *Applied Linguistics*, *36*(1), 23-50. https://doi.org/10.1093/applin/amt020

Floccia, C., Delle Luche, C., Lepadatu, I., Chow, J., Ratnage, P., & Plunkett, K. (2020). Translation equivalent and cross-language semantic priming in bilingual toddlers.

Journal of Memory and Language, *112*, 104086. https://doi.org/10.1016/j.jml.2019.104086.

Floccia, C., Sambrook, T.D., Delle Luche, C., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J. & Plunkett, K. (2018), II: METHODS. *Monographs of the Society for Research in Child Development, 83*(1),30-42. https://doi.org/10.1111/mono.12349

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: processing known words in picture contexts. *Journal of Cognitive Neuroscience*, *16*(8), 1465-1477. https://doi.org/10.1162/0898929042304705

Friedrich, M., & Friederici, A. D. (2005). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, *17*(11), 1785-1802. https://doi.org/10.1162/089892905774589172

Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: fast encoding-weak retention. *Journal of Cognitive Neuroscience*, *23*(11), 3228-3240. https://doi.org/10.1162/jocn_a_00002

Friedrich, M., & Friederici, A. D. (2017). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Developmental Science*, *20*(2). https://doi.org/10.1111/desc.12357

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 302-309. https://doi.org/10.3758/BF03202556

Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, *45*(2), 220-266. https://doi.org/10.1016/s0010-0285(02)00003-8

Gelman, S. A., & Wellman, H. M. (1991). Insides and essence: Early understandings of the non-obvious. *Cognition*, *38*(3), 213-244. https://doi.org/10.1016/0010-0277(91)90007-Q

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*(1), 23-45. https://doi.org/10.1017/S030500090001271X

Gollan, T. H., Salmon, D. P., Montoya, R. I., & Galasko, D. R. (2011). Degree of bilingualism predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated Hispanics. *Neuropsychologia*, *49*(14), 3826-3830. https://doi.org/10.1016/j.neuropsychologia.2011.09.041

Grainger, J., Midgley, K., & Holcomb, P. (2010). Chapter 14. Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). *Language*

Acquisition Across Linguistic and Cognitive Systems. https://doi.org/10.1075/lald.52.18gra

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67-81. https://doi.org/10.1017/S1366728998000133

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language, 27*(3), 689–705. https://doi.org/10.1017/s0305000900004414 Harzing, A. W. (2007). *Publish or Perish*. http://www.harzing.com/pop.htm

Hees, J., Bauer, R., Folz, J., Borth, D., & Dengel, A. (2016). *Edinburgh Associative Thesaurus as RDF and DBpedia Mapping*. In: Sack, H., Rizzo, G., Steinmetz, N., Mladenić, D., Auer, S., Lange, C. (eds) The Semantic Web. ESWC 2016. Lecture Notes in Computer Science(), vol 9989. Springer, Cham. https://doi.org/10.1007/978-3-319-47602-5_4 Heinze, H. J., Muente, T. F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology*, *47*(2), 121-135. https://doi.org/10.1016/s0301-0511(97)00024-0

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, *1*(3), 213-229. https://doi.org/10.1017/S1366728998000364

Hillis, A. E., Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, *7*(3), 191-243. https://doi.org/10.1080/02643299008253442

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: contextual diversity in early word learning. *Journal of Memory and Language*, *63*(3), 259-273. https://doi.org/10.1016/j.jml.2010.06.002

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729-739. https://doi.org/10.1111/j.1467-9280.2009.02365.x

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1-27. https://doi.org/10.1017/S0305000910000759

Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, *8*(4), 379-411. https://doi.org/10.1080/01690969308407583

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*(4), 281-312. https://doi.org/10.1080/01690969008407065

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23-32. https://doi.org/10.1016/j.cognition.2004.10.003

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*(4), 460-482. https://doi.org/10.1016/j.jml.2007.02.001

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151-171. https://doi.org/10.1016/j.actpsy.2010.11.003

Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C.-S., Yap, M. J., Bengson, J. J., Niemeyer, D., & Buchanan, E. (2013). The semantic priming project. *Behavior Research Methods,* *45*(4), 1099–1114. https://doi.org/10.3758/s13428-012-0304-zJardak, A., & Byers-Heinlein, K. (2019). Labels or concepts? The development of semantic networks in bilingual two-year-olds. *Child Development*, *90*(2), e212-e229. https://doi.org/10.1111/cdev.13050

Jarema, G., & Libben, G. (2007). *The Mental Lexicon: Core Perspectives*. Brill. https://doi.org/https://doi.org/10.1163/9780080548692

Jenkins, J. J., & Russell, W. A. (1960). Systematic changes in word association norms: 1910-1952. *The Journal of Abnormal and Social Psychology*, *60*(3), 293–304. https://doi.org/10.1037/h0042234

Johnson-Laird, P. N. (1987). The mental representation of the meaning of words. *Cognition*, *25*(1), 189-211. https://doi.org/https://doi.org/10.1016/0010-0277(87)90009-6

Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: constraints on bilingual lexical activation. *Psychological Science*, *15*(5), 314-318. https://doi.org/10.1111/j.0956-7976.2004.00675.x

Jung, C. G. (1910). The Association Method. *The American Journal of Psychology*, *21*(2), 219-269. https://doi.org/10.2307/1413002

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The timecourse of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning*, *57*(1), 119-163. https://doi.org/10.1111/j.1467-9922.2007.00401.x

Keil, F. C. (1983). On the emergence of semantic and conceptual distinctions. *Journal of Experimental Psychology: General*, *112*(3), 357-385. https://doi.org/10.1037/0096-3445.112.3.357

Kent, G. H., & Rosanoff, A. J. (1910). A study of association in insanity. *American Journal of Psychiatry*, *67*(1), 37-96. https://doi.org/10.1176/ajp.67.1.37

Kiss, G. R. (1975). An associative thesaurus of English. Structural analysis of a large relevance network. In A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory*. London: Wiley. Kiss, G. R., Armstrong, C., Milroy, R. & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey & N. Hamilton-Smith (ed.), *The Computer and Literary Studies*. Edinburgh University Press .Koff, R. H. (1965). Systematic changes in children's word-association norms 1916-63. *Child Development, 36*(1), 299–305. https://doi.org/10.2307/1126798Kohonen, T. (1995) *Self-Organizing Maps*. Third Edition, Springer, Berlin. https://doi.org/10.1007/978-3-642-97610-0

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416-430. https://doi.org/https://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*(2), 119-135. https://doi.org/10.1017/S1366728906002483

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205. https://doi.org/10.1126/science.7350657

Lalor, E., & Kirsner, K. (2001). The representation of "false cognates" in the bilingual lexicon. *Psychonomic Bulletin & Review*, *8*(3), 552–559. https://doi.org/10.3758/BF03196191

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. https://doi.org/10.2307/2529310

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2016). Vocabulary size, translation equivalents, and efficiency in word recognition in very young bilinguals. *Journal of Child Language*, *43*(4), 760-783. https://doi.org/10.1017/s0305000915000252

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, *19*(5), 585-611. https://doi.org/10.1080/01690960444000007

Lemhöfer, K., & Schriefers, H. (2018). The second language interferes with picture naming in the first language: Evidence for L2 activation during L1 production. *Language,* *Cognition and Neuroscience*, *33*(7), 867-877. https://doi.org/10.1080/23273798.2018.1430837

Levelt, W. J. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223-232. https://doi.org/10.1016/s1364-6613(99)01319-4

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-75. https://doi.org/10.1017/s0140525x99001776

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, *7*(4), 618-630. https://doi.org/10.3758/BF03212999

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*, *45*(2), 516-526. https://doi.org/10.3758/s13428-012-0267-0

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: Transcription format and programs, Vol. 1, (3rd ed*.). Lawrence Erlbaum Associates Publishers.

Malpass, D., & Meyer, A. S. (2010). The time course of name retrieval during multiple-object naming: Evidence from extrafoveal-on-foveal effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 523-537. https://doi.org/10.1037/a0018522

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake--but only for skilled producers. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(4), 843-847. https://doi.org/10.1037/a0029284

Mani, N., Mills, D. L., & Plunkett, K. (2012). Vowels in early words: an event-related potential study. *Developmental Science*, *15*(1), 2-11. https://doi.org/10.1111/j.1467-7687.2011.01092.x

Mani, N., & Plunkett, K. (2010). In the infant's mind's ear: evidence for implicit naming in 18-month-olds. *Psychological Science*, *21*(7), 908-913. https://doi.org/10.1177/0956797610373371

Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, *121*(2), 196-206. https://doi.org/10.1016/j.cognition.2011.06.013

Marian, V., & Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*(2), 173–193. https://doi.org/10.1017/S0142716403000092

Marian, V., & Spivey, M. (2003b). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97-115. https://doi.org/DOI: 10.1017/S1366728903001068

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190. https://doi.org/https://doi.org/10.1016/j.jneumeth.2007.03.024

Martin, C. D., Costa, A., Dering, B., Hoshino, N., Wu, Y. J., & Thierry, G. (2012). Effects of speed of word processing on semantic access: the case of bilingualism. *Brain and Language*, *120*(1), 61-65. https://doi.org/10.1016/j.bandl.2011.10.003

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press. https://doi.org/10.4324/9780203338001

McNeill, D. (1970). Language before symbols: Very early children's grammar. *Interchange*, *1*(3), 127-133. https://doi.org/10.1007/BF02214683

McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39-78.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 558-572. https://doi.org/10.1037/0278-7393.24.3.558

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547-559. https://doi.org/10.3758/BF03192726

Meints, K., Plunkett, K., & Harris, P. L. (1999). When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology*, *35*(4), 1072-1078. https://doi.org/10.1037//0012-1649.35.4.1072

Merck, C., Noël, A., Jamet, E., Robert, M., Salmon, A., Belliard, S., & Kalénine, S. (2020). Overreliance on thematic knowledge in semantic dementia: Evidence from an eye-tracking paradigm. *Neuropsychology*, *34*(3), 331-349. https://doi.org/10.1037/neu0000616

Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition*, *35*(3), 494-503. https://doi.org/10.3758/BF03193289

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(6), 1146-1160. https://doi.org/10.1037/0278-7393.17.6.1146

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227-234. https://doi.org/10.1037/h0031564

Mills, D. L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of vocabulary size and experience. *Cognitive Development*, *20*(1), 19-31. https://doi.org/10.1016/j.cogdev.2004.07.001

Mirman, D., & Graziano, K. M. (2012). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General*, *141*(4), 601. https://doi.org/10.1037/a0026451

Moss, H., Older, L., & Older, L. J. (1996). *Birkbeck word association norms*. Psychology Press.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: evidence from semantic priming. *Cognition*, *70*(1), B1-13. https://doi.org/10.1016/s0010-0277(99)00004-9

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In *Basic processes in reading: Visual word recognition.* (pp. 264-336). Lawrence Erlbaum Associates, Inc.

Nelson, D. L., Dyrdal, G. M., & Goodmon, L. B. (2005). What is preexisting strength? Predicting free association probabilities, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review*, *12*(4), 711-719. https://doi.org/10.3758/bf03196762

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, *28*(6), 887-899. https://doi.org/10.3758/BF03209337

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407. https://doi.org/10.3758/BF03195588

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, *84*(1), 93-116. https://doi.org/10.1037/0033-2909.84.1.93

Newman, M. A. (1970). Word associations from preschool children collected by the continued sentence associations method. *Journal of Reading Behavior*, *3*(3), 13-23. https://doi.org/10.1080/10862967009546946

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436-469. https://doi.org/10.1111/infa.12186

Oldfield, R. C. (1966). Things, Words and the Brain*. *Quarterly Journal of Experimental Psychology*, *18*(4), 340-353. https://doi.org/10.1080/14640746608400052

Palermo, D. S. (1971). Characteristics of word association responses obtained from children in grades one through four. *Developmental Psychology*, *5*(1), 118. https://doi.org/10.1037/h0031075

Palermo, D. S., & Jenkins, J. J. (1964). Word association norms: Grade school through college (NED-New edition). University of Minnesota Press. http://www.jstor.org/stable/10.5749/j.cttttpz4.

Palermo, D. S., & Jenkins, J. J. (1965). Changes in the word associations of fourth-and fifth-grade children from 1916 to 1961. *Journal of Verbal Learning and Verbal Behavior*, *4*(3), 180-187. https://doi.org/10.1016/S0022-5371(65)80018-4.

Palermo, D. S., & Jenkins, J. J. (1966). *Oral word association norms for children in grades one through four*. Department of Psychology, Pennsylvania State University.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, *66*(3), 180-194. https://doi.org/10.1007/s00426-002-0086-5

Persici, V., Vihman, M., Burro, R., & Majorano, M. (2019). Lexical access and competition in bilingual children: The role of proficiency and the lexical similarity of the two languages. *Journal of Experimental Child Psychology*, *179*, 103-125. https://doi.org/https://doi.org/10.1016/j.jecp.2018.10.002

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, *17*(1), 57-70. https://doi.org/10.1177/1367006911431198

Poulin-Dubois, D., Kuzyk, O., Legacy, J., Zesiger, P., & Friend, M. (2018). Translation equivalents facilitate lexical access in very young bilinguals. *Bilingualism: Language and Cognition*, *21*(4), 856-866. https://doi.org/10.1017/S1366728917000657

Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology*, *28*(3), 406-413. https://doi.org/10.1037/0012-1649.28.3.406

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

Roelofs, A. (2003). Shared phonological encoding processes and representations of language in bilingual speakers. *Language and Cognitive Processes*, *18*(2), 175-204. https://doi.org/10.1080/01690960143000515

Rugg, M. D. (1985). The Effects of Semantic Priming and Word Repetition on Event-Related Potentials *Psychophysiology*, *22*(6), 642-647. https://doi.org/https://doi.org/10.1111/j.1469-8986.1985.tb01661.x

Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18- and 24-month old children. *Brain and Language*, *125*(1), 1-10. https://doi.org/https://doi.org/10.1016/j.bandl.2013.01.009

Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology*, *4*(4), 293–310. https://doi.org/10.1080/09541449208406189

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, *37*(5), 569-586. https://doi.org/10.3758/mc.37.5.569

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*(1), 86-102. https://doi.org/https://doi.org/10.1016/0749-596X(90)90011-N

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, *22*(1), 106-129. https://doi.org/10.1080/01690960500463920

Scott, K., Chu, J., & Schulz, L. (2017). Lookit (Part 2): Assessing the Viability of Online Developmental Research, Results From Three Case Studies. *Open Mind*, *1*(1), 15-29. https://doi.org/10.1162/OPMI_a_00001

Scott, K., & Schulz, L. (2017). Lookit (Part 1): A New Online Platform for Developmental Research. *Open Mind*, *1*(1), 4-14. https://doi.org/10.1162/OPMI_a_00002

Seiger-Gardner, L., & Schwartz, R. G. (2008). Lexical access in children with and without specific language impairment: a cross-modal picture-word interference study. *International Journal of Language & Communication Disorders*, *43*(5), 528-551. https://doi.org/10.1080/13682820701768581

Shelton, J. R., & Caramazza, A. (1999). Deficits in lexical and semantic processing: implications for models of normal language. *Psychonomic Bulletin & Review*, *6*(1), 5-27. https://doi.org/10.3758/bf03210809

Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1191-1210. https://doi.org/10.1037/0278-7393.18.6.1191

Sheng, L., McGregor, K. K., & Marian, V. (2006). Lexical-semantic organization in bilingual children: evidence from a repeated word association task. *Journal of Speech, Language, and Hearing Research*, *49*(3), 572-587. https://doi.org/10.1044/1092-4388(2006/041)

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for comprehension of speech. *Bilingualism: Language and Cognition*, *16*(2), 304-324. https://doi.org/10.1017/S1366728912000466

Shook, A., & Marian, V. (2019). Covert co-activation of bilinguals' non-target language: Phonological competition from translations. *Linguistic Approaches to Bilingualism*, *9*(2), 228-252. https://doi.org/10.1075/lab.17022.sho

Simmons, S., & Estes, Z. (2008). Individual differences in the perception of similarity and difference. *Cognition*, *108*(3), 781-795. https://doi.org/10.1016/j.cognition.2008.07.003

Singh, L. (2013). One world, two languages: cross-language semantic priming in bilingual toddlers. *Child Development*, *85*(2), 755-766. https://doi.org/10.1111/cdev.12133

Sirri, L., & Rämä, P. (2015). Cognitive and neural mechanisms underlying semantic priming during language acquisition. *Journal of Neurolinguistics*, *35*. https://doi.org/10.1016/j.jneuroling.2015.01.003

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*(3), 214-241. https://doi.org/10.1037/h0036351

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281-284. https://doi.org/10.1111/1467-9280.00151

Styles, S., & Plunkett, K. (2009). What is 'word understanding' for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *Journal of Child Language*, *36*(4), 895-908. https://doi.org/10.1017/S0305000908009264

Styles, S. J., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, *1*(1), 1-24. https://doi.org/10.1515/LANGCOG.2009.001

Szagun, G., Stumper, B., & Schramm, A.S. (2009). *Fragebogen zur frühkindlichen Sprachentwicklung (FRAKIS) und FRAKIS-K (Kurzform)*. Frankfurt: Pearson Assessment.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences USA*, *104*(30), 12530-12535. https://doi.org/10.1073/pnas.0609927104

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172-175. https://doi.org/10.1111/1467-9280.00127

Torkildsen, J. v. K., Syversen, G., Simonsen, H. G., Moen, I., & Lindgren, M. (2007). Brain responses to lexical-semantic priming in children at-risk for dyslexia. *Brain and Language*, *102*(3), 243-261. https://doi.org/10.1016/j.bandl.2006.11.010

Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. A. M. (2007). Onset of word form recognition in English, Welsh, and English–Welsh bilingual infants. *Applied Psycholinguistics*, *28*(3), 475-493. https://doi.org/10.1017/S0142716407070269

Von Holzen, K., Fennell, C. T., & Mani, N. (2019). The impact of cross-language phonological overlap on bilingual and monolingual toddlers' word recognition. *Bilingualism: Language and Cognition*, *22*(3), 476-499. https://doi.org/10.1017/S1366728918000597

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569-586. https://doi.org/10.1016/j.jecp.2012.08.001

Von Holzen, K., & Mani, N. (2014). Bilinguals implicitly name objects in both their languages: an ERP study. *Frontiers in Psychology*, *5*, 1415. https://doi.org/10.3389/fpsyg.2014.01415

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1-25. https://doi.org/https://doi.org/10.1016/S0749-596X(03)00105-0

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: first steps in perception and comprehension. *Trends in Cognitive Sciences*, *12*(4), 144-151. https://doi.org/10.1016/j.tics.2008.01.008

White, H. (1985). The syntagmatic-paradigmatic shift and long-term memory activation. *The Journal of Genetic Psychology*, *146*(4), 561-562. https://doi.org/10.1080/00221325.1985.10532477

Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, *18*(6). https://doi.org/10.1111/infa.12026

Wojcik, E. H. (2017). 2.5-year-olds' retention and generalization of novel words across short and long delays. *Language Learning and Development*, *13*(3), 300-316. https://doi.org/10.1080/15475441.2016.1275644

Wojcik, E. H., & Kandhadai, P. (2020). Paradigmatic associations and individual variability in early lexical-semantic networks: Evidence from a free association task. *Developmental Psychology*, *56*(1), 53-69. https://doi.org/10.1037/dev0000844

Woodrow, H., & Lowell, F. (1916). Children's association frequency tables. *The Psychological Monographs*, *22*(5), i–110. https://doi.org/10.1037/h0093111

Wu, Y. J., & Thierry, G. (2011). Event-related brain potential investigation of preparation for speech production in late bilinguals. *Frontiers in Psychology*, *2*(114). https://doi.org/10.3389/fpsyg.2011.00114

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 505-524. https://doi.org/10.1080/13670050.2010.488284

Zhao, X., & Li, P. (2013). Simulating cross-language priming with a dynamic computational model of the lexicon. *Bilingualism: Language and Cognition*, *16*(2), 288-303. https://doi.org/10.1017/S1366728912000624

Zink, I., & Lejaegere, M. (2002). N-CDIs: Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. Acco, Leuven (Belgium)/Leusden (Netherlands).