Developmental trajectories of bilingual word recognition

Gonzalo Garcia-Castro

Serene Siow

Kim Plunkett

Nuria Sebastian-Galles

# Introduction

Building a mental lexicon is a major achievement in the development of an infant: by storing representations of how familiar words sound and what they mean, an infant is able to make sense of their linguistic input. The foundations of an initial lexicon are in place before their first year of life ([Bergelson & Swingley, 2012](#ref-bergelson20126), [2015](#ref-bergelson2015early); [Hallé & Boysson-Bardies, 1994](#ref-halle1994emergence); [Parise & Csibra, 2012](#ref-parise2012electrophysiological); [Tincoff & Jusczyk, 1999](#ref-tincoff1999some); [Vihman, 2004](#ref-vihman2004cross)). This initial lexicon consists of only a few items; mainly words for people, interjections, body parts, and food ([Tardif et al., 2008](#ref-tardif2008baby); [Tincoff & Jusczyk, 2012](#ref-tincoff2012six)), but it undergoes rapid growth during the second year of life ([Bloom, 2002](#ref-bloom2002children); [Ganger & Brent, 2004](#ref-ganger2004reexamining); [Goldfield & Reznick, 1990](#ref-goldfield1990early); [McMurray, 2007](#ref-mcmurray2007defusing)). According to parental reports, the average 15-month-old infant already understands more than 100 words, and by two years, they understand more than 400 words ([Frank et al., 2021](#ref-frank2021variability)). This accelerated lexical developmental is reflected in infants’ trajectories of word recognition: infants recognise familiar words faster and more efficiently as they approach their second birthday ([Fernald et al., 1998](#ref-fernald1998rapid), [2001](#ref-fernald2001half); [Hurtado et al., 2007](#ref-hurtado2007spoken)).

Despite being exposed to a more complex linguistic input, bilinguals show equivalent trajectories of word acquisition and word recognition to their monolingual peers’ ([Byers-Heinlein et al., 2023](#ref-byers-heinlein2023sometimes); [De Houwer et al., 2014](#ref-de2014bilingual); [Hoff et al., 2012](#ref-hoff2012dual); [Legacy et al., 2018](#ref-legacy2018vocabulary); [Pearson & Fernández, 1994](#ref-pearson1994patterns); [Vihman et al., 2007](#ref-vihman2007onset)). This is a remarkable deed for two reasons. First, bilingual infants receive a relative impoverished linguistic input in each of their languages compared to monolinguals ([Costa & Sebastián-Gallés, 2014](#ref-costa2014does)). Second, they face a more complex referential context, since they often learn two labels for each referent, one in each language. The mechanisms that allow bilingual’ trajectories of lexical developmental to keep up with monolinguals’ are still unclear.

Monolinguals and bilinguals share broadly common word-recognition mechanisms. The first step towards spoken word-recognition is the acoustic-phonetic processing of the speech signal, which results in the activation of lexical representations whose associated phonological forms match the acoustic signal. The activated word-forms then compete for selection: the lexical system uses the bottom-up information provided by the auditory input, together with top-down context—provided by grammatical, supra-segmental, and semantic constraints—to select the candidate lexical representation with the best match. This sequence of events occurs in a *cascaded* fashion. As the recognition system accumulates information about the unfolding acoustic signal, matching phonological representations are activated ([Grosjean, 1980](#ref-grosjean1980spoken); [Marslen-Wilson, 1987](#ref-marslen1987functional); [Marslen-Wilson & Welsh, 1978](#ref-marslen1978processing)), and in turn activate their associated semantic representations ([Neely, 1977](#ref-neely1977semantic)). By the time lexical selection occurs, many non-selected phonological and semantic representations have been activated, potentially modulating the dynamics of word recognition.

A considerable amount of evidence for such cascaded account of lexical processing has been provided by studies recording eye movements during word recognition. For example, Allopenna et al. ([1998](#ref-allopenna1998tracking)) designed a word-recognition task in which participants were presented, in each trial, with four objects in the screen. Participants would then listen to a command like “Pick up the beaker; now put it below the diamond”, and perform the action. The authors manipulated the phonological relationship between the target object’s label (e.g., *casket*), and the label of each of the three distractors. Two of the distractors were phonologically related to the target, sharing onset (e.g., *castle*) or offset (e.g., *basket*), respectively. The other distractor was phonologically unrelated to the target (e.g., *nickel*). After the onset of the auditory target label, phonological distractors attracted participants’ eye fixations more than unrelated distractors. Distractors sharing phonological onset did so at an earlier time window than those sharing phonological offset. This time course of distractor fixations suggests that the auditory presentation of the target label activated no only its (subsequently selected) phonological form in participants’ lexicon, but also that of phonologically related word-forms, leading to competition for selection between the word forms. These results converge with cascaded models of auditory word recognition, in which the unfolding acoustic input activates multiple phonological word-forms, which in turn compete for selection, like TRACE ([McClelland & Elman, 1986](#ref-mcclelland1986trace)) and Shortlist ([Norris, 1994](#ref-norris1994shortlist)).

The developing mental lexicon also seems to operate in this cascaded fashion: like adults, infants activate non-selected lexical representations. For instance, priming paradigms of word recognition have provided a strong body of evidence for the existence of phonological and semantic links between early lexical representations between 18 and 21 months of age.. Arias-Trejo & Plunkett ([2009](#ref-arias2009lexical)) found that infants’ spoken word recognition (e.g., *dot*) was interfered by the previous presentation of a semantically related word (e.g., *cat*). Other studies have provided electrophysiological evidence of semantic priming at earlier ages ([Friedrich & Friederici, 2005](#ref-friedrich2005lexical); e.g., [Rämä et al., 2013](#ref-rama2013development)). More recent studies have provided evidence for the emergence of such semantic links, even in the absence of visual referents during the experimental task ([Willits et al., 2013](#ref-willits2013toddlers)). Phonological priming effects have also been reported in the initial lexicon. Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) created an implicit naming task in which each trial started with the *silent* presentation of a prime picture. Then, a target-distractor picture pair was presented side-by-side. Finally, the target picture’s label was presented auditorily. The authors manipulated the phonological overlap between the prime label and the target label, so that in half of the trials both labels were phonologically related, sharing phonological onset (*cat*-*cup*), or phonologically unrelated (*ball*-*comb*). Prime and target-distractor pairs were semantically unrelated. Interestingly, infants showed different target looking preference patterns after phonologically related prime pictures, compared to after phonologically unrelated primes. This suggests that infants implicitly named prime pictures—despite such pictures having been presented in silence—and that the phonology of the resulting labels interacted with subsequent target.

Chow et al. ([2017](#ref-chow2017spoken)) provided further evidence of semantic *and* phonological links in the early lexicon. The authors adapted the Visual World Paradigm from Huettig & McQueen ([2007](#ref-huettig2007tug)) to explore 24- to 30-month-old toddlers’ visual fixation patterns during a word recognition when presented with phonological and semantic distractors. In each trial, the authors presented participants with four semantically and phonologically pictures. Four seconds after pictures onset, a word-form was auditorily presented. The word did not refer to any of the pictures displayed on the screen, but was phonologically related to one of them (both labels shared phonological onset), and semantically related to another one of the pictures (both referents belonged to the same taxonomic category. For instance, participants might be presented with the pictures of a sandwich, a bus, a cat, and a dress. Then they would hear the carrier phrase “Look at the *bee*!”. The authors registered participants fixations to the phonological and semantic distractors, and found evidence of a preference for the phonological distractor at earlier stages of the post-naming phase, and a preference for the semantic distractor at later stages of the trial. These results support a cascaded activation account of lexical access during the first stages of lexical development, paralleling previous findings in adults ([Allopenna et al., 1998](#ref-allopenna1998tracking)).

One of the most interesting pieces of evidence in support of a cascaded account of lexical processing is provided by bilinguals. A critical property the bilingual lexicon is that its representations are accessed in a language-non selective way. Even during monolingual situations, bilinguals access their lexical representations in both of their languages ([Duñabeitia et al., 2009](#ref-dunabeitia2009masked); [Duyck et al., 2007](#ref-duyck2007visual); [Hoshino & Kroll, 2008](#ref-hoshino2008cognate); [Strijkers et al., 2010](#ref-strijkers2010tracking); e.g., [Thierry & Wu, 2007](#ref-thierry2007brain)). This is reflected in bilinguals’ performance in word recognition and production tasks. For instance, Marian & Spivey ([1999](#ref-marian1999activation)) presented Russian-English bilinguals with a series of trials that started with an auditory instruction like “Poloji marku nije krestika” [“Put the stamp below the cross”]. The instruction would be present in Russian (dominant language for most participants). The target object (*marku* [*stamp*], in this case) would be present in the screen, together with an object whose English label shared phonological onset with the target label (e.g., *mark*), and two unrelated distractors whose labels in Russian and English did not shared phonological relationship with the target label. In line with Allopenna et al. ([1998](#ref-allopenna1998tracking))’s results in monolingual adults, bilinguals fixated the cross-language phonological distractor significantly more than the unrelated distractors after hearing the target label. These findings suggest that bilinguals activated phonologically related word-forms in both Russian *and* English, which ultimately affected their overt visual exploration patterns during the task.

Language non-selectivity in the bilingual lexicon has also been shown to affect top-down lexical access during word production. Costa et al. ([2000](#ref-costa2000cognate)) presented Catalan-Spanish bilinguals with a series of pictures of familiar objects. For each object, participants were asked to name it in Catalan (their dominant language). Unbeknownst to participants, the authors manipulated the phonological overlap between the pictures’ labels in Catalan and their translations in Spanish. In half of the trials, the labels associated to the pictures were cognates, that is, phonologically similar translation equivalents (e.g., *gat*-*gat* [*cat*]) in Catalan and Spanish, whereas in the other half of the trials the labels were non-cognates (e.g., *taula*-*mesa* [*table*]). Participants named pictures faster in cognate trials than in non-cognate trials. Critically, Spanish monolinguals who completed the same task, naming the pictures in Spanish, showed equivalent naming times in both conditions. These results reveal that bilinguals activated their Spanish phonology, despite performing the naming task exclusively in Catalan: the visual recognition of the presented pictures led to the activation of its associated phonological forms in both languages in parallel, which influenced the subsequent the dynamics of word production. This differential processing of cognates and non-cognates has generated a large body of literature providing support for the language-non selective account of lexical access ([Ana I. Schwartz & Diaz, 2007](#ref-schwartz2007reading); [Dijkstra et al., 2010](#ref-dijkstra2010cross); [Dufour & Kroll, 1995](#ref-dufour1995matching); e.g., [**degrootandnas1992?**](#ref-degrootandnas1992); [**schwartzkroll2007?**](#ref-schwartzkroll2007); [**dijkstragrainger1999?**](#ref-dijkstragrainger1999)).

Further evidence in favour of the parallel retrieval of phonological forms during bilingual word production is given by implicit naming paradigms. Using an adaptation of Mani & Plunkett ([2010](#ref-mani2010infant))‘s task, Von Holzen & Mani ([2014](#ref-von2014bilinguals)) found that bilingual adults generate implicit labels in both labels for visually fixated pictures. The authors presented German-English bilinguals with 120 prime-target pairs. Primes were presented as familiar pictures in silence. After prime picture offset, target words were presented auditorily. Participants’ N400 ERP components were recorded from target word onset. The authors manipulated the phonological relationship between the prime and target word-forms within participants’ L1 and across L1 and L2. When prime and target were phonologically identical (*Affe*-*Affe*, German for *monkey*), or similar (*Fahne*-*Sahne*, German for *flag* and *ice-cream*, respectively) within L1, participants’ showed a reduced N400 amplitude, compared to when prime and target were phonologically unrelated (*Messer*-*Seil*, German for *knife* and *rope*). Critically, a similar effect was found when prime and target were phonologically related through translation (*Rustsche*-*Kleid*, German for *slide* and *dress*). This suggests that participants activated prime labels in both languages in parallel, and that both labels impacted the dynamics of target words recognition.

Phonological priming paradigms have also revealed language non-selectivity in the initial bilingual lexicon ([Bosma & Nota, 2020](#ref-bosma2020cognate); [Jardak & Byers-Heinlein, 2019](#ref-jardak2019labels); [Poarch & Van Hell, 2012](#ref-poarch2012cross); [Singh, 2014](#ref-singh2014one); [Von Holzen et al., 2019](#ref-von2019impact)). For instance, Von Holzen & Mani ([2012](#ref-von2012language)) found evidence of cross-language phonological priming in 21- to 42-months-old children learning German and English. At the beginning of each trial, the authors auditorily presented an English prime word embedded in a carrier phrase. Then, the target label was also auditorily presented in German. Finally, the target and distractor pictures were presented side-by-side. In some trials, the auditorily presented prime and target labels were phonologically related (e.g., *slide*-*Klide* [*dress*]). In some other trials, both labels were phonologically related through translation: the auditorily presented prime label did *not* overlap with the target label (e.g., *leg*-*Stein* [*stone*]), but its translation in German did [*Bein*]. In the rest of the trials prime and target labels were phonologically unrelated in both languages. The authors found a facilitation effect of cross-language priming when both prime and target auditory labels overlapped phonologically, as revealed by a stronger target picture looking preference, as compared to that of unrelated trials. Interestingly, participants showed a weaker target preference in priming through translation trials.

Using a similar task, Floccia et al. ([2020](#ref-floccia2020translation)) provided conflicting results to those of Von Holzen & Mani ([2012](#ref-von2012language)). The authors tested 27-month-old simultaneous bilinguals in a cross-language priming paradigm. The two languages participants were learning were English, and an Additional Language (either Cantonese, Dutch, French, German, Greek, Italian, Mandarin, Polish, Portuguese, or Spanish). A prime word was embedded at the end of a carrier sentence that participants listened to. Then, participants were presented with the auditory label of the target word, and two pictures were shown side-by-side, the target picture, and a distractor picture. In Study 1, the target word was the translation equivalent of the prime word (*cheese*-*fromage* for an English-French bilingual). In other trials, prime and target words were unrelated (e.g., *sock*-*fromage*). There was no phonological overlap at onset between prime and target or distractor labels in either of the two conditions. Participants were presented with two blocks of related and unrelated trials. In one of the blocks, participants were presented with English prime words and Home Language target words (*sock*-*fromage*), and in the other block they were presented with Additional Language prime words and English target words (*fromage*-*sock*). The authors found a cross-linguistic priming effect, in which participants showed a stronger preference for the target picture in the post-naming phase in related trials, as compared to unrelated trials. This effect was found regardless of the language of the prime (English or Home Language), or the dominance of the language of the prime (i.e., whether the prime word belonged to the dominant language). Target preference exceeded chance level in both conditions, pointing to successful word recognition regardless of the experimental manipulation.

In a second study, Floccia et al. ([2020](#ref-floccia2020translation)), the authors tested an equivalent group of bilinguals in a cross-linguistic semantic priming task, in which participants were primed with words in English, and tested with semantically related words in the Additional Language (or vice versa). This time, prime and target words were not translation equivalents, but rather words whose referents shared semantic features. Results paralleled those from their first study. The authors found a strong semantic priming effect across languages, regardless of the language of the prime or the dominance of the language of the prime. Overall, these results suggest that bilinguals recognised target words in both conditions, and benefited from a cross-linguistic facilitation effect from translation equivalents. The symmetry of such effect across both languages suggests that participants lexical access occurred in parallel for the two members of translation equivalents. As the authors noted, the stimuli set of their second study contained multiple cognates due to design constraints, especially in languages with a larger degree of phonological overlap across translation pairs (e.g., English and Dutch). The authors run *post hoc* analyses exploring the effect of cognateness on the cross-language priming effects found, which showed no evidence of a cognateness effect on target preference. However, the cross-language effects reported in this second study might still reflect the confounding influence of cognateness, as the resulting effect might have been generated by an interaction by the phonological and semantic overlap between cognate primes and targets.

Previous studies had reported similar cross-language priming effects. The conclusion that such effects occurred at the lexical level, as opposed to being the result of mere acoustic similarity, was nonetheless unwarranted. In studies where the auditorily presented prime label was phonologically overlapped with the target label (e.g., [Marian & Spivey, 1999](#ref-marian1999activation)), priming effects might be explained by the activation of the phonological segments of the prime label, which facilitated the recognition of the target label, which shared such phonological segments. The priming-through-translation results by Von Holzen & Mani ([2012](#ref-von2012language)) and Floccia et al. ([2020](#ref-floccia2020translation)) demonstrate that cross-language activation in the developing lexicon occurs at the lexical level. Since participants’ target recognition was modulated by the non-presented prime translation, such an inhibitory effect must stem at the lexical level. The fact that phonological priming facilitated the recognition of the target words (both labels belong to different languages), and that phonological priming through translation interfered target recognition (translation and target belong to the same language) points to cross-language connections playing an excitatory role, and to within-language connections playing an inhibitory role.

The priming-through-translation paradigm provides a convenient framework to investigate parallel activation in the bilingual lexicon, but its design prevents strong interpretations about the extent to which parallel activation is the result of the experimental manipulations of interest. The presentation of primes and targets in both languages in the same task—or even within the same trial—presents participants with a bilingual context that might lead to a higher baseline activation of lexical representations of both languages, contrary to a context in which parallel activation results from genuine cross-language activation ([Grosjean, 1997](#ref-grosjean1997bilingual)). Another limitation of priming-through-translation studies is that, when prime words are auditorily presented, the task constrains the naming context to one of the languages. The consequence is a potential asymmetry between the activation dynamics of the lexical representation of both labels: while the participant has heard one label, the activation of its translation must spread across semantic links. This might lead to a lower strength, and higher latency of the activation of the translation.

In the present study, we put the hypothesis that bilingual infants activate both languages in parallel to a more severe test. We used an extension of the implicit naming paradigm by Mani & Plunkett ([2010](#ref-mani2010infant)), in which participants were primed with a picture presented in silence. This provides participants with an exclusively monolingual context in which they perform the task in only one of their languages. If differences in task performance are found associated with covert manipulation of cross-linguistic properties of the stimuli, such differences may be more conclusively linked to the parallel activation of lexical representations in participants’ lexicons. By using an implicit naming paradigm, the activation of the lexical representation of the prime may occur in both languages as the result of lexical retrieval after visual recognition of the prime picture, since participants are not presented with any auditory prime label. This contrasts with priming through-translation tasks, in which the prime label in one of the languages is directly activated by auditory word recognition, while the label of its translation is expected to be activated as the result of cascaded activation spreading through their semantic association. This may lead to an asymmetry in the time course of the activation of both labels, which might in turn affect the dynamics of phonological priming in subsequent word recognition.

In summary, the present study was aimed at investigating the developmental trajectories of word recognition in bilingualism, and how (if at all) language non-selectivity shapes them. Using an adaptation of ([Mani & Plunkett, 2010](#ref-mani2010infant)), we presented 20- to 32-months-old toddlers with a cross-linguistic phonological priming paradigm. We manipulated both the phonological overlap between the prime word and target word, and the cognate status of the prime word, to test the hypothesis that bilingual toddlers implicitly name familiar pictures in both languages. If so, phonological priming effect should be larger when the prime label in both languages shares phonological overlap with the target words (*Related/Cognate* condition), compared to when only the label of the prime in one language overlaps phonologically with the target (*Related/Non-cognate*). We also also included a control *Unrelated* condition in which regardless of the cognate status of the prime word, neither of the prime labels in any language overlapped phonologically with the target word.

To test cross-language priming effects, we registered participants’ time course of target picture looking in the three conditions. We predicted within-language priming effects, revealed as differences in target looking between *Unrelated* and *Related/Non-cognate* trials, as previously reported Mani & Plunkett ([2011](#ref-mani2011phonological)). We further predicted cross-language priming effects, revealed as differences in target looking between *Related/Non-cognate* and *Related/Cognate* trials. This effect should be stronger in bilinguals, if not completely absent in monolinguals, who should be indifferent to the phonological properties of the stimuli in the other language.

Following previous studies associating the emergence of phonological links between lexical items to the size of children’s vocabulary ([Mani & Plunkett, 2011](#ref-mani2011phonological)), we also compared the explanatory power of two indices of receptive vocabulary size (dominant-language vocabulary size and total vocabulary size) against participants’ age. We used parental reports of vocabulary checklist to estimate participants’ vocabulary size. We computed vocabulary size in the dominant language (L1 vocabulary) as the proportion of words in the dominant language reported by their caregivers to be understood by the child in a vocabulary checklist, and the total vocabulary size as the proportion of words reported as understood by the child in *both* languages. If within-language or cross-language phonological associations in the lexicon emerge as a function of vocabulary growth, the size of the differences between conditions should larger in participants with larger vocabulary sizes. A model including participants’ L1 or total vocabulary size as a predictor should then show better performance than a model that instead includes participants age. In the case of bilinguals, if the emergence of cross-language priming effects is associated with the growth of their vocabulary is *both* languages, a model including total vocabulary size as a predictor should perform better than a model that includes L1 vocabulary size. The opposite pattern would indicate that cross-language priming effects emerge as a function of the growth of L1 vocabulary.

Testing bilingual populations at early ages presents the limitation that the vocabulary size of participants’ non-dominant language is relatively small, which substantially constrains the number of items that can be used to as stimuli in the task, as well as the number of trials that can be constructed to generate stimuli lists. In order to circumvent the problem of limited vocabulary knowledge in the non-dominant language, tested participants only in their dominant language, therefore directing our research to the question of how possessing a second language affects processing of the dominant one in bilingual toddlers ([Costa & Sebastián-Gallés, 2014](#ref-costa2014does) provide a fuller description of the advantages of such approach).

# Study 1

## Methods

All materials, data, and reproducible code can be found at the OSF ([https://osf.io/hy984/](https://osf.io/ckydb/)) and GitHub (<https://github.com/gongcastro/cognate-priming>) repositories. This study was conducted according to guidelines laid down in the Declaration of Helsinki, and was approved by the Drug Research Ethical Committee (CEIm) of the IMIM Parc de Salut Mar, reference 2020/9080/I. Before every testing session, caregivers were asked to read and sign an informed consent form, and were given a token of appreciation at the end of it.

### Participants

We collected data from 162 children living in the Metropolitan Area of Barcelona (Spain), tested at the Laboratori de Recerca en Infància at the Universitat Pompeu Fabra. Families were recruited from maternity rooms in private hospitals and social media, and contacted via phone when the child’s age spanned between 20 and 32 months. From the 162 children that participated, 81 participated once, 55 participated twice, and 26 participated three times. Recurrent participants were tested with at least XXX months of difference. We gathered a total of 269 testing sessions. Participants were divided into monolinguals and bilinguals based on their relative degree of exposure to Catalan and Spanish, estimated using the Language Exposure Questionnaire (LEQ, [Bosch & Sebastián-Gallés, 2001](#ref-bosch2001evidence)). We categorised participants as monolingual if exposed to more than 80% or more of the time to their dominant language, and as bilingual otherwise. Eighty-three of the participants were categorised as monolinguals (49 female, 34 male) and 80 as Catalan/Spanish bilinguals (34 female, 48 male) (see **?@tbl-participants** for a detailed summary of participants’ age and language profile). Participants’ vision was normal, none used glasses or any other type of vision corrector.

|  |  |  |  |  | *Degree of Exposure (%)* | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *Age (months)* | | *Spanish* | *Catalan* | *English* |
| *N sessions* | *N participants* | *M (SD)* | *Range* | *M (SD)* | *M (SD)* | *M (SD)* |
| Monolingual | | | | | | | |
| Catalan | 95 | 53 | 25.7 (3.8) | 20.0 (32.3) | 4.3 (5.8) | 95.3 (6.0) | 0.3 (2.1) |
| Spanish | 53 | 30 | 25.4 (3.8) | 20.0 (32.0) | 91.2 (6.5) | 8.4 (6.6) | 0.3 (0.9) |
| Bilingual | | | | | | | |
| Catalan | 72 | 51 | 25.3 (3.8) | 19.4 (31.7) | 37.5 (10.4) | 61.8 (10.7) | 0.2 (0.8) |
| Spanish | 49 | 33 | 25.1 (3.4) | 20.2 (31.3) | 61.0 (11.1) | 38.7 (10.6) | 0.3 (1.4) |

**?(caption)**

### Vocabulary size

We collected vocabulary data using parental responses to the Barcelona Vocabulary Inventory (BVQ, [Garcia-Castro et al., 2023](#ref-garcia-castro2023bvq)), an online vocabulary checklist inspired in several adaptations of the the Communicative Developmental Inventory (CDI, [Fenson et al., 1994](#ref-fenson1994variability)) developed to assess the vocabulary size of Catalan-Spanish bilingual toddlers. Families received a link to the BVQ immediately after each experimental session, and were given two weeks to fill it. We calculated two measures of receptive vocabulary size for each participant: *L1 vocabulary size* (proportion of words reported as acquired in the checklist of the dominant language), and *total vocabulary size* (proportion of the words in both checklists reported as acquired). 132 (49%) Families failed to provide a complete response to the BVQ within the two-week time limit, or did not provide a successful response to the questionnaire. For missing questionnaire responses, we imputed the vocabulary size of the participant using single imputation, using the vocabulary size scores of a pool of 586 additional participants for which a successful response for the questionnaire had been gathered. We used participants age in months and their language profile (monolingual/bilingual) as predictors. We used the mice R package ([Van Buuren & Groothuis-Oudshoorn, 2011](#ref-van2011mice)) to perform imputation using the Bayesian linear regression method.

Overall, there was a substantial increase in both total and L1 vocabulary sizes associated to age (see [Figure 1](#fig-vocabulary)). Monolinguals showed overall higher L1 vocabulary sizes than monolinguals: they were reported to understand an average of 56.9% (*SD* = 17.74%) of the words at around 21 months, 75.9% (*SD* = 14.79%) at 25 months, and 82.6% (*SD* = 14.95%) at 30 months, while bilinguals were reported to understand an average of 49.1% (*SD* = 22.51%) of the words at around 21 months, 68.0% (*SD* = 20.33%) at 25 months, and 79.0% (*SD* = 13.80%) at 30 months. Total vocabulary sizes showed the opposite pattern, with monolinguals infants being reported to understand 40.50% (*SD* = 15.989%) of the words at around 21 months, 59.83% (*SD* = 19.619%) at 25 months, and 63.30% (*SD* = 21.128%) at 30 months, and bilinguals being reported to understand 47.93% (*SD* = 23.261%) of the words at around 21 months, 64.05% (*SD* = 17.150%) at 25 months, and 75.83% (*SD* = 17.170%) at 30 months.

|  |
| --- |
| Figure 1: Participant vocabulary sizes across ages and language profiles. Vocabulary size scores are presented using two measures: L1 vocabulary size (proportion of words marked as *Understands* in the vocabulary checklist of the dominant language), and total vocabulary size (proportion of words marked as *Understands* in vocabulary checklists of both dominant and non-dominant languages). Grey lines connect vocabulary size scores from the recurrent participants. For visualisation purposes, participants were group into categorical age groups at 21, 25, 30 months, the ages around which most participants were tested. Only participants included in the final data set are shown. |

### Stimuli

We used 62 distinct words included in the BVQ to create the stimuli lists. We created six stimuli lists: three in Catalan, and three in Spanish. Each list contained 32 trials, each involving a prime-target-distractor group. Each word played a role as either prime, *or* as target and distractor across the three lists in their corresponding language. For instance, the Catalan word *cadira* appeared as *prime* in the three lists, but never as *target* or *distractor*; the Catalan word *bici* appeared as *target* and *distractor* across the three lists, but never as a prime. Target-distractor pairings were held constant across the three lists in each language. For instance, in all Catalan lists the word *bici* was paired with the word *porta*. Target-distractor pairings were also yoked, so that each member of the same target-distractor pair appeared once as target and once as distractor in each list. For instance, the *bici*-*porta* paired appeared twice in each of the three Catalan lists: once with *bici* as target and *porta* as distractor, and once with *porta* as target and *bici* as distractor. This counterbalancing avoided participants encountering looking at the target word guided solely by that word having being named in a previous trial. Finally, prime words appeared only once in each list: each target-distractor pair was associated with a different prime word in both appearances. In each list, the same prime word was presented alongside a different target-distractor pair. For instance, the Catalan prime word *barret* was presented with the *bici*-*porta* target-distractor pair in one list, with the *bici*-*porta* pair in another list, and with *berenar*-*amanida* in the remaining list. To avoid competition between target and distractor pictures, semantically related target-distractor pairs were perceptually distinct ([Arias-Trejo & Plunkett, 2009](#ref-arias2009lexical); [Floccia et al., 2020](#ref-floccia2020translation)). The order of the trials was randomised across experimental session, so that each time a participant was tested, the order in which the prime-target-distractor was presented was randomised. Each participant was randomly assigned to one of the three lists in the corresponding (dominant) language, and always the same list across their experimental sessions in the case of a recurrent participant.

In 16 of the 32 trials of the same list (henceforth *related* trials), the prime and the target words were phonologically related, sharing phonological onset (at least first phoneme). In the other 16 trials (*unrelated* trials), prime and target did not share phonological onset. 8 of the 16 *related* trials included a cognate prime (*cognate* trials), and the other 8 included a non-cognate trials (*non-cognate* trials). A prime word was considered cognate if its Catalan and Spanish translation shared phonological onset. Especial attention was paid to avoiding semantic or taxonomic relationships between prime and target words, and between prime and distractor words. Target and distractor word pairs were phonologically unrelated (did not share phonological onset). Some of them shared semantic features or a taxonomic relationship. This is the case of words associated with especially salient referents such as animals or food. To avoid infants guiding their gaze to these objects based on their saliency, we paired animals and food items together. The position of the target and distractor pictures (right or left) for each target-distractor pair was alternated, so that in one list the target would appear on the left, in another list it would appear on the right, and so on.

We examined the overall equivalence of the three trial types by comparing them across three variables relating to the target word: lexical frequency, word prevalence, animacy. **?@tbl-stimuli** shows a detailed summary of the stimuli properties, broken down by trial type and testing language. Lexical frequencies were extracted from the Catalan and Spanish corpora of the CHILDES database ([MacWhinney, 2000](#ref-macwhinney2000childes); [Sanchez et al., 2019](#ref-sanchez2019childes)) as counts per million words, and transformed into Zipf scores for easier cross-language comparison ([Van Heuven et al., 2014](#ref-van2014subtlex); [Zipf, 1945](#ref-zipf1945meaning)). We defined word prevalence as the proportion of same-aged infants who were reported to understand the word in the BVQ database.

#### Auditory stimuli

The auditory stimuli were natural exemplars of the selected target words, spoken by Central Catalan-Spanish proficient bilingual female speaker who was instructed to pronounce each word in a toddler-directed manner. Recordings were made with an Audio-Tecnica 328 microphone (AT2050) at a sampling rate of 44100 Hz, in a soundproof room at the *Laboratori de Recerca en Infancia* at University Pompeu Fabra. We used the Audacity and Praat ([Boersma & Van Heuven, 2001](#ref-boersma2001speak)) software packages to record and edit the audio files. The speaker was presented with a list of words in Catalan. The order of the words was pseudo-randomised, and each word was produced three times in a row before moving to the next word in the list. After going through all the words in the list, the speaker went through the word list again generating three tokens for each word, now in an inverse order (from bottom of the list to the top). We then repeated the same procedure for the list of Spanish words. The resulting audios were manually chunked into individual word-forms. For each of the six tokens produced for each word, the most adequate was selected for further processing. The audios were then transformed to stereo by duplicating them into two channels, denoised, and finally normalised. The mean duration of the final audios was 1.23 (*SD* = 0.17) and 1.08 (*SD* = 0.14) seconds for the Catalan and Spanish lists.

To make the pronunciation of the words as familiar as possible to each infant, we generated additional pronunciation variants for some words in Catalan and Spanish. Catalan words involving the // phoneme in their Central Catalan variant (e.g., /) were also recorded with such phoneme replaced by /j/ (e.g., /), a phonological process common in the Metropolitan Area of Barcelona. Spanish words involving the // phoneme were also generated replacing such phoneme with // to better accommodate Latin variants of Spanish. Before every experimental session, caregivers were asked to utter three written words involving the // phoneme (in the case of participants tested in Catalan) or the // phoneme (in the case of participants tested in Spanish). Each token contained the critical phoneme at onset, inter-vocalic position, and coda. The experimenter assigned the participant to the Catalan or Spanish stimuli list involving the closest variant to that of caregivers’.

#### Visual stimuli

Each word was depicted by a realistic phonographic representation of a typical exemplar. Image backgrounds were removed from the original pictures using the GNU Image Manipulation Program (GIMP), resized to a rectangle of a maximum of 400 pixels hight or wide, and finally placed in the centre of a 50% grey rectangle square of 500$$13.23 cm (11.613º visual angle from participants’ perspective).

|  | Prevalence (%) | Frequency | Animacy (%) | Duration (s) |
| --- | --- | --- | --- | --- |
| Catalan | | | | |
| Cognate | 36.8 (17.6) | 3.9 (0.6) | 22.9 | 1.2 (0.2) |
| Non-cognate | 37.9 (15.9) | 3.8 (0.5) | 25.0 | 1.3 (0.2) |
| Unrelated | 35.4 (14.8) | 3.9 (0.5) | 21.9 | 1.2 (0.2) |
| Spanish | | | | |
| Cognate | 30.0 (13.4) | 4.2 (0.5) | 12.5 | 1.1 (0.1) |
| Non-cognate | 31.4 (14.0) | 4.3 (0.5) | 14.6 | 1.1 (0.1) |
| Unrelated | 28.9 (14.8) | 4.2 (0.5) | 16.7 | 1.1 (0.1) |

**?(caption)**

### Procedure

Testing took place in a sound-proof room. Participants sat on their caregivers’ lap in a dimly lit testing booth while the experimenter conducted the experiment from outside. Caregivers were instructed to keep their eyes shut (to avoid recording their gaze, instead of the participant’s), to be still, and to avoid interacting with the participant verbally or non-verbally. Participants sat at approximately 65 cm from the eye-tracker and a XX-in screen of screen resolution. We used a custom Matlab XXXX script using the PsychToolbox-3 extension ([Brainard & Vision, 1997](#ref-brainard1997psychophysics); [Kleiner et al., 2007](#ref-kleiner2007s); [Pelli & Vision, 1997](#ref-pelli1997videotoolbox)) to present the stimuli, and the Tobii Analytics SDK 3.0 to interact with the eye-tracking while the experiment was running. Sampling rate was set at 120 Hz. A 5-point calibration was performed before every experimental session, in which the picture of a colourful beach ball was presented. We set a 55% grey background for the calibration and stimuli presentation. Auditory stimuli were presented through two loudspeakers located behind the screen, one to each side. The experimenter monitored the experimental from outside the room using a centrally located video camera place above the screen. After a successful calibration the experimenter triggered the onset of the first trial. Trials were presented uninterruptedly and without intervention of the experimenter until the 32 trials were presented, or the experimental session had to be stopped because of the participant’s behaviour.

|  |
| --- |
| Experimental task design with examples in Catalan. In each trial, the prime image is presented in silence for 1,500 ms. The the auditory target label is presented, and finally the target and distractor pictures are presented side-by-side for 2,000 ms. In cognate trials (*n* = 8), Catalan *and* Spanish prime labels shared phonological onset with the target label. In non-cognate trials (*n* = 8), only the Catalan prime label shared phonological onset with the target label. In unrelated trials (*n* = 32), none of the prime labels shared phonological onset with the target label. |

Each trial started with the presentation of an attention getter for 3,000 milliseconds. Then, the prime picture was presented in silence in the centre of the screen for 1,500 milliseconds. Fifty milliseconds after the offset of the prime image, an auditory label was played from the loudspeakers and, 700 milliseconds after the onset of the auditory label, the target and distractor pictures were presented side-by-side during 1,000 milliseconds until the end of the trial. After this, the attention getter of the next trial was immediately presented. Each experimental session took approximately 10 minutes.

### Data analysis

#### Data processing

We defined a time windows of interest from 200 ms after target and distractor pictures onset until the end of the trial at 2,000 ms when both pictures disappeared from screen. To avoid modelling fixations driven by processes other than auditory word recognition ([Fernald et al., 1998](#ref-fernald1998rapid), [2001](#ref-fernald2001half)). Missing eye-tracker samples were interpolated using the last-observation-carried-forward (see [Zettersten et al., 2022](#ref-zettersten2022peekbank) for a similar approach), with a maximum of 20 maximum consecutive missing samples being interpolated (an equivalent of 166.67 ms).

We gathered data from 8,608 trials from 269 testing sessions, generated from 162 distinct participants. We excluded trials in which participants failed to provide 50% valid eye-tracking samples (equivalent to 750 ms) during the prime phase (*n* = 1,815) or 50% valid samples (equivalent to 1,000 ms) during the target-distractor phase (*n* = 1,262). We also excluded trials in which participants did not provide at least 5% of valid samples (equivalent to 100 ms) of target or distractor looking in the test phase (*n* = 2,461) (see [Floccia et al., 2020](#ref-floccia2020translation) for a similar approach).

After trials that matched any of the aforementioned exclusion criteria from the dataset, we excluded participants who did not provide at least two valid trials in the *Cognate* condition, the *Non-cognate* condition, or the *Unrelated* condition (*n* = 139), and participants with a L1 vocabulary size lower than 10% (*n* = 3). The final dataset included 2,280 trials from 130 testing sessions, generated by 90 distinct participants. Of those participants, 57 provided data from one experimental session, 26 provided data from two experimental sessions, and 7 provided data from three experimental sessions. **?@tbl-attrition-trials** shows a detailed description of the trial attrition.

|  |  | Trials | | |
| --- | --- | --- | --- | --- |
| *Participants* | *Cognate* | *Non-cognate* | *Unrelated* |
| Bilingual | 49 (72) | 233 (735) | 241 (727) | 449 (1,487) |
| Monolingual | 81 (67) | 416 (768) | 418 (766) | 770 (1,598) |
| *N* | 130 | 649 | 659 | 1,219 |

**?(caption)**

#### Modelling approach

We used Bayesian Hierarchical General Additive Models (HGAMs) to model our data and test our hypotheses ([Pedersen et al., 2019](#ref-pedersen2019hierarchical)). First, we defined our response variable as the empirical logit of target looking using [Equation 1](#eq-elogit) across all eye-tracking samples for each each participant in each condition ([Agresti, 2012](#ref-agresti2012categorical); [Barr, 2008](#ref-barr2008analyzing)). We used a Gaussian distribution to model our response variable. To test our hypotheses, we then included *Condition*, *Group*, and either *Age* (), *L1 vocabulary* (), or *Total vocabulary* () as fixed effects in the model. We also added all two-way and three-way interactions between the predictors. We set two *a priori* contrasts for the *Condition* predictor: one comparing *Unrelated* and *Related/Non-cognate* trials (sum-coded as -0.5 and +0.5, with *Related/Cognate* trials coded as 0), and another comparing *Related/Non-cognate* and *Related/Cognate* trials (sum-coded as -0.5 and +0.5, with *Unrelated* trials coded as 0). We also set two *a priori* contrasts for the *Group* predictor: one comparing *Monolingual (ENG)* with *Monolingual (CAT/SPA)* participants (sum-coded as -0.5 and +0.5, with *Bilingual (CAT-SPA)* participants coded as 0), and another comparing *Monolingual (ENG)* with *Bilingual (CAT-SPA)* participants (sum-coded as -0.5 and +0.5, with *Monolingual (CAT/SPA)* participants coded as 0). Before entering the model, the numeric predictors *Age*, *L1 vocabulary*, and *Total vocabulary* were standardised by subtracting from each observation the mean of the predictor, and dividing the result from the standard deviation of the predictor.

We included *Session*—which indexes individual testing sessions that may belong to the same participant—as grouping variable, nested within *Participant* grouping variable—which indexes a distinct participant. This nested random effects structure incorporates the longitudinal design of data collection, in which multiple participants were tested more than once at different ages. We added by by-session intercepts and *Condition* slopes, and by-participant intercepts and *Age* slopes. To model the time course of target looking across time bins, we included B-splines for the main effect of *Time*, and one for an adjustment of the *Time* cubic spline by *Group* and *Relatedness* ([Wood, 2017](#ref-wood2017generalized)). For both splines, we specified basis functions or *knots*. [Equation 2](#eq-model) shows a formal implementation of the model.

The comparison between the three models including *Age*, *L1 vocabulary*, *Total vocabulary* as predictors was done using leave-one-out cross-validation (LOO-CV) as a benchmark of model performance, using Pareto-smoothed importance sampling (PSIS) to approximate it ([Vehtari et al., 2017](#ref-vehtari2017practical)). We implemented the model using brms ([Bürkner, 2017](#ref-burkner2017brms)), an R interface to the Stan probabilistic language (2.33.0) ([Carpenter et al., 2017](#ref-carpenter2017stan)). We ran four iteration chains using the by-default No U-Turn Sampler algorithm with 1,000 iterations each and an additional 1,000 warm-up iterations per chain.

#### Statistical inference

We assessed the practical relevance of the estimated regression coefficients of the model following Kruschke & Liddell ([2018](#ref-kruschke2018bayesian)). First, we specified a region of practical equivalence (ROPE) from NA to +NA, in the probability scale. This region indicates the range of values that we considered equivalent to zero. We then summarised the posterior distribution of each regression coefficient with the 95% highest density interval (HDI). This interval contains the true value of this coefficient with 95% probability, given the data. Finally, we calculated the proportion of posterior samples in the 95% HDI that fell into the ROPE, noted as , which indicates the probability that the true value of the regression coefficient falls into the ROPE (and therefore should be considered equivalent to zero). For example, indicates that, given our data, there is a 80% probability that the true value of the coefficient falls within the ROPE, and can therefore be considered equivalent to zero.

# Study 2

In this study, we conducted a conceptual replication of Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)). Using the same task from Study 1, we tested a same-aged group of monolinguals being exclusively raised in British English speaking homes, the same population tested in the original studies. As in Study 1, we manipulated the phonological relationship between the prime and target labels at word onset, as well as the cognate status of the prime label. Evidence from implicit naming of prime images would be revealed by differences in target looking or its time course across the trials between *Unrelated* trials, and *Related/Non-cognate* and *Related/Cognate* trials. Given that participants in this study are monolinguals—and contrary to the monolinguals in Barcelona, have received no exposure to any other language—we would not expect the cognate status of the prime labels to exert any effect on subsequent target recognition.

## Methods

### Participants

We collected data from 112 children (41 female, 68 male, 3 sex not reported, age: *Mean* = 26.36 months, *SD* = 4.01, *Range* = [20.03-32.5]), living in the Oxfordshire area (United Kingdom), tested at the Oxford BabyLab at the University of Oxford. Families were recruited from maternity rooms in private hospitals and social media, and contacted via phone when the child’s age spanned between 20 and 32 months. From the 112 children that participated, 97 participated once, and 15 participated twice. Recurrent participants were tested with at least XXX months of difference. We gathered a total of 127 testing sessions. All participants were being raised in exclusively British English monolingual homes.

### Vocabulary size

We collected vocabulary data using parental responses to the Oxford Communicative Development Inventory (OCDI) ([Hamilton et al., 2000](#ref-hamilton2000infant)). Families were sent the questionnaire immediately after each experimental session, and were given two weeks to fill it. Since all vocabulary sizes of English monolingual infants were only assessed in English, both *L1 vocabulary* and *total vocabulary* sizes were identical (see [Figure 2](#fig-vocabulary-oxf)).

|  |
| --- |
| Figure 2: Participant vocabulary sizes across ages and language profiles. Vocabulary size scores are presented using two measures: L1 vocabulary size (proportion of words marked as *Understands* in the vocabulary checklist of the dominant language), and total vocabulary size (proportion of words marked as *Understands* in vocabulary checklists of both dominant and non-dominant languages). Vocabulary size scores from the recurrent participants are shown linked. |

### Stimuli

New stimuli lists were generated to adapt the experimental procedure to English-learning infants. Stimuli lists were created following the same constraints as in Study 1. Auditory stimuli were recorded following the same procedure as in Study 1, produced by Southern British English native female speaker in a toddler-directed manner. Audio and picture processing were performed following the same workflow as in Study 1.

### Procedure and apparatus

Same as in Study 1. The study was run using a custom Matlab script, PresentMate, based on the PsychToolbox-3 extension ([Brainard & Vision, 1997](#ref-brainard1997psychophysics); [Kleiner et al., 2007](#ref-kleiner2007s); [Pelli & Vision, 1997](#ref-pelli1997videotoolbox)). Visual fixations were recorded using a Tobii TX 300 eye-tracker and a 23-in screen of 1920 1080 resolution. The Tobii Analytics SDK 3.0 was used to interact with the eye-tracking while the experiment was running. Sampling rate was set at 120 Hz. A 9-point calibration was performed before every experimental session, in which the picture of a colourful beach ball was presented.

### Data analysis

Same as in Study 1, but removing the *Group* predictor from the models, as only monolingual infants were tested.

We defined a time windows of interest from 200 ms after target and distractor pictures onset until the end of the trial at 2,000 ms when both pictures disappeared from screen. To avoid modelling fixations driven by processes other than auditory word recognition ([Fernald et al., 1998](#ref-fernald1998rapid), [2001](#ref-fernald2001half)). Missing eye-tracker samples were interpolated using the last-observation-carried-forward (see [Zettersten et al., 2022](#ref-zettersten2022peekbank) for a similar approach), with a maximum of 20 maximum consecutive missing samples being interpolated (an equivalent of 166.67 ms).

We gathered data from 3,484 trials from 110 testing sessions, generated from 97 distinct participants. We excluded trials in which participants failed to provide 50% valid eye-tracking samples (equivalent to 750 ms) during the prime phase (*n* = 829) or 50% valid samples (equivalent to 1,000 ms) during the target-distractor phase (*n* = 650). We also excluded trials in which participants did not provide at least 5% of valid samples (equivalent to 100 ms) of target or distractor looking in the test phase (*n* = 1,003) (see [Floccia et al., 2020](#ref-floccia2020translation) for a similar approach).

After trials that matched any of the aforementioned exclusion criteria from the dataset, we excluded participants who did not provide at least two valid trials in the *Cognate* condition, the *Non-cognate* condition, or the *Unrelated* condition (*n* = NA), and participants with a L1 vocabulary size lower than 10% (*n* = 2). The final dataset included 1,105 trials from NA testing sessions, generated by 67 distinct participants. Of those participants, 62 provided data from one experimental session, 4 provided data from two experimental sessions, and NA provided data from three experimental sessions. From the trials included in the final dataset dataset, 554 were *Unrelated* trials (564 previously excluded), were 324 *Related/Non-cognate* trials(235 previously excluded), and 227 were *Related/Cognate* trials (333 previously excluded).

Agresti, A. (2012). *Categorical data analysis* (Vol. 792). John Wiley & Sons.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439.

Ana I. Schwartz, J. F. K., & 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>

Arias-Trejo, N., & Plunkett, K. (2009). Lexical–semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3633–3647.

Barr, D. J. (2008). Analyzing ‘visual world’eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474.

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*, *109*(9), 3253–3258.

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*(4), 369–380.

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

Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glot International*, *5*(9/10), 341–347.

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29–49.

Bosma, E., & Nota, N. (2020). Cognate facilitation in frisian–dutch bilingual children’s sentence reading: An eye-tracking study. *Journal of Experimental Child Psychology*, *189*, 104699.

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

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28.

Byers-Heinlein, K., Gonzalez-Barrero, A. M., Schott, E., & Killam, H. (2023). Sometimes larger, sometimes smaller: Measuring vocabulary in monolingual and bilingual infants and toddlers. *First Language*, *0*(0), 01427237231204167. <https://doi.org/10.1177/01427237231204167>

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*.

Chow, J., Davies, A. 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.

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.

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, *15*(5), 336–345.

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children’s vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*(6), 1189–1211.

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.

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, *23*(2), 166–180.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2009). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i–185.

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.

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.

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.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The wordbank project*. MIT Press.

Friedrich, M., & Friederici, A. D. (2005). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport*, *16*(6), 653–656.

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, *40*(4), 621.

Garcia-Castro, G., Ávila-Varela, D. S., & Sebastian-Galles, N. (2023). *Bvq: Barcelona vocabulary questionnaire database and helper functions*. <https://gongcastro.github.io/bvq>

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171–183.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*(4), 267–283.

Grosjean, F. (1997). The bilingual individual. *Interpreting*, *2*(1-2), 163–187.

Hallé, P. A., & Boysson-Bardies, B. de. (1994). Emergence of an early receptive lexicon: Infants’ recognition of words. *Infant Behavior and Development*, *17*(2), 119–129.

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.

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.

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*(1), 501–511.

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.

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by latino children learning spanish as their first language. *Journal of Child Language*, *34*(2), 227–249.

Jardak, A., & Byers-Heinlein, K. (2019). Labels or concepts? The development of semantic networks in bilingual two-year-olds. *Child Development*, *90*(2), e212–e229.

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What’s new in psychtoolbox-3?*

Kruschke, J. K., & Liddell, T. M. (2018). The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, *25*, 178–206.

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2018). Vocabulary size and speed of word recognition in very young french–english bilinguals: A longitudinal study. *Bilingualism: Language and Cognition*, *21*(1), 137–149.

MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.

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.

Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, *121*(2), 196–206.

Marian, V., & Spivey, M. (1999). Activation of russian and english cohorts during bilingual spoken word recognition. *Proceedings of the 21st Annual Conference of the Cognitive Science Society*, 349–354.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*(5838), 631–631.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, *23*(7), 728–733.

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning*, *44*(4), 617–653.

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876.

Pelli, D. G., & Vision, S. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Poarch, G. J., & Van Hell, J. G. (2012). Cross-language activation in children’s speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *111*(3), 419–438.

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.

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). Childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, *51*, 1928–1941.

Singh, L. (2014). One world, two languages: Cross-language semantic priming in bilingual toddlers. *Child Development*, *85*(2), 755–766.

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928.

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

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172–175.

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432–444.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, *45*, 1–67.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for british english. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*, 1413–1432.

Vihman, M. (2004). Cross-linguistic experiments in word-form recognition. *The Journal of the Acoustical Society of America*, *115*(5\_Supplement), 2502–2502.

Vihman, M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in english, welsh, and english–welsh bilingual infants. *Applied Psycholinguistics*, *28*(3), 475–493.

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.

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569–586.

Von Holzen, K., & Mani, N. (2014). Bilinguals implicitly name objects in both their languages: An ERP study. *Frontiers in Psychology*, *5*, 1415.

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), 1053–1075.

Wood, S. N. (2017). *Generalized additive models: An introduction with r*. CRC press.

Zettersten, M., Yurovsky, D., Xu, T. L., Uner, S., Tsui, A. S. M., Schneider, R. M., Saleh, A. N., Meylan, S. C., Marchman, V. A., Mankewitz, J., et al. (2022). Peekbank: An open, large-scale repository for developmental eye-tracking data of children’s word recognition. *Behavior Research Methods*, 1–16.

Zipf, G. K. (1945). The meaning-frequency relationship of words. *The Journal of General Psychology*, *33*(2), 251–256.