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 the end the 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 ([Bergelson, 2020](#ref-bergelson2020comprehension); [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 of age, they understand more than 400 ([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’ ([Bialystok, 2009](#ref-bialystok2009bilingualism); [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 relatively reduced linguistic input in each of their languages, compared to monolinguals ([Cattani et al., 2014](#ref-cattani2014much); [Costa & Sebastián-Gallés, 2014](#ref-costa2014does); [Thordardottir, 2011](#ref-thordardottir2011relationship)). Second, they face a more complex referential context: they often learn two labels for each referent (one in each language), which additionally may not be direct translations of each other ([Au & Glusman, 1990](#ref-au1990principle); [Bilson et al., 2015](#ref-bilson2015semantic); [De Houwer et al., 2006](#ref-de2006early); [Tsui et al., 2022](#ref-tsui2022translation)). The mechanisms that allow bilingual’ trajectories of lexical developmental to keep up with monolinguals’ are still unclear.

Previous studies have pointed to the similarity between the two languages of exposure as a facilitator of lexical acquisition in bilinguals ([Blom et al., 2020](#ref-blom2020cross); [Floccia et al., 2018](#ref-floccia2018vocabulary); [Gampe et al., 2021](#ref-gampe2021does)). Floccia et al. ([2018](#ref-floccia2018vocabulary)) reported larger vocabulary sizes in bilingual toddlers leaning two languages that shared high lexical similarity. The authors collected parental reports of vocabulary data from a sample of 367 bilingual children living in the United Kingdom, who were learning English and an additional language (out of a diverse pool of 13 languages). The authors then calculated the average phonolexical similarity between English and each of the additional languages. English and Dutch shared the highest similarity, while English and Mandarin shared the lowest. Overall, children’s vocabulary sizes in the additional language was positively associated with the amount of language similarity between their two languages. For instance, English-Dutch bilinguals showed larger vocabulary sizes in Dutch than English-Mandarin bilinguals did in Mandarin. The authors suggested that the acquisition of words in the additional language might be facilitated by their cognate status (i.e., being phonologically similar to their translation equivalent). If this is the case, larger vocabulary sizes might then be expected in bilinguals learning two languages sharing a high proportion of cognates. This would be consistent with available evidence of an earlier acquisition of cognate words ([Bosch & Ramon-Casas, 2014](#ref-bosch2014first); [Garcia-Castro, Avila-Varela, et al., 2023](#ref-garcia2023cognate); [Mitchell et al., 2022](#ref-mitchell2022cognates); [Schelletter, 2002](#ref-schelletter2002effect)).

The facilitation effect of cognateness is in line with the language non-selective account of bilingual lexical access. This account proposes that bilinguals activate both languages in parallel, even during monolingual situations. In adults, there is robust evidence in favour of this language non-selective account of lexical access ([Dijkstra et al., 1999](#ref-dijkstra1999recognition), [2010](#ref-dijkstra2010cross); [Dufour & Kroll, 1995](#ref-dufour1995matching); [Groot, 1992](#ref-groot1992determinants); [Marian & Spivey, 1999](#ref-marian1999activation); [Schwartz et al., 2007](#ref-schwartz2007reading); [Spivey & Marian, 1999](#ref-spivey1999cross)). Costa et al. ([2000](#ref-costa2000cognate)) presented highly-proficient Catalan-Spanish bilinguals with a series of pictures of familiar objects. Participants were asked to name each picture in Spanish. Unbeknownst to participants, the authors manipulated the cognate status pictures’ labels in Catalan and Spanish. In half of the trials, the labels associated to the pictures were cognates (e.g., *cat*-*gat* [cat]), 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. Spanish monolinguals showed equivalent naming times in both conditions. These results revealed that bilinguals activated their Catalan phonology, despite performing the naming task exclusively in Spanish: the visual recognition of the presented pictures led to the parallel activation of its associated phonological forms in both languages, which influenced the subsequent dynamics of word production.

Parallel activation has also been reported in the developing lexicon ([Bosma et al., 2019](#ref-bosma2019longitudinal); [Bosma & Nota, 2020](#ref-bosma2020cognate); [Floccia et al., 2020](#ref-floccia2020translation); [Jardak & Byers-Heinlein, 2019](#ref-jardak2019labels); [Poarch & Van Hell, 2012](#ref-poarch2012cross); [Singh, 2014](#ref-singh2014one); [Von Holzen et al., 2019](#ref-von2019impact)). Von Holzen & Mani ([2012](#ref-von2012language)) found evidence of cross-language phonological priming in a sample of 20 German-English bilinguals aged 21 to 43 months. In their experimental task, each trial begun with the auditory presentation of an English prime word, followed by a target word in German, and a pair of target and distractor pictures. The authors recorded participants’ target picture looking as a measure of target word recognition. The authors manipulated the cross-linguistic phonological overlap between the prime and the target labels. In a *priming through translation* condition, the English prime labels (leg) did not overlap with the German target labels (*Stein* [stone]), but with their German translations (*Bein*) did. In the *unrelated* condition, prime and target labels were not phonologically related in either German or English. If participants accessed their lexicon in a language non-selective way, the auditory presentation of the prime label in English should lead to the co-activation of its German translation. If this is the case, target word recognition should be interfered by the prior activation of a phonologically related German prime label. Under this hypothesis, the authors anticipated an delayed target looking in priming through translation trials, compared to unrelated trials. The results supported this hypothesis. In spite of the relevance of Von Holzen & Mani ([2012](#ref-von2012language)) study, some methodological issues deserve some consideration. First, for most participants, exposure to English (the less prevalent language) was lower than the minimal amount conventionally considered the threshold for bilingual exposure ([Byers-Heinlein et al., 2021](#ref-byers-heinlein2021multilab); [Rocha-Hidalgo & Barr, 2023](#ref-rocha-hidalgo2023defining)). Second, some of the prime labels in the priming through translation condition were cognates. If both English and German labels overlap phonologically with the German target label, priming effects can be explained by interference between words from the same language, as opposed to cross-language interference. Third—and most critically—participants were exposed to both English and German word in a by-trial basis. This creates a context in which interference effects may not have arised from the competition between the prime translation and the target words, but between the target word and any other word in the other language. Paradigms in which the experimental task is conducted exclusively in one language, while cross-linguistic features are covertly manipulated, offer a methodologically stronger basis for studying language non-selectivity in the developing lexicon ([Grosjean, 1997](#ref-grosjean1997bilingual)).

Mani & Plunkett ([2010](#ref-mani2010infant)) designed an implicit naming task, in which primes consisted of pictures presented in silence, instead of auditory labels. In each trial, English monolingual infants were first presented with pictures of familiar objects for 1,500 ms. Then, a target-distractor picture pair was presented for 2,000 ms, and then the auditory label of the target picture was presented. Post-naming target looking was recorded for another 2,000 ms until the end of the trial, as a measure of target word recognition. The authors manipulated the phonological overlap between the prime and the target labels, so that in half of the trials both labels were phonologically related, sharing phonological onset (cat-cup), or phonologically unrelated (ball-comb). Prime, target and distractor were unrelated otherwise. At 18 months of age, participants showed a stronger looking preference for the target pictures after phonologically related prime pictures, compared to after phonologically unrelated primes. Since the prime pictures were presented in silence, their results suggested that infants implicitly named the prime pictures, and that the phonology of the resulting word interacted with the subsequent recognition of the auditory target word. Later, Mani & Plunkett ([2011](#ref-mani2011phonological)) tested 21-month-old infants in the same task. This time, priming effects were observed in the opposite direction: when prime and target labels were phonologically related, participants showed significantly weaker target looking preference, compared to unrelated trials. The size of this interference effect was associated to participants’ vocabulary size, and to the cohort size of the prime label. The authors interpreted this finding as indicating a developmental shift. At 18 months, participants’ word recognition might have experienced a sub-lexical facilitation effect, in which the prior activation of the shared phonological onset between the prime and target labels boosted word recognition. In older participants, the lexicon might have reached a critical size at which the recognition of the target was delayed by the activation of its phonological cohort.

The implicit naming paradigm provides an ideal experimental paradigm study the developing bilingual lexicon. By covertly manipulating the cross-linguistic relationship between the prime and target labels, parallel activation can be tested while participants are presented with auditory stimuli (target labels) exclusively in one of their languages (see [Von Holzen & Mani, 2014](#ref-von2014bilinguals) for a similar approach in bilingual adults). Capitalizing on the language non-selective account of lexical access, we exploited the implicit naming to investigate the mechanisms behind the emergence of phonological priming effects in the bilingual developing lexicon. We tested a cohort of monolingual and bilingual infants learning Catalan and Spanish between 20 and 32 months of age. This allowed us to compare the performance of participants with differing vocabulary sizes in the word recognition task. In order to circumvent the problem of limited vocabulary knowledge in the non-dominant language, we tested participants only in their dominant language ([Costa & Sebastián-Gallés, 2014](#ref-costa2014does)).

Following Mani & Plunkett ([2010](#ref-mani2010infant)), each trial in the task started with the silent presentation of a prime picture. Both monolingual and bilingual infants were expected to implicitly name the prime picture. According to the language non-selective hypothesis of lexical access, bilinguals should generate two labels for the prime picture, one in each language. To test this prediction, we manipulated the phonological similarity between the prime and the target words in both languages (see [Figure 1](#fig-hypotheses)). In *Related/Non-cognate* trials, prime and target labels shared phonological onset in only the language of test. For instance an infant tested in Catalan would be presented with a chair as prime picture () and with [spoon] as target label. In line with Mani & Plunkett ([2011](#ref-mani2011phonological)), we anticipated that the phonological overlap between prime and target should modulate target word recognition in both monolinguals and bilinguals. This should be reflected in a delayed target looking preference, compared to *Unrelated* trials, in which prime and target did not share phonological onset. In *Related/Cognate* trials, the prime shared phonological onset with the target in both languages. For instance, the same infants tested in Catalan would be presented with a car as prime picture () and with [spoon] as target label. In bilinguals, parallel activation of the prime in both languages should increase the cohort of the target word, leading to stronger interference effects in this condition, compared to *Related/Non-cognate* and *Unrelated* trials.

|  |
| --- |
| Figure 1: Predicted priming effects (or their absence) in the *Related/Cognate*, *Related/Non-cognate*, and *Unrelated* conditions, with examples for a participant tested in Catalan. Words represent lexical representations. Lexical representations of the task-relevant language (Catalan) are depicted inside grey boxes. Solid arrows indicate within-language priming effects, and dashed lines indicate cross-language priming effects. In *Related/Cognate* (A) and *Related/Non-cognate* (B) trials, the recognition of the Catalan target word [spoon] is predicted to be modulated by the prior activation of the prime label in Catalan. In *Related/Cognate* trials, the parallel activation of the prime label in Spanish is predicted to increase the strength of the priming effect. |

In line with previous studies in monolinguals, we further predicted that the strength of the lexical interference effects in the *Related/Non-cognate* and *Related/Cognate* conditions would be associated to participants’ vocabulary size. Target word recognition should be delayed by the activation of a larger cohort of phonologically related words ([Chow et al., 2017](#ref-chow2017spoken); [Mani & Plunkett, 2011](#ref-mani2011phonological)). In the case of bilinguals, two measures of vocabulary size are relevant: *total vocabulary size* and *dominant-language vocabulary size*. Total vocabulary size refers to the number of words that a child has acquired, regardless of the language they belong to. Dominant-language vocabulary size (L1 vocabulary size, henceforth) refers to the number of acquired words in the dominant language of the child (in which they were tested), the language in which they know more words. If the size of interference effects are more tightly associated to bilinguals’ total vocabulary size, it would support the assumption that lexical representations from both languages enter a unique cohort in bilinguals.

Because of the short-lived effects of cross-language activation on lexical processing, and to maximise the probability of detecting priming effects, we introduced a change in the sequence of the trials relative to the original implementation by Mani & Plunkett ([2010](#ref-mani2010infant)). We presented target auditory labels immediately after the offset of the prime picture, and before the onset of the target and distractor pictures. By presenting prime pictures and target auditory labels closer in time, implicit naming of the prime picture should be more likely to influence the recognition of the target word. To test the effects of this methodological change, we first run a control experiment, Study 1, in which we tested a group of same-aged English monolinguals. In Study 2, we tested a group of monolinguals and bilinguals learning Catalan and Spanish.

# Study 1

In this study, we conducted a conceptual replication of Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)). We tested a group of English monolinguals aged 20 to 32 months, living in the Oxfordshire area (United Kingdom). As highlighted in the introduction, participants were tested exclusively in English, their dominant language. As in the original studies, we manipulated the phonological relationship between the prime and the target label. In line with Mani & Plunkett ([2011](#ref-mani2011phonological)), we predicted participants’ target looking to change as a function of the phonological relatedness between the prime and target English labels. This would reveal that participants implicitly named the prime pictures, generating a phonological label that influenced the subsequent recognition of a phonologically related word. To establish a monolingual baseline for all conditions in Study 2, we also manipulated the cognate status of the prime in English and Spanish: in half of the trials the English prime label was phonologically related to its Spanish translation. Given participants’ lack of familiarity with Spanish (or any language other than English), participants’ performance was predicted to be unaffected by the cognate status of the primes.

## 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 and the Medical Sciences Research Ethics Board at the University of Oxford, reference R60939/RE009. 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 112 children (41 female, 68 male, with three additional participants’ sex not being reported; Age: *Mean* = 26.36 months, *SD* = 4.01, *Range* = [20.03-32.5]) (see **?@tbl-participants** for a detailed summary of participants’ age and language profile), living in the Oxfordshire area (United Kingdom). Participants were 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 2.82 months of difference. We gathered a total of 127 testing sessions. All participants were being raised in exclusively British English monolingual homes. Participants’ vision was normal, none used glasses or any other type of vision corrector.

|  |  |  |  |  | *Degree of Exposure (%)* | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Sample size* | | *Age (months)* | | *Spanish* | *Catalan* | *English* |
| *Sessions* | *Participants* | *M (SD)* | *Range* | *M (SD)* | *M (SD)* | *M (SD)* |
| Study 1 - Monolingual | | | | | | | |
| English dominant | 89 (21) | 79 (21) | 26.5 (4.1) | 20.0-32.1 | 0.0 (0.0) | 0.0 (0.0) | 100.0 (0.0) |
| Study 2 - Monolingual | | | | | | | |
| Catalan dominant | 87 (8) | 50 (7) | 25.8 (3.9) | 20.0-32.3 | 4.5 (6.0) | 95.1 (6.2) | 0.4 (2.2) |
| Spanish dominant | 46 (7) | 28 (6) | 25.2 (3.8) | 20.0-32.0 | 90.8 (6.4) | 8.7 (6.4) | 0.2 (0.8) |
| Study 2 - Bilingual | | | | | | | |
| Catalan dominant | 62 (10) | 45 (9) | 25.3 (3.8) | 19.4-31.7 | 37.9 (10.5) | 61.7 (10.3) | 0.2 (0.8) |
| Spanish dominant | 42 (7) | 31 (7) | 25.6 (3.4) | 20.2-31.3 | 61.1 (11.0) | 38.5 (10.4) | 0.4 (1.6) |
| Total | 326 | 233 | 25.7 | 19.9 | — | — | — |

**?(caption)**

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. In the case that a complete response to the OCDI was not provided within the two-week limit, the participants’ testing session was excluded from the analyses (*n* = 3). [Figure 2](#fig-vocabulary-oxf) shows the distribution of participants’ vocabulary sizes across ages.

|  |
| --- |
| Figure 2: Participant receptive vocabulary sizes across ages and language profiles. Vocabulary size scores from the recurrent participants are shown linked. For visualisation purposes, testing sessions were aggregated at three age points: 21, 25, and 30 months. Points and intervals indicate mean and standard deviations, respectively. |

### Stimuli

Participants were presented with 32 trials in random order, which belonged to three conditions: *Unrelated* (16 trials), *Related/Non-cognate* (8 trials), and *Related/Cognate* (8 trials). In *Unrelated* trials, target was phonologically unrelated with the English and Spanish labels of the prime picture (e.g., ). In *Related/Non-cognate* trials, the target shared phonological onset with the English prime label, but not with the Spanish prime label (e.g., ). In *Related/Cognate* trials, the target shared phonological overlap with both English and Spanish prime labels (). Target-distractor pairs were phonologically unrelated, and held constant for all participants. Especial attention was paid to avoiding semantic or taxonomic relationships between prime and target words, and between prime and distractor words. Each target distractor pair was shown twice for each participant, switching roles. Distractors were always phonologically unrelated to the prime and target labels in the same trial.

We created fours lists of trials, across which the same target-distractor pair appeared with a different prime, counterbalancing the condition to which it belonged. For instance, in list one the *ball*-*trousers* pair was preceded by *bike* (*Related/Cognate*), by *butterfly* (*Related/Non-cognate*) in list two, and by *star* and *nose* (*Unrelated*) in lists three and four. **?@tbl-stimuli** shows a detailed summary of the stimuli properties, broken down by trial type and testing language. Lexical frequencies were extracted from the English corpora from 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)).

|  | *Phonemes* | | *Frequency (Zipf)* | | *Phon. similarity [0-1]* | |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Prime* | *Target* | *Prime* | *Target* | *Prime-Prime* | *Prime-Target* | *Duration (s)* |
| Study 1: English | | | | | | | |
| Cognate | 5.0 (1.2) | 4.5 (1.3) | 4.5 (0.5) | 4.7 (0.5) | 0.2 (0.1) | 0.3 (0.1) | 0.9 (0.2) |
| Non-cognate | 5.8 (2.7) | 4.5 (1.3) | 4.7 (0.2) | 4.7 (0.5) | 0.0 (0.1) | 0.3 (0.1) | 0.9 (0.1) |
| Unrelated | 5.2 (2.3) | 4.3 (1.5) | 4.7 (0.4) | 4.7 (0.5) | 0.1 (0.1) | 0.1 (0.1) | 0.9 (0.1) |
| Study 2: Catalan | | | | | | | |
| Cognate | 4.5 (1.3) | 4.9 (1.3) | 5.0 (0.4) | 4.6 (0.6) | 0.5 (0.2) | 0.3 (0.2) | 1.2 (0.2) |
| Non-cognate | 4.9 (1.5) | 5.2 (1.3) | 5.0 (0.3) | 4.4 (0.7) | 0.0 (0.1) | 0.3 (0.1) | 1.3 (0.2) |
| Unrelated | 5.0 (1.5) | 5.0 (1.3) | 4.6 (0.5) | 4.2 (0.8) | 0.4 (0.3) | 0.1 (0.1) | 1.2 (0.2) |
| Study 2: Spanish | | | | | | | |
| Cognate | 4.5 (0.9) | 6.1 (1.6) | 4.9 (0.4) | 4.1 (0.8) | 0.6 (0.2) | 0.3 (0.1) | 1.1 (0.1) |
| Non-cognate | 5.2 (1.2) | 5.9 (1.5) | 4.9 (0.4) | 4.2 (0.8) | 0.1 (0.1) | 0.3 (0.1) | 1.1 (0.1) |
| Unrelated | 4.6 (1.1) | 5.7 (1.5) | 4.9 (0.3) | 4.2 (0.7) | 0.3 (0.3) | 0.2 (0.1) | 1.1 (0.1) |
| Total | *4.97* | *5.12* | *4.80* | *4.42* | *0.23* | *0.23* | *1.06* |

**?(caption)**

The auditory stimuli were natural exemplars of the selected target words, spoken by a Southern British English female speaker who was instructed to pronounce each word in a toddler-directed manner. We used the Audacity and Praat ([Boersma & Van Heuven, 2001](#ref-boersma2001speak)) software packages to trim, denoised, and normalise their amplitude. The visual stimuli were realistic photographic representations of a typical exemplars of the prime, target, and distractor words. 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 500 500 pixels. The final stimuli had a resolution of 72 dpi. When presented in the eye-tracker screen, the areas of interest (AOI) occupied an area of 13.23 13.23 cm (11.613 visual angle from participants’ perspective).

### Procedure and apparatus

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. 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. We set a 55% grey background for the screen during 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.

[Figure 3](#fig-task) shows a illustrates of the task design. Each trial started with the presentation of an attention getter for 3,000 ms. 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.

|  |
| --- |
| Figure 3: Experimental task design with examples in Catalan. In each trial, the prime image is presented in silence for 1,500 ms. Then 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* = 16), none of the prime labels shared phonological onset with the target label. |

### Data analysis

#### Data processing

We defined a time window 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. The first 200 ms of the test phase were discarded 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). Target looking probability was calculated as the empirical logit, using the number of samples inside the time bin in which the participant was looking at the target and distractor AOI (see [Equation 1](#eq-elogit)) ([Agresti, 2012](#ref-agresti2012categorical); [Barr, 2008](#ref-barr2008analyzing)). This was our response variable.

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); [Mani et al., 2012](#ref-mani2012activation) 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* = 19), and participants with a L1 vocabulary size lower than 10% (*n* = 3). The final dataset included 1,861 trials from 78 testing sessions, generated by 79 distinct participants. Of those participants, 69 provided data from one experimental session, 10 provided data from two experimental sessions, and NA provided data from three experimental sessions. From the trials included in the final dataset dataset, 915 were *Unrelated* trials (502 previously excluded), were 468 *Related/Non-cognate* trials(241 previously excluded), and 478 were *Related/Cognate* trials (229 previously excluded).

|  |  | Trials | | |
| --- | --- | --- | --- | --- |
| *Participants* | *Cognate* | *Non-cognate* | *Unrelated* |
| Study 2: Catalan/Spanish (Barcelona) | | | | |
| Bilingual | 104 (17) | 596 (372) | 555 (413) | 1,140 (796) |
| Monolingual | 133 (15) | 741 (443) | 747 (437) | 1,506 (862) |
| Study 1: English (Oxford) | | | | |
| Monolingual (English) | 89 (21) | 510 (359) | 511 (360) | 1,011 (733) |
| *N* | 326 | 1,847 | 1,813 | 3,657 |

**?(caption)**

#### Modelling approach

We used Bayesian Hierarchical General Additive Models (HGAMs) to model the data ([Pedersen et al., 2019](#ref-pedersen2019hierarchical)), using a Gaussian distribution to model the the logit of target looking. First, we fit a model () that included the main effects of *Condition* and *Age* as fixed effects in the model. We set two *a priori* contrasts for the *Condition* predictor ([Schad et al., 2020](#ref-schad2020capitalize)): 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). Before entering the model, the *Age* predictor was standardised by subtracting from each observation the mean of the predictor, and dividing the result from the standard deviation of the predictor.

We included the variable *Session*—which indexes individual testing sessions that may belong to the same participant—as grouping variable, nested within the *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-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 the interaction between *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.

We implemented this 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 two 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.

## Results

### Priming effects

We tested the differences between the conditions of interest in two ways. First, we examined the posterior distribution of the regression coefficients of the linear predictors in model (see [Equation 2](#eq-model)). We assessed the practical relevance of the coefficients following Kruschke & Liddell ([2018](#ref-kruschke2018bayesian)). We specified a region of practical equivalence (ROPE) from -0.1 to +0.1, in the logit 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.

Overall, the average participant’ target looking time exceeded chance levels, as indicated by the fact that the 95% HDI of the intercept term excluded zero ( = 0.400, 95% HDI = [0.311, 0.497]) and all of its posterior samples fell outside of the ROPE. The 95% HDI of the coefficient of *Age* had a positive sign, but did not exclude zero ( = -0.017, 95% HDI = [-0.097, 0.058]), and overlapped completely with the ROPE, indicating that participants from all ages showed equivalent overall target word recognition. The 95% HDI of the first contrast of the *Condition* predictor—comparing *Unrelated* and *Related/Non-cognate* trials—included zero ( = 0.126, 95% HDI = [-0.056, 0.284]), and 45.88% of its posterior samples fell within the ROPE. The 95% HDI of the second contrast—comparing *Related/Non-cognate* and *Related/Cognate* trials—also included zero ( = 0.128, 95% HDI = [-0.053, 0.310]), and 42.02% of its posterior samples fell within the ROPE. The overall target preference was equivalent across both pairwise condition comparisons.

**?(caption)**

Second, we examined the differences between the priming conditions across the time course of the trial, incorporating the smooth functions of the HGAMs to generate marginal posterior predictions for each condition across for each time point. [Figure 4](#fig-epreds-oxf) shows the posterior predictions of the model for each condition, and a summary of the difference between the *Unrelated* and *Related/Non-cognate* conditions, and between those for the *Related/Non-cognate* and *Related/Cognate* conditions, at each time point to test the practical relevance of these differences, were compared their 95% HDI against the [-0.1, +0.1] ROPE. This analysis revealed a similar pattern of results to the previously shown: predicted target looking for the three conditions overlaps across the full time course of the trial.

|  |
| --- |
| Figure 4: A) Posterior mean predictions of the time course of target fixation in the test phase. Thin lines represent 100 individual posterior predictions. Thick lines indicate the mean of the posterior predictions. B) Posterior mean prediction of the time course of the differences in target looking time between conditions. Thin lines represent 100 individual posterior predictions. Black lines indicate the mean of the posterior predictions. |

### Age and vocabulary size effects

To test our hypotheses regarding the role of age, dominant-language vocabulary size, and total vocabulary size, on the emergence of priming effects, we compared the fit of model against the fit of other models including the two-way interaction between *Condition* and *Age* (), or the two-way interaction between *Condition* and *L1 vocabulary* (), and all main effects involved. We compared the models using one-out cross-validation (LOO-CV) as a benchmark of model performance, using a Pareto-smoothed importance sampling (PSIS) approximation ([Vehtari et al., 2017](#ref-vehtari2017practical)). A better performance by models or over would point to *Age* or *L1 vocabulary*, respectively, playing a substantial role in participants’ word-recognition, or on the emergence of priming effects. **?@tbl-loos-oxf** shows a summary of the predictive performance of the models, as quantified by the expected log-predictive density (*ELPD*), and its standard error (*SE*, a measure of uncertainty around the *ELPD*). Overall, all models, performed equivalently, as shown by the small difference in *ELPD*, relative to the uncertainty of the estimates. This suggests that participants’ target looking during the test phase can be predicted with relative accuracy without taking into account the age or vocabulary size of the participants.

|  | *ELPD* | *SE ELPD* | *ELPD (diff.)* | *SE ELPD (diff)* |
| --- | --- | --- | --- | --- |
| Model 0 | -8,431.71 | 73.85 | --- | --- |
| Model 1 (Age) | -8,432.69 | 73.71 | -0.97 | 1.33 |
| Model 2 (L1 vocabulary) | -8,433.24 | 73.78 | -1.53 | 2.43 |
| \*ELPD\*: sum of expected log pointwise predictive density for a new dataset. \*SE \(ELPD\)\*: standard error of the \*ELPD\*, which indictes the uncertainty about the predictive performance for unknown future data. \*ELPD (diff)\*: pairwise difference in \*ELPD\* for two models. The difference is computed relative to the model with lowest \*ELPD\* (best fitting model). \*SE ELPD (diff.)\*: standard error of component-wise differences of \*ELPD\* between two models. | | | | |

**?(caption)**

## Discussion

In this study, we attempted to replicate the original study by Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)), using an extension of the implicit naming paradigm. adapted to the bilingual case. The main divergence between both implementations was the order in which target labels and target-distractor picture pairs were presented. In the original study, participants were presented with the target and distractor pictures immediately after the prime picture and the target auditory label was presented 2,000 ms after the onset of the prime and distractor pictures. In our adaptation of the paradigm, we presented the target auditory label immediately after the prime picture, before the target and distractor pictures were displayed in the screen. We then presented the target and distractor pictures 700 ms after the onset of the target label. The rationale behind this modification was to maximise the probability of observing any priming effects. By presenting the prime picture and the auditory target labels closer in time, the prime word might have a higher chance to influence the lexical selection of the target word. By testing a sample English monolingual infants with equivalent demographic and linguistic characteristics to the original studies, we expected to set a baseline to later compare bilingual infants in Study 2 against.

Although we found strong evidence of successful word recognition across participants of all ages, results did not provide any evidence in favour of phonological priming. English monolingual participants from all ages showed an equivalent pattern of target looking in both related (*Related/Cognate* and *Related/Non-cognate*) and *Unrelated* trials. The absence of a phonological priming effect suggest that either English monolinguals did not generate implicit labels for the prime pictures presented in silence, or that, if generated, such labels did not interact with the subsequent recognition of the target word. Both candidate explanations conflict with both Mani and Plunkett’s studies, and also with previous studies suggesting that infants 12-months and older already generate internal labels when presented with pictures of familiar objects ([Duta et al., 2012](#ref-duta2012erp); [Styles et al., 2015](#ref-styles2015infant)).

Adding the predictors *Age* or *Vocabulary size* as predictors in the model, in interaction with *Condition* did not increase the fit of the model. This points to neither variable having a substantial influence in participants’ target looking behaviour across conditions. These results diverge from previous studies reporting an increment in word recognition speed ([Fernald et al., 1998](#ref-fernald1998rapid); [Marchman & Fernald, 2008](#ref-marchman2008speed)), and stronger phonological priming effects in children with larger vocabulary sizes ([Chow et al., 2017](#ref-chow2017spoken); [Mani & Plunkett, 2011](#ref-mani2011phonological)). Overall, the outcomes of the Study 1 point to the absence of phonological priming effects in English monolingual toddlers, in contrast with previous studies. We discuss further candidate explanations for these results, and their implications for the lexical development literature in the General discussion.

# Study 2

Due to time constraints induced by the COVID-19 lockdown between 2020 an 2021, data collection in Study 1 and Study 2 was conducted simultaneously. In this study, we present data for monolinguals and bilinguals tested in Barcelona.

## Methods

### 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 2.06 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.

We collected vocabulary data using parental responses to the Barcelona Vocabulary Inventory (BVQ, [Garcia-Castro, Ávila-Varela, et al., 2023](#ref-garcia-castro2023bvq)), an online vocabulary checklist developed to assess the vocabulary size of Catalan-Spanish bilingual toddlers, and inspired in several adaptations of the the Communicative Developmental Inventory (CDI, [Fenson et al., 1994](#ref-fenson1994variability)). 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). One hundred thirty-two (49%) families failed to provide a complete response to the BVQ within the two-week time limit. We imputed missing vocabulary size scores using single imputation, taking 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 or 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 (see Appendix B).

### Stimuli

For Study 2, we created six stimuli lists: three in Catalan, and three in Spanish. Lists were created following the same constraints as in Study 1, but now considering the cross-linguistic phonological relationship between Catalan and Spanish, instead of English and Spanish. Extracting lexical frequencies from the Catalan and Spanish corpora in the CHILDES database was not possible, given the low number of participants and tokens included in them. Instead, we mapped the English lexical frequencies onto their Catalan and Spanish translation equivalents Garcia-Castro, Avila-Varela, et al. ([2023](#ref-garcia2023cognate)). The auditory stimuli were natural exemplars of the selected target words, spoken by a Southern British English female speaker who was instructed to pronounce each word in a toddler-directed manner. New visual stimuli were created to accommodate the words included in the new stimuli lists, and possible cultural differences in the typicality of the exemplars shown in the pictures.

### Procedure

Same as in Study 1. We used a custom Matlab XXXX script using the PsychToolbox-3 extension to present the stimuli, and the Tobii Analytics SDK 3.0 to interact with the eye-tracking while the experiment was running.

### Data analysis

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 10% of valid samples (equivalent to 100 ms) for both the target *and* the distractor (*n* = 2,461).

After excluding trials that matched any of the aforementioned 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* = 32), and participants with a L1 vocabulary size lower than 10% (*n* = 3). The final dataset included 5,019 trials from 237 testing sessions, generated by 150 distinct participants. Of those participants, 82 provided data from one experimental session, 49 provided data from two experimental sessions, and 19 provided data from three experimental sessions. **?@tbl-attrition** shows a detailed description of the trial attrition.

We modelled the data following a similar approach as in Study 1, with the main difference that participants’ language profile (*Group*) was now included as a predictor in the model. In Study 2, the base model included the the main effects of *Age*, *Condition*, and *Group*, and the two-way interaction between the *Condition* and *Group* predictors. Contrast coding of the *Condition* predictor was the same as in Study 1. We set one *a priori* contrasts for the *Group* predictor, comparing *Monolingual* with *Bilingual* participants (sum-coded as -0.5 and +0.5, respectively).

## Results

### Priming effects

Overall, the average participants’ looking time exceeded chance levels, as indicated by the fact that the 95% HDI of the intercept term excluded zero ( = 0.219, 95% HDI = [0.179, 0.260]), and that all of its posterior samples fell outside of the ROPE. The coefficient of *Age* had a positive sign, but its 95% HDI overlapped completely with the ROPE ( = 0.021, 95% HDI = [-0.016, 0.060]), indicating that participants from all ages showed equivalent overall target word recognition. The 95% HDI of the coefficient of *Group* also included zero ( = -0.010, 95% HDI = [-0.087, 0.077]) and completely overlapped with the ROPE, indicating an equivalent overall target preference in monolinguals and bilinguals,

The 95% HDI of the first contrast of the *Condition* predictor—comparing *Unrelated* and *Related/Non-cognate* trials—included zero ( = 0.058, 95% HDI = [-0.032, 0.144]), and 75.18% of its posterior samples overlapped with the ROPE. The 95% HDI of the second contrast, comparing *Related/Non-cognate* and *Related/Cognate* trials, also included zero ( = -0.015, 95% HDI = [-0.120, 0.098]), and 90.67% of its posterior samples overlapped with the ROPE. The overall target preference was equivalent across both pairwise condition comparisons. The interaction term between the first *Condition* contrast contained zero ( = -0.009, 95% HDI = [-0.208, 0.153]), with 55.26% of its posterior samples overlapping with the ROPE. The interaction term between the second *Condition* contrast also contained zero ( = 0.090, 95% HDI = [-0.127, 0.294]), and 75.18% of its posterior samples fell within the ROPE. The outcomes of this model provide strong evidence against differences between monolinguals and monolinguals, and inconclusive evidence for differences in overall target looking time across conditions.

**?(caption)**

An analysis of the time course of target looking revealed a similar pattern of results (see [Figure 5](#fig-epreds)). Posterior mean prediction for the three conditions overlap across the full time course of the trial in both language groups.

|  |
| --- |
| Figure 5: Marginal posterior mean predictions. A) Posterior mean predictions of the time course of target fixation in the test phase. B) Posterior mean prediction of the time course of the differences in target looking time between conditions. Lines and intervals indicate the mean and 95% credible interval of the posterior predictions. Red lines indicate the region of practial equivalence (ROPE). |

### Age and vocabulary size effects

A comparison between models including *Age* (), *L1 vocabulary* (), and *Total vocabulary* () against model , which only included *Age* as a main effect is shown in **?@tbl-loos**. Overall, all models performed equivalently, with the model showing slightly better performance. The equivalent performance of all models suggests that participants’ target looking during the test phase can be predicted with relative accuracy without taking into account *L1 vocabulary*, or *Total vocabulary* sizes. We now report the median and 95% HDI of the coefficients of , the best-fitting model.

|  | *ELPD* | *SE ELPD* | *ELPD (diff.)* | *SE ELPD (diff)* |
| --- | --- | --- | --- | --- |
| Model 2 (L1 vocabulary) | -20,593.64 | 135.96 | --- | --- |
| Model 0 | -20,593.70 | 135.87 | -0.06 | 2.23 |
| Model 1 (Age) | -20,594.33 | 135.85 | -0.69 | 2.22 |
| Model 3 (Total vocabulary) | -20,594.42 | 135.90 | -0.77 | 2.32 |
| \*ELPD\*: sum of expected log pointwise predictive density for a new dataset. \*SE \(ELPD\)\*: standard error of the \*ELPD\*, which indictes the uncertainty about the predictive performance for unknown future data. \*ELPD (diff)\*: pairwise difference in \*ELPD\* for two models. The difference is computed relative to the model with lowest \*ELPD\* (best fitting model). \*SE ELPD (diff.)\*: standard error of component-wise differences of \*ELPD\* between two models. | | | | |

**?(caption)**

## Discussion

In Study 2, tested a sample of 20-to-32-months-old bilinguals in our adaptation of the implicit naming paradigm to investigate the potential role of within- and cross-language phonological priming. We also tested a same aged monolingual control. As in Study 1, we within-language priming effects in both *Related/Non-cognate* and *Related/Cognate* trials, as the result of the phonological similarity between the prime and target labels in the language of test. This would suggest that participants lexicalised the prime pictures (presented in silence), and that the resulting word interacted with the auditory recognition of the subsequently presented target label. We also predicted bilingual participants to lexicalise the prime picture in both languages in parallel, in line with the language non-selective hypothesis of lexical access. If this was the case, participants should show differences in target looking between *Related/Non-cognate* trials—in which the target word shared phonological onset with only the prime label in the test language—and *Related/Cognate* trials—in which the target label also shared phonological overlap with the prime label in the other language.

Paralleling the results from Study 1 participant’ looking behaviour suggested robust word recognition, regardless of experimental condition, participant language profile, age, or vocabulary size. Overall, participants’ performance was equivalent across all priming conditions suggesting that no phonological priming took place within or across languages in either monolinguals or bilinguals. These results conflict with previous findings in several ways. First, and as discussed in the Discussion section of Study 1, previous investigations using implicit naming paradigms had found evidence of within-language priming effects in same-aged monolinguals learning English (and younger). In this study monolingual infants learning Catalan or Spanish did not show evidence of such priming effect, as suggested by their equivalent target looking behaviour in both *Unrelated* and *Related/Non-cognate* trials. Given that English monolinguals in Study 1 did not show priming effects either, it cannot be concluded that the diverging results from Study 2 and those available in the literature are due to cross-language differences between Catalan and Spanish, and English.

Second, our results are also in contrast with the non-selective account of lexical access. Bilinguals in Study 1 did not show evidence of cross-language priming, as suggested by the lack of differences in target looking between *Related/Non-cognate* and *Related/Cognate* trials. Under the non-selective account, we had predicted bilinguals in this study to lexicalise the prime pictures in both languages, which should have exerted a stronger priming effect on target recognition in *Related/Cognate* trials than in *Related/Non-cognate trials*. Overall, results in Study 2 did not provide evidence of phonological priming effects in monolingual and bilingual infant learning Catalan and Spanish.

# General discussion

In this paper, we investigated the developmental trajectories of cross-language co-activation in the initial lexicon. We tested a large cohort of monolingual and bilingual toddlers in an implicit naming paradigm, in which we designed three experimental conditions to manipulate the phonological overlap between the prime and target words within and across languages. In *Unrelated* trials, prime and target were phonologically unrelated in both languages. In *Related/Non-cognate* trials, prime and target labels shared phonological onset only in the dominant language of participants, in which they were tested. In *Related/Cognate* trials, the prime label was a cognate: prime and target labels shared phonological onset in both languages. In Study 1, we attempted to replicate the original findings by Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) in a same-aged English monolingual cohort. We found a null pattern of results in which we did not find any evidence of phonological priming. In Study 2, we tested a cohort of monolingual and bilingual infants learning Catalan, Spanish, or both, and found similar results, with no evidence of phonological priming effect in either monolinguals or bilinguals. In none neither study was participants’ word recognition influenced by their age or vocabulary size.

Previous studies on word recognition had provided evidence phonological priming effects in English monolingual infants 18-month-old and older. Mani & Plunkett ([2010](#ref-mani2010infant)) had shown a facilitation effect in word recognition when 18-month-olds were previously presented with a familiar picture whose label shared phonological onset with the target word. Mani & Plunkett ([2011](#ref-mani2011phonological)) extended these results to 21-month-olds, showing an effect in the opposite direction: phonologically related primes interfered with target word recognition. The results from both studies suggested that participants lexicalised the prime pictures—even if presented in silence—and that the resulting phonological form interacted with subsequent auditory word recognition. The shift from facilitation at 18 months to inhibition at 21 months was interpreted as a shift from sub-lexical facilitation at the phonetic level. The authors suggested that at 18 months, target word recognition was facilitated by the prior activation of an overlapping set of phonological segments. At older ages, participants with larger vocabulary sizes showed an interference effect generated by the increased size of the prime cohort, which led to slower selection of the phonologically related target word. This last conclusion was supported by the fact that the size of the interference effect was associated with the increase in participants’ vocabulary size, and also with the cohort size of the prime label—primes belonging to a larger neighbourhood of phonologically related words were subject to stronger interference effects during the task.

Even earlier evidence of implicit naming had been provided through electrophysiological measurements at 14 months of age. Duta et al. ([2012](#ref-duta2012erp)) presented name-known pictures in silence for one second. After one second, the picture’s label was auditorily presented, and the infants’ N400 component was registered. In some trials the auditory label corresponded to a canonical realisation of the word-form. In other trials, the word was mispronounced. Infants ERP signal differed when presented with correct pronunciations or mispronunciations, indicating that infants had generated expectations at the phonological level about the auditory labels of the picture before they were presented with its auditory label. A later study by Styles et al. ([2015](#ref-styles2015infant)) re-analysed the first 1,000 ms of the trials in the previous study (while the picture was being presented before the auditory label was presented), and found that the ERPs of name-known pictures (according to parental reports of receptive vocabulary) differed from those of name-unknown pictures (which infants would not be able to lexicalise). This provides further support to the account that at around 14 months of age, infants already name familiar pictures, even if presented in silence. Evidence of implicit naming in bilinguals is, however, scarce.

To our knowledge, only one study has used an implicit naming paradigm to study cross-language priming in bilinguals. Von Holzen & Mani ([2014](#ref-von2014bilinguals)) adapted Mani & Plunkett ([2010](#ref-mani2010infant)) procedure to test German-English bilingual adults. The authors manipulated the phonological overlap between the prime label in English (presented in silence) and a German target (auditorily presented) in both German and English. Participants’ evoked potentials revealed differences between trials in which the German prime label overlapped phonologically with the English target label, compared to trials in which none of the labels overlapped. These results suggest that adult bilinguals implicitly named prime pictures, and that the resulting label interacted with subsequent word recognition.

In contrast with previous studies in monolingual infants, and bilingual adults, we found no evidence of implicit naming was found. We consider three scenarios under which implicit naming might have occurred, in line with previous literature, but our design failed to capture it. First, it is be possible that infants in both Studies 1 and 2 implicitly generated phonological labels for the primes, but such labels lacked the phonological detail to interact with the subsequent recognition of a phonologically related target word. This is unlikely, given that both monolinguals and bilinguals have been shown to encode lexical representations with high phonological detail from early ages ([Bailey & Plunkett, 2002](#ref-bailey2002phonological); [Ramon-Casas et al., 2009](#ref-ramon2009vowel); [Swingley & Aslin, 2000](#ref-swingley2000spoken); [Tamási et al., 2016](#ref-tamasi2016measuring), [2017](#ref-tamasi2017pupillometry)).

A second possibility is that participants successfully retrieved a detailed phonological form of the prime labels, but such forms failed to interact with target recognition. This would be explained by the lack of strong associations between phonologically related lexical representations at these ages. But even if one considers the possibility that participants in Study 1 failed to show priming effects for this reason (for instance, the emergence of phonological associations might follow different trajectories in Catalan-Spanish infants, compared to English infants), the fact that English monolingual infants in Study 2 failed to show such priming effects contradicts previous findings on the same population, reporting priming phonological priming effects in even younger infants ([Mani & Plunkett, 2010](#ref-mani2010infant), [2011](#ref-mani2011phonological)).

Third, it is possible that the modifications of the implicit naming task in the present investigation might have reduce the chances of detecting the anticipated effects. The most critical difference between the original design of the implicit naming task by Mani and Plunkett and that of the present study is the absence of a pre-naming phase during the test phase. Target auditory labels were presented immediately after the offset of the prime picture. It is possible that such time interval was too short for participants to retrieve the phonological label of the prime picture before the target was presented. Further research should investigate the role of prime-target inter-stimulus interval on the time course of lexical access and selection in toddlers.

The null pattern of results found in the present investigation would have important implications for current models of bilingual lexical development. Recent accounts of bilingual word acquisition pointed to a facilitative role of phonological similarity between the two languages. The findings reported by Floccia et al. ([2018](#ref-floccia2018vocabulary)) pointed to a language distance effect, in which bilinguals learning two lexically similar languages (sharing more cognates) knew more words in their less dominant language than those learning two language with lower lexical similarity. Later work by Mitchell et al. ([2022](#ref-mitchell2022cognates)), Siow et al. ([2022](#ref-siow2022effect)), and Garcia-Castro, Avila-Varela, et al. ([2023](#ref-garcia2023cognate)) provided evidence for a more direct facilitation effect of cognateness. The main candidate mechanism behind the cognateness facilitation effect in word acquisition is parallel activation: the co-activation of form-similar translation equivalents during linguistic exposure might lead to a faster accumulation of learning instances for cognate words, which ultimately results in an earlier age of acquisition. In the present study, we provide evidence against such cross-language effects during word recognition. These outcomes are not in line with parallel activation taking place during speech exposure in bilinguals. It is possible that the mechanisms involved in the cognate facilitation in bilingual word acquisition are not captured by the task used in the present study. The implicit naming task used in the present investigation might be tapping into lower-level functions, related to acoustic and phonological processing, that are related to—but not critical for—word acquisition. Future research should contemplate alternative mechanisms for the cognateness facilitation effect on bilingual lexical acquisition.

The present study introduces several methodological contributions. First, this is, to our knowledge, the first study investigating cross-language priming using an implicit naming paradigm in toddlers. Second, the longitudinal design of the data collection provided a more suitable design for drawing conclusions about developmental change in word recognition trajectories, compared to studies that tested participants from a single age group ([Floccia et al., 2020](#ref-floccia2020translation)), or studies that with an exclusively cross-sectional design ([Von Holzen & Mani, 2012](#ref-von2012language)). Third, the modelling approach in the present study incorporates the double source of repeated measures simultaneously into the random effects structure of the multilevel model. The models in Studies 1 and 2 accounts for the time course of target looking across the trials and its associated autocorrelation using Hierarchical General Additive Models ([Barr, 2008](#ref-barr2008analyzing); [Pedersen et al., 2019](#ref-pedersen2019hierarchical)). The models also incorporate the observations from the same testing session as a nested grouping variable inside participants, so that the test session-level parameters in each model are estimated taking into account the estimations of other testing sessions from the same participants. Overall, this approach allowed us to simultaneously model the participant-level variability, the testing session-level variability, and the time series consisting of participants’ gaze behaviour during the task. Finally, the adoption of a Bayesian approach to implement and estimate the model allowed the incorporation of prior knowledge about the distribution of the parameters to generate stable estimates, despite the complexity of the model.

In summary, we tested the language non-selective hypothesis of lexical access in bilingual toddlers using an implicit naming paradigm. In Study 1, we did not find evidence of within-language or cross-language priming effects, suggesting that participants might have not generated covert labels for the prime pictures presented in silence. In Study 2, we attempted to replicated the original study showing evidence of implicit naming in younger infants, but found no evidence of such effects: monolingual infants did not show evidence of implicit naming. Overall, our results allude to more complex state of the investigation about early priming effects than anticipated from the literature.

# References

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

Au, T. K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development*, *61*(5), 1474–1490.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*(2), 1265–1282.

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

Bergelson, E. (2020). The comprehension boost in early word learning: Older infants are better learners. *Child Development Perspectives*, *14*(3), 142–149.

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.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11.

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.

Blom, E., Boerma, T., Bosma, E., Cornips, L., Heuij, K. van den, & Timmermeister, M. (2020). Cross-language distance influences receptive vocabulary outcomes of bilingual children. *First Language*, *40*(2), 151–171.

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., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers’ expressive vocabulary: Does form similarity matter? *International Journal of Behavioral Development*, *38*(4), 317–322.

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

Bosma, E., Blom, E., Hoekstra, E., & Versloot, A. (2019). A longitudinal study on the gradual cognate facilitation effect in bilingual children’s frisian receptive vocabulary. *International Journal of Bilingual Education and Bilingualism*, *22*(4), 371–385.

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>

Byers-Heinlein, K., Tsui, A. S. M., Bergmann, C., Black, A. K., Brown, A., Carbajal, M. J., & Wermelinger. (2021). A multilab study of bilingual infants: Exploring the preference for infant-directed speech. *Advances in Methods and Practices in Psychological Science*, *4*(1).

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*.

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.

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., & De Coster, S. (2006). Early understanding of two words for the same thing: A CDI study of lexical comprehension in infant bilinguals. *International Journal of Bilingualism*, *10*(3), 331–347.

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., 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/<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.

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.

Duta, M., Styles, S., & Plunkett, K. (2012). ERP correlates of unexpected word forms in a picture–word study of infants and adults. *Developmental Cognitive Neuroscience*, *2*(2), 223–234.

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.

Floccia, C., Sambrook, T. D., Delle Luche, C., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., et al. (2018). Vocabulary of 2-year-olds learning english and an additional language: Norms and effects of linguistic distance. *Monographs of the Society for Research in Child Development*, *83*(1), 7–29.

Fourtassi, A., Bian, Y., & Frank, M. C. (2020). The growth of children’s semantic and phonological networks: Insight from 10 languages. *Cognitive Science*, *44*(7), e12847.

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

Gampe, A., Quick, A. E., & Daum, M. M. (2021). Does linguistic similarity affect early simultaneous bilingual language acquisition? *Journal of Language Contact*, *13*(3), 482–500.

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

Garcia-Castro, G., Avila-Varela, D., Castillejo, I., & Sebastian-Galles, N. (2023). *Cognate beginnings to bilingual lexical acquisition*.

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.

Groot, A. M. de. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5).

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.

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., Durrant, S., & Floccia, C. (2012). Activation of phonological and semantic codes in toddlers. *Journal of Memory and Language*, *66*(4), 612–622.

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.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16.

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.

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

Mitchell, L., Tsui, R. K., & Byers-Heinlein, K. (2022). *Cognates are advantaged in early bilingual expressive vocabulary development*.

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.

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, *59*(1), 96–121.

Rocha-Hidalgo, J., & Barr, R. (2023). Defining bilingualism in infancy and toddlerhood: A scoping review. *International Journal of Bilingualism*, *27*(3), 253–274. <https://doi.org/10.1177/13670069211069067>

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.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038.

Schelletter, C. (2002). The effect of form similarity on bilingual children’s lexical development. *Bilingualism: Language and Cognition*, *5*(2), 93–107.

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.

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

Siow, S., Gillen, N. A., Lepadatu, I., Avila-Varela, D. S., Garcia-Castro, G., Sebastian-Galles, N., & Plunkett, K. (2022). *The effect of cognates on bilingual infant vocabulary trajectories: A study using bilingual CDIs of english and one additional language*.

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.

Styles, S. J., Plunkett, K., & Duta, M. D. (2015). Infant VEPs reveal neural correlates of implicit naming: Lateralized differences between lexicalized versus name-unknown pictures. *Neuropsychologia*, *77*, 177–184.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166.

Tamási, K., McKean, C., Gafos, A., Fritzsche, T., & Höhle, B. (2017). Pupillometry registers toddlers’ sensitivity to degrees of mispronunciation. *Journal of Experimental Child Psychology*, *153*, 140–148.

Tamási, K., Wewalaarachchi, T. D., Hoehle, B., & Singh, L. (2016). Measuring sensitivity to phonological detail in monolingual and bilingual infants using pupillometry. *Proceedings of the 16th Speech Science and Technology Conference*.

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

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, *15*(4), 426–445.

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.

Tsui, R. K.-Y., Gonzalez-Barrero, A. M., Schott, E., & Byers-Heinlein, K. (2022). Are translation equivalents special? Evidence from simulations and empirical data from bilingual infants. *Cognition*, *225*, 105084.

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.

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.