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 to 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. We compared 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 ( [chair]) 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 ([Avila-Varela et al., 2021](#ref-avila2021longitudinal); [Chow et al., 2017](#ref-chow2017spoken); [Mani & Plunkett, 2011](#ref-mani2011phonological); [Mayor & Plunkett, 2014](#ref-mayor2014infant)). We defined vocabulary size as the amount of words participants were reported to understand in their dominant language by their caregivers. The choice of the dominant language for calculating vocabulary sizes is due to several reasons. First, it allows a more fair comparison between monolinguals (who do not know any language other than their dominant language) and bilinguals (who may know words in a second language). Second, since participants were tested exclusively in their dominant language, their vocabulary size in the dominant language is more likely to be associated to participants’ performance in the task. Third, previous work on word recognition in bilinguals suggests that vocabulary size in the dominant language predicts participants’ performance better than total vocabulary (in which vocabulary sizes in both languages are summed together) ([Marchman et al., 2010](#ref-marchman2010vocabulary)).

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 just said, participants were tested exclusively in English. As in the original studies, we manipulated the phonological relationship between the prime and the target label. We expected 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. In half of the trials the English prime label was phonologically related to its Spanish translation, that it they were cognates; in the other half they were not phonologically related (non-cognates)[[1]](#footnote-25). Given participants’ lack of knowledge of Spanish (or any language other than English), participants’ performance was predicted to be unaffected by the cognate status of the primes.

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* |
| *Test 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 | 65 (7) | 46 (6) | 25.2 (3.8) | 19.4-31.7 | 37.6 (10.5) | 62.0 (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 | 329 | 234 | 25.6 | 19.9 | — | — | — |

**?(caption)**

We collected vocabulary data using parental responses to the Oxford Communicative Development Inventory (OCDI) ([Hamilton et al., 2000](#ref-hamilton2000infant)). The OCDI is an adaptation of the MacArthur-Bates Communicative Development Inventory ([Fenson et al., 1994](#ref-fenson1994variability)) to British English. The OCDI includes a vocabulary checklist containing 418 words from 21 semantic-functional categories (e.g., action words, animals, household objects, adverbs, etc.). For each word, caregivers are asked to answer if they child is able to *understand*, *understand and say* or does not understand or say the word. We calculated participants’ receptive vocabulary size scores as the number of words that caregivers marked as *understands* or *understands and says*. 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. |

### Design

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 (). Especial attention was paid to avoiding semantic or taxonomic relationships between prime and target words, and between prime and distractor words. Distractors were always phonologically unrelated to the prime and target labels in the same trial.

[Figure 3](#fig-task) illustrates the sequence of a trial. 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, 700 milliseconds after the onset of the word, 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 lasted 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. |

### Stimuli

We created four 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 (see Appendix A for a detailed description of the stimuli). **?@tbl-stimuli** shows a detailed summary of the stimuli properties, broken down by trial type and testing language. Trials included in each condition had equivalent length (number of phonemes) and lexical frequency. 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)). In addition to sharing phonological onset, prime and target shared higher overall phonological similarity—as measured using the Levenshtein distance—between their phonological transcriptions—in the *Related/Cognate* and the *Related/Non-cognate* conditions, compared to the *Unrelated* condition. In *Related/Cognate* trials, the prime label in the language of test shared a higher overall phonological similarity with its translation, compared to *Related/Non-cognate* trials. Since the number of cognate and non-cognate primes was balanced in *Unrelated* trials, overall phonological similarity between prime labels prime labels in the *Unrelated* condition had an intermediate overall phonological similarity. Audio durations were also equivalent across conditions.

|  | *# 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) | 4.6 (0.5) | 4.7 (0.4) | 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 (0.4) | 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 (0.4) | 0.1 (0.1) | 0.1 (0.1) | 0.9 (0.1) |
| Study 2: Catalan | | | | | | | |
| Cognate | 4.5 (1.3) | 4.9 (1.3) | 5.1 (0.3) | 4.8 (0.3) | 0.5 (0.2) | 0.3 (0.2) | 1.2 (0.2) |
| Non-cognate | 4.9 (1.5) | 5.2 (1.3) | 5.0 (0.4) | 4.8 (0.2) | 0.0 (0.1) | 0.3 (0.1) | 1.3 (0.2) |
| Unrelated | 5.0 (1.5) | 5.0 (1.3) | 4.9 (0.3) | 4.9 (0.3) | 0.4 (0.3) | 0.1 (0.1) | 1.2 (0.2) |
| Study 2: Spanish | | | | | | | |
| Cognate | 4.5 (0.9) | 6.1 (1.6) | 5.1 (0.3) | 4.8 (0.3) | 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) | 4.7 (0.3) | 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) | 4.7 (0.2) | 0.3 (0.3) | 0.2 (0.1) | 1.1 (0.1) |
| Total | *4.97* | *5.12* | *4.89* | *4.76* | *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 normalised 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 height 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

Testing took place in a sound-proof room at the BabyLab of the University of Oxford. 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 23-inches screen with resolution. The study was run on Windows 7 (64-bit), using a custom Matlab script, PresentMate, based on the PsychToolbox-3 extension (3.0.10, 32 bit) ([Brainard & Vision, 1997](#ref-brainard1997psychophysics); [Kleiner et al., 2007](#ref-kleiner2007s); [Pelli & Vision, 1997](#ref-pelli1997videotoolbox)). Visual fixations were recorded using a Tobii TX300 eye-tracker (Tobii Technology, Stockholm, Sweden) 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.

### 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)), as follows:

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 data set, we excluded participants who did not provide at least two valid trials in each condition (*n* = 19), and participants with a vocabulary size lower than 42, which corresponds to 10% of the words in the OCDI vocabulary checklist (*n* = 3). The final data set 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 data set, 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* | *Unrelated* | *Non-cognate* | *Cognate* |
| Study 1: English (Oxford) | | | | |
| Monolingual (English) | 89 (21) | 1,011 (733) | 511 (360) | 510 (359) |
| Study 2: Catalan/Spanish (Barcelona) | | | | |
| Bilingual | 107 (14) | 1,140 (796) | 555 (413) | 596 (372) |
| Monolingual | 133 (15) | 1,506 (862) | 747 (437) | 741 (443) |
| *N* | 329 | 3,657 | 1,813 | 1,847 |

**?(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. Given that the cognate/non-cognate distinction is meaningless for monolinguals, for the following analyses we lumped the *Related/Cognate* and *Related/Non-cognate* into a single level, *Related*. We set an *a priori* contrast for the *Condition* predictor ([Schad et al., 2020](#ref-schad2020capitalize)), comparing *Unrelated* and *Related* trials (sum-coded as -0.5 and +0.5. 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 ( = 0.381, 95% HDI = [0.292, 0.464]) 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 ( = -0.023, 95% HDI = [-0.110, 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 contrast of the *Condition* predictor—comparing *Unrelated* and *Related* trials—included zero ( = 0.093, 95% HDI = [-0.035, 0.235]), and 50.07% of its posterior samples fell within the ROPE.

Second, we examined the differences between the priming conditions in 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* 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. B) Posterior mean prediction of the time course of the differences in target looking time between conditions. Intervals represent the 95% CrI of the posterior predictions. Lines indicate the mean of the posterior predictions. |

### Age and vocabulary size effects

To test our hypotheses regarding the role of age and 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 *Vocabulary* (), and all main effects involved. As with *Age*, the *Vocabulary* predictor was standardised before entering the model. 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 *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 1 (Age-by-condition) | -8,469.02 | 73.01 | --- | --- |
| Model 0 (Age as main effect) | -8,469.15 | 73.02 | -0.13 | 1.11 |
| Model 2 (Vocabulary by-condition) | -8,469.37 | 73.09 | -0.35 | 1.58 |
| \*ELPD\*: sum of expected log pointwise predictive density for a new data set. \*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

We found strong evidence of successful word recognition across participants of all ages, but we did not observe any evidence 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. In conclusion, we failed to replicate the original studies by Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)). 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 ([Avila-Varela et al., 2021](#ref-avila2021longitudinal); [Chow et al., 2017](#ref-chow2017spoken); [Mani & Plunkett, 2011](#ref-mani2011phonological)). Overall, this results suggest that our modification of the implicit naming task resulted in the loss of the originally reported effect.

# Study 2

The original planning of the present investigation was to run Study 1 first and once the procedure had been validated to start Study 2. However, right at the beginning of data collection, the outbreak of COVID-19 pandemic took place. At this point it was decided to run both experiments in parallel. Data collection at the Barcelona site proceeded at a faster rate than at Oxford. It was not until data collection was well advanced in Barcelona that the results of study 1 were available. This is the reason why Study 2 was run with the same procedure as Experiment 1.

## 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 an adaptation of 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 Questionnaire (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)). This questionnaire has four versions, each including a different but overlapping subset of words, from a total pool of 542 words from 26 functional-semantic categories. Each version included a Catalan and a Spanish vocabulary checklist. Catalan checklists contained between 343 and 349 words, and Spanish checklists contained between 349 and 349 words (see the Methods section of Chapter 1 for a detailed description of the questionnaire. Participants were randomly allocated to one of the four versions. Recurrent participants were always allocated to the same version. Families received a link to the BVQ immediately after each experimental session, and were given two weeks to fill it.

Following the same procedure as in Study 1, we calculated participants’ vocabulary size as the number of words that caregivers reported their child to *understand* or *understand and say* in the dominant language of the child (i.e., the language of test). One hundred thirty-six (51%) families failed to provide a complete response to the vocabulary checklist within the two-week time limit. We imputed missing vocabulary size scores using single imputation, taking the vocabulary size scores of a pool of 542 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).

|  |
| --- |
| Figure 5: Participant receptive vocabulary sizes across ages and language profiles. Points and intervals indicate mean and standard deviations, respectively. |

### Stimuli

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. 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. 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 proficient female bilingual speaker of Catalan (Central variety) and Castilian Spanish, 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 (see Appendix A for a detailed description of the stimuli).

### Procedure

Same as in Study 1. We run the study on Windows 10 64-bit, using custom Matlab (2018a 64-bit) script using the PsychToolbox-3 extension (3.0.15 64-bit) to present the stimuli on a 23-inches screen with resolution, and the Tobii Analytics SDK 3.0 to interact with the eye-tracker (Tobii TX300 and Tobii Pro Sprectrum, Tobii Technology, Stockholm, Sweden) 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 data set, we excluded participants who did not provide at least two valid trials in each experimental condition (*n* = 29), and participants with a dominant-language vocabulary size lower than 10% (which depending on the version of the vocabulary questionnaire they were allocated to, varied from 34 and 37) (*n* = 0). The final data set included 5,072 trials from 240 testing sessions, generated by 151 distinct participants. Of those participants, 81 provided data from one experimental session, 51 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 interaction with the *Condition* predictor. 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). 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 ( = 0.215, 95% HDI = [0.176, 0.256]), 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 ( = 0.025, 95% HDI = [-0.013, 0.064]), indicating that participants from all ages showed equivalent overall target word recognition. The 95% HDI of the coefficient of *Group* also included zero ( = -0.014, 95% HDI = [-0.100, 0.066]) 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 = [-0.033, 0.143]), and 75.73% 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 ( = -0.011, 95% HDI = [-0.121, 0.092]), and 90.16% 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 ( = -0.013, 95% HDI = [-0.182, 0.157]), with 59.00% of its posterior samples overlapping with the ROPE. The interaction term between the second *Condition* contrast also contained zero ( = 0.093, 95% HDI = [-0.106, 0.301]), and 75.73% 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.

An analysis of the time course of target looking revealed a similar pattern of results (see [Figure 6](#fig-epreds)). Posterior mean prediction for the three conditions overlap across the full time course of the trial in both language groups.

|  |
| --- |
| Figure 6: 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. Intervals represent the 95% CrI of the posterior predictions. Lines indicate the mean of the posterior predictions. |

### 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 0 | -20,864.87 | 136.77 | --- | --- |
| Model 1 (Age-by-condition) | -20,865.66 | 136.72 | -0.79 | 1.97 |
| Model 0 (Age as main effect) | -20,866.36 | 136.79 | -1.49 | 1.18 |
| \*ELPD\*: sum of expected log pointwise predictive density for a new data set. \*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

Paralleling the results from Study 1 participant’ looking behaviour suggested robust word recognition, regardless of experimental condition, participant language profile, age, or vocabulary size. Monolinguals and bilinguals showed equivalent target looking behaviour in *Unrelated*, *Related/Non-cognate*, and *Related/Cognate* trials, suggesting that no phonological priming took place, either within languages or across languages. These results contrast with those of previous studies using a similar paradigm, which reported within-language priming effects in same-aged monolinguals ([Avila-Varela et al., 2021](#ref-avila2021longitudinal); [Mani & Plunkett, 2011](#ref-mani2011phonological)) and younger ([Duta et al., 2012](#ref-duta2012erp); [Mani & Plunkett, 2010](#ref-mani2010infant); [Styles et al., 2015](#ref-styles2015infant)), and cross-language priming in adults ([Von Holzen & Mani, 2014](#ref-von2014bilinguals)).

We anticipated participants’ sensitivity to phonological priming to increase with the size of their lexicon, in the light of previous studies in which the maturation of the lexicon was associated with larger phonological interference in word recognition ([Chow et al., 2017](#ref-chow2017spoken); [Mani & Plunkett, 2011](#ref-mani2011phonological)). In Study 2, incorporating participants’ age as a predictor in the model in interaction with the two contrasts of the *Condition* predictor did not increase the predictive performance of the model. Neither did vocabulary size. This suggests that the lack of evidence of phonological priming in participants in this study, either within or across languages, did not depend of participants lexical development status.

# General discussion

In this chapter, 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, indicating no 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. We did not find any effect of participants’ age or vocabulary size.

The null pattern of results in Studies 1 and 2 contrasts with previous findings of within- and cross-language priming using an implicit naming paradigm. In their seminal study, Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) reported within-language priming effects in English monolingual infants. In bilinguals, evidence of cross-linguistic priming in a implicit naming paradigm was available in adults ([Von Holzen & Mani, 2014](#ref-von2014bilinguals)). The priming effects shown in these studies reveal that infants and adults retrieve phonologically detailed word-forms when presented with pictures of familiar objects, which later interact with the subsequent auditory recognition of phonologically related words. Evidence of such implicit naming is available in infants as young as 14 months. Electrophysiological evidence reported by Duta et al. ([2012](#ref-duta2012erp)) and Styles et al. ([2015](#ref-styles2015infant)) suggests that, at this, age, infants lexicalise name-known pictures presented in silence, and that the generated phonological form is sensitive to subsequent mispronunciations of the word. The possibility that infants in the present investigation failed to retrieve phonological word forms is therefore unlikely.

We consider three scenarios under which implicit naming might have occurred in the experiments presented in this chapter, 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 ([Bailey & Plunkett, 2002](#ref-bailey2002phonological); [Swingley & Aslin, 2000](#ref-swingley2000spoken); [Tamási et al., 2017](#ref-tamasi2017pupillometry)) and bilinguals ([Ramon-Casas et al., 2009](#ref-ramon2009vowel); [Tamási et al., 2016](#ref-tamasi2016measuring)) have been shown to encode lexical representations with high phonological detail from early ages.

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, and most likely, 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. Such failure to generate phonological word-forms for prime labels would have prevented participants in Studies 1 and 2 from being affected by phonological priming effects during target word recognition.

|  |
| --- |
| Figure 7: Distribution of lexical frequencies and word familiarity at 18 months for the stimuli in Mani and Plunkett (2010, 2011), and in Study 1. A) Lexical frequencies (expressed as Zipf scores) for prime and target words. Lexical frequences are shown for related and unrelated primes separately. B) Word familiarity scores for prime and target words, calculated as the proportion of participants in the OCDI norms that were reported by their caregivers to understand the word. Word familiarity scores are shown for related and unrelated primes separately. |

A difference in the difficulty of the stimuli might have influenced the results in the present study, compared to those of the original studies. When designing the stimuli lists, we considered three variables as indices of word difficult during recognition: lexical frequency, age of acquisition, and number of phonemes. The distribution of the three variables was equivalent across the three experimental conditions (see **?@tbl-stimuli**), so it is unlikely that such differences cancelled out a possible priming effect. However, the stricter limitations under which we build the stimuli lists, might have lead to out stimuli lists including more difficult words than in the original study by Mani & Plunkett ([2010](#ref-mani2010infant)). We examined this possibility by comparing the distributions of the English words included in Study 1, with those included in Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) (stimuli lists are identical in both studies). [Figure 7](#fig-mp) shows a comparison between the stimuli lists of the three studies in lexical frequency and word familiarity at 18 months. Overall, words included in Mani and Plunkett and in Study 1 have equivalent lexical frequency, and are know by a similar proportion of English monolinguals infants, according to the OCDI norms. This is the case for prime and target words across the related and the unrelated conditions. It is therefore unlikely that the lack of priming effects in Study 1 is due to an increased difficulty in the items included.

Another possibility is that participants in the present study had smaller vocabulary sizes than those of participants in the original studies. This hypothesis is not easy to investigate, as quantitative vocabulary sizes were not reported in original studies. A more recent study by Avila-Varela et al. ([2021](#ref-avila2021longitudinal)), in which phonological priming effects were found associated to participants vocabulary size, did provide summary statistics for participants vocabulary size scores. This study tested a cohort of monolingual German infants in a word recognition task, in which participants in which participants were presented with auditory primes and targets, which were phonologically related or unrelated. The authors estimated participants’ receptive vocabulary sizes using the *Fragebogen zur frühkindlichen Sprachentwicklung* [szagun2009fragebogen], an adaptation of the CDI to German. Their cohort of participants showed receptive vocabulary sizes larger than those of participants in the present study. Participants in Avila-Varela et al. ([2021](#ref-avila2021longitudinal)) knew an average of 405.24 (*SD* = 96.29) at 21 months and 501.97 (*SD* = 73.41) at 24 months, which contrast with receptive vocabulary sizes of participants in Study 1: 293 at 21 months (*SD* = 71.24), and 304.33 (*SD* = 134.63) at 25 months. Estimates from both questionnaires should be compared with caution—if at all—given their different lengths. It possible that the apparent smaller vocabulary size of English participants in Study 1 is a measurement artefact, resulting from a lower upper bound in the possible number of words that can be marked acquired by caregivers.

Despite the unsuccessful adaptation of the original procedure, the present study introduces several methodological contributions. First 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)). Second, 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 aimed to test the language non-selective hypothesis of lexical access in bilingual toddlers using an adaptation of the implicit naming paradigm. This adaptation involved target auditory labels immediately after the offset of prime pictures, instead of presenting the target labels after a baseline period of 2,000 after the offset of the prime pictures. In Study 1, we tested English monolinguals (same population as in the original studies) to establish a baseline to later test bilingual participants. We attempted to replicate the previously reported within-language phonological priming effect. We did not find evidence of such effect, suggesting that our modification of the original task was unsuccessful. Because data collection was conducted simultaneously for Studies 1 and 2, data in Catalan-Spanish monolinguals and bilinguals was available despite the failed replication in Study 1. In Study 2, we also found null pattern of results, in which neither monolinguals nor bilinguals showed evidence of within- or cross-language priming effects. Overall, our results suggest that the change in the timing of the trial disrupted the dynamics of word recognition in such way that priming effects were no longer detactable in our adaptation of the paradigm.

# 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.

Avila-Varela, D. S., Arias-Trejo, N., & Mani, N. (2021). A longitudinal study of the role of vocabulary size in priming effects in early childhood. *Journal of Experimental Child Psychology*, *205*, 105071.

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.

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young spanish–english bilinguals. *Journal of Child Language*, *37*(4), 817–840.

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.

Mayor, J., & Plunkett, K. (2014). Infant word recognition: Insights from TRACE simulations. *Journal of Memory and Language*, *71*(1), 89–123.

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.

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.

1. It was initially planned to collected data from English monolinguals and English-Spanish bilinguals in Oxford, therefore the manipulation of the cognate status of the words in English and Spanish. Due to time limitations imposed by the COVD-19 lockdown between 2020 and 2022, collecting data from bilinguals was not possible. We report the available data from English monolinguals as a control for Catalan and Spanish monolinguals and bilinguals from Barcelona in Study 2. [↑](#footnote-ref-25)