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 relative impoverished quantitative 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, as they often learn two labels for each referent, one in each language ([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 cross-language dynamics as a candidate mechanism boosting lexical acquisition in bilinguals ([Blom et al., 2020](#ref-blom2020cross); [Floccia et al., 2018](#ref-floccia2018vocabulary); [Gampe et al., 2021](#ref-gampe2021does)). In particular, lexical similarity between the two languages has been associated with larger vocabulary sizes in bilingual toddlers. Floccia et al. ([2018](#ref-floccia2018vocabulary)) 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 used an edit distance-based metric of phonological similarity to compute the average lexical similarity between the translation pairs in English and each of the additional languages. English and Dutch shared the highest lexical similarity, while English and Mandarin shared the lowest. Children’s vocabulary sizes in the additional language was positively associated with the amount of language similarity between their two languages. For instance, children learning English and Dutch showed larger vocabulary sizes than those learning English and 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)).

The facilitative 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, as suggested by robust experimental evidence ([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)). For instance Costa et al. ([2000](#ref-costa2000cognate)) presented Catalan-Spanish bilinguals with a series of pictures of familiar objects. Participants were asked to name each picture in Catalan, their dominant language. 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 (i.e., phonologically similar translation equivalents, e.g., *gat*-*gat* [*cat*]) in Catalan and Spanish, whereas in the other half of the trials the labels were non-cognates (e.g., *taula*-*mesa* [*table*]). Participants named pictures faster in cognate trials than in non-cognate trials. Critically, Spanish monolinguals—who completed the same task in Spanish—showed equivalent naming times in both conditions. These results revealed that bilinguals activated their Spanish phonology, despite performing the naming task exclusively in Catalan: 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)). For instance, Von Holzen & Mani ([2012](#ref-von2012language)) found suggestive 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 German label and a pair of target and distractor pictures. The authors registered participants’ target picture looking (interpreted as a proxy of target word recognition) under three conditions. In the *phonological priming* condition, the auditorily presented prime and target labels were phonologically related (e.g., *slide*-*Klide* [*dress*]). In the *priming through translation* condition, prime and target were phonologically related through translation: the auditorily presented prime label did *not* overlap with the target label (e.g., *leg*-*Stein* [*stone*]), but its German translation did [*Bein*]. In the *unrelated* condition, prime and target labels were phonologically unrelated in both languages. Interestingly, participants showed a weaker target preference in priming through translation trials, compared to unrelated trials. This was interpreted evidence of the lexical representations of the prime labels being activated in both English and German, leading to interference effects in later target recognition. Previous studies had reported similar cross-language priming effects, but the conclusion that such effects occurred at the lexical level, as opposed to being the result of mere acoustic similarity, was nonetheless unwarranted. In studies where the auditorily presented prime label was phonologically overlapped with the target label (e.g., [Marian & Spivey, 1999](#ref-marian1999activation)), priming effects could be explained by sub-lexical facilitation across phonological segments. This might explain the facilitation effect found in *phonological priming* trials in Von Holzen & Mani ([2012](#ref-von2012language)), in which English prime and German target labels shared overt phonological overlap. In contrast, the interference effect found in *priming through translation* trials points to cross-language competition at the lexical level. Since participants’ target recognition was modulated by the phonology of the German translation of the prime (which participants did *not* hear), cross-language priming can be explained by the lexical co-activation of the English *and* German prime labels.

Von Holzen & Mani ([2012](#ref-von2012language)) extended the available knowledge on the developing bilingual lexicon, but some methodological challenges prevent drawing strong conclusions from their findings. First, bilingual participants’ quantitative exposure to English as a second language was lower than what is conventionally considered as bilingualism (20-50% of cumulative exposure, [Byers-Heinlein et al., 2021](#ref-byers-heinlein2021multilab); [Rocha-Hidalgo & Barr, 2023](#ref-rocha-hidalgo2023defining))[[1]](#footnote-20). Results should therefore be taken with caution when generalised to other populations of bilinguals. A second consideration is the presence of cognates in a substantial proportion of the trials in the *phonological priming* and *priming thorugh translation* conditions. Because of this, some prime labels in *phonological priming* trials shared substantial cross-language phonological overlap thought translation with the target words. Conversely, some trials in the *phonological priming through translation* shared within-language overlap with the target. Third, the observed pattern of results was dependent on the inclusion of a single item, *Buch* [*book*, in German]. When this item was included, monolinguals showed a similar interference effect to bilinguals’ in the *priming through translation* condition, in which only bilinguals with knowledge of the translation should show an effect. Fourth, data from monolinguals and bilinguals were not compared directly (analyses were conducted separately for both groups). Differences in statistical significance between monolinguals and bilinguals across conditions do not necessarily guarantee significant differences between the two groups. As a consequence, whether the interference effects observed in the *priming through translation* condition is exclusive to bilinguals (which would suggest that the priming effect were not due to the intrinsic properties of the stimuli) or not is unclear. In summary, evidence of cross-language activation at the lexical level in infants is, to date, inconclusive.

In the present study, we use an extension of the implicit naming task ([Mani & Plunkett, 2010](#ref-mani2010infant), [2011](#ref-mani2011phonological)) to put the language non-selectivity in the developing lexicon to a more severe test. Previous studies have successfully used this task to investigate priming effects in word recognition in monolingual infants and bilingual adults. In the original implementation of the task, Mani & Plunkett ([2010](#ref-mani2010infant)) primed infants with a picture presented in silence. At test, a target-distractor picture pair were presented side-by-side. Finally, the target picture’s label was presented auditorily. 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 and target-distractor pairs were semantically unrelated. Eighteen-month-old infants showed a stronger looking preference for the target picture after phonologically related prime pictures, compared to after phonologically unrelated primes. This suggests that infants implicitly named prime pictures despite such pictures having been presented in silence, and that the phonology of the resulting label interacted with the subsequent auditory recognition of the target label.

Von Holzen & Mani ([2014](#ref-von2014bilinguals)) adapted this paradigm to investigate cross-language activation in bilingual adults. The authors presented German-English bilinguals with 120 prime-target pairs. Primes were presented as familiar pictures in silence. After prime picture offset, target words were presented auditorily. Participants’ eletrophysiological response was recorded, and N400 evoked potentials to the auditory target label was measured. The authors manipulated the phonological relationship between the prime and target word-forms within participants’ L1 and across L1 and L2. When prime and target were phonologically identical (*Affe*-*Affe*, German for *monkey*), or similar (*Fahne*-*Sahne*, German for *flag* and *ice-cream*, respectively) within L1, participants’ showed a reduced N400 amplitude, compared to when prime and target were phonologically unrelated (*Messer*-*Seil*, German for *knife* and *rope*). Critically, a similar effect was found when prime and target were phonologically related through translation (*Rustsche*-*Kleid*, German for *slide* and *dress*). This suggests that participants activated prime labels in both languages in parallel, and that both labels impacted the dynamics of target words recognition. In summary, the implicit naming paradigm has been successful used in the past to investigate the emergence of phonological priming effects in monolingual infants, and to investigate cross-language priming effects in bilingual adults.

The implicit naming paradigm provides several advantages over the priming through translation task. In the priming through translation task, primes and targets are auditorily presented in both languages in the same trial. This presents participants with a bilingual context that might lead to a higher baseline activation of lexical representations of both languages, contrary to a context in which parallel activation results from genuine cross-language activation ([Grosjean, 1997](#ref-grosjean1997bilingual)). Also, when prime words are auditorily presented, the task imposes an asymmetry between the activation dynamics of the lexical representation of the prime in each language: while participants hear prime labels in one language, the activation of prime labels in the other must occur via cascaded activation across semantic links with their translations. This might lead to a lower strength, and higher latency of the activation of prime labels in one language, compared to the presented labels in the other. In contrast, the implicit naming paradigm provides participants with an exclusively monolingual context in which they perform the task in only one of their languages. At the same time, the silent presentation of prime pictures allows participants to activate prime labels in both languages in a truly parallel fashion. If differences in task performance are found associated with covert manipulation of cross-linguistic properties of prime labels, such differences may be more conclusively linked to the parallel activation of lexical representations in both languages.

In the present investigation, we manipulated both the phonological overlap between the prime word and target word and the cognate status of the prime word, to test the whether bilingual toddlers implicitly name familiar pictures in both languages. If this is the case, phonological priming effects should be larger when the prime label in both languages shares phonological overlap with the target words (*Related/Cognate* condition), compared to when only the label of the prime in one language overlaps phonologically with the target (*Related/Non-cognate*). We also also included a control *Unrelated* condition in which regardless of the cognate status of the prime word, neither of the prime labels in any language overlapped phonologically with the target word. We registered participants’ target looking during target-distractor presentation, and that, if within-language priming effects take place, differences in target looking between *Unrelated* and *Related/Non-cognate* trials should be observed, as previously reported Mani & Plunkett ([2010](#ref-mani2010infant)). We further predicted cross-language priming effects, revealed as differences in target looking between *Related/Non-cognate* and *Related/Cognate* trials. This effect should be stronger in bilinguals than monolinguals, if not completely absent in the latter, who should be indifferent to the phonological properties of the stimuli in the other language.

Following previous studies associating the emergence of phonological links between lexical items to the size of children’s vocabulary ([Chow et al., 2017](#ref-chow2017spoken); [Mani & Plunkett, 2011](#ref-mani2011phonological)), we also compared the explanatory power of two indices of receptive vocabulary size against participants’ age: dominant-language vocabulary size and total vocabulary size. If within-language or cross-language phonological associations in the lexicon emerge as a function of vocabulary growth, the size of the differences between conditions should larger in participants with larger vocabulary sizes. A model including participants’ dominant-language or total vocabulary size as a predictor should then show better performance than a model that instead includes participants age. In the case of bilinguals, if the emergence of cross-language priming effects is associated with the growth of their vocabulary is *both* languages, a model including total vocabulary size as a predictor should perform better than a model that includes dominant-language vocabulary size. The opposite pattern would indicate that cross-language priming effects emerge as a function of the growth of dominant-language vocabulary.

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

# Study 1

## Methods

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

### Participants

We collected data from 162 children living in the Metropolitan Area of Barcelona (Spain), tested at the Laboratori de Recerca en Infància at the Universitat Pompeu Fabra. Families were recruited from maternity rooms in private hospitals and social media, and contacted via phone when the child’s age spanned between 20 and 32 months. From the 162 children that participated, 81 participated once, 55 participated twice, and 26 participated three times. Recurrent participants were tested with at least 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 79 as Catalan/Spanish bilinguals (34 female, 45 male) (see **?@tbl-participants** for a detailed summary of participants’ age and language profile). Participants’ vision was normal, none used glasses or any other type of vision corrector.

|  |  |  |  |  | *Degree of Exposure (%)* | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *Age (months)* | | *Spanish* | *Catalan* | *English* |
| *Sessions* | *Participants* | *M (SD)* | *Range* | *M (SD)* | *M (SD)* | *M (SD)* |
| Monolingual | | | | | | | |
| Catalan | 96 | 53 | 25.8 (3.9) | 20.0-32.3 | 4.5 (6.0) | 95.2 (6.2) | 0.3 (2.1) |
| Spanish | 53 | 30 | 25.4 (3.8) | 20.0-32.0 | 91.2 (6.5) | 8.4 (6.6) | 0.3 (0.9) |
| Total | 149 | 83 | 25.6 | 20.0 | 47.8 | 51.8 | 0.3 |
| Bilingual | | | | | | | |
| Catalan | 71 | 50 | 25.2 (3.8) | 19.4-31.7 | 37.7 (10.3) | 61.5 (10.6) | 0.2 (0.8) |
| Spanish | 49 | 33 | 25.1 (3.4) | 20.2-31.3 | 61.0 (11.1) | 38.7 (10.6) | 0.3 (1.4) |
| Total | 120 | 83 | 25.2 | 19.8 | 49.4 | 50.1 | 0.2 |
| Total | 269 | 166 | 25.4 | 19.9 | 48.6 | 51.0 | 0.3 |

**?(caption)**

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).

Overall, there was a substantial increase in both total and L1 vocabulary sizes associated to age (see [Figure 1](#fig-vocabulary)). Monolinguals showed higher L1 vocabulary sizes than monolinguals: they were reported to understand an average of 56.9% (*SD* = 17.736%) of the words at around 21 months, 75.9% (*SD* = 14.790%) at 25 months, and 82.7% (*SD* = 14.839%) at 30 months, while bilinguals were reported to understand an average of 49.1% (*SD* = 22.511%) of the words at around 21 months, 68.0% (*SD* = 20.326%) at 25 months, and 79.0% (*SD* = 13.972%) at 30 months. Total vocabulary sizes showed the opposite pattern, with monolinguals infants being reported to understand 40.5% (*SD* = 15.99%) of the words at around 21 months, 59.8% (*SD* = 19.62%) at 25 months, and 63.0% (*SD* = 21.07%) at 30 months, and bilinguals being reported to understand 47.9% (*SD* = 23.26%) of the words at around 21 months, 64.1% (*SD* = 17.15%) at 25 months, and 75.6% (*SD* = 17.34%) at 30 months.

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

### Stimuli

We used 94 distinct words included in the BVQ to create the stimuli lists. We created six stimuli lists: three in Catalan, and three in Spanish. Each list contained 32 trials, each involving a prime-target-distractor triplet. Words played a role as primes *or* targets/distractors across the three lists in their corresponding language. For instance, the Catalan word *cadira* appeared as *prime* in the three lists, but never as *target* or *distractor*; the Catalan word *bici* appeared as *target* and *distractor* across the three lists, but never as a prime. Target-distractor pairings were held constant across the three lists in each language. For instance, in all Catalan lists the word *bici* was paired with the word *porta*. Target-distractor pairings were also yoked, so that each member of the same target-distractor pair appeared once as target and once as distractor in each list. For instance, the *bici*-*porta* paired appeared twice in each of the three Catalan lists: once with *bici* as target and *porta* as distractor, and once with *porta* as target and *bici* as distractor. This counterbalancing avoided participants encountering looking at the target word guided solely by that word having being named in a previous trial. Finally, prime words appeared only once in each list: each target-distractor pair was associated with a different prime word in both appearances. Across lists, the same prime word was presented alongside a different target-distractor pair. For instance, the Catalan prime word *barret* was presented with the *bici*-*porta* target-distractor pair in one list, with the *bici*-*porta* pair in another list, and with *berenar*-*amanida* in the remaining list. To avoid competition between target and distractor pictures, semantically related target-distractor pairs were perceptually distinct ([Arias-Trejo & Plunkett, 2009](#ref-arias2009lexical); [Floccia et al., 2020](#ref-floccia2020translation)). The order of the trials was randomised across experimental sessions, so that each time a participant was tested, the order in which the prime-target-distractor was presented was random. Each participant was randomly assigned to one of the three lists in the corresponding (dominant) language, and always the same list across their experimental sessions in the case of a recurrent participant.

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

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

The auditory stimuli were natural exemplars of the selected target words, spoken by Central Catalan-Spanish proficient bilingual female speaker who was instructed to pronounce each word in a toddler-directed manner. Recordings were made with an Audio-Tecnica 328 microphone (AT2050) at a sampling rate of 44,100 Hz, in a soundproof room at the *Laboratori de Recerca en Infancia* at University Pompeu Fabra. We used the Audacity and Praat ([Boersma & Van Heuven, 2001](#ref-boersma2001speak)) software packages to record and edit the audio files. The speaker was presented with a list of words in Catalan. The order of the words was pseudo-randomised, and each word was produced three times in a row before moving to the next word in the list. After going through all the words in the list, the speaker went through the word list again generating three tokens for each word, now in an inverse order (from bottom of the list to the top). We then repeated the same procedure for the list of Spanish words. The resulting audios were manually chunked into individual word-forms. For each of the six tokens produced for each word, the most adequate was selected for further processing. The audios were then transformed to stereo by duplicating them into two channels, denoised, and finally normalised. The mean duration of the final audios was 1.23 (*SD* = 0.17) and 1.08 (*SD* = 0.13) seconds for the Catalan and Spanish lists.

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

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

|  | *Phonemes* | | *Frequency (Zipf)* | | *Phon. similarity [0-1]* | |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Prime* | *Target* | *Prime* | *Target* | *Prime-Prime* | *Prime-Target* | *Audio duration (s)* |
| Cognate | 4.5 (1.1) | 5.5 (1.5) | 4.9 (0.4) | 4.4 (0.8) | 0.5 (0.2) | 0.3 (0.1) | 1.1 (0.2) |
| Non-cognate | 5.1 (1.4) | 5.5 (1.5) | 4.9 (0.4) | 4.3 (0.8) | 0.1 (0.1) | 0.3 (0.1) | 1.2 (0.2) |
| Unrelated | 4.8 (1.3) | 5.4 (1.5) | 4.8 (0.5) | 4.2 (0.7) | 0.3 (0.3) | 0.1 (0.1) | 1.2 (0.2) |
| Total | *4.79* | *5.47* | *4.88* | *4.27* | *0.30* | *0.24* | *1.16* |

**?(caption)**

### Procedure

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

|  |
| --- |
| Experimental task design with examples in Catalan. In each trial, the prime image is presented in silence for 1,500 ms. 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. |

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.

### 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).

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) (see [Floccia et al., 2020](#ref-floccia2020translation); [Mani et al., 2012](#ref-mani2012activation) for a similar approach).

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.

|  |  | Trials | | |
| --- | --- | --- | --- | --- |
| *Participants* | *Cognate* | *Non-cognate* | *Unrelated* |
| Bilingual | 103 (17) | 589 (371) | 548 (412) | 1,127 (793) |
| Monolingual | 134 (15) | 748 (444) | 754 (438) | 1,519 (865) |
| *N* | 237 | 1,337 | 1,302 | 2,646 |

**?(caption)**

#### Modelling approach

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

We included 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 the models using brms ([Bürkner, 2017](#ref-burkner2017brms)), an R interface to the Stan probabilistic language (2.33.0) ([Carpenter et al., 2017](#ref-carpenter2017stan)). We ran four iteration chains using the by-default No U-Turn Sampler algorithm with 500 iterations each and an additional 500 warm-up iterations per chain.

#### Statistical inference

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. Second, we examined the differences between the priming conditions across the time course of the trial. We generated marginal posterior predictions from the model for the three conditions, for monolinguals and monolinguals. These predictions included the non-linear relationship between target looking and the time course of the trial. We then computed the difference between the marginal posterior predictions for the *Unrelated* and *Related/Non-cognate* conditions, and between those for the *Related/Non-cognate* and *Related/Cognate* conditions, at each time bin. Finally, we compared the 95% HDI of the differences against the [-0.1, +0.1] ROPE.

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 three-way interaction between *Condition*, *Group*, and *Age* (), *L1 vocabulary* (), and *Total vocabulary* (), and all two-way interactions 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*, *L1 vocabulary*, or *Total vocabulary*, respectively, playing a substantial role in participants’ word-recognition, or on the emergence of priming effects.

## 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 ( = 0.218, 95% HDI = [0.178, 0.258]), 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 = [-0.019, 0.057]), 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 = [-0.097, 0.069]) 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 ( = 0.059, 95% HDI = [-0.025, 0.149]), and 71.79% 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 ( = -0.017, 95% HDI = [-0.134, 0.086]), and 84.68% 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 ( = -0.010, 95% HDI = [-0.183, 0.158]), with 58.79% of its posterior samples overlapping with the ROPE. The interaction term between the second *Condition* contrast also contained zero ( = 0.083, 95% HDI = [-0.104, 0.311]), and 71.79% 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 2](#fig-epreds)). Posterior mean prediction for the three conditions overlap across the full time course of the trial in both language groups, and also in both extrema of the age range of participants (20 and 32 months).

|  |
| --- |
| Figure 2: 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). |

### The role of age and vocabulary size on priming 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) | -20,591.28 | 135.85 | --- | --- |
| Model 0 | -20,593.32 | 135.88 | -2.05 | 2.11 |
| Model 3 (Total vocabulary) | -20,593.71 | 135.92 | -2.44 | 2.20 |
| Model 1 (Age) | -20,594.59 | 135.89 | -3.31 | 2.11 |
| \*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)**

### Word recognition times

In an exploratory analysis, we compared the model-estimated time points at which participants’ gaze behaviour showed evidence of target looking above chance level. We defined the recognition time of the trial as the earliest time point at which the lower bound of the 95% HDI posterior mean prediction of the target looking excluded zero. As expected, the recognition times were on average shorter for older infants than for younger infants. Differences between conditions were larger in bilinguals than in monolinguals, especially in the older group. The recognition time for monolingual 20-month-olds was 687.88 ms in the *Unrelated* condition, 655.56 ms in the *Related/Non-cognate* condition, and 639.39 ms in the *Related/Cognate* condition. In monolingual 30 month-olds, the recognition point was 639.39 ms in the *Unrelated* condition, 590.91 ms in the *Related/Non-cognate* condition, and 590.91 ms in the *Related/Cognate* condition. In bilingual 20 month-olds, the recognition point was 687.88 ms in the *Unrelated* condition, 655.56 ms in the *Related/Non-cognate* condition, and 639.39 ms in the *Related/Cognate* condition. In bilingual 30 month-olds, the recognition point was 687.88 ms in the *Unrelated* condition, 655.56 ms in the *Related/Non-cognate* condition, and 639.39 ms in the *Related/Cognate* condition.

|  |
| --- |
| Figure 3: Model-predicted recognition times for monlingual and bilinguals at 20 and 30 months of age. |

## Discussion

In Study 1, we adapted an implicit naming paradigm to investigate the potential role of within- and cross-language phonological priming in 20- to 32-month-old children. Both the overall target looking and its time course across the trial suggested robust word recognition between 550 ms and 700 ms after the onset of the test phase, regardless of priming condition, participant language profile and age. Overall, participants’ performance was equivalent across all priming conditions. An exploratory analyses on word recognition times suggested a complex pattern of results, showing an age effect—in which older participants looked at the target above chance earlier than younger infants—and a marginal interaction between priming conditions, age, and language group. Overall, we did not find evidence of phonological priming within or across languages.

These results conflict with previous findings in several ways. First, previous investigations using implicit naming paradigms had found evidence of infants and toddlers generating internal labels when presented with pictures of familiar objects. Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) reported priming effects in 18 and 24-month-old infants. In primed trials, participants’ target looking preference was influenced by the prior presentation of a phonologically related referent in silence. In control trials—in which prime and target labels were unrelated—no priming effect was found. This suggested that infants implicitly named the prime label, and that the resulting internal phonological form was detailed enough to interact with the subsequent auditory recognition of the target word. In the present study, we did not find evidence in favour of differences in target looking preference between related and unrelated trials, regardless of the age of participants or their language profile.

Both Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)), and the present study, share a similar design, but sampled different populations. Participants in Mani and Plunkett’s studies were raised in English speaking households in Oxford (United Kingdom), while participants in the present study were learning Catalan or Spanish, if not a combination of both in Spain. We identify two ways in which this might have influenced the results. One is that the stimuli lists in Catalan and Spanish might have differed in some critical dimension to the English stimuli from Mani and Plunkett. This might have interfered infants’ ability to generate implicit labels for prime pictures. For example, words in the present study were predominantly bi-syllabic, while those from Mani and Plunkett’s studies were mostly monosyllabic, and contained less phonemes on average. Longer word-forms might have been more difficult to retrieve in time for a prime effect to take place in target recognition. Another possibility is that exposure to a second language might have delayed the developmental trajectories of word production in bilinguals, in such way implicit naming emerges later ages. Given that 50.51% of monolinguals in Barcelona were exposed to a second language (although less than 20% of the time), they might as well been affected by such delay in implicit naming trajectories.

The results from the Study 1 also conflict with previous reports of cross-language priming in toddlers. Particularly, Von Holzen & Mani ([2012](#ref-von2012language)) and Floccia et al. ([2020](#ref-floccia2020translation)) had provided evidence of phonological priming through translation. One critical different between both studies and the present one is the silent or auditory presentation of the prime labels. In priming through translation studies, the prime is auditorily presented to participants, and cross-language priming effects take place through the subsequent activation of the prime translation. In the present study, the prime label was not named, and participants were expected to generate both the prime label and its translation implicitly. It may be the case that infants in the present study faced a more challenging task, having to internally retrieve phonological forms without any auditory prompt like the one provided by priming through translation studies. Another possibility is that participants in the present study did in fact general prime labels, but such labels did not interfere over target recognition. This could be because they were not phonologically specified enough. This is unlikely in the light of previous work on implicit naming and mispronunciation perception. First, Duta et al. ([2012](#ref-duta2012erp)) and Styles et al. ([2015](#ref-styles2015infant)) had shown electrophysiological evidence of implicit naming in English learning infants as young as 14 months of age. Second abundant studies have shown that early lexical representations are encoded in high phonological detail, in both monolinguals and bilinguals ([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)). It could be the case then, that participants in the present study retrieved prime phonological representations, but phonological links between lexical representations are still weak or not present, so that priming effects are not possible until later in age.

In summary, the possibly higher difficulty of the task induced by the properties of the stimuli, together with differences in task design and populations might have contributed to the null pattern of results shown in Study 1. Given the complexity of the picture drawn by these factors, we conducted a second study with the aim of replicating the original findings by Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) in English monolinguals. If phonological priming effects are observed in a same-aged population using the adaptation of the task from Study 1, this would provide evidence in favour of cross-language differences in the developmental trajectories of implicit naming. If, on the contrary, a similar patter of null results is found, this would suggest either the existence of finer differences in task implementation compared to the original studies, or point to the effect of interest being less robust than suggested by the literature.

# Study 2

In this study, we conducted a conceptual replication of Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)). Using the same task from Study 1, we tested a group of monolinguals aged 20 to 32 months being exclusively raised in British English speaking homes, the same population tested in the original studies. For consistency with Study 1, we also manipulated the phonological relationship between the prime and target labels at word onset, as well as the cognate status of the prime label. Evidence from implicit naming of prime images would be revealed by differences in target looking or its time course across the trials between *Unrelated* trials and *Related/Non-cognate*, and between *Related/Non-cognate* and *Related/Cognate* trials. Given that participants in this study are monolinguals—and contrary to the monolinguals in Barcelona, have received no exposure to any other language—we should not find the cognate status of the prime labels to exert any effect on subsequent target recognition.

## Methods

### Participants

We collected data from 112 children (41 female, 68 male, with three additional participants’ sex not being reported; Age: *Mean* = 26.36 months, *SD* = 4.01, *Range* = [20.03-32.5]), 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.

We collected vocabulary data using parental responses to the Oxford Communicative Development Inventory (OCDI) ([Hamilton et al., 2000](#ref-hamilton2000infant)). Families were sent the questionnaire immediately after each experimental session, and were given two weeks to fill it. Since all vocabulary sizes of English monolingual infants were only assessed in English, both *L1 vocabulary* and *total vocabulary* sizes were identical (see [Figure 4](#fig-vocabulary-oxf)).

|  |
| --- |
| Figure 4: Participant receptive vocabulary sizes across ages and language profiles. Vocabulary size scores from the recurrent participants are shown linked. |

### Stimuli

New stimuli lists were generated to adapt the experimental procedure to English-learning infants. Stimuli lists were created following the same constraints as in Study 1. Auditory stimuli were recorded following the same procedure as in Study 1, produced by Southern British English native female speaker in a toddler-directed manner. Audio and picture processing were performed following the same workflow as in Study 1. **?@tbl-stimuli-oxf** shows a detailed summary of the properties of the English stimuli in Study 2.

|  | *Phonemes* | | *Lexical frequency (Zipf)* | | *Phonological similarity [0-1]* | |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Prime* | *Target* | *Prime* | *Target* | *Prime-Prime* | *Prime-Target* | *Audio duration (s)* |
| 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) |
| Total | *5.33* | *4.44* | *4.64* | *4.69* | *0.09* | *0.21* | *0.86* |

**?(caption)**

### Procedure and apparatus

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

### Data analysis

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

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).

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* = 18), and participants with a L1 vocabulary size lower than 10% (*n* = 2). The final dataset included 1,861 trials from NA testing sessions, generated by 80 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).

## Results

### Priming effects

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 ( = 0.399, 95% HDI = [0.308, 0.490]) 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 ( = -0.020, 95% HDI = [-0.097, 0.059]), 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 ( = 0.118, 95% HDI = [-0.050, 0.282]), and 45.26% 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 ( = 0.126, 95% HDI = [-0.065, 0.297]), and 45.50% of its posterior samples fell within the ROPE. The overall target preference was equivalent across both pairwise condition comparisons.

**?(caption)**

An analysis of the time course of target looking revealed a similar pattern of results (see [Figure 5](#fig-epreds-oxf)). Posterior mean prediction for the three conditions overlap across the full time course of the trial in both language groups, and also in both extrema of the age range of participants (20 and 32 months).

|  |
| --- |
| Figure 5: 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. |

### The role of age and vocabulary size on priming effects

**?@tbl-loos-oxf** shows a comparison between models including *Age* (), and *L1 vocabulary* ()—and their two-way and two-way interaction with *Condition*—against model , which only included *Age* as a covariate suggested that all models performed equivalently, with model showing a slightly better performance. Overall, this 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.

|  | *ELPD* | *SE ELPD* | *ELPD (diff.)* | *SE ELPD (diff)* |
| --- | --- | --- | --- | --- |
| Model 2 (L1 vocabulary) | -8,431.72 | 73.85 | --- | --- |
| Model 0 | -8,431.76 | 73.79 | -0.04 | 2.36 |
| Model 1 (Age) | -8,432.66 | 73.73 | -0.94 | 1.91 |
| \*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)**

### Word recognition times

As expected, the recognition times were on average shorter for older infants than for younger infants. Differences between conditions were larger in bilinguals than in monolinguals, especially in the older group. The recognition time for monolingual 20-month-olds was 526.22 ms in the *Unrelated* condition, 493.90 ms in the *Related/Non-cognate* condition, and 396.92 ms in the *Related/Cognate* condition. In monolingual 30 month-olds, the recognition point was 526.22 ms in the *Unrelated* condition, 558.55 ms in the *Related/Non-cognate* condition, and 542.39 ms in the *Related/Cognate* condition. In bilingual 20 month-olds, the recognition point was 526.22 ms in the *Unrelated* condition, 493.90 ms in the *Related/Non-cognate* condition, and 396.92 ms in the *Related/Cognate* condition. In bilingual 30 month-olds, the recognition point was 526.22 ms in the *Unrelated* condition, 493.90 ms in the *Related/Non-cognate* condition, and 396.92 ms in the *Related/Cognate* condition.

|  |
| --- |
| Figure 6: Model-predicted recognition times for monlingual and bilinguals at 20 and 30 months of age. |

## Discussion

Study 2 was aimed at investigating whether the phonological priming effects previously reported Mani & Plunkett ([2010](#ref-mani2010infant)) and Mani & Plunkett ([2011](#ref-mani2011phonological)) could be found using the adaptation of the implicit naming task form Study 1 in a sample with equivalent demographic characteristics to the original studies. Results from Study 2 did not provide any evidence in favour of phonological priming. English monolingual participants from all ages showed an equivalent pattern of target looking across both related 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)).

One of the main differences between the implicit naming paradigm used by Mani & Plunkett ([2010](#ref-mani2010infant)) and the one in Studies 1 and 2 is the fact that target labels were auditorily presented before target and distractor pictures were presented on screen. The rationale behind this decision was to minimise the amount of time elapsed between the presentation of the prime and target labels, in order to facilitate any potential priming effects. By presenting the auditory target label immediately after the offset of the prime picture, the internally generated prime phonological form would be more likely to affect the activation of the target lexical representation. It is possible that this difference in task design might have led to the contrasting outcomes of Mani and Plunkett and the present study. Given that participants were not presented with the target and distractor pictures until after the target label was auditorily presented, one possibility is that participants faced the additional task of exploring both pictures to find the target referent, which infants in Mani & Plunkett ([2010](#ref-mani2010infant)) did not have to do, because of the pre-naming presentation of target and distractor in the test phase. This visual exploration of the screen might have increased the cognitive load of the task, rendering the potential effect of phonological priming to a weaker, undetectable effect. Overall, the results from Study 2 support a more methodological explanation of the null pattern of results in Study 1, in which the lack of a phonological priming effect—both within and across languages—might be explained by differences in task design or replicability concerns.

# General discussion

In this paper, we investigated the developmental trajectories of cross-language co-activation in the developing lexicon. We tested a large cohort of monolingual and bilingual toddlers in an implicit naming paradigm. 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 tested a cohort of monolingual and bilingual infants learning Catalan, Spanish, or both, and found no phonological priming effect in monolinguals or bilinguals. In Study 2, we attempted to replicate the original findings by Mani & Plunkett ([2010](#ref-mani2010infant)) in a same-aged English monolingual cohort, and found the same null pattern of results.

To our knowledge this is the first study investigating cross-language priming using an implicit naming paradigm. This paradigm presents a critical asset for studying bilingual lexical processing: since the prime pictures are presented in silence, the task can be presented exclusively in the dominant language of participants. This provides a monolingual context in which differences in performance across experimental conditions that involve cross-linguistic interactions can be more easily associated to the experimental manipulation in a causal way ([Grosjean, 1997](#ref-grosjean1997bilingual)). This contrasts with previously used experimental paradigms in which participants are presented with items from both languages. Such covert manipulation of the similarity between stimuli from the two languages has allowed previous studies to provide stronger support for the language-non selective account of lexical access than studies presenting an overtly bilingual task to participants (e.g., [Costa et al., 2000](#ref-costa2000cognate); [Thierry & Wu, 2007](#ref-thierry2007brain); [Von Holzen & Mani, 2014](#ref-von2014bilinguals)).

Previous studies had provided evidence of English monolingual infants 18-month-old and older being sensitive to phonological priming effects during auditory word recognition. 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 shift from facilitation to inhibition was interpreted as a shift from sub-lexical facilitation at the phonological level—in which target word recognition is facilitated by the prior activation of an overlapping set of phonological segments—to a lexical interference effect in which the activation of the lexical representation of the phonologically similar prime led to competition dynamics with the subsequently activated representation of the 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 ERP 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 presented auditorily, 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. In bilinguals, only one study had used an implicit naming paradigm to study cross-language priming. Von Holzen & Mani ([2014](#ref-von2014bilinguals)) adapted Mani & Plunkett ([2010](#ref-mani2010infant)) procedure to test German-English bilinguals. 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’ ERPs 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.

As mentioned in the Discussion of Study 1, 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 might 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 the 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)). Finally, it is possible that the adjustments carried out during the adaptation of the task 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. Prime and target were presented very close in time. The onset of the target occurred 1,550 ms after prime picture onset. It is possible that such time interval was too short for participants to retrieve the phonological label of the prime picture, so that the auditory presentation of the target disrupted prime lexical access. Further research should investigate the role of prime-target inter-stimulus interval on the time course of lexical access and selection in toddlers.

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.

The findings in the present investigation 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.

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 ancitipated from the literature.

# References

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

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

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.

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.

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.

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*.

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.

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

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. Except for four participants who were also exposed to English at home, participants in Von Holzen & Mani ([2012](#ref-von2012language)) received their English input exclusively at school from a German native speaker. Participants’ exposure to each language ranged between three and four hours per day. This implies that participants’ exposure to English consisted of an average of 3-4 hours per day, during the time they attended daycare (reported as an average 12.35 months). Assuming that infants were exposed to a daily average of 12 total hours of speech, and that they attended daycare five days per week, the estimated cumulative to English in the life of an average child in the sample was between 10.50% and 14.00% for the 21 month-olds (the youngest in the sample), and between 5.13% and 6.84% for the 43-month-olds. These percentages were calculated as:

   . [↑](#footnote-ref-20)