# The Effect of Cognates on Bilingual Infant Vocabulary Trajectories: A Study Using Bilingual CDIs of English and One Additional Language

**Serene Siow, Nicola A. Gillen, Irina Lepadatu, Daniela S. Avila-Varela, Gonzalo Garcia-Castro, Nuria Sebastian-Galles, and Kim Plunkett**

## 1. Introduction

There are individual differences in vocabulary acquisition trajectories between bilingual children, and the same child may even have differing trajectories for their two languages (Pearson & Fernández, 1994). A recurring question in bilingual research is whether a bilingual's two languages are interconnected or independent during early acquisition. One way to measure the degree of interdependence between languages is by studying how many translation equivalents are known by the child (e.g., Pearson, Fernández, & Oller, 1995). Translation equivalents refer to a dual-language word pair with shared meaning. The term *doublet* has been used by various studies in the bilingual literature (Pearson & Fernández, 1994; De Houwer, Bornstein, & De Coster, 2006) to refer to translation equivalents that a child has learnt in both languages. To give an example, if an English-Spanish bilingual child knows the English word "dog" as well as its Spanish translation equivalent "perro", the child is said to know a doublet. One method for studying infant vocabulary is through parent-report vocabulary questionnaires, where parents mark out whether their child says, understands or does not understand a word on a list of words commonly known to young children. Vocabulary questionnaires have been widely used in bilingual research (Pearson, Fernández, & Oller, 1993; Pearson & Fernández, 1994; Cattani et al., 2014; Floccia et al., 2018;

* Serene Siow, University of Oxford, serene.siow@psy.ox.ac.uk; Nicola A. Gillen, University of Oxford; Irina Lepadatu, University of Oxford; Daniela S. Avila-Varela, Universitat Pompeu Fabra; Gonzalo Garcia-Castro, Universitat Pompeu Fabra; Nuria Sebastian-Galles, Universitat Pompeu Fabra; Kim Plunkett, University of Oxford.

Acknowledgements: This work was supported by an Economic and Social Research Council grant [CQR01830] to KP and NSG. Data collection was conducted by SS, IL and NG. We would like to thank all families who took part in this study. Huge thanks to associates who helped with the translation of the CDIs (Dutch: Prof Paula Fikkert; German: Dr Marlene Spangenberg; French: Prof Caroline Floccia; Italian: Dr Davide Volpi; Polish: Zuzanna Patryas; Portuguese: Andreia Correia Brecha; Spanish: Prof Nuria Sebastian-Galles, Daniela Avila-Varela and Gonzalo Garcia-Castro).

 *Proceedings of the 46th annual Boston University Conference on Language Development*, ed. Ying Gong and Felix Kpogo, 741-753. Somerville, MA: Cascadilla Press.

O'Toole et al., 2017; O'Toole & Hickey, 2017). When comparing across different language pairs, various methods have been used to standardise the composite measure of vocabulary size, as the number of words in each questionnaire often differs between adaptations. For example, O'Toole et al. (2017) calculated z-scores for conceptual vocabulary sizes as a proportion of the total possible concepts for each language pair. Using a different approach, Floccia et al. (2018) compared the subset of concepts that overlap across all the languages in their sample. In addition to vocabulary size, other studies have looked specifically at translation equivalents (De Houwer et al., 2006). These studies investigate the degree of overlap in vocabulary development between a young bilingual's two languages.

### 1.1. Word-level similarity (cognates) makes bilingual learning easier

It is acknowledged that by the second year of life, bilingual children understand words in both of their languages, including many translation equivalents (De Houwer et al., 2006; Legacy et al., 2017). However, predicting which words and translation equivalents are learnt first by bilinguals is no easy task, and is dependent on an array of intertwined factors. Some translation equivalents are phonologically more similar to each other than others. Translation equivalent pairs with shared etymology and high similarity in phonology/orthography are known as cognates (for example, English "train" and Spanish "tren"), and those with low similarity (e.g. English "train" and German "Bahn") are known as non-cognates.

Cross-linguistic phonological similarity is one of the factors hypothesised to influence the ease of bilingual word acquisition. Given existing knowledge of a word in one language, it should be easier to learn a word's translation equivalent if it sounds similar to the known word. Various studies have found advantages for cognates over non-cognates in terms of children's performance on receptive vocabulary tasks (Bosma, Blom, Hoekstra, & Versloot, 2019; Pérez, Peña, & Bedore, 2010) and word production tasks (Bosch & Ramon-Casas, 2014). This cognate advantage has also been found to facilitate the acquisition of a second language by adult learners (Otwinowska & Szewczyk, 2017) and school-age children (Tonzar, Lotto, & Job, 2009). Moreover, kindergarten and school-age children who grew up with spoken Arabic but later learnt standard Arabic in school were found to have stronger phonological representations for cognates than non-cognates when tested in the standard dialect (Saiegh-Haddad & Haj, 2018). Identical cognates exhibited the strongest effect, and there was an advantage for partially-overlapping cognates over non-cognates. Taken together, these studies suggest that phonological overlap between languages can facilitate the acquisition of words and the strength of encoded phonological representations.

### 1.2. Language-level similarity makes bilingual learning easier

Similarity between languages has been defined using various methods. Three common dimensions are phonological overlap, morphological complexity and word

order. We will focus on measures of phonological similarity in this paper. Floccia et al. (2018) used the average phonological overlap (Levenshtein distance) between translation equivalents in the CDI to define language distance. Gampe, Quick, and Daum (2021) applied a similar method, using a score derived from the Levenshtein distance for 40 common words. Blom et al. (2020) used two measures to justify their binary split of close and distant language pairs - average Levenshtein distance between translation equivalents and the number of cognates between languages (as judged by native speakers).

If cognates are easier to learn than non-cognates, languages with more cognates should be easier to learn simultaneously than languages with fewer cognates. Advantages of similarity between languages on vocabulary development has been observed in bilinguals of different ages. Blom et al. (2020) found better receptive vocabulary performance using a picture vocabulary task in 6 and 7-year-old bilingual children whose languages were more similar compared to bilingual children whose languages were less similar. Floccia et al. (2018) measured the vocabulary of 24 month old bilingual toddlers using vocabulary questionnaires, finding that children learning phonologically closer languages had larger expressive vocabularies than those learning more distant ones. In a similar study, Gampe et al. (2021) studied toddlers (18 to 30 months) learning Swiss German and one other language. They found larger expressive vocabulary sizes in toddlers whose other language was more similar to Swiss German.

### 1.3. The Present Study

Our study is an extension of Floccia et al. (2018)'s study with a wider age range of participants, expanding beyond their sample of 24 month old bilinguals and using full-length CDIs. The expanded age range allows us to investigate whether the relationship between predictors and vocabulary trajectories changes over early development. To this end, we collected data online from bilingual families with children between 12 and 36 months old. These families spoke English and one of 7 other languages. As all of our bilinguals had the common language of English, we will henceforth follow the convention used by Floccia et al. (2018) and refer to the non-English language as the additional language (AL).

Parents completed two vocabulary questionnaires: one in English and one in the other language. The vocabulary questionnaires had high conceptual overlap, allowing us to study not just total vocabulary size but also whether a child knows the word for a given concept in one language or both. The CDIs used in our study are adapted from the 418-word full-length Oxford CDI (with 365 concept overlap across all 8 languages), providing more vocabulary data than the 100-word short-form Oxford CDI used by Floccia et al. (2018). This allowed us to conduct more fine-grained analysis on the relationship between cognates and vocabulary growth.

### 1.3.1. Characteristics of the sample

Our sample was primarily recruited from the UK, with smaller numbers recruited from the Netherlands, Germany and Spain. The children in our UK sample had at least one parent who was a native speaker of the AL, while the children in our Germany/Netherlands/Spain samples had at least one parent who was a native speaker of English. In this manner, all toddlers were exposed to native input of the non-community language in the home.

Our demographics are sociolinguistically different from the widely-bilingual communities studied by some bilingual researchers, for example Catalan-Spanish bilinguals in Barcelona (e.g., Bosch & Ramon-Casas, 2014), who are exposed to similar quantities of Catalan and Spanish. Instead, our sample is comparable to that of Floccia et al. (2018) who similarly recruited bilingual families in the UK. In the UK community, the language input received outside the home is likely to predominantly be the community language (i.e., English), with the bilingual's other language spoken only by subgroups in the community (e.g. family members). For our Netherlands/Germany/Spain sample, the language spoken at home would be predominantly English, with the community language heard outside the home.

### 1.3.2. Hypothesis

An imbalance in the quantity of language exposure and the variety of contexts that each language is heard in may result in the two languages growing at different rates. Cognates may be particularly useful in such a learning environment. If more words are known in one language than the other, the existence of cognates may allow easier word learning via the comparison of the novel word form to an existing phonolexical representation. Indeed, past research with bilingual children has found stronger effects of cognate facilitation for receptive vocabulary performance in less proficient bilinguals (Bosma et al., 2019) and when tested in the less-dominant language (Pérez et al., 2010).

The aim of our study is to investigate whether languages with more cognates are easier to learn in this population of bilingual toddlers, and particularly if this can be associated with advantages in learning translation equivalents. We put forward two main hypotheses. Firstly, we predict that learners of languages with a higher percentage of cognates will have larger total vocabularies than learners of languages with less cognates. Total vocabulary size sums the respective vocabulary sizes in the bilingual's two languages as judged using a bilingual vocabulary questionnaire. Secondly, we predict that learners of languages with many cognates will have a higher proportion of doublets in their vocabulary than learners of languages with less cognates.

## 2. Methods

### 2.1. Participants

Our sample comprised of 12- to 36-month-old bilingual toddlers ($N$ = 397, cross-sectional data; 205 females) growing up with English and one AL (Dutch, French, German, Italian, Polish, Portuguese or Spanish).

Participants were recruited via advertisements posted on social media or emails sent to families who had registered their interest to take part in research at the Oxford University BabyLab. As a thank you for completing the study, participants were offered a choice between an e-gift voucher or a BabyLab t-shirt for their child.

All children in our UK sample had at least one parent who was a native speaker of the AL. Children in our Germany/Netherlands/Spain samples had at least one parent who was a native speaker of English. Of their two languages, our sample were exposed at least 25% of the time to each language (and therefore no more than 75%). Toddlers who were exposed more than 10% of the time to a third languages were defined as trilinguals, and thus were excluded from our analyses. An additional 18 families were excluded because the parent completing the questionnaire explicitly expressed uncertainty in the comments box about their vocabulary judgements (e.g., only the AL is spoken at home and they are unsure of their child's English vocabulary). The characteristics of the final sample is shown in Table 1.

A linear model comparing vocabulary size across samples from the different countries showed no significant difference after controlling for child's age, allowing us to merge the data.

**Table 1: Sample size per language, with mean age and mean percentage of overall English exposure (standard deviations in brackets).**

| language | country | N | N female | mean age | mean Eng exp |
|---|---|---|---|---|---|
| Dutch | Netherlands | 9 | 6 | 26.1 (7.07) | 51.1 (12.9) |
| Dutch | UK | 17 | 11 | 24.3 (6.62) | 62.1 (13.3) |
| French | UK | 50 | 26 | 23.8 (7.50) | 56.5 (12.6) |
| German | Germany | 16 | 11 | 26.2 (6.74) | 45.9 (12.9) |
| German | UK | 28 | 12 | 26.8 (6.04) | 53.6 (14.9) |
| Italian | UK | 53 | 28 | 23.1 (6.16) | 54.3 (13.3) |
| Polish | UK | 46 | 24 | 24.7 (7.52) | 53.6 (16.0) |
| Portuguese | UK | 27 | 12 | 23.4 (6.55) | 50.5 (15.3) |
| Spanish | Spain | 45 | 20 | 24.6 (6.81) | 50.3 (14.2) |
| Spanish | UK | 106 | 55 | 24.8 (7.02) | 49.7 (14.9) |

## 2.2. Materials

### 2.2.1. Demographics questionnaires

We collected demographics information about each child's age, gender and parent education level via a online questionnaire. We also asked if the child had been born premature and/or had hearing problems. Children born more than 6 weeks premature and those with reported hearing problems were excluded from the final sample. To obtain an estimate of bilingual language environment, we used a language exposure questionnaire adapted from the questionnaire used by Bosch and Sebastián-Gallés (2001). Our questionnaire asked about: (1) the ratio of overall language exposure; (2) the ratio of home language exposure; (3) each parent's native language and preferred language(s) to use with the child; (4) child's birth order; (5) age the child started nursery and the language(s) used at nursery; and (6) the cumulative number of months that the child has spent in a country/community that speaks the additional language.

### 2.2.2. Vocabulary checklist

We collected online parent-report vocabulary questionnaires in English and the AL for each child. Parents mark whether their child understands and says, understands but does not say, or does not understand each word in the list. For English, we used the Oxford Communicative Development Inventories (CDI) (Hamilton, Plunkett, & Schafer, 2000). The Oxford CDI is a vocabulary questionnaire for British English infants, normed using data from monolingual infants aged 12–25 months old. The Oxford CDI was chosen for use due to the UK demographic of the present study. Other bilingual studies conducted in the UK have similarly used the Oxford CDI (or its short form) to measure English vocabulary knowledge (Cattani et al., 2014; Floccia et al., 2018).

For each additional language, a CDI was adapted by translating the words in the Oxford CDI. Native speakers of each language were consulted for the creation of these adaptations. Not all words in these adaptations were direct translations of the Oxford CDI words – this is due to some cultural and linguistic differences requiring us to substitute some words with more appropriate replacements. These involved cases where two Oxford CDI entries of the same category had identical AL words as their translation. For example, both "clock" and "watch" are items in the Oxford CDI under the "household items" category, but both words are translated as "Uhr" in German. Rather than keeping both entries in the German adaptation of the CDI, and thus potentially confusing respondents, we chose to only keep one entry and mapped the translation to the earlier learnt meaning (as judged by the percentage of 18-month-old monolingual toddlers who understood the English word in the Oxford CDI dataset). We then replaced the later learnt word with another word from the same semantic category and of similar word knowledge norms in that language (percentage produced by 18-month-olds obtained from Wordbank (Frank, Braginsky, Yurovsky, & Marchman, 2017; Szagun, Stumper, & Schramm,

2009; Von Holzen, Nishibayashi, & Nazzi, 2018; Caselli et al., 1995; López-Ornat et al., 2005)). Continuing the example, the German word "Uhr" was mapped to the English word "clock", while the German word "Topf" (translation: "pot") replaced "watch". On the other hand, in cases where words were identical but were of different category, as in the animal "Fisch" ("fish") and food item "Fisch" ("fish"), both entries were kept because the referent for each entry was clear.

For the analyses reported in this paper, we used the subset of concepts where translations were present in all our CDIs. There were 358 concepts that fit this criteria, after excluding onomatopoeia. We can derive four measures of vocabulary knowledge from this CDI data:

1. English vocabulary: Number of English words understood by child
2. AL vocabulary: Number of AL words understood by child
3. Conceptual vocabulary: Number of concepts understood by child (knowledge of a concept is defined by understanding the word in English, in the AL or in both languages)
4. Doublets: Number of entries where both the English word **and** the AL word are understood

The two dependent variables in our study, total vocabulary size and proportion of doublets, are derived from these measures. *Total vocabulary size* (Equation 1) is used as a measure of ease of vocabulary acquisition, where a larger vocabulary size is associated with greater learning ease.

$$Total\ vocabulary\ =\ English\ vocabulary + AL\ vocabulary \tag{1}$$

*Proportion of doublets* (Equation 2) is used to measure the degree of conceptual overlap between a child's vocabulary in their two languages, where a higher proportion suggests that children find it easier to learn both translations.

$$Proportion\ of\ doublets\ =\ \frac{no.\ of\ doublets}{conceptual\ vocabulary} \tag{2}$$

### 2.3. Calculating percentage of cognates

We defined phonological similarity between languages using the percentage of cognates. We selected this definition to investigate our hypothesis that advantages in vocabulary size may be driven by the existence of words that are sound similar across languages. We first calculated the word-level phonological similarity of 236 word pairs common across all CDIs by applying Levenshtein distance (Levenshtein, 1966) on their phonological transcriptions. Levenshtein distance calculates the number of edits needed to convert one word to another. We standardised this score for different word lengths by dividing each pair's edit distance by the phoneme count of the longer word. We defined cognates as word pairs

with a standardised score of 0.6 or less. This is an arbitrary threshold intended to provide a standard cut-off for all languages. Dutch and German shared the largest number of cognates with English by a large margin. This was followed in order by French, Italian, Spanish, Portuguese and Polish (Table 2). Percentage of cognates was centred on the median (11%) and scaled by intervals of 10% for the analysis.

**Table 2: Percentage of cognates in CDI between English and each AL, with log transformed values (centered on median).**

| AL | language family | % cognates | % cognates (scaled) |
|---|---|---|---|
| Dutch | Germanic | 30.5% | 1.95 |
| German | Germanic | 27.5% | 1.65 |
| French | Romance | 11.9% | 0.08 |
| Italian | Romance | 11% | 0 |
| Spanish | Romance | 7.6% | -0.34 |
| Portuguese | Romance | 6.3% | -0.47 |
| Polish | Slavic | 5.5% | -0.55 |

## 3. Results

We expect that the total vocabulary size of toddlers learning language pairs with many cognates (e.g., English-Dutch) will be larger than that of children learning languages with few cognates (e.g., English-Polish). Additionally, we predict that the availability of cognates would make it easier for toddlers to learn translation equivalents, which should manifest as a higher proportion of doublets in the vocabularies of toddlers learning languages with more cognates. We expect that a toddler learning English and Dutch would know more words and have more doublets in their vocabulary than a toddler learning English and Polish. To test these predictions, we ran two mixed effects models using lme4 package (Bates, Mächler, Bolker, & Walker, 2015), the first with total vocabulary size as the dependent variable and the second with percentage of doublets. Both models had cognate percentage (centred on median and scaled by intervals of 10%) as the main predictor, with child's age (centred on mean age and scaled by SD) and overall English exposure (centred on 50% and scaled by SD) as covariates. Overall English exposure was added as a covariate because there is imbalance in the distribution of this statistic between languages. Language pair was added as a random effect. The model syntax is shown below:

```
lmer(total_vocabulary_size ~ age +
    overall_English_exposure +
    percentage_of_doublets + (1|language_pair))
```

P-values were obtained using lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017) via Satterthwaite's approximation degrees of freedom. Median cognate percentage was 11% and and a 1 unit change in scaled cognate percentage represents 10% more cognates across the language pair. As predicted, we found a significant positive main effect of cognate percentage on total vocabulary size (Table 3) (Figure 1). The statistics show that languages pairs with 10% more cognates are associated with an average total vocabulary size that is 3.02% larger. We likewise found a significant positive main effect of cognate percentage on the percentage of doublets (Table 4) (Figure 2). This means that children learning languages pairs with 10% more cognates understand on average 2.82% more doublets.

**Table 3: Mixed effect model for total vocabulary size, with age, English exposure and percentage of cognates as predictors, and language pair as a random effect.**

| Predictor | Estimate | Std Error | $t$ | $p$ |
|---|---|---|---|---|
| (Intercept) | 57.4 | 0.951 | 60.4 | $<.001$ |
| Age | 23.7 | 0.938 | 25.3 | $<.001$ |
| English exposure | 1.38 | 0.950 | 1.45 | .147 |
| **% cognates** | 3.54 | 1.18 | 3.02 | .003 |

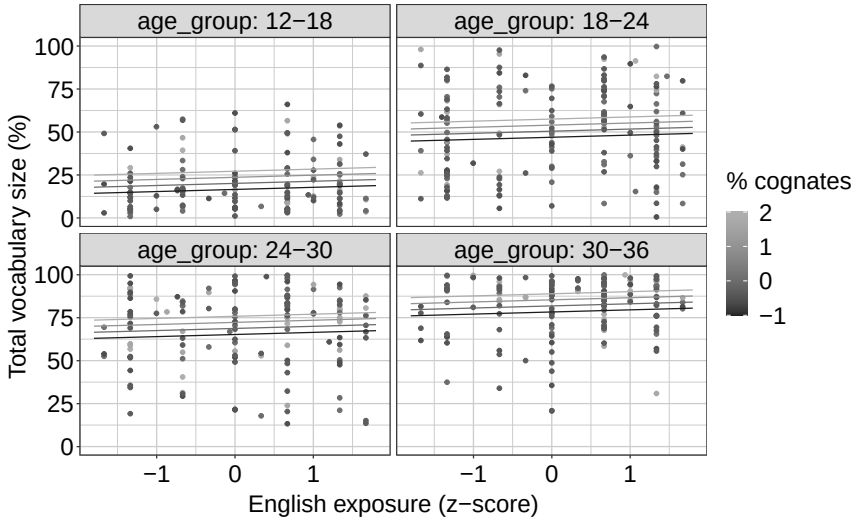


**Figure 1: Model predictions for total vocabulary size with the main effect of language similarity (defined by log-transformed percentage of cognates) as lines of best fit, and with child's age and English exposure percentage as covariates.**

**Table 4: Mixed effect model for percentage of doublets in vocabulary, with age, English exposure and percentage of cognates as predictors, with language pair as a random effect.**

| Predictor | Estimate | Std Error | $t$ | $p$ |
|---|---|---|---|---|
| (Intercept) | 58.8 | 1.39 | 42.2 | $<.001$ |
| Age | 15.6 | 0.993 | 15.7 | $<.001$ |
| English exposure | 3.05 | 1.01 | 3.02 | .003 |
| **% cognates** | 4.35 | 1.54 | 2.82 | .036 |

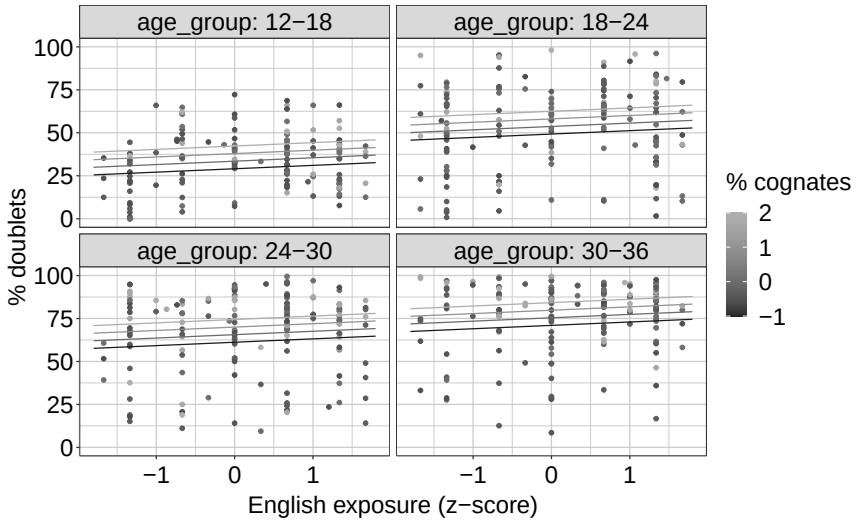


**Figure 2: Model predictions for percentage of doublets in child's vocabulary with the main effect of language similarity (defined by log-transformed percentage of cognates) as lines of best fit, and with child's age and English exposure percentage as covariates.**

## 4. Discussion

Our study presented cross-sectional vocabulary size data from bilingual toddlers between 12 and 36 months of age. Data was collected from families who speak English and one of 7 ALs. To sum up our results, we found that toddlers learning languages that have more cognates had larger vocabulary sizes and knew more doublets than toddlers learning languages with less cognates, even after controlling for child's age and overall English exposure. Our findings support the idea that languages with more cognates are easier to learn than those with less cognates. We theorise that phonological similarity between translation equivalents helps learners to disambiguate the meaning of words more easily, facilitating

the process of learning two languages in parallel. Our results complement existing literature on the relationship between cross-linguistic similarity and bilingual toddlers' vocabulary size (Floccia et al., 2018; Gampe et al., 2021).

The languages collected in our study are languages that have sizeable populations in the UK, making the study of these languages relevant to local researchers and families. The ALs cover a breadth of 3 different language families (2 Germanic, 4 Romance and 1 Slavic) and have varying numbers of cognates with English (a Germanic language), allowing us to study the effect of language similarity on vocabulary size in both languages. The CDIs for each AL were adapted from the 418-word full-length Oxford CDI, providing us with vocabulary data that has high conceptual overlap between languages. By using the subset of concepts that occur in all of our 8 CDIs for our analysis, we were able to study the effect of language similarity on the same list of concepts, thus avoiding the possible confound that observed differences in vocabulary knowledge may be caused by variation in concept complexity across groups.

### 4.1. Future directions

We are in the process of using this dataset to study how cognate status influences words' age of acquisition. If the vocabulary size differences between groups that was observed in this paper is driven by cognates, we would expect cognate words to be learnt earlier than non-cognate words after controlling for other word acquisition factors like frequency and word complexity. In addition to comparing words within a language, we can also test the hypothesis across languages – i.e. investigating if the English word "train" is acquired earlier by a Spanish-learning bilingual child than a German-learning one, because the Spanish equivalent "tren" has more phonological overlap with the English word than the German equivalent "Bahn". These analyses will provide a complementary perspective of the cognate facilitation effect on vocabulary acquisition.

## 5. Conclusion

Bilingual toddlers learning language pairs with more cognates were found to have larger total vocabulary sizes and know more doublets than those learning language pairs with fewer cognates. We suggest that toddlers are able to capitalise on the strong phonological overlap between cognates to facilitate learning of translation equivalents. Understanding the cognate facilitation effect on bilingual vocabulary acquisition can help guide strategies for teaching bilingual vocabulary.

## References

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Blom, Elma, Boerma, Tessel, Bosma, Evelyn, Cornips, Leonie, van den Heuij, Kirsten, & Timmermeister, Mona. (2020). Cross-language distance influences receptive vocabulary outcomes of bilingual children. *First Language*, *40*(2), 151–171.

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

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

Bosma, Evelyn, Blom, Elma, Hoekstra, Eric, & Versloot, Arjen. (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.

Caselli, Maria Cristina, Bates, Elizabeth, Casadio, Paola, Fenson, Judi, Fenson, Larry, Sanderl, Lisa, & Weir, Judy. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*(2), 159–199.

Cattani, Allegra, Abbot-Smith, Kirsten, Farag, Rafalla, Krott, Andrea, Arreckx, Frédérique, Dennis, Ian, & Floccia, Caroline. (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.

De Houwer, Annick, Bornstein, Marc H, & De Coster, Sandrine. (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.

Floccia, Caroline, Sambrook, Thomas D, Delle Luche, Claire, Kwok, Rosa, Goslin, Jeremy, White, Laurence, … Plunkett, Kim (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), 1–135.

Frank, Michael C, Braginsky, Mika, Yurovsky, Daniel, & Marchman, Virginia A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of child language*, *44*(3), 677–694.

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

Hamilton, Antonia, Plunkett, Kim, & Schafer, Graham. (2000). Infant vocabulary development assessed with a british communicative development inventory. *Journal of child language*, *27*(3), 689–705.

Kuznetsova, Alexandra, Brockhoff, Per B., & Christensen, Rune H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. doi: 10.18637/jss.v082.i13

Legacy, Jacqueline, Reider, Jessica, Crivello, Cristina, Kuzyk, Olivia, Friend, Margaret, Zesiger, Pascal, & Poulin-Dubois, Diane. (2017). Dog or chien? translation equivalents in the receptive and expressive vocabularies of young french–english bilinguals. *Journal of child language*, *44*(4), 881–904.

Levenshtein, Vladimir I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady* (Vol. 10, pp. 707–710).

López-Ornat, Susana, Gallego, Carlos, Gallo, Pilar, Karousou, Alexandra, Mariscal, Sonia, & Martínez, María. (2005). Macarthur: Inventario de desarrollo comunicativo. *Madrid: TEA Ediciones*.

Otwinowska, Agnieszka, & Szewczyk, Jakub M. (2017). The more similar the better? factors in learning cognates, false cognates and non-cognate words. *International Journal of Bilingual Education and Bilingualism*.

O'Toole, Ciara, Gatt, Daniela, Hickey, Tina M, Miękisz, Aneta, Haman, Ewa, Armon-Lotem, Sharon, … Kern, Sophie (2017). Parent report of early lexical production in bilingual children: a cross-linguistic cdi comparison. *International Journal of Bilingual Education and Bilingualism*, *20*(2), 124–145.

O'Toole, Ciara, & Hickey, Tina M. (2017). Bilingual language acquisition in a minority context: using the irish–english communicative development inventory to track acquisition of an endangered language. *International Journal of Bilingual Education and Bilingualism*, *20*(2), 146–162.

Pearson, Barbara Zurer, & Fernández, Sylvia C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language learning*, *44*(4), 617–653.

Pearson, Barbara Zurer, Fernández, Sylvia C, & Oller, D Kimbrough. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language learning*, *43*(1), 93–120.

Pearson, Barbara Zurer, Fernández, Sylvia C, & Oller, D Kimbrough. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of child language*, *22*, 345.

Pérez, Anita Méndez, Peña, Elizabeth D, & Bedore, Lisa M. (2010). Cognates facilitate word recognition in young spanish-english bilinguals' test performance. *Early Childhood Services (San Diego, Calif.)*, *4*(1), 55.

Saiegh-Haddad, Elinor, & Haj, Lina. (2018). Does phonological distance impact quality of phonological representations? evidence from arabic diglossia. *Journal of child language*, *45*(6), 1377–1399.

Szagun, Gisela, Stumper, Barbara, & Schramm, Satyam Antonio. (2009). *Fragebogen zur frühkindlichen sprachentwicklung (frakis) und frakis-k (kurzform)*. Frankfurt: Pearson Assessment.

Tonzar, Claudio, Lotto, Lorella, & Job, Remo. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning*, *59*(3), 623–646.

Von Holzen, Katie, Nishibayashi, Leo-Lyuki, & Nazzi, Thierry. (2018). Consonant and vowel processing in word form segmentation: An infant erp study. *Brain sciences*, *8*(2), 24.

# Proceedings of the 46th annual Boston University Conference on Language Development

edited by Ying Gong
and Felix Kpogo

Cascadilla Press Somerville, MA 2022



ISSN 1080-692X
ISBN 978-1-57473-077-7 (2 volume set, paperback)

**Ordering information**

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com