

The Influence of Bilingualism on Statistical Word Learning: A Registered Report

Matilde E Simonetti*

Chair of Cognitive and Experimental Psychology
Institute of Psychology
RWTH Aachen University, Germany
Jaegerstr. 17/19, D-52066 Aachen, Germany
E-mail: simonetti@psych.rwth-aachen.de

Megan G Lorenz

Department of Psychology and Neuroscience
Augustana College, USA

Iring Koch

Chair of Cognitive and Experimental Psychology
Institute of Psychology
RWTH Aachen University, Germany

Tanja C Roembke

Chair of Cognitive and Experimental Psychology
Institute of Psychology
RWTH Aachen University, Germany

*Corresponding author

Abstract

While statistical word learning has been the focus of many studies on monolinguals, it has received little attention in bilinguals. The results of existing studies on statistical word learning in bilinguals are inconsistent, with some research reporting a bilingual advantage over monolinguals but others finding no difference between groups. Thus, our study will investigate statistical learning using the Cross-Situational Statistical Learning paradigm **between** two groups: English-German bilinguals and English monolinguals. Participants will learn 1:1 mappings (one word maps onto one object) and 1:2 mappings (one word maps onto two objects). In contrast to previous studies, we will measure learning continuously and analyse trial-by-trial behaviour closely to understand fine-grained learning differences across language groups. We predict that it will generally be easier to acquire 1:1 than 1:2 mappings. More importantly, we predict that bilinguals will outperform monolinguals for 1:2 mappings only, consistent with a limited bilingual advantage.

Keywords: Bilingualism, cross-situational statistical learning, cross-situational word learning, statistical learning, language acquisition

The Influence of Bilingualism on Statistical Word Learning: A Registered Report

Almost half of the world's population speaks more than one language (Grosjean, 2010). For example, in Europe, according to a recent report (Eurostat, 2019), only about one-third of adults between 25 and 64 years old speaks just one language, likely due to migration and schooling. Given the large number of multilingual people worldwide and bilinguals' unique language experience, researchers have devoted much attention to understanding whether being bilingual changes cognition or brain structure (e.g., Bialystok & Craik, 2022). Indeed, several studies reported that bilingualism confers some cognitive benefits (e.g., Adesope et al., 2010; Antoniou, 2019; Grundy, 2020; Grundy & Timmer, 2017; Hartanto & Yang, 2019; Van den Noort et al., 2019; Warmington et al., 2019). In particular, a bilingual advantage (i.e., better performance for bilinguals than monolinguals) has been found for measures of metalinguistic awareness (e.g., Bialystok & Barac, 2012; Eviatar et al., 2018), phonetic perception (e.g., Antoniou et al., 2015), cognitive flexibility (e.g., Seçer, 2016), inhibitory control (e.g., Hartanto & Yang, 2019), selective attention (e.g., Chung-Fat-Yim et al., 2017; Olguin et al., 2019) and working memory (e.g., Anderson et al., 2021). Nevertheless, nowadays, the bilingual advantage has been questioned by a growing body of literature that found no difference between **monolingual** and bilingual populations (e.g., Dick et al., 2019; Donnelly et al., 2019; Lehtonen et al., 2018; Nichols et al., 2020). A recent meta-analysis that considered more than 150 papers claims that any cognitive advantage for bilinguals over monolinguals is task- and age-dependent (Ware et al., 2020).

In novel word learning, a bilingual advantage was found for bilinguals **with equal proficiency in their two best known languages** (i.e., balanced bilinguals; Kaushanskaya & Marian,

2009b; Warmington et al., 2019), and bilinguals with unequal proficiencies in their two best known languages (i.e., unbalanced bilinguals; e.g., Bogulski et al., 2019; Nair et al., 2016; Poepfel & Weiss, 2016; Singh et al., 2018; Van Hell & Mahn, 1997). Bilinguals were observed to acquire words more easily if novel words were explicitly paired with their meaning (e.g., Bogulski et al., 2019; Hirosh & Degani, 2021; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a) as well as in statistical word learning experiments where participants have to acquire the meanings more implicitly (Escudero et al., 2016; Poepfel & Weiss, 2016). However, there is only a small number of studies with bilingual participants (e.g., Bogulski et al., 2019; Nair et al., 2016; Poepfel & Weiss, 2016), and their results are inconsistent, with some studies failing to find a bilingual word learning advantage (e.g., Benitez et al., 2016).

As a result, it is not clear whether bilingualism impacts word learning and whether previous findings that did report a bilingual advantage for word learning are robust across different paradigms.¹ Thus, the goal of this study is twofold: first, to clarify whether there is a difference in statistical word learning between monolinguals and bilinguals, and second, to better characterise bilinguals' statistical word learning by using more detailed autocorrelation analyses of trial-by-trial learning behaviour.

1. Recent meta-analyses often include no or only a small number of word learning studies (e.g., Donnelly et al., 2019; Ware et al., 2020).

Word Learning in Bilinguals

To learn a word, one has to encode a referent and its label and form an association between the two (a mapping; McMurray et al., 2012). Creating an association between a word and a referent is not easy. As Quine pointed out (1960), word learning inherently involves *referential ambiguity*, as any heard word can be associated with all the concrete or abstract objects presented in a scene. Thus, the learner may not know onto which object or concept a word maps. Several heuristics exist to overcome referential ambiguity, such as the mutual exclusivity assumption (Markman & Wachtel, 1988). This assumption refers to learners preferring to map a new word on a referent that does not already have a label.

Word learning may be more complex for bilinguals than monolinguals: Consistent with this, some studies reported bilingual children to have a smaller vocabulary than monolingual ones (e.g., De Houwer et al., 2014; Montanari et al., 2018). Nevertheless, more recently, some authors did not find a lexical deficit (i.e., smaller vocabulary and slower word retrieval compared to monolinguals) in balanced bilingual children (e.g., Bylund et al., 2022); others have argued that vocabulary tests were created for testing monolinguals and that they, therefore, do not accurately represent bilinguals' word knowledge (Ehl et al., 2020).

One reason why word learning may be more complex in bilinguals is that the same word can be associated with multiple meanings (e.g., interlingual homographs; for example, "*pie*" means a type of food in English and foot in Spanish), and the same object can be mapped onto multiple words (e.g., synonyms, translation-equivalent words). In particular, a one-to-one (1:1) mapping refers to a word that maps onto one object. A two-to-one (2:1) mapping refers to two words mapping onto the same object (e.g., "*cat*" and "*kitty*" both refer to the same animal).

Similarly, a one-to-two (1:2) refers to one word mapping onto two different objects (e.g., “*crane*” refers to a bird and the machine). *Multiple mappings* (e.g., 1:2 or 2:1) are present within one language, but they are the norm for bilinguals, who are forced to map each object with multiple words and many words with multiple objects. Indeed, bilinguals have to track more than one input stream and acquire more than one vocabulary while potentially receiving less exposure to each language than monolinguals. Despite the higher complexity of bilingual word learning or maybe *because* of it (Bogulski et al., 2019), bilinguals may be able to outperform monolinguals when learning new words. Bogulski et al. (2019) distinguished between four different hypotheses why this might be the case: First, the L1 regulation hypothesis proposes that bilinguals’ advantage over monolinguals is due to having more regulatory skills in their L1 (first acquired **and often most proficient** language) than their L2 (second acquired language) because of a daily experience with processing costs. Second, the phonological awareness hypothesis suggests that bilinguals’ advantage is due to their major experience with two phonological systems and the subsequent facilitation in phonological learning. Third, the transfer of learning context hypothesis postulates that bilinguals have more experience with language learning, and a learning advantage over monolinguals is present if the new learning conditions are similar to the previous ones. Fourth, the bilingual cognitive control hypothesis proposes that bilinguals’ advantage on monolinguals is due to bilinguals’ better cognitive control on executive function tasks. **Therefore, bilinguals may have better cognitive** control than monolinguals (e.g., Bialystok & Barac, 2012; Costa et al., 2009; Sullivan et al., 2014; Verreyt et al., 2016), which may in turn facilitate word learning (e.g., higher inhibition may reduce interference from competitor referents). **Another possibility is that** bilinguals have adapted their

learning assumptions to facilitate the acquisition of multiple mappings; we will refer to this as the learning adaptation hypothesis. Consistent with this hypothesis, previous research has shown that bilingual children have relaxed their use of the mutual exclusivity assumption (e.g., Byers-Heinlein & Werker, 2009; Davidson et al., 1997; Kalashnikova et al., 2015). In line with these results, in their computational model, McMurray et al. (2012) found that the use of mutual exclusivity strongly depended on the number of multiple mappings per word encountered. According to this hypothesis, bilinguals should outperform monolinguals specifically when multiple mappings are to be learned.

Hypotheses 1-4 predict that the more balanced between languages someone is (i.e., the closer they are to being a perfectly balanced bilingual), the better language learners they are. This is particularly clear in the context of the L1 regulation hypothesis, because it refers directly to regulatory skills to explain the possible difference between monolinguals and bilinguals. For the learning adaptation hypothesis, it is also true that higher balancedness should be associated with better language learning, but it additionally predicts a specific advantage when what has to be learned is more complex (i.e., multiple mappings). One of the goals of the current study is to investigate the general prediction that higher balancedness is associated with better word learning of multiple mappings.

Consistent with these hypotheses that propose a bilingual advantage, bilinguals have indeed been found to outperform monolinguals when learning new words. These experiments are typically divided into two phases; in the first phase (familiarisation), new words (either in a novel third language or non-words) are typically presented together with their translation in L1 or L2 (e.g., Bogulski et al., 2019; Hirosh & Degani, 2021; Kaushanskaya, 2012; Kaushanskaya &

Marian, 2009a, 2009b; Papagno & Vallar, 1995; Van Hell & Mahn, 1997), or with just the object they refer to (e.g., Bartolotti & Marian, 2012; Borragan et al., 2021; Nair et al., 2016). During the second phase (testing), participants' learning can be tested through different methods, including recognition (e.g., Bogulski et al., 2019; Hirosh & Degani, 2021; Kaushanskaya, 2012), production (e.g., Hirosh & Degani, 2021; Kaushanskaya & Marian, 2009a), and recall task (e.g., Kaushanskaya, 2012; Kaushanskaya & Marian, 2009b). Within this paradigm, a bilingual advantage over monolinguals in word learning was found both in children (e.g., Borragan et al., 2021) and adults (e.g., Bartolotti & Marian, 2012; Kaushanskaya & Marian, 2009b; Nair et al., 2016), with balanced (e.g., Bartolotti & Marian, 2012; Borragan et al., 2021) and unbalanced bilinguals (e.g., Hirosh & Degani, 2021; Papagno & Vallar, 1995) (but see Bogulski et al., 2019; Borragan et al., 2021 for exceptions).

To summarise, there are reasons to believe that vocabulary acquisition may proceed differently in bilinguals than monolinguals. At this point, it is still unclear how such differences impact forming a word-object-mapping, as previous research has been contradictory: Some research has found no differences between monolinguals and bilinguals, whilst others have found bilinguals to outperform monolinguals and vice versa.

Statistical Word Learning

Referential ambiguity only exists if learning is restricted to a single situation (Siskind, 1996; Yu & Smith, 2007). However, if a word and its meaning have above baseline probability of co-occurring, this information can be used across situations to learn the correct mappings. Yu and Smith (2007) tested whether monolingual adult participants can use co-occurrence statistics in

the Cross-Situational Statistical Learning (CSSL; sometimes called Cross-Situational Word Learning) paradigm. In their study, monolingual participants heard two to four words and saw two to four objects in each trial. Thus, while every single situation by itself was ambiguous, information could be extracted across trials. Yu and Smith (2007) found that participants could learn the correct word-object-mappings with brief but cross-situational exposure and without feedback. Since then, several studies have corroborated Yu and Smith's (2007) findings (see Roembke et al., 2023, for a focused review), as learning via CSSL was found in adults (e.g., Bulgarelli et al., 2021; Chen et al., 2017; Mulak et al., 2019; Roembke & McMurray, 2016; Tuninetti et al., 2020; Wang, 2020), children (e.g., Mangardich & Sabbagh, 2022; McGregor et al., 2022; Suanda et al., 2014; Vlach & DeBrock, 2019) and even infants (e.g., Smith & Yu, 2008).

As the name suggests, Cross-Situational *Statistical* Learning is often considered a form of statistical learning (e.g., Roembke & McMurray, 2016; Weiss et al., 2020). That is, the mechanism underlying learning is thought to be the gradual accumulation of co-occurrence statistics, where mappings between words and meanings are strengthened over time. Recent accounts propose that while such associative learning may be a core feature of CSSL, it likely interacts with more explicit types of learning (e.g., McMurray et al., 2012; Roembke & McMurray, 2016; Yurovsky & Frank, 2015), such as generating and remembering a hypothesis about what object a word maps onto (so-called hypothesis-testing; Trueswell et al., 2013).

One way to investigate CSSL is to look at what characteristics of preceding trials predict performance on a current trial. In these so-called trial-by-trial analyses (or autocorrelation analyses), accuracy on a current trial is predicted by characteristics or behaviours on previous trials with the same word or objects. **Importantly, in these analyses, learning can be predicted**

by behaviour on previous trials (e.g., whether a competitor object was correctly mapped with its word the last time it was encountered) or more general trial descriptors (e.g., how often a word has been encountered). For example, we can look at whether participants are more likely to be correct on a current trial late in the experiment versus early or whether accuracy on a preceding trial predicts performance on a current trial. Using these analyses, participants have been found to be more accurate later on (effect of target count) and to be more likely to be correct on a current trial if they also selected the target the last time they encountered the same word (effect of last-target accuracy, Dautriche & Chemla, 2014; Roembke & McMurray, 2016; Trueswell et al., 2013). The effect of target count has been argued to be a measure of statistical or more implicit learning processes, while the effect of last-target accuracy has been seen as an indicator of more explicit learning (Roembke & McMurray, 2021; Trueswell et al., 2013). Moreover, the use of mutual exclusivity can be estimated by how accurate participants are on a current trial based on whether they had selected the correct referents for the competitor objects the last time they were the target object (last-competitor accuracy; Roembke et al., 2018; Roembke & McMurray, 2016). That is, are participants able to rule out competitor objects as potential targets if they had correctly mapped them on a word before?

Previous research has shown that these trial-by-trial analyses can identify differences in CSSL learning patterns across different participant groups (Roembke et al., 2018): For example, Roembke and colleagues (2018) used trial-by-trial analyses to ask whether children and adults differed in CSSL learning for different stimulus types. They found that while children's and adults' CSSL performance was higher for targets they had been correct for previously, only

adults benefited from last-competitor accuracy. This suggests that adults were able to use mutual exclusivity to select referents on a current trial, while children were not.

It has been argued that CSSL is a common way to acquire words early in development but also later in life (e.g., when encountering novel words during reading; Roembke et al., 2023). To learn more than one language, bilinguals are prone to track the co-occurrence of multiple words in multiple contexts; it is likely that much of this learning does not occur through explicit teaching but also more implicit learning processes as proposed in CSSL. However, to date, research on CSSL in bilinguals has been very limited. There is a small number of studies that investigated cross-situational statistical word learning in bilinguals, some using only 1:1 mappings (Crespo et al., 2023; Crespo & Kaushanskaya, 2021; Escudero et al., 2016), with others using multiple mappings (Benitez et al., 2016; Li & Benitez, in press; Poepsel & Weiss, 2016). Poepsel and Weiss (2016) compared statistical word learning with multiple mappings in English monolinguals and Chinese–English, and English–Spanish unbalanced bilinguals. Their first experiment found no difference between the three groups' learning rates of 1:1 mappings. Thus, in their second experiment, participants had to acquire 1:1 and 1:2 mappings. For both mappings, participants saw three objects while hearing three words during the familiarisation phase. Then, they had to choose which word they thought mapped onto which object in a separate two Alternative Forced Choices (AFC) test. The authors found that bilinguals were better than monolinguals in learning the correct associations with 1:2 mappings, while there was again no significant difference between groups with 1:1 mappings. These findings were explained as a relaxation of mutual exclusivity for bilinguals, making it easier for them to acquire 1:2 mappings (c.f., Byers-Heinlein & Werker, 2009; Kalashnikova et al., 2015; McMurray et al.,

2012), and are consistent with the bilingual advantage found in other word learning tasks (e.g., Bartolotti & Marian, 2012; Bogulski et al., 2019; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a; Nair et al., 2016).

In another study with multiple mappings, Benitez et al. (2016) tested learning of 1:1 and 2:1 mappings across three experiments with monolingual and bilingual participants. In the familiarisation phase, participants heard four words while four objects were presented on the screen, and after that, they had to select the target object in a 4AFC test. In all their studies, Benitez and colleagues (2016) found that both groups learned 1:1 mappings better than 2:1 mappings, whilst the interaction between **language background** and mappings was not significant.

In a recent study, Li and Benitez (in press) tested monolinguals and two groups of bilinguals (Chinese-English and Spanish-English) on how they learn 2:1 word-object-mappings cross-situationally. The focus of that study was the use of a linguistic cue (one of the two target words had a tonal cue). They did not find a difference between the CSSL accuracy of the three groups in the uncued condition, whilst in the cued condition, Chinese-English bilinguals outperformed the other two groups. These results are consistent with the phonological awareness hypothesis (Bogulski et al., 2019). In another recent study (Aguasvivas et al., 2024), Spanish monolinguals and two groups of bilinguals (Spanish-Bask and Spanish-English) with all three types of mappings (1:1, 2:1 and 1:2). Interestingly, they did find a significant difference between monolinguals and bilinguals, but only for 1:1 mappings. Aguasvivas and colleagues (2024) suggested that the reason why they found a difference for 1:1 mappings and not more

complex ones (as e.g., Poepfel & Weiss, 2016) may be that their participants had to learn more words than is typically true in these kinds of experiments.

Other studies that explored only 1:1 mappings found a difference between monolingual and bilingual populations (Crespo et al., 2023; Crespo & Kaushanskaya, 2021; Escudero et al., 2016), but results were mixed in how exactly groups differed. Escudero and colleagues (2016) compared the learning rate of Australian monolinguals and Singaporean balanced bilinguals adults. They found that bilinguals were more accurate than monolinguals when learning 1:1 mappings. A similar, but more narrow, effect was found by Crespo and colleagues (2023). In this study, English monolingual and English-Spanish bilingual 7-year-old children completed a CSSL experiment with multiple conditions (one with no variability and three with: multiple exemplars, multiple speakers, and combined cue). Bilingual children's performance was better than monolingual ones when they were exposed to conditions with variability, especially in the combined cue condition (where both multiple exemplars and speakers were used for the words). Instead, Crespo and Kaushanskaya (2021) found the opposite pattern. They investigated English monolingual and English-Spanish balanced bilingual children between 4 and 7 years old. Here, monolingual children outperformed bilingual ones in learning 1:1 mappings with the CSSL procedure.

Taken together, these findings indicate that the role of bilingualism in statistical learning of single or multiple mappings is still unclear, and the literature about it is scarce and conflicting. A bilingual advantage was found in four studies (Aguasvivas et al., 2024; Crespo et al., 2023; Escudero et al., 2016; Poepfel & Weiss, 2016), but with different mappings (1:1, 1:2, or 2:1), with different populations (different combination of languages and different measures for

bilingualism) and in one case only when multiple word exemplars and speakers were used (Crespo et al., 2023). A language-specific advantage was found in one study (Li & Benitez, in press). Additionally, the sample sizes in two studies were very small (around 16 participants per group; Escudero et al., 2016; Poepfel & Weiss, 2016). Studies with bigger sample sizes (35-50 participants/group) failed to report any difference between the language groups (also using different mappings, Benitez et al., 2016) or even found a monolingual learning advantage (using a different population, Crespo & Kaushanskaya, 2021). All studies used relatively coarse measures of learning behaviour (accuracy at test), potentially missing out on subtler differences between groups. Thus, at this point, it is not clear whether language learning history can impact statistical word learning and, if so, under which circumstances or language histories it does. Moreover, if there are learning differences across language groups, we currently do not have a good understanding of why they arise.

The Current Study

The current study aims to clarify whether one's language history (monolingualism vs bilingualism) impacts statistical word learning. This will also help situate the bilingual advantage observed in some word learning studies using other paradigms (e.g., Bartolotti & Marian, 2012; Bogulski et al., 2019; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a; Nair et al., 2016). To do so, we will implement a CSSL experiment similar to the one conducted by Poepfel and Weiss (2016) but with some different critical components to capture potentially existing learning differences across groups better.

First, our study will be conducted online to increase our sample size and to collect data between different language groups². In particular, we will compare a monolingual population (like Poepsel and Weiss (2016)) with an English-German bilingual group. Therefore, we will be able to see if the effect can be found with different language combinations.

Second, we will look at to what extent balancedness (i.e., how balanced between languages someone is) impacts CSSL word learning performance. To measure balancedness, we will use language entropy (i.e., relative balance between the two or more languages throughout daily life: Gullifer & Titone, 2020) of participants. Language entropy is most directly a measure of balancedness, but it has been found to be associated with age of acquisition and proficiency (Gullifer & Titone, 2020). By measuring language balancedness and relating it CSSL word learning performance, we can probe more specifically the hypotheses proposed by Bogulski and colleagues (2019) and the learning adaptation hypothesis. Third, in contrast to Poepsel and Weiss (2016) and Benitez and colleagues (2016), we will continuously assess learning³. On every

² It is difficult to gain data from a monolingual population outside an English-speaking country.

³ One way of measuring CSSL is by collecting an explicit response on each trial (Roembke et al., 2023). Forcing participants to explicitly select a response on each trial may increase explicit processing. We chose for participants to make a response on each trial because it, however, provides a continuous measure of learning (that is also needed for trial-by-trial analyses) and increases task engagement (which may help in an online study). Importantly, previous research

trial, participants will see be presented with word and three objects. Thus, instead of having a separate familiarisation and test phase, participants will choose the object they think maps onto the word on every trial (Roembke & McMurray, 2016). Fourth, by continuously assessing learning, we will have the opportunity to analyse the data using trial-by-trial analyses to better understand CSSL (Roembke et al., 2018). As described previously, these fine-grained analyses can characterise learning more closely (such as the use of the mutual exclusivity assumption). More specifically, we will look at last-competitor accuracy. This measure indicates whether one was able to select the objects used as competitors on a current trial as the targets the last time they were the correct response. This measure reflects the use of the mutual exclusivity assumption (knowledge of competitors' referents allows for its application): If one knows what words the competitor objects map onto, it should be possible to use that information to exclude these competitors as possible referents on a current trial. To our knowledge, bilingual CSSL has never been investigated using trial-by-trial analyses.

We hypothesise that—consistent with previous research—learning 1:1 mappings will be easier than learning 1:2 mappings for all participants and that performance differences will become more apparent throughout the experiment (hypothesis H1).

In addition, based on the results by Poepsel and Weiss (2016), we predict that bilinguals will outperform monolinguals when acquiring 1:2 but not 1:1 mappings (H2), consistent with the

suggests that implicit learning occurs during this version of the CSSL paradigm as well (Roembke & McMurray, 2021).

learning adaptation hypothesis. We will also explore a possible effect of language entropy (Gullifer & Titone, 2020) on this interaction (exploratory hypothesis E1).

When analysing trial-by-trial behaviour more closely, we predict, consistent with previous studies (Dautriche & Chemla, 2014; Roembke & McMurray, 2016; Trueswell et al., 2013), accuracy on a current trial will be higher if participants were also accurate on the preceding trial with the same word and the more often a word has been encountered (as indicated by the effect of last-target accuracy and target count; H3). In addition, we predict that the use of the mutual exclusivity bias (as indicated by the effect of last-competitor accuracy) will be reduced for bilinguals for all mapping types compared to monolinguals (H4). Furthermore, we predict that the use of the mutual exclusivity bias will be reduced for 1:2 compared to 1:1 mappings (H5).

Method

Participants

We will recruit **two** groups of participants—one group of English monolinguals and one group of English-German bilinguals. **The bilingual group** will speak be **proficient** in English and German (**proficiency** will be assessed through self-ratings as part of a slightly modified LEAP-Q (Kaushanskaya et al., 2020; see appendix A) and LexTALE (Lemhöfer & Broersma, 2012) score, standard measures of bilingual experience and proficiency.

Table 1. *Overview of language groups and active Prolific users.*

Language Group	L1	L2	Classification	N _{active Prolific users} (02.11.2023)
1	English	n.a.	Monolingual	45,980
2	English	German	Bilingual	10,535

Participants will be recruited through the online recruitment platform Prolific (www.prolific.com; see Table 1 for an overview of existing active members per language group). This study was approved by the Ethical Committee of the RWTH, Aachen (protocol number 2021_14_FB7_RWTH Aachen). This experiment has minimal risk, and informed consent will be obtained for each participant.

Power analysis. We will recruit **between 75 and 100** participants per group (**monolingual and bilingual**). **A minimum of 150 participants emerged from multiple power analyses (Kumle et al., 2021) based on pilot data and Poepsel and Weiss's (2016) data; the results of these power analyses are summarised in Table 2. Even so, we were not able to determine an adequate**

number of participants for H4. We decided to recruit 200 participants overall at maximum, since it is the upper boundary for our possibilities. We will decide when to stop between 150 and 200 participants, using the optional stopping practice (Rouder, 2014). We will calculate Bayesian Factor (BF; calculated using the Bf function as in Silvey et al., 2021) in intervals of ten participants periodically and we will stop when we reach a BF above 6 or below 1/6 in hypothesis H4.

Table 2. Summary of simulation-based power estimations.

Hypothesis	Model	Effect of interest	Data	Participants needed for an effect size of $>.8$
H1: Is it easier to learn simple (1:1, one word maps onto one object) or complex (1:2, one word maps onto two objects) mappings?	DV= Accuracy; IV= Block, Mapping Type and Language Group	Main effect of mapping type	Pilot data	50 participants (.999)
H2: Do bilinguals learn words more easily than monolinguals under some circumstances?	DV= Accuracy; IV= Test, Mapping Type and Language Group	Interaction between Mapping Type and Language Group	Poepsel and Weiss (2016)	150 participants (0.842)
H3: Does accuracy on a previous trial and target count (how often a word has been encountered) impact accuracy on a current trial?	DV= Accuracy; IV= Mapping Type, Language Group, Last-target-accuracy and target count	Main effect of Last-target-accuracy	Pilot data	50 participants (1.000)
		Main effect of Target count		50 participants (1.000)
		Interaction between Last-target-accuracy and Target count		50 participants (.894)
H4: Will the use of mutual exclusivity	DV= Accuracy; IV= Mapping Type,	Interaction between Last-competitor	Pilot data	200 participants (0.584)

bias be reduced for bilinguals?	Language Group, Last-competitor accuracy and target count	accuracy and Language group		
H5: Does the use of mutual exclusivity change between mapping types?	DV= Accuracy; IV= Mapping Type, Language Group, Last-competitor accuracy and target count	Interaction between Last-competitor accuracy and Mapping type	Pilot data	100 participants (.934)

For hypotheses H1, H3, H4 and H5, we performed several Generalized Mixed Models analyses (on data from a pilot experiment ($N_{\text{monolinguals}} = 25$; $N_{\text{bilinguals}} = 57$; an online experiment with a similar procedure). For H2, we performed again a Generalized Mixed Models analysis on the data from Poepel and Weiss (2016)⁴. We then performed five power analyses using simulation-based power estimation (Kumle et al., 2021). Since for H2 the power analysis revealed that 150 participants are needed to reach a power superior to .8, we concluded that 150 participants will be our minimum.

⁴ Due to the fact that in our own data, the interaction of Language Group \times Mapping Type was not significant, and it has a very low beta coefficient ($B = 0.087$), we had to use Poepel and Weiss' data to perform a meaningful power analysis. We would like to thank the editor for this suggestion.

Inclusion criteria. All participants must self-report having a normal or correct-to-normal vision and never having been diagnosed with a learning or language disorder. For the monolingual group, participants have to self-report as monolingual (“English is my mother tongue. **Yes/no**” and “Do you speak other languages? **Yes/no**”) as well as obtain an English LexTALE (Lemhöfer & Broersma, 2012) score of above 80% correct (consistent with an advanced proficient user, see Lemhöfer & Broersma, 2012). Bilinguals will have to confirm that they also speak English as their mother tongue (“English is my mother tongue. **Yes/no**”) and that they also speak German (“Do you speak German?” **Yes/no**); they will need to obtain a LexTALE score of above 80% correct in English (consistent with an upper intermediate/advanced user) and at least 50% in German (consistent with a lower intermediate user). Participants will be excluded if they fail to meet our data quality control standards (described later).

Stimuli

12 bisyllabic CVCV nonwords were created for this experiment: BERNAL, DARLON, GLANKE, GRINTER, MALFEN, MURLER, RAUPLER, STAUNKER, THERNUS, VARTION, WILTEN and SUMPER. All nonwords were plausible words in English and German, as they followed phonotactic and orthographic rules of both languages. Stimuli were selected through a survey, where ten English monolinguals and ten unbalanced German-English bilinguals evaluated 29 bisyllabic nonwords on a five points Likert scale (from “definitely not English/German” to “definitely English/German”; see appendix B). We then calculated the average score for each word for both groups and the difference between the two. We excluded the words with a difference greater than one point and selected stimuli with the greatest average acceptability score. On average,

selected stimuli were rated as 3.11 (SD = 0.22). 16 coloured photographs of unusual objects from the NOUN database (Horst & Hout, 2016) will be used as visual referents (see Figure 1 for examples). We choose the objects with the highest novelty score (calculated by Horst and Hout (2016) by subtracting the familiarity score they collected from 1).

Figure 1. Examples of to-be-used visual referents (Horst & Hout, 2016).



Procedure

All parts of the experiment will be implemented on Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020). We created two slightly different versions for the two groups (one for monolinguals and one bilingual). Firstly, we ask participants if English is their mother tongue, if not, they will be excluded. After this question, only for the bilingual group, we ask about their second language. If they claim to speak only English, they will be redirected to the monolingual experiment. If they can speak German, they will continue the experiment. In both experiments, then, participants' learning history will be measured using an extensive language questionnaire based on the LEAP-Q (for up to four languages; Kaushanskaya et al., 2020). We modified the LEAP-Q slightly to be more appropriate for our purposes (e.g., by adding a question about language use in different contexts; see version for bilinguals in

Appendix A1 and monolinguals in A2). The answers to LEAP-Q will also be used to calculate language entropy for each participant (as it is done in: Gullifer & Titone, 2020).

Additionally, participants' proficiency in German and English (only English for monolinguals) will be assessed with the LexTALE test (Lemhöfer & Broersma, 2012). In contrast to the web-based LexTALE, participants will be given up to five seconds to respond instead of unlimited time, to ensure participants do not have time to check. They will press the "s" key to indicate that a word exists and the "k" key to indicate that it does not. Subsequently, each participant will acquire twelve word-object-mappings in a 3AFC CSSL paradigm. Two-thirds of the words (eight) will be randomly assigned to only have one target referent (1:1 mapping), whilst the other third (four) will have two targets (1:2 mapping). For 1:1 mappings, the target referent and word will always co-occur, whilst the competitor objects will appear together with the word 20% of the time. For 1:2 mappings, half of the trials will include Target 1 and the other Target 2. The pilot experiment confirmed that participants were able to learn in this condition (overall learning rate = 59%; chance = 33%, $t(9719)=51.35$, $p < .001$).

Figure 2: Examples of trial sequence. After the third trial, participants are now able to deduce the meaning of "Murler" (the blue object).



The order of the trials will be pseudo-randomised within a set of twelve trials (one repetition of each word), so that no word appears in two consecutive trials (that is across sets of

repetitions). In addition, for 1:2 mappings, trials will take turns in whether they include Target 1 or Target 2. On each trial, three objects will simultaneously appear on the screen (one on top and two at the bottom for half trials and vice versa in the others: see Figure 2), together with a written word located in the centre of the screen. The word and (one of) its target object(s) will always be present on the same trial. Object competitors will be selected randomly without replacement; objects that are the target on one trial act as competitors in others. On each trial, participants will first see three objects and a fixation cross in the middle of the screen for 1000 msec. The fixation cross is then replaced by the written word in capital letters. In contrast to previous studies, nonwords will be presented in written form instead of spoken form to guarantee that stimuli are ambiguous language-wise and to reduce possible confounding variables (c.f., Escudero et al., 2023)⁵. Participants will click on one of the three objects to

⁵ Escudero and colleagues (2023) was the first CSSL study to use written words as stimuli; they found that participants were able to map written words onto objects as well as spoken ones. However, in their experiment two objects were presented with two words on each trial, while in the current experiment, three objects and only one word will be presented. This change in the methodology may potentially affect how easily it is to map a written word onto an object (though it is unclear whether it could facilitate learning or make it harder). This methodological decision was made to allow for a continuous assessment of learning and the trial-by-trial analyses. It is in line with previous literature using successfully only one spoken word as referent per trial (e.g., Roembke & McMurray, 2016; Yurovsky & Frank, 2015).

proceed to the subsequent trial. Thus, there will be no independent learning and testing phase. Exact sizes of objects and written word will differ across devices that participants use. Participants will only be allowed to access the experiment with a desktop or laptop computer (but not a smartphone or tablet).

During the study, participants will never receive any feedback on whether they selected the correct response or not. They will be instructed that it is their task to learn which object goes with which word, but they will not be told that certain words and objects co-occur, consistent with an implicit learning task. In total, there will be 360 trials divided into five blocks. In each block, each word will be seen an equal number of times. Participants will take approximately 30 minutes to complete the experiment.

Data quality control. At the end of each experiment, we will allow participants to disclose any strategies they may have used during the experiment that would make their data non-usable (consistent with other online studies, e.g. Hulme et al., 2022). We will ask the following questions: 1) Did you try to the best of your abilities to learn the words? Yes/no; 2) Did you use any “cheating” tactics to learn the words, such as taking notes, during the experiment? If yes, please explain. Yes/no; 3) Is there anything else you would like us to know about the experiment?

In our pilot experiment, we noticed a relatively high proportion of non-learners across all different language groups (i.e., participants that never scored above 40% correct; chance is at 33%), even as CSSL experiments have been successfully conducted online (Escudero et al., 2023). As in any learning experiment, it is unclear whether non-learners’ performance is meaningful (e.g., it is possible that if there is a bilingual advantage, monolinguals will have a

harder time acquiring 1:2 mappings and, therefore, never learn the mappings) or not (e.g., participants may not be motivated enough to actually engage in the task; this may be an increased issue in an online experiment). To identify people who did not engage with the task in the way intended, we will exclude participants if they indicated that they did not try their best (Did you try to the best of your abilities to learn the words? Yes/**no**) or that they used any “cheating” strategies (Did you use any “cheating” tactics to learn the words, such as taking notes, during the experiment? **Yes/no**).

Design

For the main analysis (H1 and H2), there will be two within-subject independent variables that are: block (1-5) and mapping type (1:1/1:2). There will also be one between-subject independent variable: language group (monolingual/bilingual). The dependent variable will be accuracy.

For the exploratory analysis (E1), we will also consider as an independent variable: language entropy. The coefficient of language entropy for each participants will be calculated using LEAP-Q answers (Kaushanskaya et al., 2020).

In addition, trial-by-trial analyses (H3, H4 and H5) will be conducted to explore learning more deeply and determine the influence of the learning group. Trial-by-trial analyses use previous trials with the same word as the current trial's predictors. We will explore the following independent variables: last-**target** accuracy, target count and last-competitor accuracy (together with mapping type and language group). The dependent variable will be accuracy on the current trial.

Analysis

We will analyse the study's data using R software (Version 4.3 or a more recent version, R Core Team, 2022). We will use the library lme4 (Bates et al., 2014) and lmerTest packages (Kuznetsova et al., 2017) to implement Generalized Mixed Models. We will calculate the Bayesian Factor (BF; using the Bf function as in Silvey et al., 2021) for each analysis and following the literature, we will consider a BF larger than 3 as moderate support for the alternative hypothesis, and a BF smaller than $\frac{1}{3}$ as moderate support for the null hypothesis (see Jeffreys, 1998; Lakens et al., 2020).

Main analyses. We will use binomial mixed models with accuracy (1/0) as the dependent variable. Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), language group (monolingual vs bilingual; both contrast coded as +0.5/-0.5), and block (from 1 to 5; centered) will always be included as fixed factors with all their interactions. As random effects, we will consider subject, target object, and word. Target and object will be nested in list (target objects and words will be randomised in five different lists; each participant will see only one of them). To choose the best random effect structure, we will use the function buildmer in the buildmer package (v. 2.8, Voeten, 2019). This function uses a backwards-fitting model selection procedure that starts from the maximal model and gives us a model that converges by systematically ruling out random slopes. The model will be run with two levels of language groups (monolingual, bilingual; contrast coded as +0.5/-0.5; H2). This model will also be used to evaluate whether 1:1 mappings are acquired more easily than 1:2 mappings (H1) and if there is a significant difference between the two language groups (H2). In contrast to that we consider as null hypotheses

respectively: no difference between the two mappings and no difference between the two groups. Consistent with previous findings (Benitez et al., 2016; Poepel & Weiss, 2016), we expect to find that participants will acquire 1:1 mappings more easily since they are less complex than 1:2 mappings. If any of these analyses result in a significant interaction, we will conduct post-hoc tests to compare performance of each language group for each mapping type separately (following Poepel and Weiss, 2016). For these analyses, we will use the same model as for the main analyses (as revealed by buildmer) with all instantiations of mapping type removed.

Exploratory analysis. To calculate language entropy we will use the package languageEntropy for R (Gullifer & Titone, 2020). In particular we will start from the results of LEAP-Q⁶ to calculate the proportion of time that participants are exposed to each language. Generally, low entropy scores (i.e., near zero) are indicative of high certainty (or low diversity) of some outcome, whereas high entropy scores (i.e., near one; the maximum reachable is

⁶ We will use percentages (of use, speaking, and reading) and Likert scales (from 1 to 10; of using with friends, family, colleagues as well as watching television/streaming, using social media, self-instructed apps and of listening to the radio/podcast/music) for all languages. We will calculate language entropy using the formula (n = number of languages, P_i = proportion of time that they are exposed to language i):

$$Entropy = \sum_{i=1}^n P_i \log_2(P_i)$$

$\log_2(\text{number of languages})$) are indicative of high diversity of some outcome. We will use a binomial mixed models with accuracy (1/0) as the dependent variable. Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), block (from 1 to 5; centered) included as fixed factors with their interaction, and with language entropy (continuous variable from 0 to $\log_2 n$, centered) as covariate. As random effects, we will consider subject, target object, and word (target objects and words were randomised across five different lists).

Trial-by-trial analyses. We will then analyse trial-by-trial behaviour. To do so, we will implement an autocorrelation analysis, examining the performance of the current trial in the function of what happened the last time a word was encountered. We will investigate the effect of last-target accuracy (H3), target count (H3, H4 and H5) and last-competitor accuracy (H4; H5). Here the null hypotheses are respectively: Participants' accuracy is not predicted by their accuracy on the preceding trial (H3), the use of mutual exclusivity bias does not differ for monolinguals and bilinguals (H4) and for mapping types (H5).

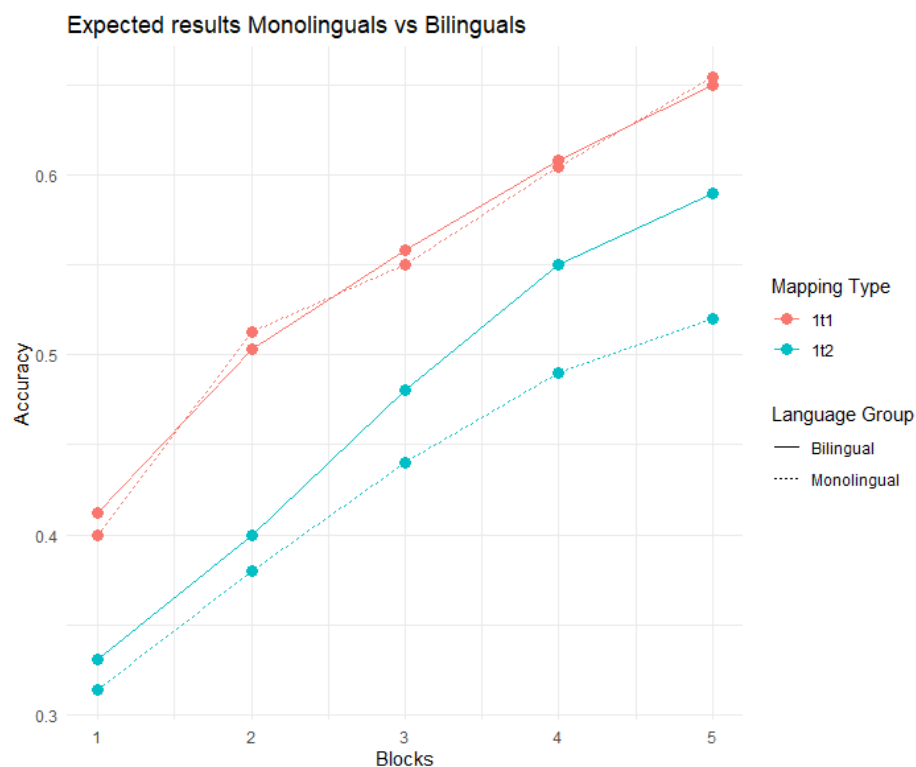
As has been done in previous trial-by-trial analyses (Roembke et al., 2018), we will first add the effect of last-target-accuracy and target count. Thus, the fixed effects will be mapping type (contrast coded), last-target-accuracy (1/0; centered) and target count (log-transformed and centered), language group (monolingual, bilingual; contrast coded as +0.5/-0.5) as well as their interactions. We will again use the buildmer function to find the most appropriate random effect structure; the resulting model will constitute the baseline model for the subsequent analysis. We will first use the baseline model to evaluate the effect of last-target-accuracy and target count (H4). Second, we will add last-competitor accuracy (0, 0.5, 1; centered) to evaluate

H5 and H6. We will here consider only participants that will reach 40% of accuracy during the last block to be sure that we have enough usable trials.

Potential Results and Implications

In the binomial linear mixed model, we predict a main effect of mapping type (with 1:1 mappings learned better than 1:2 mappings, H1) and a main effect of block (accuracy increasing over blocks). We also predict a significant two-way interaction of mapping type and block, where 1:1 mappings are acquired more quickly than 1:2 mappings (H1). This interaction can be due to the inherent difficulty of learning 1:2 instead of 1:1 or to the frequency of exposure (1:1 mappings are shown twice as often as 1:2). Based on Poepel and Weiss' (2016) results, we expect a significant two-way interaction between language group and mapping type. Namely, bilinguals are hypothesised to learn 1:2 mappings better than monolinguals; however, we predict no such difference for 1:1 mappings (H2).

Figure 3: Mock-up of the expected results (H1 and H2)



If we find the expected interaction between language group and mapping type in the model, this can be considered more conclusive evidence that bilinguals outperform monolinguals in statistical word learning but that such bilingual advantage is limited to specific learning circumstances consistent with the bilingual learning experience. Therefore, these findings would be consistent with the notion that bilinguals adopt different learning strategies (such as relaxing the mutual exclusivity assumption), which helps them acquire statistical relationships compatible with multiple mappings per word. These results would be inconsistent with a more general bilingual advantage that extends to all possible mapping types.

If we do not find a difference in statistical word learning between language groups, this could be the result of various reasons: The observed interaction of mapping type \times language group by Poepel and Weiss (2016) could be language-dependent (they used Chinese-English

and English-Spanish bilinguals whilst we use English-German bilinguals), modality-dependent (the stimuli were spoken words whilst here are written words), or due to post-learning evaluation (they did not assess learning continuously, as there was a testing phase after the learning phase). However, there are no a priori reasons why a language group advantage should be observed with one methodological configuration but not another; if anything, the changes made in this experiment should facilitate observing a small learning difference across groups. Thus, if we do not observe a performance difference between language groups (either as an interaction with mapping type or as a main effect), we will conclude that bilinguals are not better at learning words statistically than monolinguals. Given that an advantage for bilinguals over monolinguals is generally observed in other word-learning paradigms (e.g., Bogulski et al., 2019; Kaushanskaya & Marian, 2009b; Nair et al., 2016), not finding a difference between bilinguals and monolinguals in statistical word learning may indicate that statistical learning is a theoretically interesting boundary condition that is unaffected by people's language learning history.

Regarding the trial-by-trial analyses, we predict, consistent with previous studies (Dautriche & Chemla, 2014; Roembke & McMurray, 2016; Trueswell et al., 2013), an effect of last-target accuracy and target-count (H3): accuracy on a current trial will be higher if participants were also accurate on the preceding trial with the same word and the more often a word has been encountered. Moreover, we predict an effect of last-competitor accuracy (H4): We expect that bilinguals will use of the mutual exclusivity bias less than monolinguals, consistent with previous research showing that bilinguals relax the use of the mutual exclusivity assumption during word learning (e.g., Byers-Heinlein & Werker, 2009; Davidson et al., 1997;

Kalashnikova et al., 2015). Lastly, we predict that the use of the mutual exclusivity bias will be reduced for 1:2 compared to 1:1 mappings (H5).

Timeline

We predict it will take two months to collect the data via Prolific. It will take us additional four months to analyse the data and finalise the Stage 2 manuscript.

Question	Hypothesis	Sampling plan	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
<p>H1: Is it easier to learn simple (1:1, one word maps onto one object) or complex (1:2, one word maps onto two objects) mappings?</p>	<p>Learning 1:1 mappings will be easier than learning 1:2 mappings for all participants, performance differences will become more apparent throughout the experiment.</p>	<p>We performed a power analysis using simulation-based power estimation (Kumle, 2021) on a Generalized Mixed Models analysis (DV= Accuracy; IV= Block, Mapping Type and Language Group) on data from a pilot experiment (Nmonolinguals = 25; Nbilinguals = 57; an online experiment with a similar procedure but only one monolingual and one bilingual group).</p> <p>Based on the power analysis to have a medium-big effect size (>.8) on this hypothesis we need</p>	<p>Binomial mixed models DV:</p> <ul style="list-style-type: none"> • Accuracy (1/0) <p>IV:</p> <ul style="list-style-type: none"> • Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), • Block (from 1 to 5; centered) <p>Random effects:</p> <ul style="list-style-type: none"> • Subject, • Target object, • Word 	<p>We will calculate the Bayesian Factor for each analysis and following the literature we will consider a Bayes factors larger than 3 as support for the alternative, and Bayes factors smaller than 1/3 as support for the null model (See Jeffreys, 1998; Lakens et al., 2020).</p>	<p>In favour of the hypothesis: Significant main effect of mapping type, participants are more accurate with 1:1 than 1:2. Effect in the same direction as the literature. 1:1 mappings are easier to learn than 1:2.</p> <p>In contrast to the hypothesis: Significant main effect of mapping type, participants are more accurate with 1:2 than 1:1. Effect in the opposite direction compared to the literature. 1:2 are easier to learn than 1:1. Very implausible.</p> <p>In contrast to the hypothesis:</p>	<p>This hypothesis is a replication of the literature. It can be considered as a check of the actual execution of the manipulation.</p>

		at least 50 participants (0.999).			There is no significant main effect of mapping type, participants are equally accurate with 1:2 and 1:1. In contrast with the literature. There is no difference in accuracy between the mappings.	
H2: Do bilinguals learn words more easily than monolinguals under some circumstances?	Bilinguals will outperform monolinguals when acquiring 1:2 but not 1:1 mappings.	We performed a power analysis using simulation-based power estimation (Kumle, 2021) on a Generalized Mixed Models analysis (DV= Accuracy; IV= Test, Mapping Type and Language Group) on the data from Poepel and Weiss (2016). Based on the power analysis to have a medium-big effect size (>.8) on this hypothesis (Significant two-way interaction of mapping type and language group) we need at least 150 participants (0.842).	Binomial mixed models DV: <ul style="list-style-type: none"> Accuracy (1/0) IV: <ul style="list-style-type: none"> Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), Language group (monolingual vs bilingual; contrast coded as +0.5/-0.5), Block (from 1 to 5; centered) Random effects: <ul style="list-style-type: none"> Subject, Target object, Word 		In favour of the hypothesis: <ul style="list-style-type: none"> Significant two-way interaction of mapping type and language group, bilinguals are more accurate than monolinguals, but only with 1:2 Effect in the same direction as Poepel and Weiss (2016). Bilinguals outperform monolinguals only with more complex (1:2) mappings. <ul style="list-style-type: none"> Significant three-way interaction between block, mapping and language groups. In particular, difference 	If our hypothesis is confirmed, we will replicate Poepel and Weiss (2016) findings with a bigger sample size, with different language combinations and assessing learning continuously. These findings would be consistent with the notion that bilinguals adopt different learning strategies (such as relaxing the mutual exclusivity assumption), which helps them acquire statistical relationships compatible with multiple mappings

					<p>between monolinguals and bilinguals with more complex mappings only at the end. Consistent with a bilingual advantage present only after training.</p> <p>In contrast to the hypothesis: Significant interaction of mapping type and language group in any other direction.</p> <ul style="list-style-type: none"> • Monolinguals are better than bilinguals only for 1:1 mappings • or only for 1:2 mappings. • Bilinguals are better than monolinguals only for 1:1 mappings. <p>Effect in contrast with the literature.</p> <p>In contrast to the hypothesis: Significant main effect of language group, bilinguals are more</p>	<p>per word. These results would be inconsistent with a more general bilingual advantage that extends to all possible mapping types.</p> <p>If, however, we do not find a (limited) bilingual advantage, this would call into question whether a bilingual word learning advantage can be extended to statistical word learning.</p>
--	--	--	--	--	--	--

					<p>accurate than monolinguals. Effect in the direction of a general bilingual advantage Bilinguals outperform monolinguals (similar to Escudero et al. (2016)'s findings).</p> <p>In contrast to the hypothesis: Significant main effect of language group, monolinguals are more accurate than bilinguals. Effect in the direction of a general bilingual disadvantage. Monolinguals outperform bilinguals (similar to Crespo and Kaushanskaya (2021)'s findings).</p> <p>In contrast to the hypothesis: No significant main effect of language group, monolinguals are as accurate as bilinguals. No difference in learning between monolinguals and</p>	
--	--	--	--	--	--	--

					bilinguals (similar to Benitez et al. (2016)'s findings).	
<p>H3: Does accuracy on a previous trial with the same word and target count (how often a word has been encountered) impact accuracy on a current trial?</p>	<p>The accuracy on a current trial will be higher if participants were also accurate on the previous trial with the same word (as indicated by the effect of last-target accuracy) and the more often a word has been encountered (target count)</p>	<p>We performed a power analysis using simulation-based power estimation (Kumle, 2021) on a Generalized Mixed Models analysis (DV= Accuracy; IV= Mapping Type, Language Group, Last-target accuracy and Target Count) on data from a pilot experiment (Nmonolinguals = 25; Nbilinguals = 57; an online experiment with a similar procedure but only one monolingual and one bilingual group). Based on the power analysis to have a medium-big effect size (>.8) on these hypotheses we need:</p> <ul style="list-style-type: none"> • Significant main effect of 	<p>Binomial mixed models DV:</p> <ul style="list-style-type: none"> • Accuracy (1/0) on current trial <p>IV:</p> <ul style="list-style-type: none"> • Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), • Language group (monolingual and bilingual; contrast coded as +0.5/-0.5), • Last-target-accuracy (1/0; centered) • Target count (log-transformed and centered) <p>Random effects:</p> <ul style="list-style-type: none"> • Subject, • Target object, • Word 		<p>In favour of the hypothesis:</p> <ul style="list-style-type: none"> • Significant main effect of Last-target-accuracy • Significant main effect of Target count • Significant interaction of Last-target-accuracy and Target Count <p>Participants are more accurate if they were accurate also in the previous trial with the same word and the more often a word has been encountered. The effect of last-target-accuracy is higher the more often a word has been encountered (in line with the literature, e.g. Dautriche & Chemla, 2014; Roembke & McMurray, 2016; Trueswell et al., 2013).</p>	<p>This hypothesis is a replication of the literature. An effect of last-target-accuracy has been interpreted to be evidence for more explicit learning processes, while an effect of target count has been interpreted as evidence of more implicit learning processes (e.g., Roembke & McMurray, 2016). An absence of these effects would indicate that very little learning occurred.</p>

		<p>Last-target-accuracy: Already with 50 participants we reached a power of 1.</p> <ul style="list-style-type: none"> • Significant main effect of Target count: Already with 50 participants we reached a power of 1. • Significant interaction of Last-target-accuracy and Target Count: Already with 50 participants we overtake a power of .8 (.894). 			<p>In contrast to the hypothesis:</p> <ul style="list-style-type: none"> • No significant main effect of Last-target-accuracy <p>Participants are not more accurate if they were accurate also in the previous trial with the same word (in contrast with the literature).</p> <p>In contrast to the hypothesis:</p> <ul style="list-style-type: none"> • No significant main effect of Target count <p>Participants are not more accurate the more often a word has been encountered (In contrast with the literature).</p> <p>In contrast to the hypothesis:</p> <ul style="list-style-type: none"> • No significant interaction of last-target-accuracy and target count <p>The effect of last-target-accuracy does not increase with</p>	
--	--	---	--	--	--	--

					higher target count (in contrast to literature).	
H4: Will the use of mutual exclusivity bias be reduced for bilinguals?	The mutual exclusivity bias (as indicated by the effect of last-competitor accuracy) will be reduced for bilinguals for all mapping types compared to monolinguals.	We performed a power analysis using simulation-based power estimation (Kumle, 2021) on a Generalized Mixed Models analysis (DV= Accuracy; IV= Mapping Type, Language Group, Last-competitor accuracy and Target Count) on data from a pilot experiment (Nmonolinguals = 25; Nbilinguals = 57; an online experiment with a similar procedure but only one monolingual and one bilingual group). Based on the power analysis to have a medium-big effect size (>.8) on this hypothesis we need:	<p>Binomial mixed models</p> <p>DV:</p> <ul style="list-style-type: none"> Accuracy (1/0) on a current trial <p>IV:</p> <ul style="list-style-type: none"> Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), Language group (monolingual and bilingual; contrast coded as +0.5/-0.5), Last-competitor accuracy (0, 0.5, 1; centered), Target count (log-transformed and centered) <p>Random effects:</p> <ul style="list-style-type: none"> Subject, Target object, Word 		<p>In favour of the hypothesis:</p> <ul style="list-style-type: none"> Significant interaction of Last-competitor accuracy and Language group <p>Bilinguals use less mutual exclusivity bias than monolinguals. In line with the literature (e.g., Byers-Heinlein & Werker, 2009; Davidson et al., 1997; Kalashnikova et al., 2015).</p> <p>In contrast of the hypothesis:</p> <ul style="list-style-type: none"> No significant interaction of Last-competitor accuracy and Language group, Or significant interaction, but with monolinguals showing less of an effect of last-competitor accuracy than bilinguals. 	<p>If bilinguals will show less of an effect of last-competitor accuracy than monolinguals, this is consistent with previous research showing that bilinguals relax the use of the mutual exclusivity assumption during word learning (e.g., Byers-Heinlein & Werker, 2009; Davidson et al., 1997; Kalashnikova et al., 2015). It has also been hypothesised that this is why a limited bilingual word learning advantage can be observed when acquiring complex word mappings (Poepsel & Weiss, 2016).</p> <p>However, if we do not find such an</p>

		<ul style="list-style-type: none"> Significant interaction of Last-competitor accuracy and Language group: 			In contrast with the literature (e.g., Byers-Heinlein & Werker, 2009; Davidson et al., 1997; Kalashnikova et al., 2015).	interaction between the effect of last-competitor accuracy and language group, it calls into question whether a relaxation of the mutual exclusivity bias underlies any existing bilingual word learning advantage.
H5: Does the use of mutual exclusivity change between mapping types?	The use of the mutual exclusivity bias (as indicated by the effect of last-competitor accuracy) will be reduced for complex (1:2) mappings compared to simple (1:1) mappings.	We performed a power analysis using simulation-based power estimation (Kumle, 2021) on a Generalized Mixed Models analysis (DV= Accuracy; IV= Mapping Type, Language Group, Last-competitor accuracy and Target Count) on data from a pilot experiment (Nmonolinguals = 25; Nbilinguals = 57; an online experiment with a similar procedure but only one monolingual and	Binomial mixed models DV: <ul style="list-style-type: none"> Accuracy (1/0) on the current trial IV: <ul style="list-style-type: none"> Mapping type (1:1 or 1:2; contrast coded as +0.5/-0.5), Language group (monolingual and bilingual; contrast coded as +0.5/-0.5), Last-competitor accuracy (0, 0.5, 1; centered), 		In favour of the hypothesis: <ul style="list-style-type: none"> Significant two-way interaction of Last-competitor accuracy and mapping type Participants use the mutual exclusivity bias less with 1:2 than with 1:1 mappings. In line with the fact that with 1:2 mappings, participants need to relax the use of mutual exclusivity bias when learning the second meaning. <ul style="list-style-type: none"> Significant three-way interaction with mapping type, language 	A possible explanation for the relaxation of the mutual exclusivity of bilinguals is that they are more exposed to complex mappings than monolinguals. If an interaction between the effect of last-competitor accuracy and mapping type can be observed, this supports that exposure to more complex mappings impacts how the mutual exclusivity bias is applied (independently of

		<p>one bilingual group). Based on the power analysis to have a medium-big effect size (>.8) on this hypothesis we need:</p> <ul style="list-style-type: none"> • Significant two-way interaction of Last-competitor accuracy and mapping type: We need 100 participants to have a power superior of .8 (.934). 	<ul style="list-style-type: none"> • Target count (log-transformed and centered) <p>Random effects:</p> <ul style="list-style-type: none"> • Subject, • Target object, • Word 		<p>group and target count. Coherent with a bilingual advantage due to training.</p> <p>In contrast to the hypothesis:</p> <ul style="list-style-type: none"> • No significant interaction of Last-competitor accuracy and mapping type, • Or significant interaction, but participants use the mutual exclusivity bias less with 1:1 than with 1:2 mappings. <p>With 1:1 mappings, participants need to relax the use of the mutual exclusivity bias when learning the second meaning. Implausible.</p>	<p>language group status).</p> <p>If, however, no such interaction is observed, this calls into question whether experience with more complex mapping types impacts how the mutual exclusivity bias is applied.</p>
--	--	--	---	--	---	---

Supplementary material:

Appendix A1: LEAP-Q Bilinguals

Language Experience and Proficiency Questionnaire (LEAP-Q) - 2 Languages

Please list all the languages you know **in order of dominance**. Start with the language you are most proficient in (Language 1).

Language 1

Please select...



Language 2

Please select...



Language 3

Please select...



Language 4

Please select...



Please list all the languages you know **in order of acquisition** (your native language first).

early

⌘: \${Dominance_Language1}

⌘: \${Dominance_Language2}

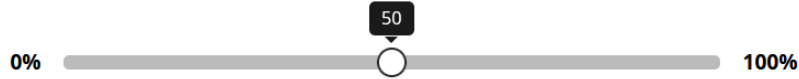
⌘: \${Dominance_Language3}

⌘: \${Dominance_Language4}

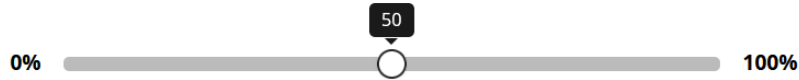
late

Please list what percentage of time you are currently on average exposed to each language. Your percentages should add up to 100%.

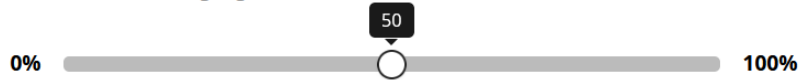
\$\$ {Dominance_Language1}



\$\$ {Dominance_Language2}



\$\$ {Dominance_Language3}

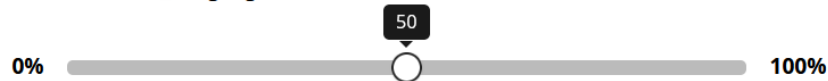


\$\$ {Dominance_Language4}

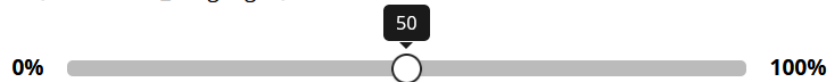


When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you. Your percentages should add up to 100%

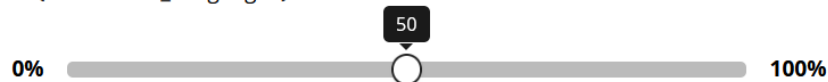
\$\$ {Dominance_Language1}



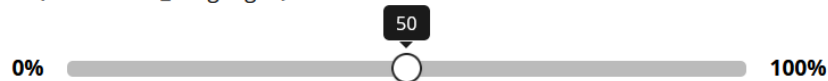
\$\$ {Dominance_Language2}



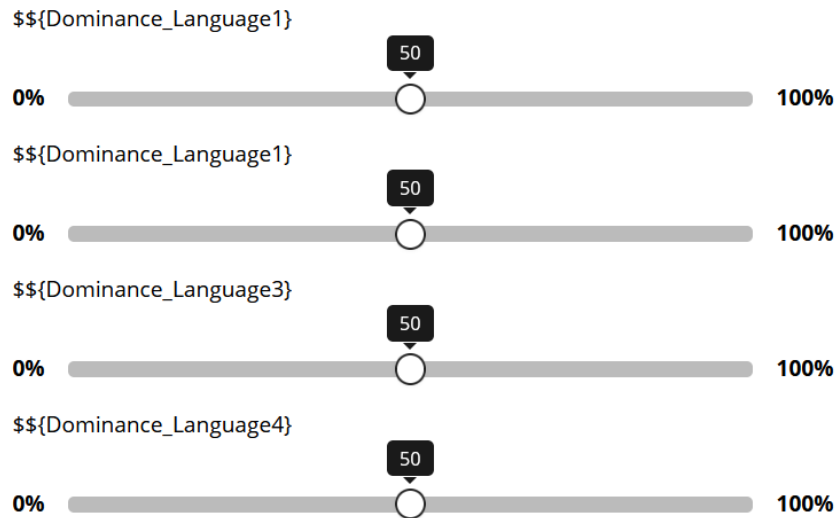
\$\$ {Dominance_Language3}



\$\$ {Dominance_Language4}



When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you. Your percentages should add up to 100%



Language Experience and Proficiency Questionnaire (LEAP-Q)

All questions below refer to your knowledge of **\$\$ {Dominance_Language1}**

\$\$ {Dominance_Language1} is my ...

Age when you began acquiring \$\$ {Dominance_Language1}

Age when you became fluent in \$\$ {Dominance_Language1}

Age when you began reading in \$\$ {Dominance_Language1}

Age when you became fluent reading in \$\$ {Dominance_Language1}

How many years and months have you spent in a country where $\\{\text{Dominance_Language1}\\}$ is spoken?

Years



Months



How many years and months have you spent in a family where $\\{\text{Dominance_Language1}\\}$ is spoken?

Years



Months



How many years and months have you spent in a school and/or working environment where $\\{\text{Dominance_Language1}\\}$ is spoken?

Years



Months



On a scale from zero to ten, please select your level of proficiency in speaking, understanding, and reading $\\{\text{Dominance_Language1}\\}$ from the scroll-down menus:

Speaking

Please Select...



Understanding spoken language

Please Select...



Reading

Please Select...



On a scale from zero to ten, please select how much the following factors contributed to you learning $\text{\$}\{Dominance_Language1\}$:

Interacting with friends

Interacting with family

Interacting with colleagues/at work

Reading

Self instruction (e.g., Duolingo)

Watching TV or streaming (e.g., Netflix)

Listening to the radio, podcast or music

Social media (e.g., Instagram or TikTok)

Please rate to what extent you are currently exposed to $\\$\\{Dominance_Language1\\}$ in the following contexts:

Interacting with friends

Please Select...



Interacting with family

Please Select...



Interacting with colleagues/at work

Please Select...



Reading

Please Select...



Self-instruction (e.g., Duolingo)

Please Select...



Watching TV or streaming (e.g., Netflix)

Please Select...



Listening to the radio, podcast or music

Please Select...



Social media (e.g., Instagram or TikTok)

Please Select...



In your perception, how much of a foreign accent do you have in $\\$\\{Dominance_Language1\\}$?

Please Select...



Please rate how frequently others identify you as non-native speaker based on your accent in $\\$\\{Dominance_Language1\\}$?

Please Select...



Appendix **A2**: LEAP-Q Monolinguals

Language Experience and Proficiency Questionnaire (LEAP-Q)

All questions below refer to your knowledge of **English**

Age when you began acquiring English

Please select age... 

Age when you became fluent in English

Please select age... 

Age when you began reading in English

Please select age... 

Age when you became fluent reading in English

Please select age... 

How many years and months have you spent in a country where English is spoken?

Years 

Months 

How many years and months have you spent in a family where English is spoken?

Years 

Months 

How many years and months have you spent in a school and/or working environment where English is spoken?

Years 

Months 

On a scale from zero to ten, please select your level of proficiency in speaking, understanding, and reading English from the scroll-down menus:


Speaking

Understanding spoken language

Reading

Appendix B:

	German	English	Together	Delta	
Patchet	4,25	1,9	2,57	2,35	
Tallot	4,25	2,2	2,79	2,05	
Dungle	4	2	2,57	2	
Jeplin	4	2,3	2,79	1,7	
Churge	3,5	1,8	2,29	1,7	
Masset	4,25	2,7	3,14	1,55	
Conick	3,75	2,2	2,64	1,55	
Lanter	4,25	2,9	3,29	1,35	
Sumper	4	3	3,29	1	*
Vartion	3,5	2,7	2,93	0,8	*
Linchet	3	2,4	2,57	0,6	
Bernal	3,75	3,2	3,36	0,55	*
Murler	3,5	3	3,14	0,5	*
Grinter	3,5	3	3,14	0,5	*
Frectop	2,25	1,9	2,00	0,35	
Staunker	3,25	3	3,07	0,25	*
Hullor	2,5	2,3	2,36	0,2	
Fassick	2,75	2,6	2,64	0,15	
Darlon	3	2,9	2,93	0,1	*
Wilten	3	2,9	2,93	0,1	*
Rauplet	2,75	2,7	2,71	0,05	*
Thernus	3,25	3,3	3,29	-0,05	*
Malfen	2,75	3,2	3,07	-0,45	*
Sterhin	2	2,7	2,50	-0,7	
Glanke	3	3,7	3,50	-0,7	*
Nintal	2,25	3,1	2,86	-0,85	
Rorgell	1,75	2,6	2,36	-0,85	
Wachelt	2	3,3	2,93	-1,3	
Zugelt	2	3,7	3,21	-1,7	

References

- Aguasvivas, J. A., Cespón, J., & Carreiras, M. (2024). Does bilingual experience influence statistical language learning? *Cognition*, *242*, 105639.
- Anderson, J. A. E., Grundy, J. G., Grady, C. L., Craik, F. I. M., & Bialystok, E. (2021). Bilingualism contributes to reserve and working memory efficiency: Evidence from structural and functional neuroimaging. *Neuropsychologia*, *163*, 108071.
<https://doi.org/10.1016/j.neuropsychologia.2021.108071>
- Antoniou, M., Liang, E., Ettliger, M., & Wong, P. C. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, *18*(4), 683–695.
<https://doi.org/10.1017/S1366728914000777>
- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407.
<https://doi.org/10.3758/s13428-019-01237-x>
- Bartolotti, J., & Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cognitive Science*, *36*(6), 1129–1147. <https://doi.org/10.1111/j.1551-6709.2012.01243.x>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*. <https://doi.org/10.48550/arXiv.1406.5823>
- Benitez, V. L., Yurovsky, D., & Smith, L. B. (2016). Competition between multiple words for a referent in cross-situational word learning. *Journal of Memory and Language*, *90*, 31–48.
<https://doi.org/10.1016/j.jml.2016.03.004>
- Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*(1), 67–73.
<https://doi.org/10.1016/j.cognition.2011.08.003>

- Bialystok, E., & Craik, F. I. M. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review*, 29(4), 1246–1269. <https://doi.org/10.3758/s13423-022-02057-5>
- Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, 22(5), 1052–1067. <https://doi.org/10.1017/S1366728918000858>
- Borragan, M., De Bruin, A., Havas, V., de Diego-Balaguer, R., Vulchanova, M. D., Vulchanov, V., & Duñabeitia, J. A. (2021). Differences in word learning in children: Bilingualism or linguistic experience? *Applied Psycholinguistics*, 42(2), 345–366. <https://doi.org/10.1017/S0142716420000594>
- Bulgarelli, F., Weiss, D. J., & Dennis, N. A. (2021). Cross-situational statistical learning in younger and older adults. *Aging, Neuropsychology, and Cognition*, 28(3), 346–366. <https://doi.org/10.1080/13825585.2020.1759502>
- Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, 12(5), 815–823. <https://doi.org/10.1111/j.1467-7687.2009.00902.x>
- Bylund, E., Antfolk, J., Abrahamsson, N., Olstad, A. M. H., Norrman, G., & Lehtonen, M. (2022). Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-022-02136-7>
- Chen, C., Gershkoff-Stowe, L., Wu, C., Cheung, H., & Yu, C. (2017). Tracking multiple statistics: Simultaneous learning of object names and categories in English and Mandarin speakers. *Cognitive Science*, 41(6), 1485–1509. <https://doi.org/10.1111/cogs.12417>
- Chung-Fat-Yim, A., Sorge, G. B., & Bialystok, E. (2017). The relationship between bilingualism and selective attention in young adults: Evidence from an ambiguous figures task. *The Quarterly*

Journal of Experimental Psychology Section A, 70(3), 366–372.

<https://doi.org/10.1080/17470218.2016.1221435>

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, 113(2), 135–149.

Crespo, K., & Kaushanskaya, M. (2021). Is 10 better than 1? The effect of speaker variability on children's cross-situational word learning. *Language Learning and Development*, 17(4), 397–410.

<https://doi.org/10.1080/15475441.2021.1906680>

Crespo, K., Vlach, H., & Kaushanskaya, M. (2023). The effects of bilingualism on children's cross-situational word learning under different variability conditions. *Journal of Experimental Child Psychology*, 229, 105621. <https://doi.org/10.1016/j.jecp.2022.105621>

Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 892–902.

<https://doi.org/10.1037/a0035657>

Davidson, D., Jergovic, D., Imami, Z., & Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of Child Language*, 24(1), 3–24.

<https://doi.org/10.1017/S0305000996002917>

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. <https://doi.org/10.1017/S0142716412000744>

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour*, 3(7), 692–701. <https://doi.org/10.1038/s41562-019-0609-3>

- Donnelly, S., Brooks, P. J., & Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review*, 26(4), 1122–1147. <https://doi.org/10.3758/s13423-019>
- Ehl, B., Bruns, G., & Grosche, M. (2020). Differentiated bilingual vocabulary assessment reveals similarities and differences compared to monolinguals: Conceptual versus single-language scoring and the relation with home language and literacy activities. *International Journal of Bilingualism*, 24(4), 715–728. <https://doi.org/10.1177/1367006919876994>
- Escudero, P., Mulak, K. E., Fu, C. S. L., & Singh, L. (2016). More limitations to monolingualism: Bilinguals outperform monolinguals in implicit word learning. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01218>
- Escudero, P., Smit, E. A., & Angwin, A. J. (2023). Investigating orthographic versus auditory cross-situational word learning with online and laboratory-based testing. *Language Learning*, 73, 543–577. <https://doi.org/10.1111/lang.12550>
- Eviatar, Z., Taha, H., & Shwartz, M. (2018). Metalinguistic awareness and literacy among semitic-bilingual learners: A cross-language perspective. *Reading and Writing*, 31(8), 1869–1891. <https://doi.org/10.1007/s11145-018-9850-9>
- Grosjean, F. (2010). Bilingual. In *Bilingual*. Harvard university press.
- Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23(2), 283–294. <https://doi.org/10.1017/S1366728919000026>
- Hartanto, A., & Yang, H. (2019). Does early active bilingualism enhance inhibitory control and monitoring? A propensity-matching analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(2), 360. <https://doi.org/10.1037/xlm0000581>

- Hirosh, Z., & Degani, T. (2021). Novel word learning among bilinguals can be better through the (dominant) first language than through the second language. *Language Learning*, 71(4), 1044–1084. <https://doi.org/10.1111/lang.12457>
- Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, 48(4), 1393–1409. <https://doi.org/10.3758/s13428-015-0647-3>
- Hulme, R. C., Shapiro, L. R., & Taylor, J. S. H. (2022). Learning new words through reading: Do robust spelling–sound mappings boost learning of word forms and meanings? *Royal Society Open Science*, 9(12), 210555. <https://doi.org/10.1098/rsos.210555>
- Jeffreys, H. (1998). *The theory of probability*. OuP Oxford.
https://books.google.com/books?hl=it&lr=&id=vh9Act9rtzQC&oi=fnd&pg=PA1&dq=Theory+of+probability&ots=ffYBGT05iZ&sig=_B1p3VoHmd6f2dLq7NlpMTVZJhY
- Kalashnikova, M., Mattock, K., & Monaghan, P. (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism: Language and Cognition*, 18(4), 626–638. <https://doi.org/10.1017/S1366728914000364>
- Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*, 15(3), 470–489. <https://doi.org/10.1017/S1366728911000472>
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2020). The language experience and proficiency questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 23(5), 945–950. <https://doi.org/10.1017/S1366728919000038>
- Kaushanskaya, M., & Marian, V. (2009a). Bilingualism reduces native-language interference during novel-word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 829–835. <https://doi.org/10.1037/a0015275>

- Kaushanskaya, M., & Marian, V. (2009b). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705–710. <https://doi.org/10.3758/PBR.16.4.705>
- Kumle, L., Vö, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, *53*(6), 2528–2543. <https://doi.org/10.3758/s13428-021-01546-0>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B*, *75*(1), 45–57.
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394. <https://doi.org/10.1037/bul0000142>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Li, Y., & Benitez, V. (in press). Lexical tone as a cue in statistical word learning from bilingual input. *Bilingualism: Language and Cognition*. <https://osf.io/9kuy4/download>
- Mangardich, H., & Sabbagh, M. A. (2022). Event-related potential studies of cross-situational word learning in four-year-old children. *Journal of Experimental Child Psychology*, *222*, 105468. <https://doi.org/10.1016/j.jecp.2022.105468>
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157. [https://doi.org/10.1016/0010-0285\(88\)90017-5](https://doi.org/10.1016/0010-0285(88)90017-5)

- McGregor, K. K., Smolak, E., Jones, M., Oleson, J., Eden, N., Arbisi-Kelm, T., & Pomper, R. (2022). What children with developmental language disorder teach us about cross-situational word learning. *Cognitive Science*, *46*(2), e13094. <https://doi.org/10.1111/cogs.13094>
- McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831–877. <https://doi.org/10.1037/a0029872>
- Montanari, E. G., Abel, R., Graßer, B., & Tschudinovski, L. (2018). Do bilinguals create two different sets of vocabulary for two domains?: Vocabulary development and overlap in the first years of schooling. *Linguistic Approaches to Bilingualism*, *8*(4), 502–522. <https://doi.org/10.1075/lab.16021.mon>
- Mulak, K. E., Vlach, H. A., & Escudero, P. (2019). Cross-situational learning of phonologically overlapping words across degrees of ambiguity. *Cognitive Science*, *43*(5), e12731. <https://doi.org/10.1111/cogs.12731>
- Nair, V. K., Biedermann, B., & Nickels, L. (2016). Consequences of late bilingualism for novel word learning: Evidence from Tamil–English bilingual speakers. *International Journal of Bilingualism*, *20*(4), 473–487. <https://doi.org/10.1177/1367006914567005>
- Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science*, *31*(5), 548–567. <https://doi.org/10.1177/0956797620903113>
- Olguin, A., Cekic, M., Bekinschtein, T. A., Katsos, N., & Bozic, M. (2019). Bilingualism and language similarity modify the neural mechanisms of selective attention. *Scientific Reports*, *9*(1), 8204. <https://doi.org/10.1038/s41598-019-44782-3>

- Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *The Quarterly Journal of Experimental Psychology Section A*, *48*(1), 98–107.
<https://doi.org/10.1080/14640749508401378>
- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, *152*, 9–19. <https://doi.org/10.1016/j.cognition.2016.03.001>
- Quine, W. (1960). *Word and object: An inquiry into the linguistic mechanisms of objective reference* (pp. xv, 294). John Wiley.
- Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, *87*, 105–127.
<https://doi.org/10.1016/j.jml.2015.09.005>
- Roembke, T. C., & McMurray, B. (2021). Multiple components of statistical word learning are resource dependent: Evidence from a dual-task learning paradigm. *Memory & Cognition*, *49*(5), 984–997.
<https://doi.org/10.3758/s13421-021-01141-w>
- Roembke, T. C., Simonetti, M. E., Koch, I., & Philipp, A. M. (2023). What have we learned from 15 years of research on cross-situational word learning? A focused review. *Frontiers in Psychology*, *14*.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10400455/>
- Roembke, T. C., Wiggs, K. K., & McMurray, B. (2018). Symbolic flexibility during unsupervised word learning in children and adults. *Journal of Experimental Child Psychology*, *175*, 17–36.
<https://doi.org/10.1016/j.jecp.2018.05.016>
- Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301–308. <https://doi.org/10.3758/s13423-014-0595-4>
- Seçer, I. (2016). Skills of cognitive flexibility in monolingual and bilingual younger adults. *The Journal of General Psychology*, *143*(3), 172–184. <https://doi.org/10.1080/00221309.2016.1200530>

Silvey, C., Dienes, Z., & Wonnacott, E. (2021). *Bayes factors for mixed-effects models*.

<https://osf.io/m4hju/download>

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1–2), 39–91. [https://doi.org/10.1016/S0010-0277\(96\)00728-7](https://doi.org/10.1016/S0010-0277(96)00728-7)

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. <https://doi.org/10.1016/j.cognition.2007.06.010>

Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology*, *126*, 395–411.

<https://doi.org/10.1016/j.jecp.2014.06.003>

Sullivan, M. D., Janus, M., Moreno, S., Astheimer, L., & Bialystok, E. (2014). Early stage second-language learning improves executive control: Evidence from ERP. *Brain and language*, *139*, 84–98.

<https://doi.org/10.1016/j.bandl.2014.10.004>

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–156.

<https://doi.org/10.1016/j.cogpsych.2012.10.001>

Tuninetti, A., Mulak, K. E., & Escudero, P. (2020). Cross-situational word learning in two foreign languages: Effects of native language and perceptual difficulty. *Frontiers in Communication*, *5*, 602471. <https://doi.org/10.3389/fcomm.2020.602471>

Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, *47*(3), 507–546. <https://doi.org/10.1111/0023-8333.00018>

Verreyt, N., Woumans, E. V. Y., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, *19*(1), 181–190.

- Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *45*(4), 700–711. <https://doi.org/10.1037/xlm0000611>
- Voeten, C. C. (2019). Using 'buildmer' to automatically find & compare maximal (mixed) models. *R package version*, *1*(6), 1–7.
- Wang, F. H. (2020). Explicit and implicit memory representations in cross-situational word learning. *Cognition*, *205*, 104444. <https://doi.org/10.1016/j.cognition.2020.104444>
- Ware, A. T., Kirkovski, M., & Lum, J. A. G. (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Frontiers in Psychology*, *11*, 1458. <https://doi.org/10.3389/fpsyg.2020.01458>
- Warmington, M. A., Kandru-Pothineni, S., & Hitch, G. J. (2019). Novel-word learning, executive control and working memory: A bilingual advantage. *Bilingualism: Language and Cognition*, *22*(4), 763–782. <https://doi.org/10.1017/S136672891800041X>
- Weiss, D. J., Schwob, N., & Lebkuecher, A. L. (2020). Bilingualism and statistical learning: Lessons from studies using artificial languages. *Bilingualism: Language and Cognition*, *23*(1), 92–97. <https://doi.org/10.1017/S1366728919000579>
- Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420. <https://doi.org/10.1111/j.1467-9280.2007.01915.x>
- Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53–62. <https://doi.org/10.1016/j.cognition.2015.07.013>