Bilingualism: Language and Cognition

cambridge.org/bil

Lexical tone as a cue in statistical word learning from bilingual input

Ye Li and Viridiana L. Benitez

Department of Psychology, Arizona State University, Tempe, United States

Research Article

Cite this article: Li, Y., & Benitez, V.L. (2024). Lexical tone as a cue in statistical word learning from bilingual input. *Bilingualism: Language and Cognition,* **27**, 886–900. https://doi.org/10.1017/S1366728923000858

Received: 4 May 2022
Revised: 13 September 2023
Accepted: 3 November 2023
First published online: 12 December2023

Keywords:
statistical word learning; lexical tones; bilingualism; language familiarity

Corresponding author:
Ye Li; Email: yeli7@asu.edu

Abstract

Learners can track word-referent co-occurrences across individually-ambiguous naming events to form correct word-referent mappings, termed statistical word learning (SWL). Prior research largely focuses on learning from a single language input, where a referent co-occurs with a single word (1:1 mapping). Here, we tested adults' SWL from a simulated bilingual environment, where one referent co-occurred with two words (2:1 mapping) and the two words were either differentiated by a linguistic cue (Mandarin lexical tones, Cued condition) or not (Uncued condition). Results showed that in the Cued condition, Chinese–English bilinguals ($N = 38$) outperformed Spanish–English bilinguals ($N = 56$) and English monolinguals ($N = 55$), while Spanish–English bilinguals and English monolinguals performed similarly. The three groups did not differ in the Uncued condition. Self-reported learning confidence and strategies showed limited conscious awareness of learning. Results demonstrate that familiarity with a linguistic cue boosts overall statistical word learning from bilingual input.

1. Introduction

Statistical learning, the ability to track probabilistic regularities in sensory input, has been proposed as key for language acquisition, including grammar learning (Gomez & Gerken, 1999), segmenting speech (Saffran et al., 1996), and linking words with referents (L. B. Smith & Yu, 2008; Yu & Smith, 2007). However, statistical learning research has predominantly addressed language learning of a single and invariant input. In a bilingual environment, everyday language experiences can vary in that learners encounter multiple languages, across changing scenes, and with linguistic variations between languages. Statistical learning theories therefore need to incorporate learners' abilities to deal with multiple, changing, and varied inputs (Benitez et al., 2016, 2020a; Byers-Heinlein, 2014; Crespo & Kaushanskaya, 2021; Crespo et al., 2023; Poepsel & Weiss, 2016; Qian et al., 2012; Tsui et al., 2021; Weiss et al., 2009, 2020). In this paper, we provide a test of adults' statistical word learning from bilingual input by investigating how word learning is affected by a linguistic cue (lexical tone) differentiating two languages and learners' language experience.

1.1. Statistical word learning (SWL)

Word learning often happens under ambiguity: words are heard in the context of a number of potential referents, with limited cues to track which words refer to which referents (Medina et al., 2011; Quine, 1960; Yu & Smith, 2007). There are many accounts for how learners can resolve the problem of referential ambiguity (Baldwin, 1993; Hollich et al., 2000; Kucker et al., 2015; Markman, 1990; Medina et al., 2011; Trueswell et al., 2013). One prominent account, termed statistical word learning (SWL), posits that learners can resolve word-referent ambiguity by employing a form of statistical calculation and aggregating the co-occurrences between words and referents across multiple individually-ambiguous learning events (L. B. Smith & Yu, 2008; Yu & Smith, 2007). In the first test of SWL, Yu and Smith (2007) instructed adults to map artificial words with novel objects. Within a trial, several auditory words were presented with an equal number of objects without a clear indication as to which word referred to which object. Across trials, however, each word occurred consistently with a single target object, and less consistently with distractor objects. Results showed that adults aggregated the word-referent co-occurrences across trials and learned the correct word-referent mappings. To date, a large literature has replicated this effect in adults, and demonstrated that children and infants also utilize such statistical co-occurrences to identify word-referent mappings from ambiguous naming events (Alt et al., 2014; Benitez & Li, 2023; Benitez et al., 2020b; Crespo & Kaushanskaya, 2021; Crespo et al., 2023; L. B. Smith

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

& Yu, 2008; K. Smith et al., 2011; Suanda et al., 2014; Vlach & DeBrock, 2017; Vlach & Johnson, 2013; Vouloumanos & Werker, 2009; Yu & Smith, 2007, 2011; Yurovsky & Frank, 2015; Yurovsky & Yu, 2008; Zettersten et al., 2018).

1.2. SWL in a bilingual environment

Critically, a majority of SWL work has focused on acquiring one-to-one word-referent mappings, where a referent co-occurs consistently with a single word (1:1 MAPPING). However, for more than half of the world's population who speaks more than one language (Romaine, 2012), learners can routinely encounter overlapping mappings such as translation equivalents, where each referent refers to two words, each from a different language (2:1 MAPPING). For example, a bilingual learner of English and Mandarin Chinese must learn that the English word "*shoe*" and the Mandarin Chinese word "*xíezi*" both refer to **shoe**. Although monolinguals may occasionally come across overlapping mappings within a language (e.g., synonyms), for bilinguals, translation equivalents occur more frequently and present linguistic variations. How do learners accommodate SWL of overlapping mappings in a bilingual environment?

Answering this question involves not only understanding how learners accommodate 2:1 mappings, but also how between-language cues may affect learning. In a bilingual environment, words from each language are recognized as distinctive units with the help of ample cues, including contextual cues such as a change of interlocutors (Evans, 2011), pauses between transitions (Bhatt, 1997; but see Lyu et al., 2010), or a shift in fundamental frequency (Keating & Kuo, 2012); and, more importantly and commonplace, linguistic cues highlighting cross-linguistic differences in phonotactic structures, phonetics, and prosody (Fabiano-Smith & Goldstein, 2010; Torres Cacoullos, 2020). Linguistic cues can be a direct, salient, and robust signal of the presence of two languages, which may in turn facilitate statistical learning from multiple language inputs (Poepsel & Weiss, 2014; Weiss et al., 2009). Understanding how statistical word learning interacts with linguistic cues is critical not only for honing theories of statistical learning under variability, but also for unveiling which properties of language inputs matter for learners' ability to track surrounding regularities.

The question of how learners accommodate SWL of multiple mappings has been addressed in a limited set of studies, demonstrating that learning 2:1 mappings is more challenging than learning 1:1 mappings (Benitez & Li, 2023; Benitez et al., 2016; Chan & Monaghan, 2019; Ichinco et al., 2009; Kachergis et al., 2012). However, these studies did not include cues to signal the presence of multiple languages. That is, the words sharing a referent were not linguistically differentiated in the above studies, such that the input was more akin to synonym learning within a language, rather than translation equivalents across languages. Benitez et al. (2016) provided preliminary evidence on how a linguistic cue may impact SWL of structure containing 2:1 mappings. Researchers presented monolingual and bilingual adults with an SWL task consisting of 1:1 and 2:1 mappings. Importantly, they examined how an artificial phonotactic cue differentiating the two words sharing a referent affected learning. For the 2:1 mappings, one word followed a constant-vowel-consonant-vowel structure (CVCV, e.g., "*gaso*"), while the other followed a consonant-vowel structure with a /k/ ending (CV-/k/, e.g., "*meek*"). Results showed that learning 2:1 mappings interacted with language experience: the phonotactic cue facilitated 2:1 learning for bilinguals but not for monolinguals. However, despite the phonotactic manipulation, both words for a referent were still English-like pseudowords, and therefore resembled input from a single language. An open question remains: how do learners aggregate 2:1 mappings in a bilingual environment when a linguistic cue signals different language sources? The current study presents the first test of SWL in a simulated bilingual environment, by employing lexical tone as a linguistic cue to differentiate words sharing a referent and mimic word inventories from two languages.

1.3. Lexical tone as a linguistic cue

Lexical tones in tonal languages refer to the pitch variation at a syllabic level to represent distinctive referential meanings (Antoniou & Chin, 2018; Wang & Saffran, 2014; Yip, 2002). For instance, the Mandarin Chinese monosyllabic "*ma*" refers to distinctive referents when embedded with different pitch contours: "*ma*" stands for **mother** with a flat tone (Tone 1), for **hemp** with a rising tone (Tone 2), for **horse** with a dip tone (Tone 3), and for **criticize** with a falling tone (Tone 4) (C. Chen et al., 2016). In short, in tonal languages, pitch variations at the syllabic level are lexically contrastive (Hay et al., 2015).

We chose lexical tone as the cue to differentiate language sources for several reasons. First, lexical tone is widely used in the world's languages: about sixty to seventy percent of the world languages are tonal (Yip, 2002), such as East Asian languages (e.g., Vietnamese and Mandarin Chinese) and a majority of African languages (e.g., Nilo-Saharan). Second, many tonal language speakers grow up bilingual with the other language being non-tonal. For instance, most Mandarin-Chinese speaking children grow up learning English (a non-tonal language) as a required second language in the educational system (Feng, 2007). Thus, lexical tone can be representative as a distinctive linguistic marker to differentiate tonal and non-tonal language input for a large group of bilinguals.

Third, although pitch changes can signal a change in semantic contexts for both tonal and non-tonal speakers, only speakers of tonal languages use lexical tone contrastively, i.e., as a signal for referential change. For instance, in English, different intonations embedded onto the same word "*car*" in an imperative ("*Give me your car!*") and in a question ("*Is this your car?*") may convey different pragmatic inference: one as a request and the other as a moderate question (Bolinger, 1989; Tomlinson & Bott, 2013). Yet, the referential meaning of **car** does not change in both cases; thus, a pitch change in this case is not lexically contrastive. A lexical tone cue is therefore a convenient, suprasegmental, and acoustic cue that can be perceived by both tonal and non-tonal speakers (S. Chen et al., 2017, 2020), but only tonal speakers use as a signal of a referential change (Hay et al., 2015; Singh & Foong, 2012).

Fourth, lexical tone can be added onto the base syllables of novel words while keeping other linguistic properties constant. This creates INCONGRUENT INVENTORIES (Gebhart et al., 2009; Weiss et al., 2009) for a more ecologically valid bilingual environment: the inventories of two languages usually share some properties (e.g., vowels or consonants) but stay distinctive in others (e.g., prosody). This allowed us to develop novel word items that shared consonants, vowels, and the syllabic structure that are acceptable across different languages (CVCV syllabic structure, which is present in English, Spanish, and Mandarin Chinese – the languages of the participants in the study) but differed in whether or not they were embedded with lexical tones (e.g., "*migu*" and "*gádi*").

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

Finally, using lexical tone additionally enabled us to examine how language experience may interact with word learning. On the one hand, being familiar with lexical tones has been found to affect learning of natural and artificial language input containing lexical tone information (Hay et al., 2015; Potter et al., 2017; Singh & Fu, 2016; Singh et al., 2016; Wang & Saffran, 2014). This suggests that experience with tonal languages may provide a LANGUAGE-SPECIFIC ADVANTAGE to learning only in conditions that contain lexical tone information. On the other hand, previous research on statistical word learning has demonstrated that bilingual experience in general provides benefits for SWL (Chan & Monaghan, 2019; Escudero et al., 2016; Poepsel & Weiss, 2016). This suggests that experience with multiple languages may generate a LANGUAGE-GENERAL ADVANTAGE on statistical word learning. To test these two possibilities, we included English monolinguals, Spanish–English bilinguals, and Mandarin Chinese–English bilinguals in the current study, which allowed us to examine how the presence of lexical tone as a cue to differentiate language inputs during statistical word learning interacts with language learning experience.

1.4. The role of conscious awareness in bilingual SWL

If a linguistic cue influences statistical word learning of input containing 2:1 mappings, a secondary question is how? One possibility is that it may influence the learning process. The broader statistical learning literature has debated whether learning is supported by implicit processes (e.g., Hamrick et al., 2012; Kim et al., 2009), explicit processes (e.g., Dale et al., 2012; Dautriche et al., 2021), or both (e.g., Batterink et al., 2015; Turk-Browne et al., 2005). The debate speaks to the learning mechanisms for statistical word learning in particular (e.g., Berens et al., 2018; Medina et al., 2011; Trueswell et al., 2013; Yu & Smith, 2007): the statistical word learning account resides more on the implicit side, suggesting that statistical word learning is a subconscious process of gradually accumulating co-occurrences between words and referents over time via associative processes (Yu & Smith, 2007). On the contrary, a hypothesis-testing account resides more on the explicit side, proposing that learning across multiple ambiguous naming events is a conscious process of proposing a word-referent link at a time, and then confirming or rejecting the hypothesis on future encounters (Berens et al., 2018; Medina et al., 2011; Trueswell et al., 2013). Other accounts suggest both mechanisms could be at play (K. Smith et al., 2011; Yurovsky & Frank, 2015).

An important question for this debate is how cues may impact the underlying learning process. If a cue benefits statistical word learning of structure containing 2:1 mappings, does it do so through implicit or explicit learning processes? One way to probe this question is by asking participants to report on how well they think they learned the word-referent links (e.g., Benitez et al., 2016; Poepsel & Weiss, 2014; Yurovsky et al., 2013). Benitez et al. (2016) demonstrated that adults were more confident in their learning of word-referent pairings that were cued (those that contained the differential phonotactic structure, CV-/k/). However, confidence ratings did not strongly predict accuracy scores, suggesting a limited role of conscious awareness on learning. In the current study, we explore how a cue may influence conscious awareness of learning by asking participants to self-report how well they learned, as well as any strategies they may have implemented to learn.

1.5. The current study

Our study aimed to examine adults' statistical word learning of structure containing 2:1 mappings in a simulated bilingual environment, assessing 1) whether a linguistic cue (lexical tone) differentiating words sharing a referent affects learning, and 2) how language experience interacts with the effect of the linguistic cue, as pre-registered on the Open Science Framework (OSF: https://osf.io/bv5ts). We presented adults of different language backgrounds - English monolinguals, Spanish–English bilinguals, and Mandarin Chinese–English bilinguals (Chinese–English bilinguals hereafter) - with two SWL conditions of 2:1 mappings. The two conditions differed on whether a linguistic cue differentiated the two words sharing a referent (Cued condition) or not (Uncued condition). Specifically, in the Cued condition, the two words differed by the presence or absence of a Mandarin lexical tonal contour such that one word was non-tonal and the other was tonal (e.g., "*migu*" and "*gádi*"). In the Uncued condition, the two words were both non-tonal (e.g., "*migu*" and "*gadi*").

We included a group of bilingual speakers with knowledge of Mandarin lexical tones (Chinese–English bilinguals), a group of bilingual speakers without tonal experience (Spanish–English bilinguals), and compared their performance to a group of monolingual speakers without tonal experience (English monolinguals). Chinese–English bilinguals were chosen as they are experienced with the lexical tones, and represent a common experience among bilingual speakers who have knowledge of a tonal and non-tonal language. English monolinguals and Spanish–English bilinguals were recruited because non-tonal monolinguals (e.g., Hao, 2012; Lee et al., 1996) and non-tonal bilinguals (Morett, 2020) are capable of discriminating foreign and/or artificial tonal contours, and because they represent the majority of the population where the study was conducted, in Phoenix, Arizona, USA (Migration Policy Institute, 2019). By recruiting the three language groups, we were able to assess 1) whether familiarity with lexical tone provides a language-specific effect on learning, and 2) whether bilingualism provides a language-general effect on learning.

Participants were asked to complete the training of both the Uncued and the Cued conditions (order counterbalanced), and were tested on their knowledge of the word-referent mappings immediately after each training. After learners completed both conditions of the word learning task, they were asked to provide ratings on how much they learned and to explicitly report any strategies they used during the learning process. We were specifically interested in exploring whether the statistical word learning process is explicit in any form and associated with learners' conscious awareness. We therefore examined if participants' rating of how much they learned predicted actual performance separately for the Cued and Uncued conditions, and whether participants reported any specific learning strategies that indicated conscious awareness of the cue or the mapping structure in the tasks.

Our study was designed to address four main questions. Our first question asked if the presence of lexical tone affects adults' SWL of 2:1 mappings. If the cue facilitates SWL for all learners, there should be an overall benefit to learning in the Cued condition compared to the Uncued condition. Our second question examined whether and how language experience interacts with the presence of lexical tone as a linguistic cue during SWL of 2:1 mappings. If bilingualism benefits SWL overall (Escudero et al., 2016; Poepsel & Weiss, 2016), then the two bilingual groups (Spanish–English bilinguals and Chinese–English bilinguals)

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

should outperform English monolinguals in both conditions. If language-specific experience with lexical tones matters, then Chinese–English bilinguals should outperform the other groups (Spanish–English bilinguals and English monolinguals) in the Cued condition only. Our third question was concerned with understanding learning in a more fine-grained fashion. If participants succeeded at learning from 2:1 structure, were learners more likely to learn one label (singlets) or two labels (doublets) for an object? Our final question assessed if participants had conscious awareness of their learning. To address this question, we explored the link between participants' subjective rating of learning and their actual performance and qualitatively examined their retrospective self-report of learning strategies.

2. Method

2.1. Participants

A total of 149 adults were included in the final sample[1] (M_{age} = 20.65, SD = 4.55, age range: 17–36). The majority of participants were recruited from the Department of Psychology's subject pool at Arizona State University located in Tempe, Arizona, USA and received course credit for participation. Bilingual participants were additionally recruited from the wider campus community via a flyer and received monetary compensation ($5) for participation. Consent was obtained according to the Institutional Review Board at Arizona State University.

Participants were grouped into three groups: 55 English monolinguals (M_{age} = 19.20, SD = 1.70, age range 17–29; 45 female, 10 male); 56 Spanish–English bilinguals (M_{age} = 20.18, SD = 3.52, age range 18-36; 42 female, 13 male, 1 non-binary); and 38 Chinese–English bilinguals (M_{age} = 22.59, SD = 5.09, age range: 18–37; 24 female, 14 male), according to responses from a Language Background Questionnaire (modified from P. Li et al., 2006). According to the pre-registration, bilinguals were functional bilinguals who self-reported their first and second languages' average proficiency (average across speaking, listening, and reading) in English and the other language (Spanish or Mandarin Chinese) higher than a 4 out of 10 (on a scale of 1-10, 10 being native-like; Poepsel & Weiss, 2016). Monolinguals were English speakers who self-reported either no knowledge of a second language, or a second language with average proficiency lower than a 4. Additional participants were tested but excluded for missing data (13), low English proficiency (1), and self-reported average proficiency in language(s) other than English, Spanish, or Mandarin Chinese at or above a 4 (32).

As for the linguistic history and experiences (see Supplementary Materials Section 1: https://osf.io/zg782), English monolinguals acquired English significantly earlier (M_{age} = .22, SD = 1.15) than Spanish–English (M_{age} = 3.55, SD = 3.72) and Chinese–English bilinguals (M_{age} = 7.39, SD = 3.93). Chinese–English bilinguals were significantly lower in self-rated English proficiency (in listening, reading, and speaking) than the other two groups. Among bilinguals, Spanish–English bilinguals self-rated a higher proficiency in the second language, the language acquired later (either English or Spanish) than that of Chinese–English bilinguals (either English or Chinese). But Spanish–English bilinguals self-rated a lower proficiency in Spanish than Chinese–English bilinguals rated in Mandarin Chinese. The age of acquisition (AoA) for the second language and the non-English language were not significantly different between the two bilingual groups.

As for the demographic information (see Supplementary Materials Section 1: https://osf.io/zg782), Chinese–English bilinguals were slightly but significantly older than the other two groups. The majority of participants reported being current college students, with 1 English monolingual, 5 Spanish–English bilinguals, and 5 Chinese–English bilinguals already having a Bachelor's degree. Several Chinese–English bilinguals (11) reported having a graduate level degree (Masters or Ph.D.). English monolinguals identified themselves as predominantly White (39), but also as Black/African American (4), Asian (5), Hispanic/Latino (3), multiple racial/ethnic categories (3), and other (1). Spanish–English Bilinguals identified themselves as predominantly Hispanic/Latino (44), but also as White (6), Asian (1), and multiple racial/ethnic categories (5). Chinese–English bilinguals all identified themselves as Asian (38).

2.2. Stimuli

Stimuli consisted of two sets of 16 novel words and two sets of 8 novel objects. The objects were drawn from the Novel Object and Unusual Name database (NOUN; Horst & Hout, 2016). Novel words were created from inventories of consonants and vowels present in English, Spanish, and Mandarin Chinese and that have been used in prior studies (Gebhart et al., 2009; Wang & Saffran, 2014). The consonant inventory constituted [d], [b], [m], [g], [k], and [t]; the vowel inventory was made up of [i] (close front vowel), [u] (close back vowel), and [a] (open back vowel). We first created all possible combinations of consonants and vowels in a consonant-vowel-consonant-vowel (CVCV) bisyllabic structure. We chose a bisyllabic structure given that a CV monosyllabic structure generated many real words in Mandarin Chinese (e.g., "*ma*"). The list of the generated CVCV base words was then assessed for real words by researchers in the lab who were native speakers of English, Spanish, or Mandarin Chinese (Mandarin Chinese speakers additionally considered each base word in all four tonal contours); real words were then removed. We then controlled the position of each syllable within the bisyllabic words such that each syllable appeared approximately the same number of times in word initial (e.g., "***bu***ka" for syllable "*bu*") and word final position (e.g., "*ti****bu***"). Our final two novel word sets are available in Supplementary Materials (Section 2: https://osf.io/zg782; the word composition by syllabic position is also accessible here). In each set, half of the words were Word 1 (W1) for objects and the other half were Word 2 (W2) for objects.

Given the bisyllabic base word structure, and that tonal contour was embedded at a syllabic level, each novel word contained two lexical tones. We chose two distinctive tones, Mandarin Tone 2 (T2, a rising tone), and Mandarin Tone 4 (T4, a dipping tone), because the T2 vs. T4 tonal contrast is acoustically dissimilar in their initial and final fundamental frequency compared with other Mandarin tonal contrasts, because the tonal contrast is salient and easier to perceive by native and non-native Mandarin listeners (Hao, 2012, 2018; So & Best, 2010), and because words embedded with the T2-T4 tonal contour are common in Mandarin Chinese (e.g., T2-T4 contour: "*mábi*" - **numbness**, and T4-T2 contour: "*gùjí*" - **consideration**).

All words were recorded by a U.S. born bilingual speaker proficient in Mandarin and English in three formats: non-tonal, T2-T4 contour, and T4-T2 contour. For the non-tonal contour recordings, the speaker was instructed to read each word in a monotone, with no pitch variation across syllables within a

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

word (e.g., "*tika*"). For the T2-T4 contour recordings, the speaker was instructed to use a rising tone (T2) for the first syllable followed by a dipping tone (T4) for the second syllable to create a Mandarin rising-falling tonal contour (e.g., "*tíkà*"). For the T4-T2 contour recordings, the speaker was instructed to use T4 for the first syllable followed by T2 for the second syllable to create a Mandarin falling-rising contour (e.g., "*tìká*"). The recording was conducted in one session in a single-walled sound attenuated booth using a Blue Snowball microphone.

Each word was saved as three audio files in three formats (non-tonal, T2-T4, and T4-T2 contour). All words were normalized for duration (.99 seconds). Analyses of the audio files demonstrate that the recorded pitch contours (T2 and T4) resembled the pitch contours in Mandarin rising and falling tones (C. Chen et al., 2016). The pitch variation of each recorded word in the three tonal formats, as well as the recorded words' acoustic properties, is depicted in the Supplementary Materials (Section 2: https://osf.io/zg782).

Judgment of stimuli

To test whether tonal and non-tonal words were perceived as words stemming from different languages, a separate group of naïve listeners without tonal experience ($N = 65$; non-tonal monolinguals and bilinguals) made judgments about the word stimuli in a three-alternative forced choice task (modified from Hopkins & Moore, 2007). In each trial, participants were auditorily presented with three words, and instructed to pick the one that was from a different language than the other two. The three words were all non-tonal (control trials), or one word differed from the other two regarding whether it was embedded with a tonal contour or not (e.g., two words were tonal and the other was non-tonal; test trials).

Results showed that naïve listeners chose the target word in test trials above chance, while choices were at random in control trials. The results support that listeners without tonal experience used the presence (or absence) of tonal information to judge words as stemming from two languages. Details of the task and data are openly accessible in Supplementary Materials (Section 3: https://osf.io/zg782).

2.3. Design

Each participant completed two SWL conditions where each object was paired with two words (modified from Yu & Smith, 2007): the Cued and Uncued conditions (order was counterbalanced across participants). In each condition, participants were first trained to learn 8 novel objects during the training phase, each consistently co-occurring with two novel words (a total of 16 words). During the testing phase, participants were tested on their knowledge of word-object links. Each participant was presented with a different set of word-referent mappings across conditions (so that no words nor objects were the same across conditions for each participant). Conditions differed only in whether or not the two words to an object were differentiated by lexical tones.

In the Cued condition, Mandarin lexical tones served as the linguistic cue to differentiate W1 and W2 as stemming from two languages. W1 was in a flat tone (e.g., "*batu*"), but W2 was embedded with one of the Mandarin lexical tonal contours, either the rising-falling contour (T2-T4 contour, e.g., "*tíkà*") or the falling-rising contour (T4-T2 contour, e.g., "*tìká*"). In the Uncued condition, W1 and W2 were both in a flat tone (e.g., W1 "*batu*" and W2 "*tika*"), resembling words from a single language.

Training

Each SWL condition presented 48 training trials, with a duration of 4.5 minutes per condition. Each training trial visually presented two objects, and auditorily played two words. See Figure 1. The two objects appeared simultaneously, side-by-side, while the words were played one at a time with a 1.5-second pause in between. The onset of the object display was 2 seconds prior to the onset of the first word presentation. The two objects were located at the left and the right of the computer screen symmetrically to the central vertical line, both centered at the central horizontal line. The word-object mapping was ambiguous within each trial, since the order of the word presentation (first and the second) did not necessarily match with the objects' spatial location (left and right). There was a 0.1-second blank screen after each training trial. Across trials, each object co-occurred 6 times with each of two words.

In the Uncued condition, each object co-occurred 6 times with a non-tonal word (W1) and 6 times with a different non-tonal word (W2) (see Figure 1 Training). In the Cued condition, each object co-occurred 6 times with a non-tonal word (W1), and 6 times with a different word embedded with lexical tones–tonal words (W2). Tonal words were embedded with either the T2-T4 contour or the T4-T2 contour, but never a mix of the two (tonal contours were counterbalanced across participants). Each word co-occurred with non-target objects less frequently, 0–3 times. The two words for an object never appeared on the same trial; the presentation of each word for an object was intermixed across the training with order of presenting W1 and W2 for an object randomized across objects and test lists. That is, in the Cued condition, the first presented word for an object could have been W1 (non-tonal) or W2 (tonal); and in the Uncued condition, the first presented word could have been W1 (non-tonal) or W2 (non-tonal). An example of one randomized order of presenting W1 and W2 is listed in Supplementary Materials (Section 4: https://osf.io/zg782).

Testing

Testing immediately followed each training in each condition. Test trials contained an auditorily presented target word, one target object, and three distractor objects. Target position was randomized across trials. Participants were instructed to click on the target word's referent after hearing it (see Figure 1 Testing). All words at training were tested for once in each condition, creating a total of 16 test trials per condition. All objects served as the target object twice, once for each word.

Participants completed one condition at a time. Order of conditions was counterbalanced across participants. Before each training phase, participants were instructed that they would hear words and see objects with the aim of figuring out which words referred to which objects. Participants were not told how many words were mapped with each object. After training and testing in the first condition, participants were instructed to proceed to the second condition and were provided a short break if needed. The SWL tasks lasted about 12 minutes (4.5 minutes for training and 1.5 minutes for testing per condition).

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

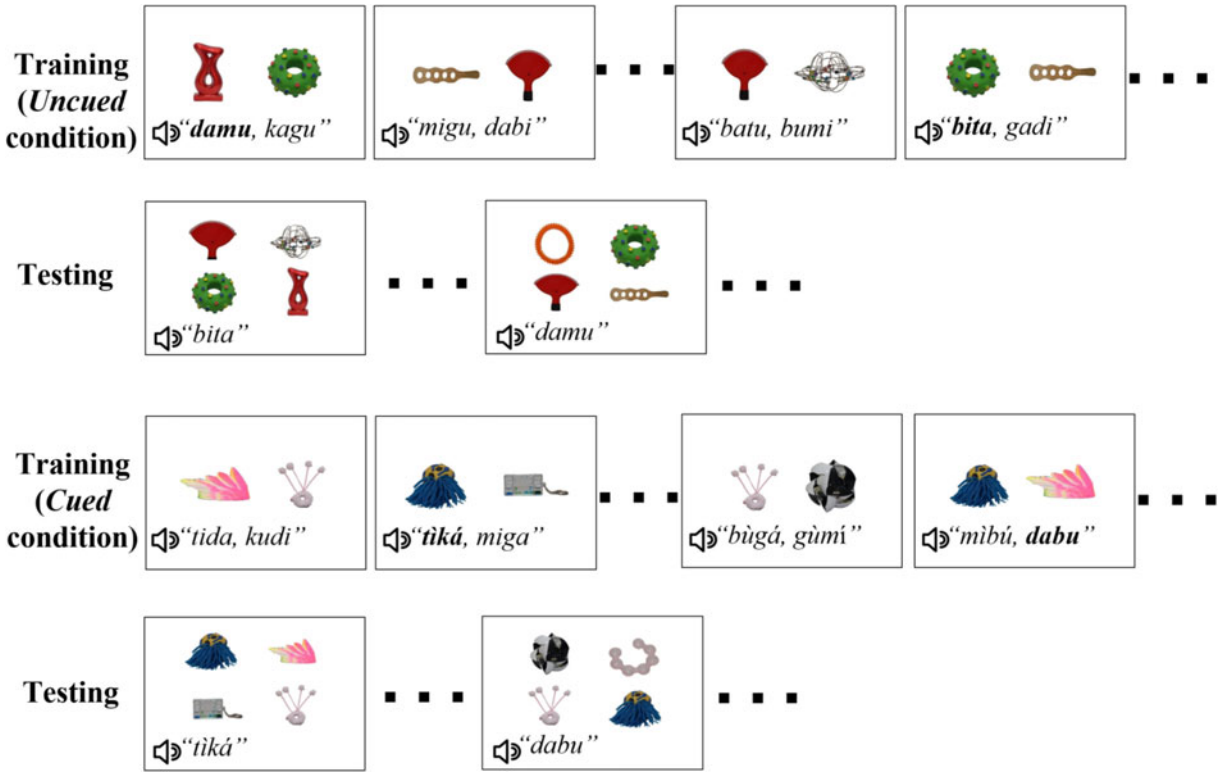


Figure 1. Statistical Word Learning (SWL) of 2:1 Mapping in the Cued and Uncued Conditions.
Note. Example training and testing trials for the Uncued and Cued conditions (condition order was counterbalanced). In training, two words (W1 and W2) co-occurred most often with a shared referent (i.e., the bold words). In the Uncued condition, W1 and W2 were non-tonal. In the Cued condition, one of the two words was non-tonal and the other was embedded with lexical tones (indicated by tonal signs) in either T2-T4 (rising-falling) or T4-T2 (falling-rising) tonal contour. In testing, each word was tested once by a four-alternative-forced-choice task. Dots represent not presented training trials (if shown in Training) and testing trials (if shown in Testing).

2.4. Questionnaires

Subjective Rating on Learning Questionnaire (SRQ)

After completing both SWL tasks, participants filled out a short survey regarding subjective rating on learning. Participants were asked "*Please subjectively rank how much you've learned, from 0 (not at all) to 5 (a great deal)*" and provided a scale bar to drag their ratings horizontally (left side 0 and right side 5). An open-ended question followed: "*What strategy did you use to learn the words for the objects? (e.g., Did you focus on tracking particular words or objects? Did you use a pen or pencil to take notes?)*". Participants answered the open-ended question in a text entry box. This portion of the study was not pre-registered.

Language Background Questionnaire (LBQ)

Participants were asked to report on their language background and demographic information using the Language Background Questionnaire (modified from P. Li et al., 2006). Demographic information included education background, socio-economic status, age, gender, and race/ethnicity. Language use covered experiences with English and language(s) other than English: age of acquisition, language proficiency in speaking, listening, and reading (based on a self-rated scale from 1 to 10), the frequency of language mixing, and the most comfortable language(s) daily.

2.5. Procedure

The SWL tasks were built in PsychoPy3 (version 2020.2.10 – Peirce, 2007) and transferred to Pavlovia for online testing (https://pavlovia.org/; Bridges et al., 2020). The questionnaires were designed in Qualtrics (https://www.qualtrics.com). Due to COVID-19 restrictions on in-person data collection, the study was conducted online via a video conference platform (Zoom: https://zoom.us/). An experimenter met participants in an online Zoom session, provided the experimental link, and instructed each participant to proceed with the experiment in a quiet space. Participants were instructed to turn on their camera to ensure better task engagement, and encouraged to inform the experimenter of any technical issues. Up to two participants were tested in the same Zoom session at a time.

Before the experiment, participants were provided with an online consent form. After consenting, tasks were distributed in this order: SWL tasks (the Cued and Uncued condition with counterbalanced order), the SRQ, and finally the LBQ. A verbal debriefing was given afterwards. The entire study lasted 30 minutes.

3. Results

All data and the analysis scripts in R (version 4.2.2) are openly accessible (OSF: https://osf.io/kq72m/). We conducted linear mixed models (generalized linear mixed models, GLMM, or linear mixed models, LMM) by using the R package *lme4* (v1.1-26 – Bates et al., 2014). Model comparisons were conducted via likelihood ratio tests (using Wald X^2 tests of best fit). We reported beta coefficients, standard errors, Wald X^2 statistics, and Wald confidence intervals where possible[2].

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

3.1. Word learning

We first examined if adults were successful at learning. We compared the trial-by-trial accuracy in each Condition (Uncued and Cued) and each Group (English monolingual, Spanish–English bilingual, and Chinese–English bilingual) against chance (0.25) using a GLMM. The final model included the dichotomous score on individual test trials as the dependent variable (0 as incorrect and 1 as correct), an offset corresponding to the logit of chance performance (0.25) applied to the intercept, and a simple random intercept for subject. The addition of a random intercept for item produced a singular fit for all models except where noted (the results of the full and the final model were the same).

Results showed that learning was above chance in the Uncued condition for English monolinguals ($M = .37$, $SD = .17$; $b = .54$, $STE = .10$, Wald $\chi^2(1) = 28.84$, $p < .001$, Wald 95% CI = [0.34, 0.74]), for Spanish–English bilinguals ($M = .40$, $SD = .12$; $b = .67$, $STE = .07$, Wald $\chi^2(1) = 93.32$, $p < .001$, Wald 95% CI = [0.53, 0.81]), and for Chinese–English bilinguals ($M = .39$, $SD = .16$; $b = .60$, $STE = .14$, Wald $\chi^2(1) = 19.52$, $p < .001$, Wald 95% CI = [0.33, 0.86]; this model additionally included a random intercept for item). Similarly, in the Cued condition, learning was above chance for English monolinguals ($M = .35$, $SD = .13$; $b = .45$, $STE = .08$, Wald $\chi^2(1) = 32.73$, $p < .001$, Wald 95% CI = [0.30, 0.61]), for Spanish–English bilinguals ($M = .37$, $SD = .14$; $b = .67$, $STE = .07$, Wald $\chi^2(1) = 93.32$, $p < .001$, Wald 95% CI = [0.36, 0.70]; this model additionally included a random intercept for item), and for Chinese–English bilinguals ($M = .45$, $SD = .15$; $b = .89$, $STE = .10$, Wald $\chi^2(1) = 83.88$, $p < .001$, Wald 95% CI = [0.70, 1.08]). Averaged across conditions and groups, the overall learning was significantly above chance ($M = .38$, $SD = .15$; $b = .60$, $STE = .04$, Wald $\chi^2(1) = 184.76$, $p < .001$, Wald 95% CI = [0.52, 0.69]). See Figure 2. Thus, all groups were successful at learning in both the Cued and Uncued conditions.

3.2. Effects of Condition and Group on learning

We next examined how Condition and Group affected word learning[3]. We assessed a GLMM model that included the dichotomous score on individual test trials as the dependent variable, the fixed effects of Group (contrast coded) and Condition (Uncued vs. Cued), and the Group×Condition interaction. For the contrast coding of Group, we compared the performance of English monolinguals (reference group) with that of Spanish–English bilinguals (contrast 1: -1/3, 2/3, -1/3) and that of Chinese–English bilinguals (contrast 2: -1/3, -1/3, 2/3). The model additionally included random intercepts for subject and item, as well as a by-subject random slope for Condition and a by-item random slope for Group. Results showed no significant effect of Condition (Wald $X^2(1) = .06$, $p = .814$) or Group (Wald $X^2(2) = 5.82$, $p = .055$). However, the Group×Condition interaction was significant (Wald $X^2(2) = 7.16$, $p = .028$).

We followed up on the Group×Condition interaction by exploring the differences among levels of Group within each level of Condition. We added the fixed effect of Group (dummy coded) and the dependent variable of score in two GLMMs for the Uncued and Cued conditions separately (the models additionally included random intercepts for subject and item). In the Cued condition, results showed that Chinese–English bilinguals ($M = .45$, $SD = .15$) performed better than English monolinguals ($M = .35$, $SD = .13$, Wald $X^2(1) = 12.15$, $p < .001$, Wald 95% CI = [0.19, 0.69]) and better than Spanish–English bilinguals ($M = .37$, $SD = .14$, Wald $X^2(1) = 8.09$, $p = .004$, Wald 95% CI = [0.11, 0.60]), while English monolinguals and Spanish–English bilinguals did not differ (Wald $X^2(1) = .53$, $p = .467$, Wald 95% CI = [-0.14, 0.31]). In contrast, in the Uncued condition, no differences were found among English monolinguals ($M = .37$, $SD = .17$), Spanish–English bilinguals ($M = .40$, $SD = .12$), and Chinese–English bilinguals ($M = .39$, $SD = .16$, $ps > .381$). The Group×Condition interaction still held when additional analyses considered the factors of Age and Education (analyses and results are available in Supplementary Materials Section 5: https://osf.io/zg782). Thus, the cue provided a benefit for SWL but only for Chinese–English bilinguals.

3.3. Effects of Tonality in the Cued condition

To further examine whether words of different tonality (i.e., tonal and non-tonal words) were learned differently, we conducted a

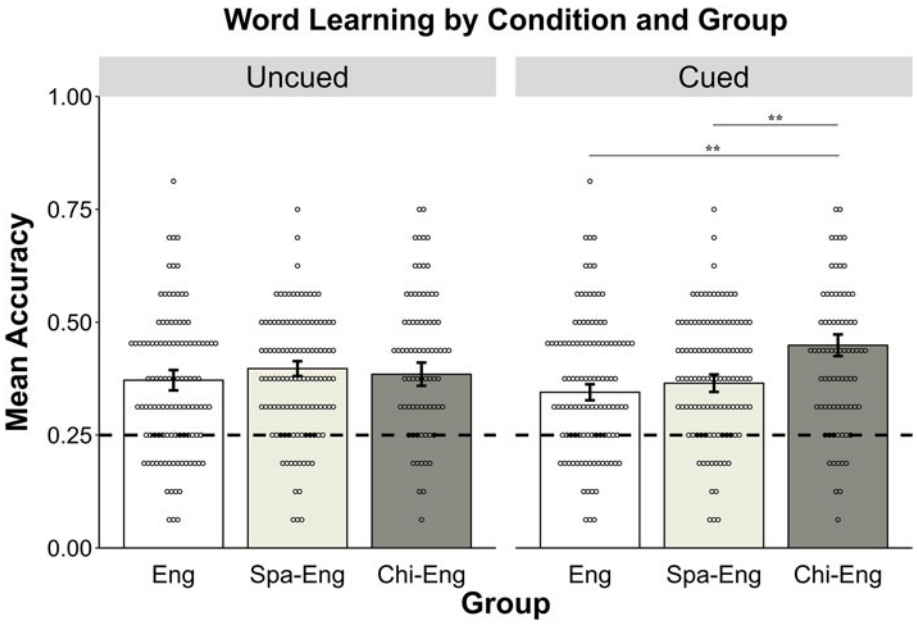


Figure 2. Mean Accuracy in the Uncued and Cued Conditions by Group.
Note. Mean accuracy (and standard error indicated by black bar) for word learning as a function of Condition (Uncued and Cued) and Group (English monolingual, Spanish–English bilingual, and Chinese–English bilinguals). Asterisks denote significant between-group differences (*$p < .05$, **$p < .01$). Dashed line denotes chance performance (0.25). Dots represent individual data points.

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

GLMM on word learning performance in the Cued condition only. We first included the dichotomous score on individual test trials as the dependent variable, and added the fixed effects of Group (contrast coded) and Tonality (Tonal vs. Non-tonal), as well as the interaction between the two. The model additionally included random intercepts for subject and item, a by-subject random slope for Tonality, and a by-item random slope for Group.

Results showed no significant effect of Tonality (Wald $X^2(1) = .77$, $p = .381$), suggesting participants learned tonal ($M = .39$, $SD = .21$) and non-tonal words ($M = .37$, $SD = .19$) similarly. Additionally, we observed a significant main effect of Group (Wald $X^2(2) = 13.15$, $p = .001$), consistent with our prior findings of an advantage for Chinese–English bilinguals over the other two language groups in the Cued condition. The Group×Tonality interaction was not significant (Wald $X^2(2) = 1.15$, $p = .563$). Thus, although Chinese–English bilinguals displayed a performance advantage in the Cued condition, participants across groups learned tonal words similarly to non-tonal words. See Figure 3.

We additionally explored if learning differed for the two tonal contours (T2-T4 contour vs. T4-T2 contour; note that these analyses were not pre-registered). We conducted a GLMM with trial-by-trial accuracy for all tonal words in the Cued condition as the dependent variable and with the fixed effects of Contour pattern (T2-T4 contour vs. T4-T2 contour) and Group (contrast coded), with a random intercept for subject. Interestingly, there was a significant main effect of Contour pattern (Wald $X^2(1) = 4.33$, $p = .037$), such that participants who were presented with the T2-T4 contour (e.g., "*bátù*", $M = .42$, $SD = .22$) performed better than participants who were presented with the T4-T2 contour (e.g., "*bàtú*", $M = .36$, $SD = .21$). The effect of Group was again significant (Wald $X^2(2) = 10.09$, $p = .006$), consistent with our main findings. The Group×Contour pattern interaction was not significant (Wald $X^2(2) = 1.64$, $p = .441$). These results suggest that the tonal contour pattern of rising-falling was easier to learn than that of falling-rising for all language groups.

3.4. Learning one or two labels

Were learners more likely to learn a single label (singlets) or both labels (doublets) for each object? This question is important, as successful learning could be achieved by predominantly learning singlets, predominantly learning doublets, or a mixture of both (Benitez & Li, 2023; Benitez et al., 2016; Ichinco et al., 2009). We first assessed if participants learned singlets and doublets above what would be expected by chance.

To assess learning singlets, we first coded if participants learned one label for each object or not (i.e., object A received a 1 if one label was learned, and a 0 if none or both labels were learned). Then we compared the likelihood of learning a singlet to chance in each condition (chance for learning singlet = ¼ = .25) using GLMMs with an offset corresponding to the logit of chance performance (0.25) applied to the intercept. The model for the Uncued condition additionally included a random intercept for subject and the model for the Cued condition included random intercepts for subject and for item. Results showed that learners were above chance for learning singlets in the Uncued ($M = .53$, $SD = .20$; $b = 1.22$, $STE = .07$, Wald $X^2(1) = 331.05$, $p < .001$, Wald 95% CI = [1.09, 1.35]) and the Cued condition ($M = .52$, $SD = .18$; $b = 1.17$, $STE = .09$, Wald $X^2(1) = 176.05$, $p < .001$, Wald 95% CI = [1.00, 1.35]).

To assess learning doublets, we coded whether participants learned both labels for an object (coded as 1) or not (coded as 0 if none or one label was learned). We then compared the likelihood of learning a doublet to chance in each condition (chance for learning doublet = ¼ × ¼ = .0625) using GLMMs with an offset corresponding to the logit of chance performance (.0625) applied to the intercept; models additionally included a random intercept for subject. Results showed that learners were above chance for learning doublets in the Uncued ($M = .12$, $SD = .13$; $b = .61$, $STE = .12$, Wald $X^2(1) = 24.22$, $p < .001$, Wald 95% CI = [0.37, 0.85]) and the Cued condition ($M = .12$, $SD = .15$; $b = .45$, $STE = .15$, Wald $X^2(1) = 9.59$, $p = .002$, Wald 95% CI = [0.16, 0.73]).

To compare learning singlets with learning doublets across Condition and Group, we calculated the proportion of objects for which adults learned one label or two labels. We fit an LMM on the proportion of learned objects with the fixed effects of Label type (Singlet vs. Doublet), Condition, and Group (contrast coded), as well as their interactions. The model additionally included a random intercept for subject. Results showed a significant main effect of Label type, such that learners were more likely

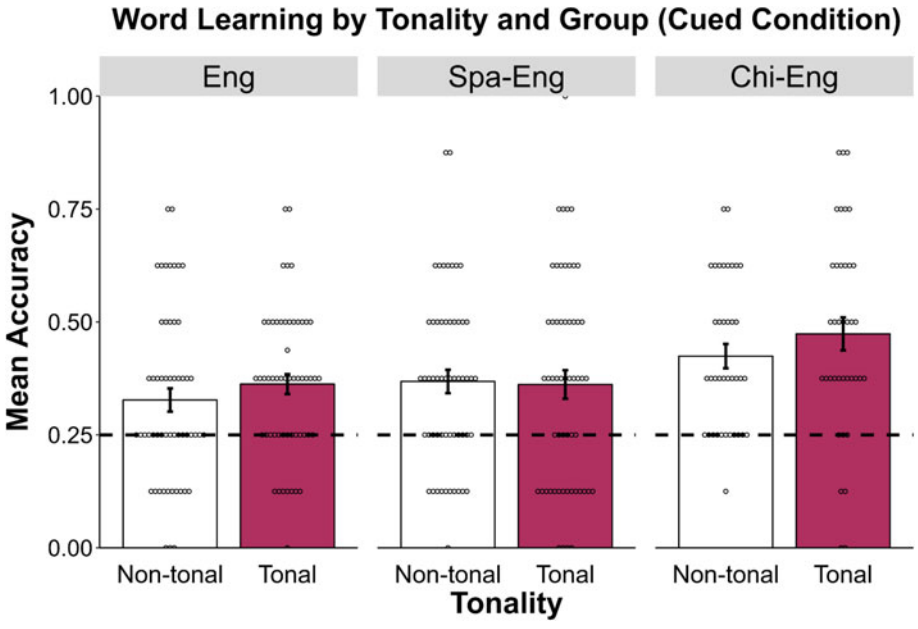


Figure 3. Mean Accuracy for the Tonal and Non-tonal Words by Group in the Cued Condition.
Note. Mean accuracy (and standard error indicated by black bar) for word learning in the Cued condition only as a function of Tonality (Non-tonal and Tonal words) and Group. Tonal words (in maroon) were embedded with Mandarin lexical tones (e.g., "*tíkà*"), while non-tonal words (in white) were not (e.g., "*batu*"). Non-tonal and tonal words were not differentiated within each language group. Dashed line denotes chance performance (0.25). Dots represent individual data points.

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

to learn singlets ($M = .52$, $SD = .19$) than doublets ($M = .12$, $SD = .14$; Wald $X^2(1) = 855.83$, $p < .001$). The main effect of Group was significant (Wald $X^2(2) = 6.23$, $p = .044$). Further, the Group×Condition interaction was significant (Wald $X^2(2) = 6.88$, $p = .032$). See Figure 4.

We followed up on the significant Group×Condition interaction by conducting post-hoc tests (using the R package *emmeans*; Lenth & Lenth, 2017) testing the differences among levels of Group within each level of Label type and each level of Condition, with *p*-values Holm-Bonferroni adjusted. In the Cued condition, Chinese–English bilinguals ($M = .60$, $SD = .21$) learned more singlets than English monolinguals ($M = .47$, $SD = .15$; $b = .12$, $STE = .03$, $t = 3.59$, $p = .001$) and Spanish–English bilinguals ($M = .51$, $SD = .18$; $b = .08$, $STE = .03$, $t = 2.36$, $p = .037$). The three groups did not differ in learning doublets in the Cued condition (English monolinguals, $M = .11$, $SD = .12$; Spanish–English bilinguals, $M = .11$, $SD = .13$; Chinese–English bilinguals, $M = .15$, $SD = .19$), $ps > .756$. In the Uncued condition, the three groups did not differ in learning singlets (English monolinguals, $M = .51$, $SD = .21$; Spanish–English bilinguals: $M = .56$, $SD = .18$; Chinese–English bilinguals, $M = .51$, $SD = .21$; $ps > .555$) or learning doublets (English monolinguals, $M = .12$, $SD = .15$; Spanish–English bilinguals, M = .12, SD = .12; Chinese–English bilinguals, $M = .13$, $SD = .12$; $ps > .164$).

These results reveal two things. First, adults had an overall tendency to link a single word with an object rather than two words. Second, the learning advantage of Chinese–English bilinguals, when the cue of lexical tone was present, manifested mainly in learning more singlets, rather than learning more doublets.

3.5. Subjective responses of learning

Confidence in learning

In order to explore the relation between SWL and conscious awareness, participants' confidence in learning was analyzed from the Subjective Rating Questionnaire (SRQ; note that the analyses in this section were not pre-registered). After completion of both tasks, participants were asked to self-report their learning from 0 (not learning at all) to 5 (learning to a great extent). Participants' overall confidence in learning was low ($M = 1.42$, $SD = 1.01$). Results from a one-way ANOVA test showed no significant differences in confidence ratings among the groups (English monolinguals: $M = 1.26$, $SD = 1.08$; Spanish–English bilinguals: $M = 1.50$, $SD = .92$; Chinese–English bilinguals: $M = 1.53$, $SD = 1.03$; $F(2, 146) = 1.13$, $p = .326$, $\eta^2 = .02$).

We also assessed whether participant's confidence in learning predicted their actual SWL performance. We conducted two simple linear regression models separately for each condition, with each model including predictors of Confidence in learning, Group, and their interactions, and including the outcome of SWL performance (see Supplementary Materials for model estimates in Section 6: https://osf.io/zg782). In the Cued condition, Confidence in learning significantly predicted SWL performance ($b = .03$, $STE = .01$, $p = .009$); the Group×Confidence in learning interaction was not significant for all group comparisons ($ps > .449$). In the Uncued condition, however, Confidence in learning did not significantly predict SWL performance ($b = .01$, $STE = .01$, $p = .242$); and such a prediction did not differ by Group, as indicated by the non-significant interactions for Group×Confidence in learning for all group comparisons ($ps > .593$). In all, the results show that participants' self-rated confidence in learning predicted actual performance in the Cued condition but not that in the Uncued condition. This suggests some conscious awareness of SWL when a lexical tone cue is present.

Learning strategies

Additionally, we qualitatively analyzed participants' self-reported strategies in word learning from the SRQ. The relevant question instructed participants to recall any learning strategies during learning, based on the question "*What strategy did you use to learn the words for the objects?*" (note that the analysis in this section was not pre-registered). We coded participants' valid responses ($n = 107$) into 13 strategy types, which were further grouped into 4 main categories. See Table 1. The types and categories were not mutually exclusive so that each response could belong to multiple types and categories. The four categories, together with percentage of responses coded for that category were: Learning mechanisms (53.28%), Memory (41.12%), Acoustic patterns of words (39.25%), and Others (18.69%). Invalid responses include: blank, vague (e.g., "*I tried to follow objects*"), or unclear (e.g., "*Phonological loop*") descriptions.

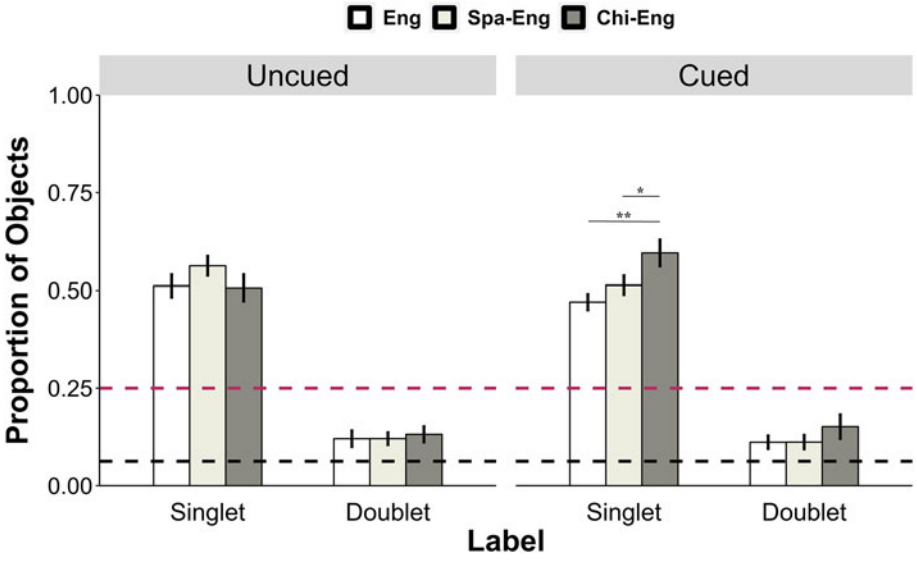


Figure 4. Learning Singlets or Doublets by Condition and Group.
Note. The figure depicts the mean proportion of the number of objects (and standard error) for which learners learned one label (singlets) or two labels (doublets) out of the total number of objects per condition by Condition and Group. Asterisks denote significant between-group differences (**$p < .01$, * $p < .05$). The dashed lines denote chance performance for learning singlets (0.25) in maroon, and for learning doublets (0.0625) in black.

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

Table 1. Qualitative Analysis of Learning Strategies (n = 107)

Strategy Type	Proportion (%)	Examples
A. Learning mechanisms	**53.28**	
1 Adopt hypothesis-testing	24.3	*"I am trying to connect the words with the objects when the objects appear twice in a row. I am listening for the same word if I see that the object appear again…"*
2 Learn one or a small set of words first	15.89	*"I focused on tracking particular words to only a few objects, and then once I established words that were associated with the few object I was following, I tried to do the process again with the other objects."*
3 Realize multiple-to-one mapping	9.35	*"…It was made more difficult because after a while I realized that symbols can have multiple sounds corresponding to them, so I would learn one word-symbol pair and then I would get confused when I would see the same symbol but the words were not what I had thought…"*
4 Adopt associative-learning	3.74	*"I tried to tracking particular words and trying to see what object appeared the most when the word was said."*
B. Memory	**41.12**	
5 Memorize	32.71	*"I tried to memorize what sounds I heard when certain images popped up."*
6 Repeat	5.61	*"…Repetition of the words plus the image help to memorize the word to the object."*
7 Read aloud	2.8	*"Follow the words and read with them."*
C. Acoustic patterns of words	**39.25**	
8 Link words with known lexicons	15.89	*"I associated the words with a word I knew in English and it helped me remember that object name more, for example "Mee- cah" reminded me of my aunt named Mika so that's how I associated it."*
9 Focus on acoustic patterns of words or abbreviate words	14.02	*"I tried to match the word with the object by focusing on the first letter of the word and seeing if it visibly matches with the object. For instance, if a sound started with a T, I tried to look for a T within the object(s) that pop up."*
10 Link words with familiar language(s)	4.67	*"…And sometimes I try to make some connections from those objects to my Chinese words."*
11 Map objects' visual features with words (Sound symbolism)	4.67	*"I listened to the words and if it sounded like a sharp should I categorized that word with the picture that had edges. So if I heard a word that was dull I categorized it with a picture that was roundish."*
D. Others	**18.69**	
12 Difficulty	15.89	*"I focused on the objects but it was very difficult."*
13 Others	2.8	*"Trying to cross match the objects with the sounds. Like guessing unknown number in math problems."*
Invalid (n = 42)		
Invalid		*"I tried to follow objects."*

Note. The valid participants (n = 107) composed of the 4 major categories (Learning mechanisms, Memory, Acoustic pattern of words, and Others), which were further decomposed into 13 types in total. The final valid participants excluded the "Invalid" responses that did not belong to any categories (n = 42). Each participant's response can be categorized into one or multiple types and categories. Proportion refers to the number of participants who used such a strategy type divided by the total number of valid participants. The examples shown above were corrected from grammatical errors and/or typos.

We were specifically interested in whether learners were consciously aware of either tonal information or many-to-one mappings in the task. A small percentage of the responses indicated strategies of linking novel words with known lexicons (15.89%) and/or familiar language inventories (4.67%) to scaffold word learning (see Acoustic patterns of words). For instance, one participant specifically noted "making some connections from those objects to their Chinese words". Further, only a few responses indicated the existence of many-to-one mappings (9.35%, see Learning mechanisms), such as "*I realized that symbols* [objects] *can have multiple sounds* [words] *corresponding to them.*"

Further, the strategy of linking novel words to learners' prior lexicons and language inventories was not unique to Chinese–English bilinguals (n = 7), compared to English monolinguals (n = 7) and Spanish–English bilinguals (n = 8). Similarly, detecting multiple-to-one mappings did not vary much by language group: 3 Chinese–English bilinguals indicated knowledge of multiple-to-one mappings, compared to 2 English monolinguals and 5 Spanish–English bilinguals. In sum, only a small number of participants reported explicitly the use of prior language knowledge or the presence of many-to-one mappings; and the few who did have such conscious awareness did not seem to come from one specific language group. These findings suggest a limited role of conscious awareness of learning.

4. Discussion

In this study, we examined statistical word learning of structure presenting 2:1 mappings in learners with different language experience across two conditions: when the two words for a referent were linguistically differentiated by a lexical tone cue (Cued condition) or not (Uncued condition). We found that adults succeeded at learning in both conditions, but did not necessarily learn better in the Cued condition over the Uncued condition.

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

Instead, learning interacted with learners' language experience: Chinese–English bilinguals outperformed English monolinguals and Spanish–English bilinguals, but only in the Cued condition. This advantage was not specific to words containing the lexical tone cue or to learning doublets (learning both words of a referent). Instead, Chinese–English bilinguals learned tonal and non-tonal words equally well, and learned more singlets (learning a single word for a referent) in comparison to English monolinguals and Spanish–English bilinguals. Finally, exploration of participants' self-reported confidence in learning and learning strategies revealed a limited role of conscious awareness. These findings demonstrate that a linguistic cue differentiating two language inputs provides a boost in overall statistical word learning only for learners familiar with that cue.

4.1. How lexical tone impacted learning

What role did the cue of lexical tone play in learning? Comparisons of the Cued and Uncued conditions revealed there was no overall learning advantage for the Cued condition. Instead, the cue only provided an advantage to Chinese–English bilinguals. Further, assessments of learning tonal and non-tonal words revealed how the cue benefited Chinese–English bilinguals. It was not the case that Chinese–English bilinguals outperformed the other two groups only on words containing lexical tone information, as some previous research has found (Potter et al., 2017; Wang & Saffran, 2014). Instead, Chinese–English bilinguals learned the artificial tonal words and the non-tonal words equally well in the Cued condition. The results support the idea that statistical word learning in general is improved due to familiarity with certain linguistic features in the input. That is, familiarity with some linguistic features in the input boosts learners' ability to track statistical regularities overall from two language inputs. This finding is in line with recent research demonstrating that familiarity with features in the input benefits learning new regularities in that input (Antoniou et al., 2015; Palmer et al., 2019; Stärk et al., 2023).

Although it is clear that Chinese–English bilinguals obtained an advantage in learning in the Cued condition, the mechanism underlying such an advantage is not clear. One possibility is that familiarity with the cue enhanced Chinese–English bilinguals' general attention during learning. Previous studies from other domains suggest that familiarity with some features in the input (e.g., familiar objects or familiar faces) heightens attentional allocation to the learning process (Christie & Klein, 1995; Ujiie et al., 2021). In studies examining language learning and listening, language familiarity is found to modulate infants' and adults' selective attention toward speakers and object naming events (Barenholtz et al., 2016; Kinzler & Spelke, 2007; Lewkowicz & Hansen-Tift, 2012; Marno et al., 2016). It is very well possible that the presence of a familiar cue may have heightened attention in Chinese–English bilinguals during learning, fostering better memories for words with and without lexical tone information (Chun & Turk-Browne, 2007; Pomper & Saffran, 2019). However, this is speculative, as we did not measure attention in our study. We suggest that future studies should examine how moment-to-moment indices of attentional processes, such as pupil size changes and eye-movements (e.g., Yu & Smith, 2011; Yu et al., 2012), are linked with statistical word learning of cued and uncued 2:1 structure.

4.2. Learning doublets was consistent across groups

Another important finding was that the boost in word learning for Chinese–English bilinguals was specific to learning singlets, rather than learning more doublets. That is, a familiar cue did not help learners in mapping two labels to an object. Instead, all groups learned doublets similarly, and to a lesser extent than singlets, and there was no evidence that the cue specifically modified the learning of doublets. This finding is consistent with previous research demonstrating that adult learners are more likely to learn singlets than doublets from SWL tasks with (Benitez et al., 2016) or without cues (Chan & Monaghan, 2019; Ichinco et al., 2009; Kachergis et al., 2012), and indicates that learning doublets is particularly challenging. This could be because the two words for an object may compete or interfere with each other during learning (Benitez et al., 2016; Degani & Tokowicz, 2010) similar to competition or interference during lexical retrieval (e.g., Kroll & Stewart, 1994). The low likelihood of acquiring two words for one referent in statistical word learning aligns with natural language research that translation equivalents only account for a small set of receptive vocabulary inventories of bilingual infants and toddlers (e.g., 25%-33%, Legacy et al., 2016).

Yet, learners do show some knowledge of multiple words for the same concept in early childhood despite the difficulty (Bilson et al., 2015; De Houwer et al., 2006; Legacy et al., 2016; Nicoladis & Laurent, 2020; Pearson et al., 1995). How do learners eventually come to learn two words for the same referent if the two words compete during learning? An alternative proposal suggests that bilinguals may acquire each novel word-referent mapping independently, which may not necessarily require a word-to-word association across languages (Genesee & Nicoladis, 2007; Patterson & Pearson, 2004). That is, learners could map each word in each language to its concept via two distinct lexical systems (e.g., "*dog*" with **dog**, and "*perro*" with **dog**) without knowing the two words are referring to the same object ("*dog*" and "*perro*"). Still other proposals suggest that one word for a meaning can facilitate learning another word for that same meaning via semantic networks (Bilson et al., 2015). Considering the possible (and complex) mechanisms underlying learning words for the same referent across- and within- languages, it will be important for future work to examine what kind of statistical input and cues may give rise to successful learning of BOTH words for a single referent across language experience, age, and timescales of learning.

4.3. No evidence for a general effect of bilingualism on SWL

Interestingly, Spanish–English bilinguals' performance did not differ from that of English monolinguals with or without the presence of a linguistic cue, suggesting that bilingual experience in general does not provide a benefit for SWL, at least under the conditions studied here. These results are in contrast to two studies which show a bilingual advantage in statistical learning of structure containing multiple mappings (Chan & Monaghan, 2019; Poepsel & Weiss, 2016). When Poepsel and Weiss (2016) presented learners with an SWL task containing 1:2 mappings (one word mapped with two objects), results showed that Chinese–English bilinguals and Spanish–English bilinguals outperformed monolinguals. The authors explained such an advantage by bilinguals' loosened reliance on the mutual exclusivity assumption (ME; that a referent by default has only one name) due to bilinguals' increased encounters with ME-violating circumstances (Houston-Price et al., 2010). Chan and Monaghan (2019) presented an SWL task to adults containing 2:1 mappings and found that bilinguals demonstrated an advantage in the learning

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

rate (but not the learning accuracy) compared to their monolingual counterparts.

What could be driving these inconsistencies? One possibility is that bilingual experience *per se* may not be a strong predictor of better statistical word learning. Instead, the differences observed between monolinguals and bilinguals in word learning in previous studies may have resulted from cognitive differences across groups. Several studies have found an advantage in bilinguals over monolinguals in cognitive skills, such as memory, attention, and inhibitory control (Bialystok et al., 2012; Brito & Barr, 2012; Costa et al., 2009; Grundy & Timmer, 2017; Kaushanskaya & Marian, 2009; Prior & MacWhinney, 2010), though the evidence for a bilingual cognitive advantage can be mixed (Gunnerud et al., 2020; Ware et al., 2020). It may well be the case that the SWL studies mentioned above with a bilingual advantage were capturing individual differences in cognitive abilities that support statistical word learning (Crespo & Kaushanskaya, 2021; Vlach & DeBrock, 2019). Thus, we suggest that examining how individual differences in cognitive processes, together with language experience, may be related to statistical word learning is a fruitful avenue for future research.

4.4. The role of conscious awareness was limited

This study also provides insight into the conscious awareness of SWL with and without a cue. First, confidence in learning for all learners predicted actual word learning ONLY WHEN a linguistic cue was presented, but not without such a cue. This suggests that participants had some awareness of how well they were learning when a lexical tone cue was present. In line with the current study results, Benitez et al. (2016) found that cued words were rated with more confidence of being learned compared with uncued words that were linked with the same referent. Similarly, Poepsel and Weiss (2014) found contextual cues (i.e., a speaker or an instruction cue) augmented adult learners' confidence in the knowledge of words, but not their actual statistical word learning performance, in an SWL task that presented 1:2 mappings. Our results suggest that the presence of a linguistic cue differentiating two language sources does not improve all learners' statistical word learning performance, but it may enhance learners' precision in gauging how well they have learned.

However, learners seemed less consciously aware of the linguistic cue, or the presence of 2:1 mappings in the task. According to the qualitative analysis of learners' self-reported learning strategies, very few learners reported familiarity with the novel words or the presence of more than one word for each object. In fact, very few learners seemed to report the existence of the tonal cue, or the difference of tonal information across conditions. Now, it is possible that this result was due to the question we presented to participants. We designed an open-ended question for learners to provide any learning strategies regarding their learning, but we did not ask them to report on the structure of the task. Thus, learners may have noticed the tone cue or the cross-condition differences in lexical tone, but reported only prominent learning strategies, e.g., memorizing. Nonetheless, the evidence that we do have suggests a limited role of conscious awareness of learning. In future studies, it will be important to not only ask about learners' learning strategies, but also to design more explicit and precise questions to probe what aspects of the statistical word learning task learners are attuned to.

What do these results mean for the processes underlying statistical word learning? Although our study was not set up to differentiate whether learning processes were implicit (Yu & Smith, 2007), explicit (Berens et al., 2018; Medina et al., 2011; Trueswell et al., 2013), or both (K. Smith et al., 2011; Yurovsky & Frank, 2015), the fact that learners had limited conscious awareness of learning suggests that more implicit processes may be playing a role. However, the presence of a cue may serve to make some learning more explicit. We suggest that making headway in this debate requires considering learning of different kinds of mapping structure as well as the incorporation of real-world linguistic cues that learners may use for statistical learning.

5. Conclusion

To conclude, we examined English monolinguals, Spanish–English bilinguals, and Chinese–English bilinguals' statistical word learning from simulated bilingual input where the two words for a referent were either differentiated by lexical tone (Cued condition) or not (Uncued condition). We found that Chinese–English bilinguals outperformed English Monolinguals and Spanish–English bilinguals only when a lexical tone cue was present; the three language groups did not differ in learning without such a cue. Further, with the presence of a familiar cue, Chinese–English bilinguals learned both tonal words and non-tonal word singlets. Finally, explorations of participants' confidence in learning and self-reported learning strategies demonstrated a limited role of conscious awareness of learning. In all, the study contributes to the current theories of statistical learning by addressing the importance of linguistic variability and the role of learners' language familiarity. Our results indicate that when learning statistics of multiple languages, FAMILIARITY WITH LINGUISTIC FEATURE(S) boosts overall statistical word learning.

Competing interests declaration. The study is approved by the IRB committee at Arizona State University (ID: STUDY00007151 Language Learning in Adults). We have no conflict of interests to disclose. Upon publication, the study materials will be shared with the public, including the auditory and/or visual stimuli of the tasks listed in the methods section, supplementary files, data, analysis scripts, and the manuscript. All data and analysis scripts in R are openly accessible (OSF: https://osf.io/kq72m/). Figures should appear in color in the online version only.

Supplementary Materials. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728923000858

For all other supplementary files accompanying this paper, visit the OSF link https://osf.io/zg782

Acknowledgements. This project was funded by Arizona State University's Graduate College Pandemic Impact Award awarded to Ye Li. We thank Arthur M. Glenberg and Gene Brewer for their helpful advice in experimental design. We especially thank Samantha Anderson for her detailed suggestions on statistical analysis. We thank Sophie Tang (Ty Tang), Christine Yu, Cassie Leedom, Haoze Zhu, and Tengteng Tang for their assistance in experimental set-up, stimuli, or figure. We thank the research assistants in the Learning and Development Lab for their help in data collection and coding: Kaede Hattori, Alejandra Garcia Hernandez, Rebeca Alvarado Ortega, and Jillian Kuo. We give our deepest appreciation to our family members, friends, and colleagues who provide consistent support.

Notes

[1] The pre-registered plan was to collect 56 participants per group (with a total of 168) according to a 3-way ANOVA estimation using the R package *easypower* (McGarvey, 2015), assuming a medium effect size. However, we had a lower response rate from Chinese–English bilinguals. After eight months of

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

recruitment, we opted to terminate data collection given time constraints. The decision was made prior to data analysis.

[2] Our original pre-registered analysis plan was to conduct traditional analyses on mean accuracy scores aggregated over trials, which included independent samples t-tests to compare learning against chance performance, ANOVA analyses to assess effects of Condition, Group, and Word Type on mean accuracy, and ANOVA analyses to examine how many words were learned per object. In accordance with reviewer feedback, we instead report the results of linear mixed models. Linear mixed models are a more conservative approach as these can account for the binary outcome variable of accuracy (GLMM), and random effects at subject and/or item level (GLMM, and LMM). The results from the less conservative, pre-registered analyses (openly accessible in Supplementary Materials Section 5: https://osf.io/zg782) were consistent with the results reported here.

[3] Our pre-registered analysis plan was to conduct a 3-factor analysis (Group×Condition×Word type). However, reviewers indicated that this analysis was likely overfactored, given that the Word Type factor (Tone vs. Non-Tonal) was not present in the Uncued condition. After consulting with an expert in quantitative statistics, we instead conducted a 2-factor analysis (Group×Condition) on learning, and then conducted a separate analysis examining the effect of Tonality on learning only for the Cued condition. The results from the 3-factor pre-registered analysis (openly accessible in Supplementary Materials Section 5: https://osf.io/zg782) are qualitatively similar to the results reported here.

References

Alt, M., Meyers, C., Oglivie, T., Nicholas, K., & Arizmendi, G. (2014). Cross-situational statistically based word learning intervention for late-talking toddlers. *Journal of Communication Disorders*, *52*, 207–220.

Antoniou, M., & Chin, J. L. (2018). What can lexical tone training studies in adults tell us about tone processing in children?. *Frontiers in psychology*, *9*, 1.

Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, *18* (4), 683–695.

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental psychology*, *29*(5), 832.

Barenholtz, E., Mavica, L., & Lewkowicz, D. J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition*, *147*, 100–105.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406*.5823.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of memory and language*, *83*, 62–78.

Benitez, V. L., & Li, Y. (2023). Cross-situational word learning in children and adults: The case of lexical overlap. *Language Learning and Development*. Advanced online publication. https://doi.org/10.1080/15475441.2023.2256713

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.

Benitez, V. L., Bulgarelli, F., Byers-Heinlein, K., Saffran, J. R., & Weiss, D. J. (2020a). Statistical learning of multiple speech streams: A challenge for monolingual infants. *Developmental science*, *23*(2), e12896.

Benitez, V. L., Zettersten, M., & Wojcik, E. (2020b). The temporal structure of naming events differentially affects children's and adults' cross-situational word learning. *Journal of Experimental Child Psychology*, *200*, 104961.

Berens, S. C., Horst, J. S., & Bird, C. M. (2018). Cross-situational learning is supported by propose-but-verify hypothesis testing. *Current Biology*, *28* (7), 1132–1136.

Bhatt, R. M. (1997). Code-switching, constraints, and optimal grammars. *Lingua*, *102*(4), 223–251.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, *16*(4), 240–250.

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition*, *140*, 122–134.

Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford, CA: Stanford University Press.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414.

Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental science*, *15*(6), 812–816.

Byers-Heinlein, K. (2014). Languages as categories: Reframing the "one language or two" question in early bilingual development. *Language Learning*, *64*(s2), 184–201.

Chan, J., & Monaghan, P. (2019). Simulating bilingual word learning: monolingual and bilingual adults' use of cross-situational statistics. In *Proceedings of the 41st annual meeting of the cognitive science society* (Vol. 41). Montreal, Canada. Cognitive Science Society.

Chen, C., Bunescu, R., Xu, L., & Liu, C. (2016). Tone classification in Mandarin Chinese using convolutional neural networks. *Interspeech* 2016, 2150–2154.

Chen, S., Zhu, Y., & Wayland, R. (2017). Effects of stimulus duration and vowel quality in cross-linguistic categorical perception of pitch directions. *PloS one*, *12*(7), e0180656.

Chen, S., Zhu, Y., Wayland, R., & Yang, Y. (2020). How musical experience affects tone perception efficiency by musicians of tonal and non-tonal speakers?. *PloS one*, *15*(5), e0232514.

Christie, J., & Klein, R. (1995). Familiarity and attention: Does what we know affect what we notice?. *Memory & cognition*, *23*(5), 547–550.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current opinion in neurobiology*, *17*(2), 177–184.

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.

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.

Dale, R., Duran, N. D., & Morehead, J. R. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Advances in Cognitive Psychology*, *8*(2), 196.

Dautriche, I., Rabagliati, H., & Smith, K. (2021). Subjective confidence influences word learning in a cross-situational statistical learning task. *Journal of Memory and Language*, *121*, 104277.

Degani, T., & Tokowicz, N. (2010). Ambiguous words are harder to learn. *Bilingualism: Language and cognition*, *13*(3), 299–314.

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

Escudero, P., Mulak, K. E., Fu, C. S., & Singh, L. (2016). More limitations to monolingualism: bilinguals outperform monolinguals in implicit word learning. *Frontiers in psychology*, *7*, 1218.

Evans, S. (2011). Hong Kong English and the professional world. *World Englishes*, *30*(3), 293–316.

Fabiano-Smith, L., & Goldstein, B. A. (2010). Phonological Acquisition in Bilingual Spanish–English Speaking Children. *Journal of Speech, Language, and Hearing Research*, *53*, 160–178.

Feng, A. (Ed.). (2007). *Bilingual education in China: Practices, policies and concepts*. Clevedon, UK: Multilingual Matters.

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive science*, *33* (6), 1087–1116.

Genesee, F., & Nicoladis, E. (2007). Bilingual first language acquisition. *Blackwell handbook of language development*, 324–342.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*(3), 325–340.

Gunnerud, H. L., Ten Braak, D., Reikerås, E. K. L., Donolato, E., & Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press

in children? A systematic review and meta-analysis. *Psychological Bulletin*, *146*(12), 1059.
Hamrick, P., Rebuschat, P., Rebuschat, P., & Williams, J. (2012). How implicit is statistical learning. *Statistical learning and language acquisition*, 365–382.
Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of phonetics*, *40*(2), 269–279.
Hao, Y. C. (2018). Second language perception of Mandarin vowels and tones. *Language and Speech*, *61*(1), 135–152.
Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child development*, *86*(1), 10–22.
Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., & Rocroi, C. (2000). Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3), i–123.
Hopkins, K., & Moore, B. C. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *The Journal of the Acoustical Society of America*, *122*(2), 1055–1068.
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.
Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15* (2), 125–150.
Ichinco, D., Frank, M. C., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. In *Proceedings of the 31st annual meeting of the cognitive science society* (Vol. 31). Austin, TX: Cognitive Science Society.
Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word–referent mappings. *Psychonomic bulletin & review*, *19*(2), 317–324.
Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705–710.
Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, *132*(2), 1050–1060.
Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit?. *Neuroscience letters*, *461* (2), 145–149.
Kinzler, K. D., & Spelke, E. S. (2007). Core systems in human cognition. *Progress in brain research*, *164*, 257–264.
Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
Kucker, S. C., McMurray, B., & Samuelson, L. K. (2015). Slowing down fast mapping: Redefining the dynamics of word learning. *Child Development Perspectives*, *9*(2), 74–78.
Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *25*(5), 527–542.
Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2016). Vocabulary size, translation equivalents, and efficiency in word recognition in very young bilinguals. *Journal of child language*, *43*(4), 760–783.
Lenth, R., & Lenth, M. R. (2017). Package 'Emmeans'. *Statistician*, *34*(4), 216–221.
Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, *109*(5), 1431–1436.
Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior research methods*, *38*(2), 202–210.
Lyu, D. C., Tan, T. P., Chng, E. S., & Li, H. (2010). An analysis of a Mandarin-English code-switching speech corpus: SEAME. *Age*, *21*, 25–8.
Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive science*, *14*(1), 57–77.
Marno, H., Guellai, B., Vidal, Y., Franzoi, J., Nespor, M., & Mehler, J. (2016). Infants' selectively pay attention to the information they receive from a native speaker of their language. *Frontiers in Psychology*, *7*, 1150.
McGarvey, A. (2015). easypower: Sample size estimation for experimental designs. *R package version*, 1(1).
Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014–9019.
Migration Policy Institute. (2019). *State immigration data profiles for Arizona State*. https://www.migrationpolicy.org/data/state-profiles/state/workforce/AZ
Morett, L. M. (2020). The influence of tonal and atonal bilingualism on children's lexical and non-lexical tone perception. *Language and Speech*, *63*(2), 221–241.
Nicoladis, E., & Laurent, A. (2020). When knowing only one word for "car" leads to weak application of mutual exclusivity. *Cognition*, *196*, 104087.
Palmer, S. D., Hutson, J., White, L., & Mattys, S. L. (2019). Lexical knowledge boosts statistically-driven speech segmentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 139.
Patterson, J. L., & Pearson, B. Z. (2004). Bilingual Lexical Development: Influences, Contexts, and Processes. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish–English speakers* (pp. 77–104). Paul H Brookes Publishing.
Pearson, B. Z., Fernández, S., & Oller, D. K. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two?. *Journal of child language*, *22*(2), 345–368.
Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13.
Poepsel, T. J., & Weiss, D. J. (2014). Context influences conscious appraisal of cross situational statistical learning. *Frontiers in psychology*, *5*, 691.
Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, *152*, 9–19.
Pomper, R., & Saffran, J. R. (2019). Familiar object salience affects novel word learning. *Child development*, *90*(2), e246-e262.
Potter, C. E., Wang, T., & Saffran, J. R. (2017). Second language experience facilitates statistical learning of novel linguistic materials. *Cognitive science*, *41*, 913–927.
Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and cognition*, *13*(2), 253–262.
Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in psychology*, *3*, 228.
Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: MIT Press.
Romaine, S. (2012). The bilingual and multilingual community. *The handbook of bilingualism and multilingualism*, 443–465.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.
Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*(2), 128–142.
Singh, L., & Fu, C. S. (2016). A new view of language development: the acquisition of lexical tone. *Child development*, *87*(3), 834–854.
Singh, L., Poh, F. L., & Fu, C. S. (2016). Limits on monolingualism? A comparison of monolingual and bilingual infants' abilities to integrate lexical tone in novel word learning. *Frontiers in Psychology*, *7*, 667.
Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.
Smith, K., Smith, A. D., & Blythe, R. A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science*, *35*(3), 480–498.
So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and speech*, *53*(2), 273–293.
Stärk, K., Kidd, E., & Frost, R. L. (2023). Close encounters of the word kind: Attested distributional information boosts statistical learning. *Language Learning*.
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.
Tomlinson, J. M., Jr., & Bott, L. (2013). How intonation contrains pragmatic inference. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).
Torres Cacoullos, R. (2020). Code-switching strategies: Prosody and syntax. *Frontiers in Psychology*, *11*, 2130.
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/10.1017/S1366728923000858 Published online by Cambridge University Press

Tsui, A. S. M., Erickson, L. C., Mallikarjunn, A., Thiessen, E. D., & Fennell, C. T. (2021). Dual language statistical word segmentation in infancy: Simulating a language-mixing bilingual environment. *Developmental Science*, *24*(3), e13050.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552.

Ujiie, Y., Kanazawa, S., & Yamaguchi, M. K. (2021). The other-race effect on the McGurk effect in infancy. *Attention, Perception, & Psychophysics*, *83*(7), 2924–2936.

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of memory and language*, *93*, 217–230.

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.

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382.

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental psychology*, *45*(6), 1611.

Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in psychology*, *5*, 953.

Ware, A. T., Kirkovski, M., & Lum, J. A. (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Frontiers in Psychology*, *11*, 1458.

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning?. *Language Learning and Development*, *5*(1), 30–49.

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.

Yip, M. (2002). *Tone*. Cambridge University Press.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological science*, *18*(5), 414–420.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental science*, *14*(2), 165–180.

Yu, C., Zhong, Y., & Fricker, D. (2012). Selective attention in cross-situational statistical learning: evidence from eye tracking. *Frontiers in psychology*, *3*, 148.

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53–62.

Yurovsky, D., & Yu, C. (2008). Mutual exclusivity in cross-situational statistical learning. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 30, No. 30).

Yurovsky, D., Yu, C., & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive science*, *37*(5), 891–921. https://doi.org/10.1111/cogs.12035

Zettersten, M., Wojcik, E., Benitez, V. L., & Saffran, J. (2018). The company objects keep: Linking referents together during cross-situational word learning. *Journal of memory and language*, *99*, 62–73.

https://doi.org/10.1017/S1366728923000858 Published online by Cambridge University Press