

# One language or two? Navigating cross-language conflict in statistical word segmentation

Dylan M. Antovich  | Katharine Graf Estes

University of California, Davis, Davis, CA, USA

## Correspondence

Dylan M. Antovich, University of California, Davis, Davis, California, USA.  
Email: dmantovich@ucdavis.edu

## Funding information

UC Davis Institute for Social Sciences; Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: R03HD084941R03HD084941; Childs Family Fund

## Abstract

Bilingual infants must navigate the similarities and differences between their languages to achieve native proficiency in childhood. Bilinguals learning to find individual words in fluent speech face the possibility of conflicting cues to word boundaries across their languages. Despite this challenge, bilingual infants typically begin to segment and learn words in both languages around the same time as monolinguals. It is possible that early bilingual experience may support infants' abilities to track regularities relevant for word segmentation separately across their languages. In a dual speech stream statistical word segmentation task, we assessed whether 16-month-old infants could track syllable co-occurrence regularities in two artificial languages despite conflicting information across the languages. We found that bilingual, but not monolingual, infants were able to segment the dual speech streams using statistical regularities. Although the two language groups did not differ on secondary measures of cognitive and linguistic development, bilingual infants' real-world experience with bilingual speakers was predictive of their performance in the dual language statistical segmentation task.

## KEYWORDS

bilingualism, statistical learning, word segmentation

## 1 | INTRODUCTION

Despite the prevalence of bilingualism (Crystal, 2004; Grosjean, 2010), there is limited work investigating how infants detect words in two languages. Infants with bilingual exposure may encounter conflicting prosodic, lexical, and co-occurrence cues to word boundaries across their languages, which in turn could disrupt word segmentation. For example, an infant learning both French and English would encounter opposing prosodic cues; the predominant stress pattern in English is trochaic (strong-weak), whereas French words tend to follow an iambic (weak-strong) pattern. Language-specific patterns that support segmentation within one language may interfere with segmentation in another (Polka, Orena, Sundara, & Worrall, 2017). However, effective separation of the languages' properties and efficient shifting between language contexts may support accurate word segmentation and

allow infants to learn from the rich information present in dual speech streams.

Early exposure to diverse regularities across languages may allow infants' language knowledge to self-organize into linguistic clusters over time (e.g. PRIMIR; Curtin, Byers-Heinlein, & Werker, 2011). This process would improve as infants learn core features that differentiate between languages (e.g. prosody, phonology, etc.; Byers-Heinlein, Burns, & Werker, 2010), allowing infants to monitor for language changes and to integrate new information by language, reducing cross-linguistic interference. This is critical, as the language in use in bilingual environments may change between, or even within, interactions and speakers, and failure to anticipate language switches can incur processing costs for both infants and adults (Byers-Heinlein, Morin-Lessard, & Lew-Williams, 2017), suggesting that change detection is important for efficient bilingual processing.

Once bilinguals acquire tools to track information for each language separately, they can more effectively learn language-specific



regularities, including those that support word segmentation. However, the utility of different segmentation strategies likely varies across languages (Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006; Saksida, Langus, & Nespor, 2017). Compounding this issue, bilingual infants receive less exposure to each language than monolinguals (Costa & Sebastián-Gallés, 2014; Werker, Byers-Heinlein, & Fennell, 2009), which could delay detection of the appropriate segmentation strategies within each language. Despite this additional challenge, bilinguals segment fluent speech around the same age as monolinguals (Bosch, Figueras, Teixidó, & Ramon-Casas, 2013; Orena & Polka, 2019; Singh & Foong, 2012; although success may depend on the method of testing, as in Polka et al., 2017). Moreover bilingual infants' vocabularies are similar in size to monolinguals, particularly when words from both languages are taken into account (Cote & Bornstein, 2014; De Houwer, Bornstein, & Putnick, 2014; Hoff et al., 2012; Pearson, Fernandez, & Oller, 1993). This further suggests that segmentation, which is requisite for much of word learning, is not significantly delayed in bilinguals relative to monolinguals.

The present study explores how syllable co-occurrence cues may support segmentation in bilingual environments. Tracking the co-occurrence of sounds and syllables via statistical learning may allow individuals to find consistent chunks in fluent speech, which likely form meaningful units, such as words and morphemes (see Saffran & Kirkham, 2018 for a recent review of statistical learning in infancy). The universal nature of statistical segmentation cues makes them relevant in many if not all languages (though see Saksida et al., 2017). Given that even newborns have the capacity to track these regularities (Bosseler, Teinonen, Tervaniemi, & Huotilainen, 2016; Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009; Teinonen & Huotilainen, 2012), it is possible that bilingual infants may use co-occurrence regularities to support segmentation across many language pairings. This may be particularly relevant early in development, when infants' lexicons of segmented words are not large enough to detect word-level patterns relevant for segmentation (e.g. stress cues, Thiessen & Saffran, 2003).

Recently, Antovich and Graf Estes (2018) demonstrated that 14-month-old infants with bilingual experience could segment words using syllable-level co-occurrence patterns in two novel speech streams, whereas monolinguals could not. However, as with other segmentation cues, syllable regularities can conflict across languages (Polka et al., 2017; Tyler & Cutler, 2009; Weber & Cutler, 2006), which could diminish their reliability as a segmentation cue. The artificial languages used by Antovich and Graf Estes (2018) did not have overlapping syllable sets, so the syllable-level statistical regularities were congruent across languages. Given this design, infants could succeed by treating the two speech streams as distinct instances of one larger language system or by differentiating the co-occurrence statistics by speech stream. In natural language pairings it is likely that overlap occurs at the level of sounds and syllables (Von Holzen, Fennell, & Mani, 2018), which can influence learning (Havy, Bouchon, & Nazzi, 2016). This overlap may result

### Research Highlights

- Bilingual, but not monolingual, 16-month-olds were able to segment two novel speech streams using syllable co-occurrence regularities.
- Bilingual infants' dual speech stream statistical segmentation is robust, overcoming the challenges introduced by conflicting regularities across speech streams.
- Bilingual infants' segmentation performance was associated with the proportion of bilingual speakers in their language environments, suggesting that this experience was related to their success.
- Monolingual and bilingual infants did not differ on measures of global cognitive development, language acquisition, or socioeconomic status, suggesting these factors cannot account for segmentation differences.

in incongruent co-occurrence regularities across languages, which could increase statistical noise and disrupt segmentation.

The present experiment tested whether infants can acquire statistical regularities across two novel speech streams in the presence of diverging information. Leveraging their experience with two language systems, dual language learners may be able to maintain separate information about each speech stream's statistical regularities, avoiding conflict introduced by incongruent regularities. In the present experiment, distinct speakers produced each artificial language, supporting differentiation, but the languages also contained overlapping syllable inventories. If infants treated the two speech streams as different samples of one language, the syllable-level overlap would reduce the pooled syllable-to-syllable transitional probabilities of the wordforms, as sounds that occurred together consistently in one artificial language co-occurred with different syllables in the second language. Given the potential difficulty of this task and the short learning timeline inherent to lab-based tests, we elected to test 16-month-olds, who are slightly older than participants in previous work (e.g. 14-month-olds; Antovich & Graf Estes, 2018).

Past work using similar materials suggests that monolingual adults can segment overlapping languages using statistical cues. Weiss, Gerfen, and Mitchel (2009) investigated monolingual adults' abilities to segment two artificial languages presented in interleaved blocks using syllable co-occurrence regularities. They found that when the artificial languages were presented by distinct speakers, participants demonstrated successful segmentation regardless of whether the languages were independent (i.e. had no overlap between the languages) or if the languages shared syllables, introducing conflicting syllable-level co-occurrence statistics. However, when the languages included conflicting statistical regularities, but were not differentiated by speaker voices, participants did not demonstrate successful segmentation. This suggests that the conflicting syllable-level regularities across



languages inhibited segmentation when the speech streams were not easily distinguished.

One obstacle for dual language statistical segmentation may be the failure to detect a change in the underlying speech structure after a language switch. A number of studies have demonstrated that learning statistical regularities for two sequentially presented artificial languages is difficult for both infants and adults (e.g. Bulgarelli, Benitez, Saffran, Byers-Heinlein, & Weiss, 2017; Bulgarelli & Weiss, 2016; Gebhart, Aslin, & Newport, 2009). Gebhart et al. (2009) found that adult learners were unable to segment one artificial language after prior exposure to another artificial language unless the switch was explicit or after prolonged exposure to the second language. This was recently replicated with infants by Benitez, Bulgarelli, Byers-Heinlein, Saffran, and Weiss (2019), who found that 8-month-old monolinguals were unable to learn two sequentially presented artificial languages with syllable overlap regardless of whether they were tested on the first or second language or whether the languages were differentiated by vocal pitch or speaker accent. This finding aligns with monolingual performance in the dual language task from Antovich and Graf Estes (2018). It is possible that failure to detect the change in speech structure, and thus to accurately learn the languages' regularities, might be due to waning attention in the presence of a predictable stimulus (Bulgarelli & Weiss, 2016; Karuza et al., 2016).

Bilinguals' experience switching between languages may render them particularly adept at detecting changes in the speech stream. Onnis, Chun, and Lou-Magnuson (2018) had bilingual adults perform a multimodal audio-visual statistical learning task, finding that they readily learned the underlying regularities present in both languages. Interestingly, there were individual differences in learning that were associated with bilinguals' language dominance (i.e. balance of language use). Specifically, more balanced bilinguals, presumably individuals who regularly shift between languages, were better able to detect the underlying regularities within each artificial language. This suggests that real-world experience with language affects statistical learning of novel dual language stimuli.

The theoretical model for contextual change detection proposed by Qian, Jaeger, and Aslin (2012) provides a framework for interpreting findings in the dual language statistical learning literature. The authors suggest that anticipating context changes can lower learners' thresholds for detecting multiple systems in novel input. For these learners, a small deviation from the expected pattern may indicate a change in context. In contrast, learners who do not anticipate contextual change will have high tolerance for variability in the stimulus. These learners may fail to detect a change in context because of their prior expectation that contextual change is rare or non-existent. In the dual language statistical segmentation literature, expectations about the number of underlying structures may have influenced learning. When learners were given explicit information about the presence of multiple languages, they were more likely to succeed than when this information was implicit (Gebhart et al., 2009). Similarly, salient perceptual information accompanying

a change, such as change in speaker voice, appears to facilitate learning (Weiss et al., 2009).

Of relevance to the present study, bilinguals' prior experience with two linguistic systems may allow minor variation in the expected pattern to cue a language change. In Antovich and Graf Estes's (2018) work, changes in speaker voice and syllable structure may have provided enough information to support bilinguals' detection of the two artificial language contexts. Conversely, this may have been insufficient evidence to cue monolinguals. Monolingual experience would likely produce a very high threshold for language change detection given their experience with a single language (for related discussion, see Poepsel & Weiss, 2016; Potter & Lew-Williams, 2019). This mirrors evidence from Onnis et al. (2018) as well as related work demonstrating that bilingual infants were more likely than monolinguals to detect and adapt to changes in a pattern presented by auditory or visual stimuli (e.g. Kovács & Mehler, 2009a, 2009b). It is possible that individual differences in bilinguals' language experience, such as how regularly they hear each language and their likelihood of interacting with monolingual or bilingual individuals, affect speech processing.

In general, there have been mixed findings regarding the effects of bilingualism on statistical learning (see Bulgarelli, Lebkuecher, & Weiss, 2018 for a recent review). Yim and Rudoy (2013) found no differences between monolinguals' and bilinguals' performance in basic auditory and visual statistical learning tasks. However, these tasks did not simulate features specific to bilingual input (e.g. dual information streams). In contrast, studies showing bilingual advantages in statistical learning tasks have incorporated features that mimic bilingual input (Antovich & Graf Estes, 2018; Bartolotti, Marian, Schroeder, & Shook, 2011; Onnis et al., 2018; Poepsel & Weiss, 2016). For example, Poepsel and Weiss (2016) found that adult bilinguals were better able to learn word-object co-occurrences using cross-situational statistics than monolinguals. However, this advantage was only present when some labels referenced two objects, which required flexibly tracking overlapping label-object pairings, a common feature in bilingual vocabularies. The advantage was not present for learning single label-object mappings. Findings from dual input statistical learning tasks with infants (Antovich & Graf Estes, 2018) and adults (Onnis et al., 2018) also suggest that bilingualism's effects on statistical learning are most evident when learning task features align with features of bilinguals' language environments.

Thus, prior work suggests that bilingualism does not influence statistical learning mechanisms directly. Rather, aspects of bilingual experience affect how dual language learners approach these tasks. With this in mind, the present work assessed whether aspects of bilingual experience, specifically balance of exposure to each language and prevalence of bilingual caregivers, were associated with performance in a dual language statistical learning task. Infants with relatively balanced exposure to each language likely experience frequent language switches. This might prepare infants to monitor speech for language changes, supporting the ability to track independent regularities for each language. Furthermore, infants' expectations about whether unfamiliar speech comprises multiple

languages may be shaped by the prevalence of bilinguals among the individuals they hear daily. Infants with a large proportion of bilingual caregivers may develop a strong implicit expectation that new interactions will involve two languages. The present study assessed whether these factors were associated with bilingual segmentation in a task mimicking a dual language exchange.

In summary, the present experiment tested whether early bilingual experience shapes the ability to track statistical regularities in dual speech streams. As in past work (Antovich & Graf Estes, 2018), infants had to track statistical information across speech intervals to successfully segment the input. However, unlike that prior work, infants in the present study had to selectively integrate across samples of the same language to preserve reliable syllable co-occurrence regularities. If infants tracked statistical regularities across, rather than within, the languages, the co-occurrence patterns would not reliably indicate word boundaries. We compared monolingual and bilingual infants' performance in the task to assess whether dual language exposure affected learning. We also examined whether success in this difficult statistical learning task was associated with aspects of bilinguals' language environments. Based on related work, we predicted that bilingual, but not monolingual, infants would successfully segment the dual speech streams. Additionally, we hypothesized that having a larger proportion of bilingual speakers in dual language learners' environments and receiving more balanced exposure to each language would benefit segmentation in the present task (e.g. Onnis et al., 2018).

## 2 | METHOD

### 2.1 | Participants

Twenty-six monolingual (females = 14) 16-month-olds ( $M = 16.2$  months; range = 15.1 – 17.1 months) and 26 bilingual (females = 16) 16-month-olds ( $M = 16.3$  months; range = 15.1–18.1 months) participated in the present experiment. Parental consent was acquired for each participant in the study. Monolingual and bilingual infants came from comparable households, with similar household incomes estimated via ZIP-code census data (monolingual: \$64,028, bilingual: \$62,512; U.S. Census Bureau, 2017). Monolingual and bilingual households' average scores on a parent

education scale (ranging from 1: 8th grade completion to 7: Doctoral degree) were equivalent to a 4-year college degree for both language groups. As shown in Table 1, independent samples  $t$  tests indicated that there were no differences between groups in income or parent education.

Monolingual infants came from English-speaking households. Five infants in the monolingual group also heard a second language for up to 5% of their overall language experience, assessed via parent interview. Monolingual results were the same with these infants excluded. Bilingual infants were exposed to English as well as a second language for 15%–85% of their overall language input, per initial parent report. This range was broader than in prior related work (i.e. Antovich & Graf Estes, 2018), which allowed for greater variability in our secondary language experience measures (for similarly broad bilingual range criteria, see Hoff et al., 2012; Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009). Bilinguals' average exposure to English was 58%, and they heard a variety of second languages: Spanish ( $n = 13$ ), Mandarin (3), Russian (3), Arabic (1), Cantonese (1), Hindi (1), Italian (1), Korean (1), Polish (1), and Tagalog (1). A detailed language interview revealed that two of the bilingual infants were exposed to a third language for up to 15% of their overall language experience and one infant's exposure to English was less than 15%. To maintain statistical power, these infants were included in subsequent analyses. Infants with history of chronic ear infection, hearing loss, developmental delay, or premature birth were not included in the final sample. An additional 27 infants were excluded due to fussiness (monolingual: 6; bilingual: 8), excessive movement (monolingual: 6; bilingual: 2), inattention during testing (monolingual: 1; bilingual: 1), experimenter error (monolingual: 0; bilingual: 2), and equipment error (monolingual: 0; bilingual: 1).

In addition to our demographic measures, we assessed infants' general language abilities and global cognitive development to ensure that our monolingual and bilingual groups did not differ on these factors, which could affect performance (see Table 1). As a measure of general cognitive development, infants completed the Cognitive Scale of the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III; Bayley, 2006) during a follow-up visit. Parents also completed the English MacArthur-Bates Communicative Development Inventory: Words and Gestures (MCDI; Fenson et al., 2007). Bilingual parents supplemented this form by reporting on whether their infant understood or produced

Variable	Language group	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>t</i> value	<i>p</i> value
Estimated Income (in USD)	Monolingual	26	64,028 (18,783)	0.374	.710
	Bilingual	25	62,067 (18,673)		
Parental Education Scale (range 1–7)	Monolingual	26	5 (1)	−0.146	.884
	Bilingual	25	5 (1)		
MCDI Conceptual Vocabulary Size	Monolingual	24	184 (97)	−0.475	.637
	Bilingual	22	199 (115)		
Bayley-III Cognitive Scale	Monolingual	24	102 (4)	0.007	.994
	Bilingual	17	102 (9)		

**TABLE 1** Secondary Measure Comparisons by Language Group

names for each English MCDI concept in a second language. From these records we calculated infants' estimated conceptual vocabulary size (i.e. total number of concepts understood in any language) as well as the number of translational equivalents understood for bilinguals (i.e. total number of concepts understood across both languages).

Table 1 shows the total conceptual vocabulary size from the MCDI as well as the age-normalized composite scores for the Bayley. A set of independent samples *t* tests revealed that, as a group, monolinguals did not differ from bilinguals on either of these measures or the socioeconomic status variables. This suggests that any differences in performance between the two language groups on the segmentation task were not driven by large-scale differences in linguistic knowledge, cognitive skills, or demographic factors. On average, bilingual infants knew 121 translational equivalents across languages ( $SD = 120$ , range = 5–395).

## 2.2 | Stimuli

### 2.2.1 | Exposure phase

During the exposure period, infants listened to the two artificial languages (L1 and L2) in eight interleaved blocks. The languages' words are presented in Table 2. Importantly, the two languages' syllable inventories overlap, so that some word-offset syllables in L1 are word-onset syllables in L2 and vice versa. With this design, the high within-word transitional probabilities are only maintained if the syllable-level co-occurrences are tracked separately for each language. If learners collapse across the two artificial languages to calculate the syllable co-occurrence statistics, the distinction between within- and between-word transitional probabilities is greatly reduced.

For example, the L1 word *dobu* overlaps with the L2 word *pudo*. If the artificial languages are maintained separately, the within-word transitional probability for each word is 1.00, as within L1 each time *do* occurs, it is followed by *bu* and similarly for *pu* and *do* in L2. Across word boundaries, there is a dip in transitional probability (0.33), as the word-final syllable of each word can precede any of the three word-initial syllables. However, if the artificial languages are collapsed, the syllable *do* can be preceded or followed by several

syllables (i.e. after *pu*, *may*, *mo*, or *ga* and before *bu*, *mo*, *ri*, and *may*), reducing the within-word transitional probability to 0.50, which may not be as reliably discriminated from between-word transitional probabilities, which fall to 0.17. Although in both cases there is a difference between word and part-word transitional probabilities, the magnitude of this difference is twice as large if the languages' regularities are maintained separately (i.e. a difference of 0.77 if kept separate and a difference of 0.33 if collapsed). Additionally, the same syllable sometimes indicated a word onset and sometimes indicated a word offset.

L1 was presented in a male voice and L2 was presented in a female voice to promote discriminability of the two artificial languages. Each word was presented 120 times during the exposure period. The languages were not designed to frequency balance the test items. This allowed for comparison with prior work using similar artificial languages (Antovich & Graf Estes, 2018). Words were presented in a pseudorandom order, with the stipulation that no word occurred twice in succession. The L1 artificial language was identical in syllable structure and speaker to the male voice for L1 used by Antovich and Graf Estes (2018). The L2 artificial language was a modified version of the L2 language from the same study, with several syllables changed to add overlap across the artificial languages and a new speaker.

The artificial languages were edited and assembled using Praat speech editing software (Boersma & Weenink, 2012). To generate the exposure language stimuli, a male native English speaker recorded all possible three-syllable sequences for L1 and a female native English speaker recorded sequences for L2. The middle syllable of each of these sequences was extracted. This method allowed for appropriate coarticulation while reducing the likelihood of systematic differences in the syllables' prosodic properties. The resulting syllable sets were edited in Praat to ensure uniformity. Specifically, the syllables were monotonized, duration was adjusted to 260 ms, median pitch was standardized (L1: 110Hz; L2: 210Hz), and volume was set to 70dB. The individual syllables were then concatenated in the appropriate order to form approximately 4 min of fluent speech in each language. In order to interleave the stimuli, the artificial languages were both divided into four roughly equal segments, ranging from 55 to 68 s in duration. The segment intervals were chosen to prevent the onset and offset of the speech intervals from being reliable cues to word boundaries. Half of the segments started in the middle of a word (i.e. beginning with syllable 2) and half started at the onset of a word (i.e. beginning with syllable 1). Similarly, half the language segments ended in the middle of a word (i.e. ending with syllable 1) and half ended at the offset of a word (i.e. ending with syllable 2). The resulting eight segments of speech were interleaved, alternating between L1 and L2 with 5 s of silence between each segment. Two exposure sets were created to allow for counterbalancing. One set began with an L1 interval and ended with an L2 interval and another began with L2 and ended with L1. The order of segments within each language was maintained to ensure intra-language continuity.

**TABLE 2** Artificial Language Words and Part-words

Language 1 (L1)		Language 2 (L2)	
Words	Part-words	Words	Part-words
<b>timay</b>	<b>moti</b>	<b>mayta</b>	<b>tamo</b>
<b>dobu</b>	<b>gado</b>	<b>pudo</b>	<b>dori</b>
<b>kuga</b>		<b>mola</b>	
<b>pimo</b>		<b>riku</b>	

*Note:* One syllable from each L1 word occurs in L2 as well but occupies a different position within the L2 word (e.g. word-final syllable from L1 is word-initial L2). Test items are listed in bold.





## 2.2.2 | Test phase

During test trials infants heard repetitions of an isolated token of either a word or part-word test item. The word test items were two of the bisyllabic words used in each of the artificial languages, whereas the two part-words were syllable sequences that occurred in the speech stream but crossed word boundaries rather than forming words (see Table 2). The test items were recorded individually by the same speakers as the artificial language stimuli. The tokens were edited for uniformity using Praat. For each test item, the clip was monotonized, the duration was adjusted to 650 ms, the median pitch was standardized (L1: 110Hz; L2: 210Hz), and the volume was set to 70dB. After this adjustment, the test items were assembled to form 20 s test trials in which a single word or part-word was repeated with 800 ms of silence between each repetition. These recordings were each paired with a video of an orange rotating petal, identical to the video used by Antovich and Graf Estes (2018).

## 2.3 | Procedure

### 2.3.1 | Word segmentation task

During the exposure period, infants listened to the two artificial languages presented in eight alternating, interleaved blocks (e.g. L1, L2, L1, L2...) during approximately 8.5 min of quiet play with a caregiver in a small, sound-attenuated booth. Past work using similar stimuli has demonstrated that interleaved exposure may benefit learning in this type of task, particularly for acquisition of the second artificial language's statistics (Zinszer & Weiss, 2013). This design also simulates aspects of language switching that occurs in many bilingual environments.

Presentation order was counterbalanced so that approximately half of the infants initially heard L1 (i.e. L1, L2, L1, L2...) and the remaining infants heard L2 first (i.e. L2, L1, L2, L1...). When the exposure period was complete, infants and parents were moved to a sound attenuated booth in which infants completed a listening preference task (e.g. Cooper & Aslin, 1990). In the booth infants sat on a parent's lap approximately 3' from a large television monitor. Parents were asked to refrain from talking or directing the infant's attention during testing. Additionally, parents wore headphones playing classical music to reduce their likelihood of influencing infant attention. Prior to the test trials, infants heard 30 s refresher trials of both L1 and L2 while viewing a silent cartoon clip. The refresher trials were designed to bridge the brief gap between exposure and test. The order of appearance of the two refresher trials was randomized. During each of the remaining trials, infants viewed a video of an orange petal rotating in a circle in the center of the screen at approximate infant eye level. Immediately following the refresher trials, infants viewed a pretest trial, in which they saw the orange petal video accompanied by repetitions of an attention-getting phrase in English. This was designed to familiarize the infant with the

visual stimulus that was paired with each test trial. Afterward infants completed 16 test trials. The test trials included two words and two part-words (see Table 2), each tested four times, randomized within blocks of four trials.

To reduce the possibility of interference across the two artificial languages at test, infants were tested using either L1 or L2 stimuli, rather than both. Test condition was counterbalanced so that approximately half of the infants were tested on L1 and half were tested on L2. Note that assignment to exposure language order and to test language were independent. Trials began when infant gaze was centered on the screen and ended when infants had looked away for more than 1 s, or after a total of 20 s. Stimuli were presented and infant listening time was recorded using Habit2 software (Oakes, Sperka, & Cantrell, 2015). Coders were blind to test trial stimuli. Between trials infants briefly viewed an attention-getting video of a cartoon until infant gaze was re-centered on the screen. Test phase stimuli were presented on a large television monitor with computer speakers.

### 2.3.2 | Language background interview

During a structured interview with parents, we asked about the individuals who regularly interacted with the infant, the amount of time spent with those individuals, the languages they used, and the proportion of time each individual used each language around the child. Orena, Byers-Heinlein, and Polka (2019) recently found that parental report of the proportion of exposure to two languages aligns well with the proportions that occur in extensive natural language recordings. This suggests that parental report of bilingual input provides a good estimate of infants' actual language exposure. For bilinguals in the present work, the interview provided a detailed account of the child's exposure to each language and the manner in which they gained this experience. The difference in the proportion of exposure to English versus an infant's second language, calculated from the language background interview, was significantly correlated with the difference in total MCDI vocabulary size for both languages, that is, the difference in the total number of words understood in English and the infants' second language (Pearson's  $r = 0.45$ ,  $p = .039$ ). This suggests consistency across reporting measures.

In addition to verifying bilingual infants' overall exposure to each language, two predictor variables were derived from this questionnaire. First, we assessed the balance of infants' language experience, calculated as the absolute value of the difference in proportion of exposure to English versus other languages. An infant with 75% exposure to English and 25% exposure to Spanish would receive a score of 0.5 whereas an infant with equal English and Mandarin exposure would receive a score of 0. This metric was selected as past work with adults has found that the balance of bilingual language use was correlated with performance on statistical learning tasks (Onnis et al., 2018). Balanced exposure to two languages could provide infants with more real-world experience switching between languages than unbalanced exposure.

If this were true, we hypothesized that balanced bilinguals might be better prepared to monitor speech for changes and thus learn each speech streams' regularities.

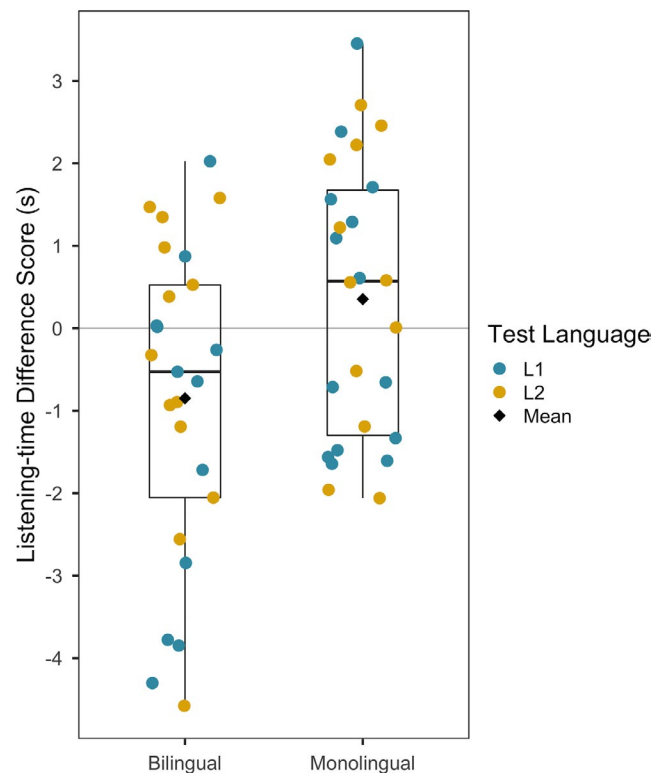
It is also possible that experience with bilingual individuals, regardless of balance of language exposure, might influence infants' implicit expectations about the nature of novel speech. We predicted that experience with bilingual speakers would develop or strengthen infants' expectations about the likelihood that speech will contain multiple structures. For example, if most individuals in an infant's environment use two languages, that infant might develop a strong implicit expectation that speech input will be bilingual in nature. In turn, this expectation could support infants' abilities to segment the speech by preventing accidental pooling of conflicting co-occurrence regularities across the two speech streams. To test this, we calculated the proportion of bilingual individuals among the infants' most frequent caregivers. This scale ranged from 0, indicating there were no frequent caregivers who interacted with the infant in two languages, to 1, indicating that each individual was reported to be bilingual. Individuals were categorized as bilingual if they used two or more languages during their regular interactions with the infant. During the interview, parents could report up to 10 individuals who interacted with the infant for at least an hour per week on average. Since parents reported on different numbers of individuals, using a proportion calculated from the total number of caregivers each parent reported allowed for comparison across families.

### 3 | RESULTS

#### 3.1 | Word segmentation task

For the initial analysis, difference scores were calculated by subtracting infants' average listening time during part-word trials from listening time during word trials. Negative scores indicate novelty (part-word) preferences and positive scores indicate familiarity (word) preferences. Preliminary analyses found that one infant from the bilingual group had a listening-time difference score greater than 3 SD from the group mean, so that infant's data were removed from all subsequent analyses (bilingual  $n = 25$ ; monolingual  $n = 26$ ). To assess whether there were any effects of exposure or testing condition on performance we conducted a 2 (exposure order: L1 first vs. L2 first)  $\times$  2 (test language: L1 vs. L2)  $\times$  2 (refresher order: L1 first vs. L2 first) ANOVA, which revealed that there were no significant main effects or interactions with language exposure order, refresher trial order, or test language in our preliminary analysis of listening-time difference scores; all  $ps > .369$ . Thus, we collapsed across these variables for all subsequent analyses.

To assess whether infants successfully discriminated between word and part-word test items and whether there was an effect of bilingualism on performance, we conducted a 2 (trial type: word vs. part-word; within subjects)  $\times$  2 (language group: monolingual vs. bilingual; between subjects) mixed ANOVA assessing infants' listening time. The analysis indicated that there were no main effects of language group or trial type but that there was a significant language



**FIGURE 1** Plot of individual infant listening-time difference scores (listening time to word trials – listening time to part-word trials) by bilingual status and test language superimposed with separate boxplots and means for monolingual and bilingual groups. There were no significant differences in listening-time difference scores across language versions (L1 versus L2) for monolinguals,  $t(24) = 0.449$ ,  $p = .657$ , or bilinguals,  $t(23) = -0.996$ ,  $p = .330$

group by trial type interaction,  $F(1, 49) = 5.678$ ,  $p = .021$ ,  $\eta^2 = 0.104$  (see Figure 1). To further probe this finding, we conducted a set of two-tailed, paired  $t$  tests to compare infant listening time to word and part-word test trials within each language group. Bilingual infants reliably discriminated between word and part-word test items,  $t(24) = -2.204$ ,  $p = .037$ ,  $d_z = -0.441$ , listening longer to the part-word items, whereas monolinguals did not,  $t(25) = 1.077$ ,  $p = .292$ ,  $d_z = 0.211$  (see Table 3).

This pattern of results suggests that bilinguals were able to use syllable-level co-occurrence patterns to detect individual words in fluent speech, whereas monolinguals were not. The difference between monolinguals' and bilinguals' performance in the present study replicates and extends the outcome of the dual language segmentation task used by Antovich and Graf Estes (2018). In contrast with that prior work, the reliability of the artificial languages' statistical structures was reduced if the co-occurrence patterns were determined globally, collapsed across the two languages, rather than locally, within each language. Bilinguals' success in the task suggests that they were able to track distinct statistical regularities for each language. It is possible that bilinguals' regular experience with two languages prepared them to anticipate and detect the multiple underlying structures within the speech and to use a salient cue, speaker voice, to discriminate between samples of each language during learning.

**TABLE 3** Listening-Time by Trial Type and Language Group

Language Group	Word <i>M</i> ( <i>SD</i> )	Part-Word <i>M</i> ( <i>SD</i> )
Bilingual	7.37(2.44)	8.22(2.64)
Monolingual	7.84(3.08)	7.48(3.08)

**TABLE 4** Descriptive Statistics for Language Experience Measures

Variable	<i>M</i> ( <i>SD</i> )	Range
Balance of Exposure	0.38(0.23)	0.11 – 0.79
Proportion of Bilingual Speakers	0.53(0.30)	0.00 – 1.00
Number of Bilingual Caregivers	2.7(1.8)	0 – 7
Number of Monolingual Caregivers	2.5(1.8)	0 – 6

### 3.2 | Language experience variables

Beyond assessing infants' segmentation of the dual language stimuli, the present study was designed to investigate the mechanisms supporting bilinguals' performance in the segmentation task. Informed by prior statistical learning work with bilingual adults (e.g. Onnis et al., 2018), we hypothesized that the nature of bilinguals' language experience could affect performance in the present segmentation task.

To investigate how bilinguals' language experience linked to performance in the segmentation task, we assessed whether listening-time difference scores (listening to words – part-word test items) were predicted by exposure balance and by the proportion of bilingual speakers among the infants' most frequent caregivers. Although listening-time scores have traditionally been used to assess group-level performance, there is evidence that the scores can be used to investigate individual differences, as in studies examining the relation between listening-time scores and cognitive abilities or language knowledge (e.g. listening measures of prosodic knowledge, Cristià & Seidl, 2011; Newman, Ratner, Jusczyk, Jusczyk, & Dow, 2006; statistical learning, Graf Estes, Gluck, & Grimm, 2016; Hoareau, Yeung, & Nazzi, 2019; Lany, Shoaib, Thompson, & Graf Estes, 2018; general word segmentation, Singh, Reznick, & Xuehua, 2012).

Five bilinguals from the primary analysis were not included in this analysis because their parents did not provide detailed language questionnaire data. Means and standard deviations for these variables, as well as infants' average number of monolingual and bilingual caregivers are presented in Table 4. In the model, we regressed bilingual infants' listening-time difference scores on our measures of bilingual speaker proportion and balance of exposure. The model accounted for a marginally significant amount of variance in listening-time difference scores,  $F(2, 17) = 3.36$ ,  $p = .059$ ,  $R^2 = 0.258$  (see Table 5). Of the two predictors, proportion of bilingual speakers, but not language balance, significantly predicted infants' performance in the dual language word segmentation task (see Figure 2 and Table 5). Having a larger proportion of bilingual caregivers was associated with a stronger novelty preference, aligning with the bilingual

**TABLE 5** Regression Analysis for Language Experience Measures

Variable	<i>B</i>	<i>SE</i>	<i>t</i> value	<i>p</i> value
Intercept	1.685	1.304	1.292	.214
Proportion of Bilingual Speakers	–3.874	1.495	–2.592	.019
Balance of Exposure	–1.592	1.945	–0.819	.424

group's overall preference. This association suggests that having a larger proportion of bilingual individuals in an infant's daily language environment supported performance in the dual language segmentation task. In contrast, balance of language exposure was not associated with performance in the segmentation task. This differs from past work demonstrating that adult bilinguals' performance in a dual stream statistical learning task was associated with balance of language use (Onnis et al., 2018).<sup>1</sup>

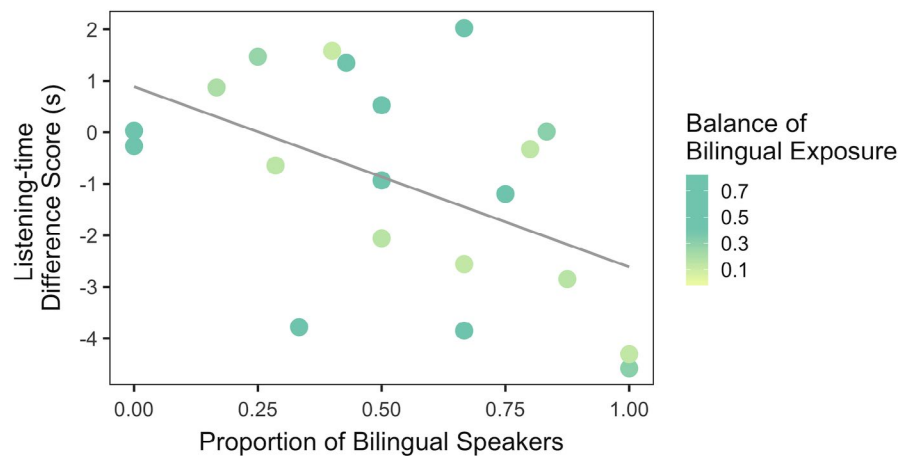
## 4 | DISCUSSION

The present study assessed whether infants were able to perform statistical word segmentation when presented with two speech streams that contain conflicting regularities, a challenge present in bilingual infants' natural language environments. We found that bilingual infants could navigate a demanding dual language segmentation task despite conflicting co-occurrence regularities across speech streams. In contrast, monolingual infants, who lack experience navigating dual language environments, displayed no evidence of segmenting the languages. The findings suggest that young bilinguals may leverage their dual language experience to detect multiple structures in novel linguistic input and to track information about this input categorically. In turn, this may help bilinguals learn the unique properties of their native languages. An additional novel contribution of this work is the examination of mechanisms supporting bilingual success in the dual language segmentation task. Past work has reported differences in monolingual and bilingual infants' performance in a dual language segmentation task without conflicting statistics (Antovich & Graf Estes, 2018), but it remained unclear what factors drove this difference.

It is possible that the bilingual group succeeded in the segmentation task due to advantages in general cognitive development, language acquisition, or family resources. Although bilinguals follow trajectories of linguistic development that are similar to monolinguals (e.g. Hoff et al., 2012), some work has suggested that bilinguals may demonstrate cognitive advantages even in infancy (e.g. Brito & Barr, 2014; Kovács & Mehler, 2009a, 2009b). Previous work also suggests that, in some cases, socioeconomic status may account for bilingual advantages in cognitive tasks (Naeem, Filippi, Periche-Tomas, Papageorgiou, & Bright, 2018). However, it is unlikely that bilinguals succeeded in the present task due to these factors; the language groups did not differ on measures of global cognitive development, conceptual vocabulary, or socioeconomic status (see Table 1).



**FIGURE 2** Plot of bilingual infants' looking-time difference scores (y axis) by proportion of bilingual speakers (x axis) and balance of bilingual exposure (color gradation). The regression line represents the relationship between the listening-time and bilingual speaker variables. Note that the balance of bilingual exposure variable is the difference in proportion of exposure to each language, thus larger values (and darker shades) represent less balanced exposure



In contrast, we did find that bilinguals' performance in the segmentation task was associated with the proportion of their caregivers who provided bilingual input. Infants with stronger novelty preferences in the word segmentation task had larger ratios of bilingual to monolingual caregivers in their daily interactions. This finding aligns with recent work from Hoareau et al. (2019), who found that infants' statistical word segmentation performance was predicted by the nature of their early language environments.

The relationship between bilingual experience and dual-input statistical word segmentation may be driven by differences in bilinguals' implicit expectations about the likelihood that unfamiliar speech will contain multiple structures. Having a large proportion of bilingual caregivers in an infant's speech environment may confer early metalinguistic knowledge about the existence of multiple language systems, suggesting that new input should be parsed to find unique speech streams. Infants who hear many bilingual speakers may have regular experience hearing individual caregivers produce and shift between multiple languages. Additionally, environments with a large proportion of bilingual speakers may provide opportunities to experience language switches across individuals. For example, bilingual parents might each favor a different language with the child. Similarly, hearing many bilingual speakers at home may provide a strong contrast between the home language(s) and those used by individuals outside the home environment. To test this idea directly, it will be important to further assess how dual language segmentation is affected by the nature of bilingual infants' language experience, such as caregiver code-switching and mixing behavior (e.g. Byers-Heinlein, 2013) or the structure of different language environments. This information can support our understanding of how bilingual infants track patterns from the same voice and across different voices.

Although bilingual experience is clearly relevant for performance in the dual language segmentation task, the predicted association between balance of bilingual exposure and segmentation was not supported. This departure from adult findings (i.e. Onnis et al., 2018) may be due to differences in balance measures or tests of learning, rather than true developmental differences. In the work by Onnis et al. (2018), the adult language dominance measure assessed the ease with which individuals used each of their languages. For infants,

we examined the estimated rates at which the languages were used in their environments. It is likely that measures of adult language dominance reflect individuals' unique language processing abilities, whereas measures of infant language exposure may not be as closely tied to infants' individual capabilities. Future research using more direct, infant-driven measures of bilingual language knowledge and processing may better test the role of language balance in infant word segmentation. Another critical difference between the adult and infant work is testing method; infants were tested using a preferential-listening paradigm, whereas adults were tested using a two-alternative forced-choice task. It is possible that the adult task captured a different aspect of segmentation performance than the infant task (see Siegelman, Bogaerts, & Frost, 2017 for discussion of individual differences in adult statistical learning).

Our findings suggest that bilingual success in the dual language statistical learning task may be driven by expectations about dual linguistic structures that come from bilingual experience. Bilingual experience, and specifically having a large proportion of bilingual caregivers, may help infants detect multiple speech streams in their input. This facilitation may be driven by increased sensitivity to variability, allowing subtle speech changes to highlight a language shift (Qian et al., 2012). In turn, this may lead bilinguals to monitor speech for changes in language. This attentional pattern would benefit bilinguals' real-world speech processing, as detecting the current language context is important for successfully acquiring each of the learners' languages. This is also reflected in bilinguals' real-world speech processing of dual language input (Byers-Heinlein et al., 2017).

Although it may generally benefit bilinguals to interpret variability as indication of language switches, it is possible that this would not be adaptive in all contexts. In particular, this strategy could lead infants to overestimate the number of distinct underlying structures present in the speech stream and thus dilute the statistical information collected for each structure (i.e. language). If so, this process may be subject to a Goldilocks effect (e.g. Kidd, Piantadosi, & Aslin, 2014), in which infants must titrate the level of variability that indicates change in language versus meaningless noise (e.g. variation in accent within one language). Future work should examine the cues that infants use to determine whether variability is meaningful



(i.e. indication of a language switch) or meaningless (e.g. change in speaker without change in language) in bilingual contexts (see Graf Estes & Lew-Williams, 2015 for related discussion in a monolingual context).

Another potential explanation for differences between monolinguals and bilinguals in the present task comes from recent research by Bulgarelli and Weiss (2016). Their work suggests that difficulty acquiring statistical regularities from sequentially presented artificial languages may be due to entrenchment during the learning process. In other words, overlearning of the first artificial language may inhibit learning of the second language. The authors presented monolingual adults with two artificial languages, testing them periodically throughout the exposure phase. Once subjects' learning of the first language reached a criterion point, some participants began exposure to the second language. Other participants were exposed to the full duration of the first artificial language before hearing the second language (see also Gebhart et al., 2009). Adults only succeeded in segmenting the second language when exposure to the first language ended immediately upon reaching criterion.

This work from Bulgarelli and Weiss (2016) suggests that it is possible for monolinguals to learn sequential sets of co-occurrence statistics but that overlearning of the first artificial language may inhibit learning of the second language. The authors suggest that this effect may be due to waning attention once sequences become highly predictable. This aligns with the hypothesis that bilinguals may have succeeded in our task due to greater sensitivity to changes in context or increased monitoring for language change. The effect of waning attention can be mitigated, at least for adults, by presenting the languages in an interleaved fashion (Zinszer & Weiss, 2013). Given that prior work examining statistical learning has found that infants rapidly learn the statistical structure of artificial languages (e.g. with 2 min of exposure; Saffran, Aslin, & Newport, 1996), it is possible that monolinguals' attention to the initially predictable speech stream rapidly declined and they subsequently failed to detect the shift between the two languages. Past work demonstrates that infants' attention to novel auditory and visual stimuli varies based on the level of structure and predictability within the stimulus (e.g. Addyman & Mareschal, 2013; Kidd et al., 2014). If monolinguals did not redirect attention to the speech stream after the change in speaker, they may have tracked the regularities across, rather than within, the languages. Thus, the conflicting statistical information may have prevented monolingual infants from discriminating between word and part-word syllable sequences at test.

If success in the dual language statistical learning task requires detection of separate structures in the speech stream, it is possible that additional cues to highlight differences between the two languages (e.g. different prosody, the addition of visual cues to language change) could scaffold monolingual infants' abilities to segment the speech. This is supported by recent work from Potter and Lew-Williams (2019), which found that monolingual infants were able to track regularities in two streams, but only when the speech

sounds used in the two streams differed; when the sounds were identical, infants failed to track both patterns even with a change in speaker voice.

In the present study, limited information (speaker, syllable set) provided cues to differentiate the artificial languages, but in natural speech, a number of additional cues support language discrimination. While bilinguals succeeded in the present experiment with speaker identity as a cue to language differentiation, it is likely that adding naturalistic cues, like speaker accent would further support learning (Tsui, Erickson, Thiessen, & Fennell, 2017, but see Benitez et al., 2019). Bilingual infants also use auditory cues (e.g. prosodic information; Bosch & Sebastián-Gallés, 2001) in conjunction with overlapping cues provided by visual information from the mouth of a speaker to discriminate between languages and process phonological information (Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012; Weikum et al., 2007). Importantly, infants can also integrate auditory and visual information from speech to learn sound properties of their native languages (Teinonen, Aslin, Alku, & Csibra, 2008). Given that statistical learning mechanisms can incorporate audio-visual input cross-modally (Mitchel & Weiss, 2011), it is possible that additional language discrimination cues in the visual modality would support monolingual infants' abilities to segment the dual speech streams. More broadly, this type of visual input may scaffold bilinguals' processing of complex, naturalistic bilingual speech, given bilingual infants' attentional preference for the mouth during face and speech processing (Ayneto & Sebastián-Gallés, 2017; Fort, Ayneto-Gimeno, Escrichs, & Sebastián-Gallés, 2018; Pons, Bosch, & Lewkowicz, 2015).

An additional alternate explanation for monolinguals' failure in the preferential listening task is that they were able to track the regularities in the speech but were unable to retain the wordforms elicited via those regularities across the brief delay between exposure and test. Past work suggests that infants' memories for newly segmented words are fragile, and thus even the short delay might have impeded their ability to recognize previously segmented wordforms at test (Benavides-Varela et al., 2011; Karaman & Hay, 2018). There is some evidence that bilinguals may demonstrate advantages related to memory (e.g. Brito & Barr, 2014; Hernández, Costa, & Humphreys, 2012), though this hypothesized benefit of bilingualism has not been systematically studied in infancy. We did not find a bilingual advantage on the Cognitive Scale of the Bayley-III, which includes items related to information encoding and retrieval (e.g. A-not-B hiding task), but we cannot rule out this alternative possibility, as we did not test bilingual infants' verbal working memory directly. Future work could assess this possibility using a similar statistical segmentation task with or without delay between exposure and test.

The present work assessed bilingual infants who were already well on their way to building a lexicon (see Table 1). If bilingual infants use syllable co-occurrence cues to segment dual language speech, it is likely that this ability develops significantly earlier than 16 months (e.g. 6- to 9-months of age). With the basic learning pattern established in the present



research, future work can address early learning directly by testing younger infants with bilingual experience. Importantly, work using artificial language paradigms tests infants' abilities to acquire novel, though concentrated, regularities over the course of mere minutes, whereas infants' real-world segmentation is supported by months of language data. Thus, it may be difficult to detect this form of statistical segmentation with younger infants in the lab. Development of more sensitive measures of segmentation would improve our ability to assess statistical learning in bilingual contexts at an earlier age and could provide a means for reliable assessment of individual differences in word segmentation and the influence of experience on segmentation performance.

## 5 | CONCLUSION

To summarize, we found that infants with bilingual experience, but not monolinguals, segmented the speech incorporating dual streams. Our results support past work from Antovich and Graf Estes (2018) and elaborate on that finding by demonstrating that bilingual infants were able to segment words from interleaved speech streams with conflicting regularities, simulating bilingual input, whereas monolinguals were not. Monolingual and bilingual infants did not differ in general cognitive, linguistic, or demographic measures, suggesting that bilinguals' advantage in the task was not driven by these factors. In contrast, infants who heard bilingual input from a greater proportion of caregivers showed stronger evidence of segmentation. This suggests that bilingual experience prepares infants to track the complex syllable co-occurrence patterns they may encounter in dual language speech. Further work is necessary to determine whether, as we suggest, bilinguals are more attuned to the presence of multiple linguistic systems than monolinguals and thus better prepared to use changes in speaker voice or phonology as indication of a change in the underlying structure of the speech.

## ACKNOWLEDGMENTS

This work was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant Number: R03HD084941R03HD084941), The UC Davis Institute for Social Sciences, and the Childs Family Fund.

## CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, Dylan Antovich (dmantovich@ucdavis.edu), upon reasonable request.

## ORCID

Dylan M. Antovich  <https://orcid.org/0000-0002-8276-0294>

## ENDNOTE

<sup>1</sup> To further explore the data, we performed additional post hoc analyses to assess the relationship between the bilinguals' vocabulary measures and performance in the statistical word segmentation task. A set of zero-order correlations revealed that segmentation performance was not associated with difference in vocabulary size across languages (Pearson's  $r = 0.21$ ,  $p = .368$ ), translational equivalents (Pearson's  $r = 0.163$ ,  $p = .493$ ), or total conceptual vocabulary size (Pearson's  $r = 0.08$ ,  $p = .727$ ) for bilingual infants.

## REFERENCES

- Addyman, C., & Mareschal, D. (2013). Local redundancy governs infants' spontaneous orienting to visual-temporal sequences. *Child Development*, 84(4), 1137–1144. <https://doi.org/10.1111/cdev.12060>
- Antovich, D. M., & Graf Estes, K. (2018). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*, 21(2), e12548. <https://doi.org/10.1111/desc.12548>
- Ayneto, A., & Sebastián-Gallés, N. (2017). The influence of bilingualism on the preference for the mouth region of dynamic faces. *Developmental Science*, 20(1), <https://doi.org/10.1111/desc.12446>
- Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, 2, <https://doi.org/10.3389/fpsyg.2011.00324>
- Bayley, N. (2006). *Bayley scales of infant and toddler development-third edition*. San Antonio, TX: Harcourt Assessment.
- Benavides-Varela, S., Gómez, D. M., Macagno, F., Bion, R. A. H., Peretz, I., & Mehler, J. (2011). Memory in the neonate brain. *PLoS ONE*, 6(11), e27497. <https://doi.org/10.1371/journal.pone.0027497>
- Benitez, V. L., Bulgarelli, F., Byers-Heinlein, K., Saffran, J. R., & Weiss, D. J. (2019). Statistical learning of multiple speech streams: A challenge for monolingual infants. *Developmental Science*, <https://doi.org/10.1111/desc.12896>
- Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer*. Amsterdam, Netherlands: University of Amsterdam. Retrieved from <http://www.praat.org>.
- Bosch, L., Figueras, M., Teixidó, M., & Ramon-Casas, M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: Evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology*, 4, <https://doi.org/10.3389/fpsyg.2013.00106>
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29–49. [https://doi.org/10.1207/S15327078IN0201\\_3](https://doi.org/10.1207/S15327078IN0201_3)
- Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS ONE*, 11(9), e0162177. <https://doi.org/10.1371/journal.pone.0162177>
- Brito, N. H., & Barr, R. (2014). Flexible memory retrieval in bilingual 6-month-old infants. *Developmental Psychobiology*, 56(5), 1156–1163. <https://doi.org/10.1002/dev.21188>
- Bulgarelli, F., Benitez, V., Saffran, J. R., Byers-Heinlein, K., & Weiss, D. J. (2017). Statistical learning of multiple structures by 8-month-old infants. *Proceedings of the Annual Boston University Conference on Language Development*. Boston University Conference on Language Development, 41, 128–139. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5576994/>.
- Bulgarelli, F., Lebkuecher, A. L., & Weiss, D. J. (2018). Statistical learning and bilingualism. *Language, Speech, and Hearing Services in Schools*, 49(3S), 740–753. [https://doi.org/10.1044/2018\\_LSHSS-STLT1-17-0139](https://doi.org/10.1044/2018_LSHSS-STLT1-17-0139)
- Bulgarelli, F., & Weiss, D. J. (2016). Anchors aweigh: The impact of over-learning on entrenchment effects in statistical learning. *Journal of*



- Experimental Psychology: Learning, Memory, and Cognition*, 42(10), 1621–1631. <https://doi.org/10.1037/xlm0000263>
- Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition*, 16(01), 32–48. <https://doi.org/10.1017/S1366728912000120>
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343–348. <https://doi.org/10.1177/0956797609360758>
- Byers-Heinlein, K., Morin-Lessard, E., & Lew-Williams, C. (2017). Bilingual infants control their languages as they listen. *Proceedings of the National Academy of Sciences of the United States of America*, 114(34), 9032–9037. <https://doi.org/10.1073/pnas.1703220114>
- Census Bureau, U. S. (2017). American Community Survey 2012–2016 5-year Estimates. Retrieved from [http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS\\_14\\_5YR\\_DP03&src=pt](http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_14_5YR_DP03&src=pt)
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61(5), 1584–1595. <https://doi.org/10.1111/j.1467-8624.1990.tb02885.x>
- Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, 15(5), 336–345. <https://doi.org/10.1038/nrn3709>
- Cote, L. R., & Bornstein, M. H. (2014). Productive vocabulary among three groups of bilingual American children: Comparison and prediction. *First Language*, 34(6), 467–485. <https://doi.org/10.1177/0142723714560178>
- Cristià, A., & Seidl, A. (2011). Sensitivity to prosody at 6 months predicts vocabulary at 24 months. In N. Danis, K. Mesh & H. Sung (Eds.), *BUCLD 35: Proceedings of the 35th annual Boston University Conference on Language Development* (pp. 145–156). Boston, MA: Cascadia Press.
- Crystal, D. (2004). *The Cambridge encyclopedia of the English language*. Cambridge, UK: Cambridge University Press.
- Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, 39(4), 492–504. <https://doi.org/10.1016/j.wocn.2010.12.002>
- De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, 35(06), 1189–1211. <https://doi.org/10.1017/S0142716412000744>
- Fenson, L., Bates, E. A., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur-Bates Communicative Development Inventories*. Baltimore, MD: Brookes.
- Fort, M., Ayneto-Gimeno, A., Escrichs, A., & Sebastián-Gallés, N. (2018). Impact of bilingualism on infants' ability to learn from talking and nontalking faces. *Language Learning*, 68(S1), 31–57. <https://doi.org/10.1111/lang.12273>
- 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. <https://doi.org/10.1111/j.1551-6709.2009.01041.x>
- Graf Estes, K., Gluck, S.-C.-W., & Grimm, K. J. (2016). Finding patterns and learning words: Infant phonotactic knowledge is associated with vocabulary size. *Journal of Experimental Child Psychology*, 146, 34–49. <https://doi.org/10.1016/J.JECP.2016.01.012>
- Graf Estes, K., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology*, 51(11), 1517. <https://doi.org/10.1037/a0039725>
- Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard University Press.
- Havy, M., Bouchon, C., & Nazzi, T. (2016). Phonetic processing when learning words: The case of bilingual infants. *International Journal of Behavioral Development*, 40(1), 41–52. <https://doi.org/10.1177/0165025415570646>
- Hernández, M., Costa, A., & Humphreys, G. W. (2012). Escaping capture: Bilingualism modulates distraction from working memory. *Cognition*, 122(1), 37–50. <https://doi.org/10.1016/j.cognition.2011.08.002>
- Hoareau, M., Yeung, H. H., & Nazzi, T. (2019). Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities. *Developmental Science*, 22(4), e12803. <https://doi.org/10.1111/desc.12803>
- Hoff, E., Core, C., Place, S., Rumiche, R., Senor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39(01), 1–27. <https://doi.org/10.1017/S0305000910000759>
- Karaman, F., & Hay, J. F. (2018). The longevity of statistical learning: When infant memory decays, isolated words come to the rescue. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(2), 221–232. <https://doi.org/10.1037/xlm0000448>
- Karuz, E. A., Li, P., Weiss, D. J., Bulgarelli, F., Zinszer, B. D., & Aslin, R. N. (2016). Sampling over nonuniform distributions: A neural efficiency account of the primacy effect in statistical learning. *Journal of Cognitive Neuroscience*, 28(10), 1484–1500. [https://doi.org/10.1162/jocn\\_a\\_00990](https://doi.org/10.1162/jocn_a_00990)
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks Effect in infant auditory attention. *Child Development*, 85(5), 1795–1804. <https://doi.org/10.1111/cdev.12263>
- Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16), 6556–6560. <https://doi.org/10.1073/pnas.0811323106>
- Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, 325(5940), 611–612. <https://doi.org/10.1126/science.1173947>
- Lany, J., Shoaib, A., Thompson, A., & Graf Estes, K. (2018). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, 45(2), 368–391. <https://doi.org/10.1017/S0305000917000253>
- Mitchel, A. D., & Weiss, D. J. (2011). Learning across senses: Cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1081–1091. <https://doi.org/10.1037/a0023700>
- Naeem, K., Filippi, R., Periche-Tomas, E., Papageorgiou, A., & Bright, P. (2018). The importance of socioeconomic status as a modulator of the bilingual advantage in cognitive ability. *Frontiers in Psychology*, 9, 1818. <https://doi.org/10.3389/fpsyg.2018.01818>
- Nazzi, T., Iakimova, G., Bertoni, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54(3), 283–299. <https://doi.org/10.1016/j.jml.2005.10.004>
- Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, 42(4), 643–655.
- Oakes, L. M., Sperka, D. J., & Cantrell, L. (2015). *Habit 2. Center for Mind and Brain*. Davis, CA: University of California, Davis.
- Onnis, L., Chun, W. E., & Lou-Magnuson, M. (2018). Improved statistical learning abilities in adult bilinguals. *Bilingualism: Language and Cognition*, 21(2), 427–433. <https://doi.org/10.1017/S1366728917000529>
- Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). What do bilingual infants actually hear? Evaluating measures of language input to bilingual-learning 10-month-olds. *Developmental Science*, 23(2), e12901. <https://doi.org/10.1111/desc.12901>





- Orena, A. J., & Polka, L. (2019). Monolingual and bilingual infants' word segmentation abilities in an inter-mixed dual-language task. *Infancy*, 24(5), 1229–1296. <https://doi.org/10.1111/inf.12296>
- Pearson, B. Z., Fernandez, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, 43(1), 93–120. <https://doi.org/10.1111/j.1467-1770.1993.tb00174.x>
- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, 152, 9–19. <https://doi.org/10.1016/j.cognition.2016.03.001>
- Polka, L., Orena, A. J., Sundara, M., & Worrall, J. (2017). Segmenting words from fluent speech during infancy – challenges and opportunities in a bilingual context. *Developmental Science*, 20(1), e12419. <https://doi.org/10.1111/desc.12419>
- Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science*, 26(4), 490–498. <https://doi.org/10.1177/0956797614568320>
- Potter, C. E., & Lew-Williams, C. (2019). Infants' selective use of reliable cues in multidimensional language input. *Developmental Psychology*, 55(1), 1–8. <https://doi.org/10.1037/dev0000610>
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in Psychology*, 3, <https://doi.org/10.3389/fpsyg.2012.00228>
- Ramon-Casas, M., Swingle, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, 59(1), 96–121. <https://doi.org/10.1016/j.cogpsych.2009.02.002>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, 69(1), 181–203. <https://doi.org/10.1146/annurev-psych-122216-011805>
- Saksida, A., Langus, A., & Nespore, M. (2017). Co-occurrence statistics as a language-dependent cue for speech segmentation. *Developmental Science*, 20(3), e12390. <https://doi.org/10.1111/desc.12390>
- Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, 23(9), 994–999. <https://doi.org/10.1177/0956797612436817>
- Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, 49(2), 418–432. <https://doi.org/10.3758/s13428-016-0719-z>
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, 124(2), 128–142. <https://doi.org/10.1016/j.cognition.2012.05.008>
- Singh, L., Reznick, J. S., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science*, 15(4), 482–495. <https://doi.org/10.1111/j.1467-7687.2012.01141.x>
- Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, 108(3), 850–855. <https://doi.org/10.1016/j.cognition.2008.05.009>
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BioMed Central Neuroscience*, 10(1), 21. <https://doi.org/10.1186/1471-2202-10-21>
- Teinonen, T., & Huotilainen, M. (2012). Implicit segmentation of a stream of syllables based on transitional probabilities: An MEG study. *Journal of Psycholinguistic Research*, 41(1), 71–82. <https://doi.org/10.1007/s10936-011-9182-2>
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology*, 39(4), 706–716.
- Tsui, A.-S.-M., Erickson, L. C., Thiessen, E. D., & Fennell, C. T. (2017). Statistical learning from accented speech: A bilingual advantage. In M. LaMendola & J. Scott (Eds.), *Proceedings of the 41st annual Boston University Conference on Language Development* (pp. 679–690). Somerville, MA: Cascadia Press.
- Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, 126(1), 367–376. <https://doi.org/10.1121/1.3129127>
- Von Holzen, K., Fennell, C. T., & Mani, N. (2018). The impact of cross-language phonological overlap on bilingual and monolingual toddlers' word recognition. *Bilingualism: Language and Cognition*, 22(3), 1–24. <https://doi.org/10.1017/S1366728918000597>
- Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America*, 119(1), 597–607. <https://doi.org/10.1121/1.2141003>
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, 316(5828), 1159. <https://doi.org/10.1126/science.1137686>
- 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. <https://doi.org/10.1080/15475440802340101>
- Werker, J. F., Byers-Heinlein, K., & Fennell, C. T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1536), 3649–3663. <https://doi.org/10.1098/rstb.2009.0105>
- Yim, D., & Rudy, J. (2013). Implicit statistical learning and language skills in bilingual children. *Journal of Speech, Language, and Hearing Research*, 56(1), 310–322. [https://doi.org/10.1044/1092-4388\(2012/11-0243\)](https://doi.org/10.1044/1092-4388(2012/11-0243))
- Zinszer, B., & Weiss, D. J. (2013). When to hold and when to fold: Detecting structural changes in statistical learning. *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35). Merced, CA: Proceedings of the Annual Meeting of the Cognitive Science Society.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Antovich DM, Graf Estes K. One language or two? Navigating cross-language conflict in statistical word segmentation. *Dev Sci*. 2020;00:e12960. <https://doi.org/10.1111/desc.12960>