Cognition 140 (2015) 60–71

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Does hearing two dialects at different times help infants learn dialect-specific rules?

Kalim Gonzales [a,*], LouAnn Gerken [b], Rebecca L. Gómez [b]

[a] National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, Guangdong Province 510420, China
[b] Department of Psychology, The University of Arizona, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 14 September 2013
Revised 28 March 2015
Accepted 30 March 2015
Available online 11 April 2015

Keywords:
Rule learning
Dynamic change
Memory development

ABSTRACT

Infants might be better at teasing apart dialects with different language rules when hearing the dialects at different times, since language learners do not always combine input heard at different times. However, no previous research has independently varied the temporal distribution of conflicting language input. Twelve-month-olds heard two artificial language streams representing different dialects—a "pure stream" whose sentences adhered to abstract grammar rules like aX bY, and a "mixed stream" wherein any a- or b-word could precede any X- or Y-word. Infants were then tested for generalization of the pure stream's rules to novel sentences. Supporting our hypothesis, infants showed generalization when the two streams' sentences alternated in minutes-long intervals without any perceptually salient change across streams (Experiment 2), but not when all sentences from these same streams were randomly interleaved (Experiment 3). Results are interpreted in light of temporal context effects in word learning.

1. Introduction

Infants commonly hear more than one dialect of their native language, such as a standard variety of the language spoken on nationally broadcasted television shows and a nonstandard variety spoken in the local community. Though typically mutually intelligible, dialects have distinct vocabulary, pronunciation, and grammar rules. Infants exposed to more than one dialect thus receive conflicting information about appropriate language usage.

Regarding grammar rules, for example, varieties of Standard English have subject–verb agreement rules that some nonstandard varieties lack. In standard varieties, the main verb of a sentence should typically take a different form when the subject is plural than when singular. Thus, whereas when the subject is singular (e.g., *The dog*) the verb *to be* in simple tense should typically be *is* or *was* (e.g., *The dog is/was fast*), when plural (e.g., *The dogs*) it should typically be *are* or *were* (e.g., *The dogs are/were fast*). In some varieties of nonstandard English (e.g., Appalachian English), in contrast, the main verb can typically take the same form when the subject is plural as when singular, such that *The dogs is/was fast* is grammatical (see Hazen, 2014). This means that infants hearing both Standard English and one of these nonstandard varieties receive conflicting information regarding what form a sentence's main verb can take when the subject is plural rather than singular. How do infants extract the rules of one particular dialect of their native language when another heard dialect lacks those rules?

* Corresponding author at: National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Baiyun District, Baiyun North Avenue #2, Guangzhou, Guangdong Province 510420, China. Tel.: +86 20 36618025 (Office).

E-mail address: kalimg@gdufs.edu.cn (K. Gonzales).

http://dx.doi.org/10.1016/j.cognition.2015.03.015
0010-0277/

1.1. Combining dialects

One possibility is that infants combine information about rules across dialects to discover general tendencies across all utterances in their native language input, irrespective of dialect. Consequently, infants' ability to learn the rules unique to a particular dialect depends on whether they can learn those rules from statistical averages across dialects. For example, whether infants converge on the rule that a sentence whose subject is plural must have a main verb that is also plural depends on whether the average probability of those units occurring together is sufficiently high across dialects. If this is true, infants' ability to extract dialect-specific rules might be rather limited, since infants show only moderate tolerance for noisy language input. There is certainly evidence that in the rare case where children receive grammatically inconsistent input exclusively from nonnative speakers, they do not represent this input veridically but rather impose consistency onto it, such that their own speech actually boasts more nativelike consistency than that of their language models (Hudson Kam & Newport, 2005, 2009; Ross & Newport, 1996; Senghas & Coppola, 2001; Singleton & Newport, 2004). However, the rules to which these children adhere do not always exist in the target language but may be of the children's own invention. Indeed, young learners may not even discriminate between rules in the target language and alternative rules if the former are violated too much. Gómez and Lakusta (2004) familiarized 12-month-olds with an artificial language containing rules such as aX bY, where the a- and b-categories comprised high frequency words and the X- and Y-categories lower frequency words distinguished by syllable length. Infants discriminated the language's target rules from opposite rules (e.g., aY bX) if only 0% or 17% of familiarization sentences conformed to the latter but not if 33% did so.

1.2. Separating dialects due to salient perceptual differences between them

Rather than infants' ability to extract the rules of a particular dialect hinging on whether they can extract those rules from statistical averages across dialects, infants might be able to compute statistics more conducive to learning individual dialects. For example, due to infants' sensitivity to differences in the rhythmic properties of two dialects (Molnar, Gervain, & Carreiras, 2013), or differences in the talker characteristics of the people speaking those dialects (Schmale & Seidl, 2009), infants might computationally separate them. Returning to our subject–verb agreement example, infants might compute within each dialect separately co-occurrence statistics between plural subjects and plural versus singular forms of the main verb. Indeed, it is often similarly suggested that bilingual learners separate their two languages based on salient perceptual differences between them such as linguistic rhythm or talker voice (e.g., Byers-Heinlein, Burns, & Werker, 2010; Curtin, Byers-Heinlein, & Werker, 2011; Mitchel & Weiss, 2010; Nazzi, Bertoncini, & Mehler, 1998; Ramus & Mehler, 1999; Sundara & Scutellaro, 2011; Weiss, Gerfen, & Mitchel, 2009; see Byers-Heinlein, 2014). Experimental support comes from Weiss et al. (2009), who familiarized adults with two artificial languages alternating in 2 min periods. Each language consisted of a stream of words whose boundaries were defined statistically: Adjacent syllables and phonemes had lower co-occurrence probabilities when spanning words than when falling within them. However, the boundaries in one language conflicted with those in the other. Results showed that adults recognized words from both languages only when the languages were spoken in different voices. Without this indexical cue, they recognized words from only one language, or from neither. The authors argued that the indexical cue helped adults pull out words from both languages by enabling them to compute statistics within each language separately, as opposed to combining statistics across them.

1.3. Separating dialects due to their temporal separation

Despite considerable interest in the notion that salient perceptual dimensions like talker voice help infants tease apart conflicting language input, infants might not necessarily rely on such dimensions so long as two languages or dialects are already contextually distinguished in terms of when they are spoken. Evidence is mounting that language learners are sensitive to the dynamics of their input such that they sometimes separate statistics heard at different times. Gebhart, Aslin, and Newport (2009) examined the conditions under which learners confronted with a change in language structure after a period of time can represent both the initial and subsequent structure. The authors presented two artificial languages in succession, the first for 5 min and the second for the same duration or longer. As in Weiss et al. (2009), statistical cues to word boundaries conflicted across languages. Results showed that adults were able to segment the first language regardless of whether they segmented the second, which they segmented only when its duration was longer or when instructed that two languages would be presented and the language transition was marked by a pause. The authors suggested that since first language learning did not depend on second language learning, the languages were represented independently. One reason adults in Gebhart et al. (2009) might not have combined languages as in Weiss et al. (2009) is that the first language's initial period was comparatively longer (5 min rather than 2), which might have permitted representation of it strong enough to resist combining with the second language. Regardless, Gebhart et al.'s (2009) results suggest that hearing two dialects at different times might be helpful for teasing them apart to learn dialect-specific rules. Note, though, that learners might have more difficulty learning rules unique to the second than the first of two dialects, given Gebhart et al.'s order effect.

Young infants might be especially sensitive to the occurrence of conflicting language input in different temporal contexts given recent evidence that the temporal range over which they integrate statistical information is subject to memory limitations. Vlach and Johnson (2013) compared 16- and 20-month-olds on a cross-situational word learning task in which a word's referent could be learned only by integrating word-object co-occurrences across multiple individually ambiguous scenes. Results showed that although 20-month-olds were able to integrate 6 word-object pairings whether these were presented successively for roughly half a minute or instead spread out over roughly 3 min with largely irrelevant pairings interleaved, 16-month-olds integrated the pairings only across the shorter period. After ruling out attentional factors, the authors argued that 16-month-olds did not aggregate word-object pairings across the longer period like 20-month-olds because they had not yet developed the memory capacity to do so. Like Gebhart et al.'s (2009) results, then, these results suggest that learners might sometimes separate statistical information received at one time from that at another and thus that hearing two dialects at different times might be helpful for learning dialect-specific rules. However, whereas it is possible that adults in Gebhart et al. (2009) separated the two languages due to strong representation of the first one, Vlach and Johnson's (2013) results suggest that in early infancy at least, learners might separate conflicting language input heard at different times simply because the window over which they integrate statistical information is subject to memory limitations. This possibility is consistent with the general view that infants' memory limitations can in certain cases actually facilitate language learning (Elman, 1993; Newport, 1990). To our knowledge, however, this view has not been considered with respect to the challenge of teasing apart conflicting sources of language input, such as different languages or dialects.

1.4. The present study

The aim of the present study was to investigate the possibility that hearing two dialects at different times can help young infants learn dialect specific rules, independent of a salient perceptual cue. Our experiments extend Gómez and Lakusta's (2004) artificial language paradigm with 12-month-olds. Importantly, this language's rules, such as aX bY, resemble natural language rules in that they permit generalization to novel utterances. Recall that the X- and Y-words differ in syllable length (they are monosyllabic and disyllabic, respectively). This means that the rules are cued by underlying morphophonological relations between the a- and b-words on the one hand and the syllable length of the X- and Y-words on the other, permitting generalization of the rules to sentences with novel X- and Y-words that retain their syllable length distinction. An example of generalizing rules in Standard English based on underlying sound patterns is recognizing subject–verb number agreement in *Some feps are fast* but not in **Some feps is fast*. Even though the plural subject *Some feps* in these sentences is novel, its plurality is cued by the adjective *some* and the suffix –*s* (***Some** feps*). So a learner who has formed associations between plural subject cues and plural rather than singular forms of the main verb can tell that *some feps* goes with the plural *are* in the first sentence but not with the singular *is* in the second.

Across three experiments, we familiarized 12-month-olds with two alternating speech streams representing different dialects of the artificial language. In a "pure stream", all sentences conformed to rules such as aX bY. In a "mixed stream", any a- or b-word could precede any X- or Y-word, and thus sentences supported the pure stream's rules no more often than they did opposite rules (e.g., aY bX). The artificial language was played over loudspeakers while the infant explored a playroom environment. Foreshadowing Experiment 1, results established that at the very least, infants are capable of extracting the pure steam's rules when the two streams are presented not only at different times but also in different voices. The results further suggested no effect of stream order. Following up, two additional experiments examined whether presenting the exact same streams at different times (Experiment 2) rather than in interleaved fashion (Experiment 3) permits infants to extract the pure stream's rules independent of the talker voice cue. In all experiments, we tested whether infants extracted the pure stream's rules such that they could generalize them to conforming sentences with novel X- and Y-words, using a variation of the Head-turn Preference Procedure administered after familiarization (Kemler Nelson et al., 1995). Generalization was operationalized as a significant difference in listening time to sentences composed of the exact same words but conforming to opposite rules.

Our study bears some resemblance to previous language learning studies that alternated two input streams dynamically (Gebhart et al., 2009; Weiss et al., 2009), but owing to our interest in whether hearing two dialects at different times facilitates rule dialect-specific rule learning in infancy, it differs in three notable ways. First, we test infants rather than adults. Second, our streams consist of sentences generated from abstract rules (i.e., aX bY and aY bX), whereas these previous studies' streams consisted of word sequences generated from stimulus-specific co-occurrence statistics (i.e., transitional probabilities between specific syllables). Accordingly, our study uniquely asks whether learners can generalize regularities to test strings with novel elements. Finally, our study includes two experiments that are identical except for whether speech streams are slowly alternated or randomly interleaved. Given these previous studies' different research aims, they did not likewise vary independently whether streams were separated in time. Weiss et al. alternated their streams only in 2-min intervals. Gebhart et al. varied whether their second language was 5 min like the first or was longer, but this manipulation simultaneously changed the relative amount of exposure to each language. Further, although one can infer as we have above that the streams' slow alternation (irrespective of the second stream's duration) was what enabled adults to represent the languages separately, one cannot be certain that interleaving the two languages would not have yielded a statistically equivalent learning outcome.[1]

[1] This possibility might seem unlikely given Gebhart et al.'s order effect, but this effect might simply reflect a bias to learn the particular language presented first, regardless of its temporal order relative to the language presented second. The authors showed that both languages were equally learnable in isolation, but this does not mean that they were equally learnable when alternated, as adults in Weiss et al. (2009) showed a bias to learn one of two alternated languages that were equally learnable in isolation.

2. Experiment 1

Experiment 1 tested whether 12-month-olds can extract the pure stream's rules when this stream is presented not only at different times but also in different voices. In principle, infants might be better at extracting the pure stream's rules from this particular stream when it is differentiated from the mixed stream along the voice dimension, the temporal dimension, or both. Differentiating the two speech streams along both dimensions in this initial experiment was thus a conservative attempt to establish that infants can extract the pure stream's rules from our bi-dialectal input under at least some circumstances. To present the two streams at different times, we alternated them every 4 min. Each 4 min period of the pure stream provided an opportunity to extract this stream's rules from the consistent support therein exclusively, as 12-month-olds in Gómez and Lakusta (2004) extracted rules consistently supported throughout a familiarization phase under 4 min. We counterbalanced stream order to examine whether, compatible with Gebhart et al.'s (2009) order effect, infants represent the first presented stream more strongly than the second.

2.1. Method

2.1.1. Participants

Participants were 16 infants (5 females) with a mean age of 12.23 months (range: 11.6–12.8 months) from the local Tucson area, recruited through birth announcements published in the local newspaper and through various informational fairs such as the Tucson Baby Fair. We tested the same age group as Gómez and Lakusta (2004) for greater confidence in our predictions derived from this previous study's results. To be included in this and all subsequent experiments, infants needed to complete at least 4 test trials, to have no family history of speech or language therapy, to have a birth weight of at least 5.8 oz, to have a birth term of at least 37 weeks, and to be from a predominantly English-speaking household, with no more than 10 h of exposure to a non-English language each week. Twelve additional infants participated but were excluded from data analyses due to technical difficulties (5), a history of language therapy in the immediate family (3), an ear infection (3), or restlessness during testing (1).

2.1.2. Stimuli

Our artificial language stimuli retained key features of Gómez and Lakusta's (2004) stimuli. First, our pure and mixed speech streams representing different dialects of this language contained a- and b-categories variably instantiated by 2 words each and X- and Y-categories by 6 words each, as shown in Fig. 1 (top panel). Second, we counterbalanced across infants whether the target rules, in this study the pure stream's rules, were aX bY or aY bX. Third, the X- and Y-words were distinguished by syllable length (monosyllabic versus disyllabic) so that the target rules contained underlying relations between the a- and b-words on the one hand and the syllable length of the X- and Y-words on the other. These rules resemble dependencies in natural language between high and low frequency grammatical units. In Standard English, for example, high frequency singular and plural forms of the

familiarization				test	
a	**b**	**X**	**Y**	**X**	**Y**
alt	erd	frin	freevit	biff	binnow
ush	ong	lorse	gelshry	fodge	gackle
		narf	masgrin	rud	meeper
		throove	nemtry	tood	nawlup
		brez	shrappo	vot	roosa
		drak	tressid	pel	bevit

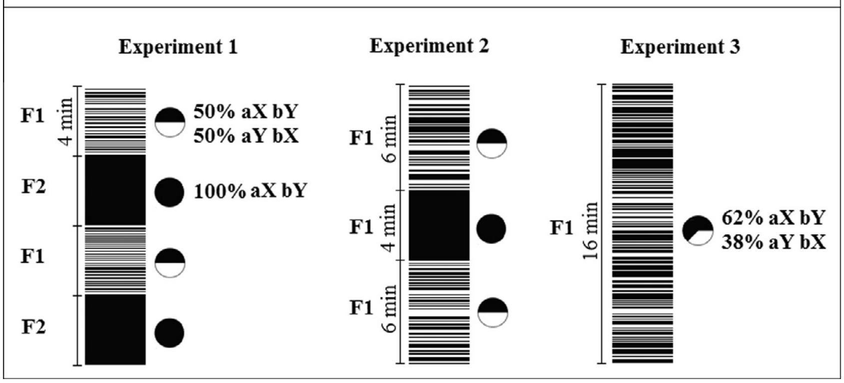

Fig. 1. Language familiarization and test materials. Top: Vocabulary for familiarization and test. a- and b-words are paired with different sets of X- and Y-words during familiarization and test. X- and Y-words are cued by syllable number (monosyllabic and disyllabic, respectively). Bottom: Examples of familiarization lists for Experiments 1, 2, and 3. For half the infants in Experiment 1, familiarization began with the mixed stream (50% aX bY, 50% aY bX) and ended with the pure stream (100% aX bY or aY bX), as illustrated; for the other half of infants, vice versa. Streams were spoken by different female talkers (F1 and F2). In Experiments 2 and 3, both streams were spoken by the same female talker (F1 or F2). These experiments were identical except that streams were blocked in the former and interleaved in the latter.

verb to be (e.g., *is/was* versus *are/were*) are paired respectively with lower frequency singular and plural sentence subjects (e.g., *A dog* versus *Some dogs*), and these agreement rules are cued by underlying sound patterns (***A*** *dog**Ø*** ***is/was*** versus ***Some*** *dog**s*** ***are/were***).

2.1.2.1. Familiarization stimuli. To counterbalance across infants whether the pure stream's rules were aX bY or the opposite (aY bX), we created two familiarization lists—a Language 1 (L1) list and a Language 2 (L2) list. An example of the L1 list is illustrated in Fig. 1 (bottom panel). For this list, we paired each of the 12 possible aX phrases ($a_{1-2}X_{1-6}$) with each of the 12 possible bY phrases ($b_{1-2}Y_{1-6}$), first to create 144 aXbY sentences and then 144 bYaX sentences. A random order of the resulting 288 sentences was then divided equally into two sublists for the pure and mixed speech streams. Each sublist was in turn divided into fourths for 4 unique periods of each speech stream. Finally, half the sentences in each fourth of the mixed stream sublist were modified so that unlike the pure stream that adhered to the aX bY rules, the mixed stream supported opposite rules with equal likelihood. Each sentence was modified by swapping the X- and Y-words within that sentence, resulting in an aYbX or bXaY sentence (e.g., $a_1\mathbf{X_4}b_2\mathbf{Y_3}$ became $a_1\mathbf{Y_3}b_2\mathbf{X_4}$). Within each resulting period of the mixed stream, no more than 3 consecutive sentences supported the pure stream's rules and no more than 4 the opposite rules. The L2 familiarization list was a mirror image of the L1 list, created by swapping X- and Y-words within every sentence of both sublists. In this list, then, the pure stream still supported a single set of rules and the mixed stream these rules and opposite rules with equal likelihood, but the rules in the pure stream were now aY bX rather than aX bY.

Each speech stream was presented by playing individual sound files of recorded words in rapid succession, using the program DMDX (Forster & Forster, 2003). The periods of the two speech streams were alternated in one of two orders to counterbalance across infants which stream came first versus last. The program played adjacent words within and across phrases in immediate succession, and separated adjacent words across sentences (both within the same stream and across streams) by approximately 425 ms. The two speech streams were formed from recordings by different female speakers (F1 and F2). Whether the pure stream was formed from F1's recordings and the mixed stream from those of F2 or vice versa was randomly determined for each infant.

The change in talker voice across speech streams was the only salient perceptual cue to the transition from one stream to the other. The fact that sentences in the pure stream invariably supported either aX bY rules or aY bX rules but sentences in the mixed stream variably supported both did not by itself differentiate these streams in any perceptually salient way. First, both sentence types were instantiated by the same vocabulary. Second, both had the same prosody. Specifically, aXbY and bYaX sentences supporting the aX bY rules are matched respectively with bXaY and aYbX sentences supporting the opposite rules in terms of the syllable length of the word in each sentence position, and all sentences have the same rise-fall-rise-fall intonation pattern spanning words.

2.1.2.2. Test stimuli. To compare listening times to the pure stream's rules versus the opposite rules in this and all subsequent experiments, we created two test lists adhering to the aX bY rules and two adhering to the aY bX rules, using a- and b-words from familiarization but all novel X- and Y-words (see Fig. 1). The aX bY test lists conformed to the pure stream's rules and the aY bX test lists to the opposite rules for infants who had received the L1 familiarization list, and vice versa for infants who had received the L2 familiarization list. To create the aX bY test lists, each one of the 12 possible aX phrases was randomly paired with a different one of the 12 possible bY phrases, first to create 12 aXbY sentences and then 12 bYaX sentences. A unique random order of the resulting 24 sentences was used for each of the two aX bY test lists. These lists were then used to create mirror image aY bX lists by swapping the X- and Y-words within each sentence (yielding aYbX and bXaY sentences). Test phrases were recorded by a novel female speaker (F3) who used the same intonation patterns as the familiarization speakers.

2.1.3. Procedure

2.1.3.1. Familiarization. Familiarization was conducted in a playroom with the infant seated on a rug surrounded by toys. The familiarization list, which lasted approximately 16 min, was played over two loudspeakers resting on a small table. The experimenter played silently with the infant and the caregiver was invited to do likewise. Similar incidental familiarization procedures have been employed previously by Saffran et al. (2008) and Lany and Gómez (2008).

2.1.3.2. Test. Immediately following familiarization, the infant was seated on its caregiver's lap in a sound-attenuated booth. Testing was conducted using the Head-turn Preference Procedure (Kemler Nelson et al., 1995). The experimenter initiated trials and recorded looking behavior from outside the booth using an experimental control program running on an Apple Powermac. Each test trial began with the onset of a blinking light directly in front of the infant. When the infant fixated this light, the experimenter pressed a button to extinguish it. One of two identical lights, each directly below a loudspeaker on either side of the infant, then began blinking. Which light began blinking was randomly determined. When the infant made a head-turn of approximately 30° toward the blinking light, the experimenter began playing one of the test lists from the loudspeaker above. The test list played until the infant looked away from the blinking light for 2 s or until the list terminated (approximately 79 s). The 4 test lists were presented in 3 random orders for a total of 12 test trials.

2.2. Results

Infants should generalize the pure stream's rules to test if they can do so when this stream is presented in a different voice than the mixed stream and in different 4 min periods. If, however, infants form a stronger representation of the first presented language than of the second, infants who first hear the pure stream should show stronger discrimination of the pure stream's rules from the opposite rules compared to infants who first hear the mixed stream, which equally supports both sets of rules. In Gómez and Lakusta (2004), infants' test discrimination manifested as a familiarity effect in that infants listened significantly longer to the rules that predominated during familiarization compared to the opposite rules. Since the present study used the same artificial language and the pure stream's rules predominated (75% of sentences conformed to them across streams), one might expect generalization of these rules to likewise manifest as a familiarity effect, though a significant preference for the opposite rules would be no less indicative of generalization.

A difference score reflecting test discrimination was computed for each infant by subtracting from the infant's average looking time to the pure stream's rules its average looking time to the opposite rules. A planned one-way ANOVA testing for an effect of stream order did not approach significance, $F(1,14)$ 2.11, $p = .17$, $\omega = .24$, indicating that discrimination did not depend on whether familiarization began or ended with the pure stream. We thus collapsed the data across this factor. As the left side of Fig. 2 shows, a one-sample t-test revealed that infants' mean difference score ($M = 2.74$, $SD = 3.82$, $Range = -1.64$ to 11.08) was significantly higher than chance (i.e., zero), $t(15) = 2.81$, $p = .01$, $r = .59$, indicating a reliable preference for the pure stream's rules, consistent with generalizing these rules to test.

2.3. Discussion

Experiment 1 indicates that when the two speech streams are presented in different voices and in alternating 4 min periods, 12-month-olds are able to extract the rules unique to the pure stream. Notably, learning did not depend on whether familiarization began with the pure stream and ended with the mixed stream or vice versa. However, this null order effect should be interpreted with caution. Possibly, Experiment 1 lacked sufficient power for detecting this effect given that only 8 infants received each presentation order.

3. Experiment 2

Experiment 2 tests the possibility that hearing the two speech streams at different times helps infants extract the pure stream's rules, independent of a perceptually salient cue like talker voice. There are at least two reasons why, however much the two streams are separated in time, infants might rely on a salient perceptual cue for extracting the pure stream's rules. One reason follows from Weiss et al.'s (2009) results suggesting that unless their artificial languages were presented in different voices, adults combined statistics across them. These results raise the possibility that infants compute statistical averages over the two speech streams when lacking a perceptually salient cue for distinguishing them, precluding extraction of the pure stream's rules when these noisy combined statistics exceed infants' tolerance threshold. Regarding Experiment 1, infants might have learned the pure stream's rules only because the talker voice cue allowed them to compute, within each speech stream separately, phrase-level co-occurrences between each a- and b-word on the one hand and the syllable length of the X- and Y-words on the other, with co-occurrences computed within the pure stream itself then serving as the basis for generalizing this stream's rules to test. Had infants instead aggregated these co-occurrence statistics across speech streams, they might have failed to generalize the pure stream's rules to test. This possibility cannot be ruled out because the pure stream's rules were violated by 25% of all familiarization sentences across both streams (72/288 sentences), a potentially intolerable percentage falling between what Gómez and Lakusta (2004) found to be tolerable (17%) and intolerable (33%) when violations were randomly distributed across a single speaker's input. It is also possible that due to decay, more recent sentences received greater weight, though this predicts stronger

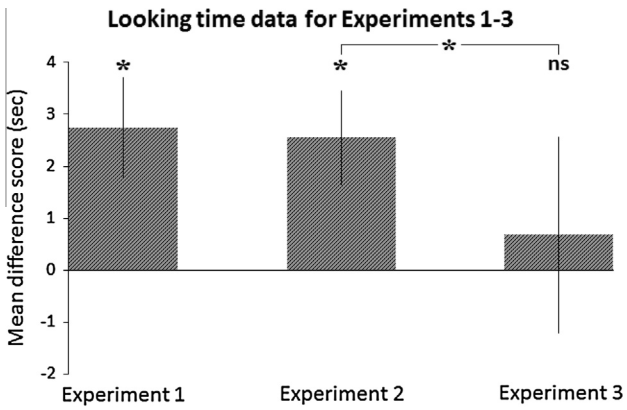

Fig. 2. Looking time data for Experiments 1–3. Infants' mean difference score in each experiment is their average listening time to the pure stream's rules minus that to the opposite rules. (Error bars denote SEM; *$p < .05$; ns = nonsignificant.)

discrimination in Experiment 1 when familiarization ends with the pure stream, contrary to the null order effect.

Second, when familiarization begins with the mixed stream, infants might fail to learn the pure stream's rules unless the transition to the pure stream is signaled by a perceptually salient change. This would be consistent with Gebhart et al.'s (2009) assertion that their adult participants had difficulty learning the second of two successively presented artificial languages due to a failure to detect the language transition. In support, adults learned the second language when its duration was doubled, allowing evidence for the language change to further accumulate. Though the authors reported that a change in talker voice did not likewise cue adults to the transition to the second language, it might nevertheless cue infants to the transition to the pure stream, since language learners sometimes show stronger sensitivity to voice changes at earlier ages (see Newman, 2008).

In Experiment 2, the familiarization input was composed such that infants were expected to extract the pure stream's rules if temporally separating the two streams enables infants to extract those rules, but not if infants rely on a salient perceptual cue for either reason discussed above. For all 12-month-olds, familiarization began and ended with separate 6 min periods of the mixed stream, interposed by a single 4 min period of the pure stream, presented in the same female voice. As in Experiment 1, then, the speech streams had a slow alternation rate that might help infants tease them apart, such as by extracting the pure stream's rules from within this stream itself. However, infants should fail to extract the pure stream's rules if they simply aggregate phrase-level statistics across the two speech streams due to the lack of a salient perceptual cue (e.g., talker voice) for differentiating them. This is because in this latter case, learning should be the same as when sentences violating the pure stream's rules are randomly distributed across a single input stream as in Gómez and Lakusta (2004), and the percentage of violations (108/288 sentences = 38%) exceeds what Gómez and Lakusta (2004) found to be intolerable (33%). Further, the chances of extracting the pure stream's rules should only decrease if more recent sentences receive greater weight, since familiarization ends with 6 min of the mixed stream. Infants should likewise fail to learn the pure stream's rules if they rely on a salient perceptual cue to detect the transition to the pure stream when familiarization begins with the mixed steam, since familiarization not only ends with an extended period of the mixed stream but also begins with an extended period of this stream, with no salient perceptual cue to the onset of the pure stream.

3.1. Method

3.1.1. Participants

Participants were 16 infants (8 females) with a mean age of 12.29 months (Range: 11.66–12.95 months) from the local Tucson area. Exclusion criteria were identical to those in Experiment 1. Six additional infants participated but were excluded from data analyses due to restlessness during testing (1), crying during familiarization (1), a history of language therapy in the immediate family (1), or over 10 h of foreign language exposure per week (3).

3.1.2. Stimuli

As in Experiment 1, we created two familiarization lists to counterbalance the two versions of the artificial language across infants. Also as in Experiment 1, we first created a L1 familiarization list consisting of 288 randomly ordered sentences supporting the aX bY rules, including 144 aXbY sentences and 144 bYaX sentences. However, rather than dividing this list in half for the two different speech streams, we designated the first fourth for the single period of the pure stream and then divided the remainder in half for the two flanking periods of the mixed stream. So that sentences in each of these latter periods supported the pure stream's rules and opposite rules with equal likelihood, half were randomly selected to be modified using the same X–Y swapping method as in Experiment 1. This resulted in no more than 6 successive sentences supporting the pure stream's rules and no more than 7 successive sentences the opposite rules. Also as in Experiment 1, the L2 familiarization list was a mirror image of the L1 list. Each infant was randomly assigned to one of the two familiarization voices in Experiment 1 (F1 or F2).

3.1.3. Procedure

The procedure was the same as in Experiment 1.

3.2. Results

To reiterate, we hypothesized that infants should extract the pure stream's rules only if hearing the two streams at different times enables infants to extract these rules. Consistent with infants extracting the pure stream's rules, a one-sample *t*-test revealed that infants' mean difference score ($M = 2.55$, $SD = 3.62$, $Range = -3.17$ to 7.52), as shown in Fig. 2 (center), was significantly greater than chance, $t(15) = 2.81$, $p = .01$, $r = .58$, reflecting reliably longer listening times to the pure stream's rules than to the opposite rules.

3.3. Discussion

Infants' ability to extract the pure stream's rules in Experiment 2 supports two claims. One is that infants were not disproportionately influenced by the mixed stream at the beginning of familiarization, despite having received no perceptually salient cue to the transition to the pure stream. If infants had been disproportionately influenced by this initial input, they should have failed to discriminate the pure stream's rules from the opposite rules, since this initial input provided equal support for both. The other claim is that hearing the two speech streams at different times helped infants extract the pure stream's rules. As explained earlier, Gómez and Lakusta's (2004) results suggest that if hearing the speech streams at different times in Experiment 2 does not help infants extract the pure stream's rules, because infants simply aggregate phrase-level statistics across streams, then infants should not be able to extract these rules from the familiarization input.

4. Experiment 3

Experiment 3 sought to confirm that infants in Experiment 2 were able to extract the pure stream's rules due to the two streams being separated in time. An alternative possibility is that infants did not rely on these streams being separated in time because they extracted the pure stream's rules from phrase-level statistics aggregated across streams. This latter possibility cannot be entirely ruled out by Gómez and Lakusta's (2004) study suggesting that the overall percentage of violations to the pure stream's rules across streams would exceed infants' tolerance threshold, regardless of whether all sentences are given equal weight or more recent sentences greater weight. The reason is that familiarization in that study consisted of 72 sentences total, whereas familiarization in the current study consisted of 288 sentences total. One might argue that the infants tested in Experiment 2 tolerated a higher overall percentage of inconsistency due to receiving 4 times as much exposure to the artificial language.

Infants in Experiment 3 heard the exact same sentences as infants in Experiment 2, but the pure stream's sentences were randomly interleaved with those from the mixed stream rather than being presented successively to form a distinct period of this stream. If infants in Experiment 2 were able to extract the pure stream's rules due to hearing the two speech streams at different times, infants in the present experiment should show weaker learning, because these streams no longer occur at different times. However, if infants in Experiment 2 learned the pure stream's rules by aggregating phrase-level statistics across the entire familiarization corpus, they should not show weaker learning, since the overall number of familiarization sentences is preserved in addition to the percentage of sentences violating the pure stream's rules. In fact, if infants in Experiment 2 gave more recent sentences greater weight due to decay, then in the present experiment they might even show stronger, rather than weaker, learning. As a result of interleaving the two streams, the distance from the beginning of familiarization averaged by sentences supporting the pure stream's rules rises (to ordinal position 139.8 from 89.73) more so than the distance averaged by sentences supporting the opposite rules (to 153.5 from 147.1). Consequently, the sum of the ordinal positions of sentences supporting the pure stream's rules differs more from that of sentences supporting the opposite rules in the present experiment (25,022 vs. 16,151) than in Experiment 2 (16,573 vs. 15,745). Thus, if summing the ordinal positions of sentences supporting each set of rules is indicative of the relative amount of support each set receives in a memory store subject to decay, the redistribution of sentences in the present experiment would appear to promote unbalanced support for them, and thus a stronger basis for discriminating them.

4.1. Method

4.1.1. Participants

Participants were 16 infants (9 females) with a mean age of 12.18 months (Range: 11.6–12.6 months). Recruiting methods and exclusion criteria were identical to those used in Experiments 1 and 2. Three additional infants participated but were not included in data analyses due to having over 10 h of foreign language exposure per week (2) or due to technical difficulties (1).

4.1.2. Stimuli

The L1 familiarization list was identical to that in Experiment 2 except that sentences from the two speech streams were randomly interleaved. This interleaving resulted in no more that 11 successive sentences supporting the pure stream's rules and no more than 5 successive sentences supporting the opposite rules. The L2 familiarization list mirrored the L1 familiarization list as in previous experiments.

4.1.3. Procedure

The procedure was the same as in Experiments 1 and 2.

4.2. Results

Test discrimination should be weaker than in Experiment 2 if infants in this previous experiment relied on the speech streams' temporal separation to extract the pure stream's rules. The mean difference scores of Experiments 2 and 3 are shown in the center and right hand side of Fig. 2, respectively. A Kolmogorov–Smirnov test indicated that the distribution of infants' difference scores in Experiment 3 deviated from a normal distribution, $D(16) = 0.343$, $p < .001$ (those in Experiments 1 and 2 did not; $p > .05$), so non-parametric tests were used for subsequent comparisons. Consistent with infants' extraction of the pure stream's rules in Experiment 2 being due to the two streams' temporal separation, a Mann–Whitney[2] test revealed a reliable tendency for the difference scores in Experiment 3 ($M = 2.55$, $SD = 3.62$, $Range = -3.17$ to 7.52) to be lower than those in Experiment 2 ($M = .68$, $SD = 7.56$, $Range = -7.06$ to 26.18), $U = 73.5$, $p = .04$, $r = -.36$. A follow up one-sample Wilcoxon Signed Rank test showed that the Experiment 3 scores were not significantly different from chance, $Ws = 64$, $p = .84$, $r = -.05$, suggesting failure to learn the pure stream's rules in this experiment.

4.3. Discussion

These results reveal that interleaving the two speech streams rather than presenting them at different times eliminated infants' ability to extract the pure stream's rules. This finding supports the notion that infants in Experiment 2 were able to extract the pure stream's rules due to these streams being presented at different times. It is not the case that separating the two streams in time was unhelpful because infants simply aggregated across these streams phrase-level statistics between the a- and b-words and the syllable length of the following X- and Y-words.

[2] A Levene's test revealed that the variances between the two experiments were not significantly different ($p > .05$). Therefore, the Mann–Whitney test was appropriate for testing for a shift in central tendency between these two experiments.

5. General discussion

The present study suggests that when infants hear two different dialects, their ability to learn rules present in only one dialect does not necessarily require detecting those rules from statistical averages computed across dialects. Rather, infants may be able to compute statistics more supportive of dialect-specific rules. Moreover, this computational approach may not depend on a salient perceptual distinction between the two dialects (such as talker voice), so long as these conflicting sources of input are sufficiently separated in time. These insights derive from Experiments 2 and 3, in which our two artificial language streams representing different dialects were not differentiated in any perceptually salient way. The experiments together suggest that alternating the two speech streams in minutes-long intervals (Experiment 2), rather than presenting these streams' sentences in random order (Experiment 3), enabled infants to learn the pure stream's aX bY (or aY bX) rules without inordinate interference from the mixed stream in which any a- or b-word could precede any X- or Y-word. Separating the streams in time should not have helped infants learn the pure stream's rules if infants invariably average statistics across streams. Importantly, the rules that infants learned were abstract, as indicated by infants' ability to generalize them to novel sentences at test based on underlying morphophonological relations between the a- and b-words and the syllable length of the following (monosyllabic) X- and (disyllabic) Y-words.

5.1. How did infants avoid aggregating across the familiarization corpus?

Our finding that infants in Experiment 2 did not simply compute statistical averages across the two speech streams even though these streams were presented in the same voice does not necessarily contradict Weiss et al.'s (2009) claim that unless their languages were presented in different voices, adults combined statistics across them. Consistent with the view that memory limitations can facilitate language learning (Newport, 1990), one plausible explanation for these contrasting findings is that 12-month-olds in Experiment 2 lacked the memory capacity to combine across speech streams, whereas adults in Weiss et al. (2009) were evidently capable of combining across languages. This potential difference between 12-month-olds and adults could have been developmental, given Vlach and Johnson's (2013) finding that 16-month-olds were unable to aggregate word-object pairings across a roughly 3-min period like 20-month-olds. Or it could have arisen from methodological differences, given that our speech streams alternated in longer intervals than Weiss et al.'s (2009) languages (4–6 min vs. 2) and thus might have required greater memory capacity to combine. Further, our streams consisted of sentences conforming to productive rules, whereas Weiss et al.'s (2009) languages consisted of word sequences conforming to transitional probabilities between sublexical segments. Research suggests that processing either type of structure is subject to memory constraints (Frank & Gibson, 2011; Frank, Goldwater, Griffiths, & Tenenbaum, 2010) but does not exclude the possibility that processing our streams was subject to greater constraints.

5.2. Why did the initial period of the mixed stream not block subsequent learning?

Our results suggest that infants' exposure to two dialects at different times might help them extract dialect-specific rules even when the dialect containing those rules is both preceded and followed by a longer period of the dialect lacking them. This is because infants in Experiment 2 extracted the pure stream's rules when a single 4 min period of the pure stream was both preceded and followed by 6 min of the mixed stream. Of particular note is infants' ability to learn the pure stream's rules from this stream notwithstanding the preceding 6-min period of the mixed stream. Since sentences in the mixed stream randomly conformed to the pure stream's rules and opposite rules (aX bY and aY bX), one might have expected infants to conclude precipitously from this preceding input that any a- or b-word can precede both monosyllabic and disyllabic words, precluding learning of the more systematic morphophonological relations underlying the pure stream's rules. Such a learning outcome would have been compatible with studies in which adults had difficulty learning other than from their initial input.

That infants were able to learn the pure stream's rules despite familiarization beginning with 6 min of the mixed stream makes more sense when considering the possibility that during this initial period, infants' sensitivity to the X- and Y-words' syllable length distinction was not sufficient for strongly representing these words' random morphophonological relations with the preceding a- and b-words. Infants might have lacked sensitivity to the X- and Y-words' syllable length distinction precisely because of these words' random relations with the preceding a- and b-words, which resulted in the X- and Y-words sharing the same speech context, inasmuch as they were preceded by the same words. Increasing evidence suggests that gaining sensitivity to contrasting speech categories may depend critically upon the categories having distinct speech contexts (see Feldman, Griffiths, Goldwater, & Morgan, 2013). For example, Thiessen (2007) found that 15-month-olds noticed when a novel object previously labeled "daw" was subsequently labeled "taw" only if they had previously heard "daw" and "taw" in distinct lexical contexts ("dawbow" and "tawgoo"), not in the same context ("dawgoo" and "tawgoo"), or in isolation. Importantly, Gómez (2006) found that 12-month-olds are not readily sensitive to the X- and Y-words' syllable length distinction. Specifically, infants were unable to generalize rules such as aX bY to test if they heard only half the X- and Y-words (3 each) rather than the full set (6 each), presumably because higher variation in X- and Y-words highlights each category's invariant syllable length (Gómez, 2006). We are suggesting here that another requirement for gaining sensitivity to the X- and Y-words' syllable length distinction may be that these words differentially covary with the a- and b-words. None of this is to argue that strong representation of the mixed stream's random relations would have necessarily prevented infants from later learning the more systematic relations underlying the pure stream's rules, only that strong representation of these initial random relations may have been one unmet requirement for such an order effect.

The above provides a theoretically motivated explanation for contrasting results from our study and that of Gebhart et al. (2009). Recall that these authors' results suggested that adults' exposure to the first of two languages presented in immediate succession blocked learning of the second unless the duration of the second was longer (15 vs. 5 min). This contrasts with our finding that infants' exposure to the initial period of the mixed stream did not block subsequent learning from the pure stream that followed, despite this latter period's shorter duration (4 vs. 6 min). In Gebhart et al. (2009), however, adults' test performance indicated strong sensitivity to first language regularities that conflicted with second language regularities. We have suggested that because the initial period of the mixed stream randomly supported the pure stream's rules and opposite rules, infants did not strongly represent the morphophonological relations in this initial input, ensuring their ability to subsequently learn the more systematic relations underlying the pure stream's rules. We would perhaps replicate Gebhart et al. (2009) order effect if the initial period of the mixed stream were modified such that rather than sentences supporting the pure stream's rules just as often as opposite rules (aX bY and aY bX), sentences consistently supported the opposite rules (e.g., aY bX), leading infants to form a strong representation of the relations underlying these opposite rules. This might then block subsequent learning of the relations underlying the pure stream's rules. Indeed, infants might be particularly susceptible to such an order effect given their relatively weak inhibitory control (e.g., Davidson, Amso, Anderson, & Diamond, 2006; Gerardi-Caulton, 2000), as Bartolotti, Marian, Schroeder, and Shook (2011) found that adults' ability to learn the second of two artificial languages with conflicting regularities correlated positively with inhibitory control.

5.3. A preliminary account of Experiment 2

The experiments reported here do not indicate precisely how 12-month-olds in Experiment 2 extracted the pure stream's rules, but based on the points discussed above, we will suggest a preliminary account comprising four key assumptions. First, infants were unable to combine statistics broadly across speech streams. This follows from Vlach and Johnson's (2013) finding that 16-month-olds were unable to aggregate word-object pairing over a roughly 3 min period, together with the fact that our speech streams alternated in relatively long intervals of 4–6 min. Our second assumption, just discussed, is that during the initial period of the mixed stream infants were unable to strongly represent the random relations between the a- and b-words and the syllable length of the following X- and Y-words. Our third assumption, also just discussed, is that weak representation of the random relations in this initial input ensured that exposure to this input did not block subsequent learning during the pure stream of the more systematic morphophonological relations underlying the pure stream's rules. To clarify, this assumption is not that the initial period of the mixed stream had no influence whatsoever on subsequent learning, only that whatever influence it might have had did not prevent subsequent extraction of the morphophonological relations underlying the pure stream's rules. Our final assumption is that infants retained this rule learning to test without inordinate interference from the final period of the mixed stream, consistent with adults in Gebhart et al. (2009) retaining the first of two languages through the second. To clarify this assumption as well, it is not that the final period of the mixed stream in no way modified infants' representation of the pure stream's rules, only that any possible modification did not prevent generalization of the pure stream's rules to test.

An alternative account is that infants in Experiment 2 formed a representation of the initial period of the mixed stream strong enough to resist combining with the period of the pure stream that followed, and a representation of this latter period strong enough to resist combining with the final period of the mixed stream. Separating the two speech streams thus, infants had little difficulty extracting the pure stream's rules. Insofar as this account explains how infants separated speech streams, it meshes well with our suggestion in the Introduction that adults in Gebhart et al. (2009) might not have combined languages like adults in Weiss et al. (2009) because the first language's initial period was longer than in Weiss et al. (2009), permitting representation of it strong enough to resist combining with the second language. As already noted, however, we would predict on the basis of Gebhart et al.'s (2009) results that strong representation of the initial period of the mixed stream would block subsequent extraction of the pure stream's rules. Modifying this alternative account slightly, one might suggest that infants combined the initial period of the mixed stream with the period of the pure stream that followed, and from this combined input formed a representation of the pure stream's rules resistant to combining with the final period of the mixed stream. Putting aside our suspicion that infants lacked the memory capacity to combine these first two periods, the pure stream's rules were violated by 30% of sentences across these periods. Our more recent data (Gonzales, Gerken, & Gómez, 2014) indicate that this percentage exceeds infants' tolerance threshold, even when the familiarization phase is longer than the combined duration of these two periods.

5.4. Broader considerations

The research discussed throughout this paper paints a complex picture of the circumstances under which language learners might succeed at learning structure unique to one of two input sources, such as two different languages or dialects. We have discussed previous work suggesting that learning may depend on whether the input source in question is differentiated from the other input source by a salient perceptual cue (Weiss et al., 2009), whether it is heard before or after the other input source (Gebhart et al., 2009), and whether its duration is longer when heard after the other one (Gebhart et al., 2009). Adding to this complexity, we have argued on the basis of our results that learning may depend on whether the two input sources are heard in distinct time periods, and perhaps also on whether the other input source's competing structure is strongly represented. However, all of these dimensions could be thought of as variations along a more general continuum of support for competing structures.[3] Such a perspective explains why at least some of these dimensions appear to interact, like the order of the two sources of input and the duration of the second source (Gebhart et al., 2009). Further, it provides a useful framework for predicting additional interactions. For example, one could predict that although infants were unable to extract the pure stream's rules when the two streams' sentences were presented in the same voice and randomly interleaved (Experiment 3), infants might extract the rules under the exact same circumstances except with each stream's sentences presented in a different voice. We are currently investigating such questions to better understand how infants negotiate linguistic diversity in their rich multidimensional language environments (Gonzales et al., 2014).

5.5. Conclusion

To conclude, the present study suggests that infants might come better prepared to negotiate bi-dialectal input than previously thought. While it is widely suspected that infants tease apart linguistically diverse input using salient perceptual cues such as rhythm or talker voice, our artificial language results suggest that infants are sensitive to the dynamics of their input such that they can tease apart two dialects independent of such cues. Specifically, it appears that for 12-month-olds at least, simply hearing two dialects at different times may be helpful for learning dialect-specific rules.

Acknowledgements

We thank Andrew Lotto, József Fiser, and two anonymous reviewers for thoughtful comments on previous versions of this manuscript. This work was supported by NIH HD42170 to LAG and RLG and by NSF CAREER Award BCS-0238584 to RLG.

References

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*, 324.

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

Byers-Heinlein, K., Burns, T. F., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science, 21*, 343–348.

Curtin, S. A., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics, 39*, 492–504.

Davidson, M., Amso, D., Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037–2078.

Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71–99.

Feldman, N., Griffiths, T., Goldwater, S., & Morgan, J. (2013). A role for the developing lexicon in phonetic category acquisition. *Psychological Review, 120*, 751–778.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.

Frank, M. C., & Gibson, E. (2011). Overcoming memory limitations in rule learning. *Language Learning and Development, 7*, 130–148.

Frank, M. C., Goldwater, S., Griffiths, T. L., & Tenenbaum, J. B. (2010). Modeling human performance in statistical word segmentation. *Cognition, 7*, 130–148.

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

Gerardi-Caulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24–36 months of age. *Developmental Science, 3*, 397–404.

Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7*, 567–580.

Gómez, R. L. (2006). Dynamically guided learning. In Y. Munakata & M. Johns (Eds.), *Attention & performance XXI: Processes of change in brain and cognitive development* (pp. 87–110). Oxford, UK: Oxford University Press.

Gonzales, K., Gerken, L. A., & Gómez, R. L. (2014). From acoustic detail to abstract grammar? How indexical variation constrains infants' grammar generalizations. *Paper presented at the XIX Biennial international conference on infant studies*, Berlin, Germany.

Hazen, K. (2014). A new role for an ancient variable in Appalachia: Paradigm leveling and standardization in West Virginia. *Language Variation and Change, 26*, 77–102.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development, 1*, 151–195.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology, 59*, 30–66.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development, 18*, 111–116.

Lany, J., & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science, 19*, 1247–1252.

Mitchel, A. D., & Weiss, D. J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes, 25*, 456–482.

Molnar, M., Gervain, J., & Carreiras, M. (2013). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy, 19*, 326–337.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 756–766.

Newman, R. S. (2008). The level of detail in infants' word learning. *Current Directions in Psychological Science, 17*, 229–232.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America, 105*, 512–521.

Ross, D. S., & Newport, E. L. (1996). The development of language from non-native linguistic input. In A. Stringfellow, D. Cahana-Amitay, E. Hughs, & A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University conference on language development* (pp. 623–645). Boston: Cascadilla.

Saffran, J. R., Hauser, M., Seibel, R. L., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by infants and cotton-top tamarin monkeys. *Cognition, 107*(2), 479–500.

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science, 12*, 583–601.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science, 12*, 323–328.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from impoverished input. *Cognitive Psychology, 49*, 370–407.

[3] We thank József Fiser and Toben Mintz for this suggestion.

Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics, 39*, 505–513.
Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language, 56*, 16–34.
Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127*, 375–382.
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*, 30–49.