Developmental Psychology

Infants' Selective Use of Reliable Cues in Multidimensional Language Input

Christine E. Potter and Casey Lew-Williams

BRIEF REPORT

Infants' Selective Use of Reliable Cues in Multidimensional Language Input

Christine E. Potter and Casey Lew-Williams
Princeton University

Learning always happens from input that contains multiple structures and multiple sources of variability. Though infants possess learning mechanisms to locate structure in the world, lab-based experiments have rarely probed how infants contend with input that contains many different structures and cues. Two experiments explored infants' use of two naturally occurring sources of variability—different sounds and different people—to detect regularities in language. Monolingual infants (9–10 months) heard a male and female talker produce two different speech streams, one of which followed a deterministic pattern (e.g., AAB, *le-le-di*) and one of which did not. For half of the infants, each speaker produced only one of the streams; for the other half of the infants, each speaker produced 50% of each stream. In Experiment 1, each stream consisted of distinct sounds, and infants successfully demonstrated learning regardless of the correspondence between speaker and stream. In Experiment 2, each stream consisted of the same sounds, and infants failed to show learning, even when speakers provided a perfect cue for separating each stream. Thus, monolingual infants can learn in the presence of multiple speech streams, but these experiments suggest that infants may rely more on sound-based rather than speaker-based distinctions when breaking into the structure of incoming information. This selective use of some cues over others highlights infants' ability to adaptively focus on distinctions that are most likely to be useful as they sort through their inherently multidimensional surroundings.

Keywords: statistical learning, infant language learning, rule learning, cognitive development

To make sense of complex perceptual environments, infants must prioritize the cues that allow them to efficiently extract relevant information. It has been widely shown that infants are endowed with powerful learning abilities and are highly sensitive to structure in their environment. Infants detect regularities across a variety of domains, modalities, and types of statistical relations (e.g., Fiser & Aslin, 2002; Gómez, 2002; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Saffran, Aslin, & Newport, 1996; see Saffran & Kirkham, 2018 for a recent review). However, the vast majority of studies exploring infants' statistical learning have used deterministic input from a single source, which is unlikely to reflect the true challenge of learning important structures such as those found in language, as natural language environments contain myriad structures and cues that vary in their relevance. In two experiments, we tested infants' ability to use different cues to discover regularities in input that contained multiple, distinct patterns.

Researchers have begun to test infants' learning from input that may better reflect infants' real experience. For example, Pelucchi, Hay, and Saffran (2009) found that infants could segment words from a corpus of real Italian speech, showing the ability to find regularities in natural input. Graf Estes and Lew-Williams (2015) introduced a different form of noise, the presence of multiple talkers, and found that infants could demonstrate learning of an artificial language produced by many different speakers. Still, even these studies contained only one set of learnable regularities, and the only cues to structure came from statistical co-occurrences between syllables, which does not capture the multidimensionality of infants' experience. Similarly, studies that have included more complex statistical relations (e.g., Gerken & Knight, 2015; Gómez, 2002; Gómez & Lakusta, 2004; Gómez & Maye, 2005) have not typically included other forms of noise, such as speaker-based indexical differences. In the present research, we focused on three separate sources of variability: different statistical regularities, different speakers, and different sets of speech sounds. We introduced these cues simultaneously to probe infants' ability to contend with some of the challenges found in real language environments, such as the need to track different aspects of linguistic structure, or even different languages. Infants heard two speakers

(one female, one male) produce two different streams of speech: a target stream with recurring repetition-based patterns (e.g., AAB, as in previous research by Marcus et al., 1999) and a nontarget stream with no overt repetition. Hereafter, we use the term "stream" to refer to a sequence of items constructed according to these constraints. These streams consisted of either nonoverlapping (Experiment 1) or fully overlapping sounds (Experiment 2) in order to test infants' use of sounds and speakers to discover reliable structure in variable input.

Statistics available to young learners are rarely deterministic, and inconsistency shapes infants' learning strategies (e.g., Tummeltshammer & Kirkham, 2013). Even adults struggle to demonstrate learning when there are multiple sets of regularities within the same materials (e.g., Bulgarelli & Weiss, 2016; Gebhart, Aslin, & Newport, 2009; Karuza et al., 2016; Pacton & Perruchet, 2008; Weiss, Gerfen, & Mitchel, 2009). The few studies that have tested infants' learning from multiple streams of information have highlighted the challenge presented by less consistent input. Both Antovich and Graf Estes (2018) and Bulgarelli, Benitez, Saffran, Byers-Heinlein, and Weiss (2017) reported that monolingual infants were unable to demonstrate learning of two different artificial languages, presented either sequentially or interleaved. However, monolingual infants learn more effectively when competing speech streams are separated in time (Gonzales, Gerken, & Gómez, 2015) and when regularities are highlighted by additional cues, such as prosody (e.g., Chun & Jiang, 1998; Gervain & Werker, 2013; Richardson & Kirkham, 2004). Therefore, infants may benefit from a range of cues that help them separate concurrently presented streams of information.

In auditory environments, in particular, many cues can facilitate learning, and infants use experience to focus on acoustic cues that are most likely to support learning (e.g., Lew-Williams & Saffran, 2012; May & Werker, 2014; Nazzi, Jusczyk, & Johnson, 2000; Nazzi, Mersad, Sundara, Iakimova, & Polka, 2014; Thiessen & Saffran, 2007; Werker & Tees, 1984). For example, different sequences of sounds and different speakers are ubiquitous in infants' environments, and their use of both phonological and indexical cues is tuned through experience (e.g., Best, Tyler, Gooding, Orlando, & Quann, 2009; Houston & Jusczyk, 2000; Werker & Tees, 1984).

Infants' ability to differentiate different types of speech sounds has been convincingly demonstrated (e.g., Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Kuhl, 1983; Werker & Lalonde, 1988). Young infants perceive contrasts that older listeners do not (e.g., Mazuka, Hasegawa, & Tsuji, 2014; Werker & Tees, 1984), suggesting that early in development, infants may be particularly focused on fine phonological details. Moreover, infants are acutely sensitive to the distribution of sounds in their input (e.g., Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002) and track the frequency of different sound combinations in their language (e.g., Friederici & Wessels, 1993; Graf Estes, Edwards, & Saffran, 2011; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Mattys, Jusczyk, Luce, & Morgan, 1999). They can also discriminate between two languages on the basis of sound properties (e.g., Bosch & Sebastián-Gallés, 2001) and can make category generalizations on the basis of phonetic information (e.g., Gómez & Lakusta, 2004). However, though past studies demonstrate that infants can separate different sets of sounds, it is not yet known if infants use this dimension to facilitate learning through multiple streams.

Similarly, infants are attuned to the presence of different speakers from an early age and readily distinguish male and female speakers (e.g., Floccia, Nazzi, & Bertoncini, 2000; Jusczyk, Pisoni, & Mullennix, 1992). They often attend to indexical properties of speakers' voices even when they are not task relevant (e.g., Graf Estes & Lew-Williams, 2015; Houston & Jusczyk, 2000; Singh, Morgan, & White, 2004). Given that adults better detect multiple speech streams that are separated by speaker (Mitchel & Weiss, 2010; Weiss et al., 2009), we expected that having two streams produced by speakers of different genders might offer a powerful cue for infants.

The current studies were designed to test how infants might exploit the combination of sounds and speakers in order to locate structure in complex input—specifically, input containing multiple streams of information. We defined structure using two patterns that have previously been tested in isolation. Infants heard three-syllable sequences that followed either an AAB pattern (e.g., *le-le-di*) or an ABA pattern (e.g., *le-di-le*) and were subsequently tested on their learning of that pattern. These kinds of regularities have been readily learned by infants across many studies (e.g., Gerken, 2006; Marcus et al., 1999; see Rabagliati, Ferguson, & Lew-Williams, 2018, for meta-analyses). Furthermore, infants are more likely to discover patterns in stimuli that are connected to their real-life experiences (Ferguson & Lew-Williams, 2016; Rabagliati et al., 2018; Saffran, Pollak, Seibel, & Shkolnik, 2007), suggesting that realistic sources of variability may, in fact, help infants identify patterns.

In two experiments, we asked how the co-occurrence of ecologically valid cues would support infants' learning of regularities embedded in input containing two streams of speech, produced by two different speakers. In prior studies, monolingual infants have been unable to contend with multiple streams of information (e.g., Antovich & Graf Estes, 2018; Bulgarelli et al., 2017; Kovács & Mehler, 2009). But given that monolingual infants constantly encounter multiple sources of information, they must be able to overcome this complexity.

We tested the hypothesis that co-occurring cues in the speech signal can support infants' ability to locate regularities in the input. We presented infants with interleaved Target and Nontarget speech streams. Half of the infants heard each stream produced by a separate speaker (Consistent Speaker condition), and half of the infants heard 50% of each stream produced by each speaker (Variable Speaker condition). Because prior studies suggest that infants struggle to learn two structured streams, we chose regularities for the Target stream that have previously been shown to be robustly learnable (i.e., AAB or ABA patterns). The Nontarget stream contained no repetitions (i.e., CDE), so the structure of the two streams did not conflict. In Experiment 1, the Target and Nontarget streams involved unique sets of sounds, meaning that phonological cues, which can be used to signal to the presence of different structures, were available in addition to speaker cues to help infants segregate the two streams. In Experiment 2, we eliminated the phonological distinction and tested whether or not speaker cues alone would support infants' learning.

Experiment 1

In Experiment 1, infants were provided with multiple cues to the presence of different structures. We predicted that the combination of sound- and speaker-based cues would help infants segregate the Target from the Nontarget stream. We also predicted infants would demonstrate better learning when regularities in the Target stream co-occurred with consistent speaker information.

Method

Participants. Experiment 1 included 40 full-term monolingual English-learning infants (17 female), ranging in age from 9.3 to 11.0 months ($M = 10.1$). A power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that a sample size of $n = 20$ per condition will have .83 power to detect a medium effect for the interaction between condition and test item ($f = .25$, based on comparable studies, e.g., Lew-Williams & Saffran, 2012) with two groups of participants. All infants were reported to have normal hearing and were exposed to English at least 85% of the time. Half of the infants were assigned to the Consistent Speaker condition; the other half were assigned to the Variable Speaker condition. Four additional infants were tested but excluded due to fussiness ($n = 3$) or performance that was more than 2.5 standard deviations from the mean ($n = 1$). The parents of all infants provided informed consent, and all participants received a small gift in exchange for their participation. All experimental protocols, including procedures for obtaining informed consent, were approved by the Princeton University Institutional Review Board (Approval no.: 0000007117A021, Language Learning: Sounds, Words, and Grammar).

Stimuli. All stimuli consisted of trisyllabic strings, produced in infant-directed speech. Two native English speakers, one male and one female, recorded the stimuli. Each syllable was recorded in isolation, and all syllables were normed to match in intensity and duration (625 ms). Sequences were created by combining syllables with 250 ms of silence between syllables, and there was 1 s of silence between three-syllable items.

Familiarization. During familiarization, infants listened to two separate streams of speech. The *Target stream* was closely modeled on materials previously demonstrated to be learnable by infants of this age (Marcus et al., 1999) and consisted of a series of strings that followed either an AAB (e.g., *le-le-di*) or ABA pattern (e.g., *le-di-le*). There were 16 unique items in the Target stream (Table 1), and infants heard each item twice. The Target stream was randomized and divided into four 8-item blocks of approximately 30 s, which were intermixed with a second, *Nontarget stream*. Half of the infants were randomly assigned to hear

Figure 1. Blocking design for the familiarization streams. Green indicates the Target stream; gray is the Nontarget stream. The *X*s denote the female speaker, while spaces without *X*s denote the male. In the Consistent Speaker condition, there was perfect correspondence between speaker and language (top row), while in the Variable Speaker condition, there was no reliable correspondence between speaker and language (bottom row). See the online article for the color version of this figure.

the AAB Target materials; the other half heard the ABA Target materials.

The Nontarget stream was also comprised of 16 different trisyllabic sequences (each repeated twice), but unlike the Target stream, each syllable within a string was unique (CDE), so there were no repetitions within strings (e.g., *foi-nae-vuh*; see Table 1). In addition, none of the phonemes used in the Target stream appeared in the Nontarget stream. Like the Target stream, the Nontarget stream was divided into four blocks of eight items.

All infants heard identical linguistic materials, with alternating 30-s blocks (8 strings) of the Target and Nontarget streams. In the *Consistent Speaker* condition, the Target stream was produced entirely by the female speaker, while the male speaker produced the full Nontarget stream. In the *Variable Speaker* condition, each speaker produced half of each stream. The speaker change always occurred midway through an eight-item block in the Variable Speaker condition, such that infants listened to 30-s blocks of a single speaker, and 30-s blocks of a single stream, as in the Consistent Speaker condition (Figure 1).

Test. In the test phase, infants listened to single items that followed either an AAB (*ko-ko-ba, po-po-ga*) or ABA (*ba-ko-ba, ga-po-ga*) pattern. Items that maintained the regularity from the Target stream were considered Familiar (e.g., AAB test items following exposure to an AAB regularity) while items that violated the Target regularity were considered Unfamiliar (e.g., ABA items following exposure to an AAB regularity). None of the phonemes used in the test items appeared in either familiarization stream. All test items were produced by the female speaker.

Procedure. Infants were tested using the head-turn preference procedure. Participants sat on their parents' lap in a darkened room with monitors on three sides. Parents listened to music through headphones to prevent them from interfering with children's behavior. During familiarization, infants listened to the intermixed Target and Nontarget streams for nearly four minutes while visual stimuli appeared on the monitors, contingent with the infants' looking behavior.

Table 1

Familiarization Stimuli for AAB Condition in Experiment 1

Target stream				Nontarget stream			
le-le-we	le-le-je	le-le-li	de-de-li	fɔɪ-nae-vʌ	ru-vae-tʌ	tae-fu-sʌ	fae-su-tɔɪ
wi-wi-je	wi-wi-di	de-de-je	le-le-di	nu-tae-rɔɪ	tu-sɔɪ-rae	vɔ-nu-fʌ	tʌ-vu-nɔɪ
ji-ji-we	ji-ji-li	ji-ji-di	wi-wi-li	sʌ-rae-vɔɪ	fʌ-tɔɪ-nae	sae-nʌ-ru	nɔɪ-sʌ-vu
de-de-di	wi-wi-we	ji-ji-je	de-de-we	vae-rʌ-tu	vʌ-fae-su	rʌ-fɔɪ-sae	nʌ-rɔɪ-fu

Note. Participants in the ABA condition heard the same syllables, but arranged to form ABA patterns instead. Infants heard all 16 items in each stream presented twice, in two different randomized orders and presented in blocks of eight.

Infants were then tested for learning of the Target pattern. On each test trial, the infant's attention was drawn to the center monitor with an interesting visual stimulus (i.e., a pinwheel). The same stimulus then appeared on one of the two side monitors, and when the infant looked to that side, a test item (either Familiar or Unfamiliar) played from a loudspeaker until the infant looked away for 1 s or a maximum of 20 s had elapsed. A trained experimenter, who wore noise-canceling headphones in order to be unaware of what the infant was hearing, controlled the stimuli using custom software. Each test item was repeated three times, for a total of 12 test trials.

Results and Discussion

Our main analysis in Experiment 1 tested infants' ability to discriminate Familiar versus Unfamiliar test items after hearing the patterns produced by a single speaker versus multiple speakers. For each participant, we calculated mean looking times for Familiar and Unfamiliar trials. Looking times were analyzed using a 2×2 mixed analysis of variance (ANOVA), with test type (Familiar vs. Unfamiliar) as a within-subjects factor and condition (Consistent vs. Variable Speaker) as a between-subjects factor. The ANOVA revealed a main effect of test type ($F(1, 38) = 13.05$, $p = .0009$, $\eta_p^2 = .008$), demonstrating that infants differentiated between Familiar and Unfamiliar items. However, there was no interaction ($F(1, 38) = .02$, $p = .88$), suggesting that learning was similar under both Consistent and Variable Speaker conditions. Planned comparisons using two-tailed paired samples t tests confirmed that infants displayed similar performance across conditions. In both conditions, infants listened significantly longer to Familiar items (Consistent Speaker: 5.2 s vs. 4.3 s, $t(19) = 2.19$, $p = .04$, Cohen's $d = .49$; Variable Speaker: 5.6 s vs. 4.6 s, $t(19) = 3.08$, $p = .006$, Cohen's $d = .69$; see Figure 2). Fourteen of 20 infants in the Consistent Speaker condition, and 15 infants in the Variable Speaker condition demonstrated this preference.
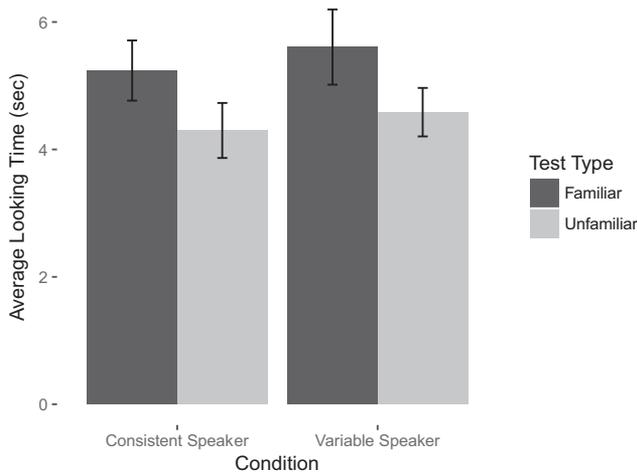

Figure 2. Mean looking times in Experiment 1, where the two familiarization streams consisted of unique phonology. In the Consistent Speaker condition, each speaker produced a single stream, while in the Variable Speaker condition, each speaker produced half of each stream. Error bars represent standard errors of the mean.

These results suggest that infants discovered patterns presented in noisy input, and there was no advantage for hearing each pattern produced deterministically by a unique speaker.

Infants' successful learning, independent of the correspondence between speaker and information stream, reveals that monolingual infants can discover structure in noisy, multidimensional input, unlike previous studies (e.g., Antovich & Graf Estes, 2018; Bulgarelli et al., 2017). Without relying on indexical cues, infants could segregate different sources of information and discovered regularities that required them to generalize across different voices. Reduplication has been suggested to be easy for infants to learn (e.g., Ota & Skarabela, 2016), and it may be that these particular regularities were highly salient, such that even when only 50% of the input conformed to a particular pattern, infants could learn. Alternatively, it could be that the phonological distinction between the two streams facilitated infants' ability to find the Target regularity. Sound differences help infants separate two natural languages (e.g., Bosch & Sebastián-Gallés, 2001; Molnar, Gervain, & Carreiras, 2014), and adults struggle to learn two artificial languages with overlapping phonology (Perruchet, Poulin-Charronnat, Tillmann, & Peereman, 2014), suggesting that salient phonological cues may be helpful in dividing different streams of information. Furthermore, phonological cues can boost adults' learning of regularities that are otherwise challenging (e.g., Onnis, Monaghan, Richmond, & Chater, 2005; Van den Bos, Christiansen, & Misyak, 2012), and concurrent statistical and phonological regularities can support infants' learning of artificial language materials (Sahni, Seidenberg, & Saffran, 2010). Indeed, infants did not need speaker cues to separate the Target and Nontarget streams when there was a phonological distinction between the two streams. In our second study, we examined whether or not infants were able to display learning when presented with two streams that were not phonologically marked.

Experiment 2

In Experiment 2, we tested whether infants could discover regularities when faced with two speech streams consisting of the same sounds. As in Experiment 1, infants were assigned to the Consistent Speaker or Variable Speaker condition and presented with interleaved Target and Nontarget streams, but in Experiment 2, the two streams involved the same syllable inventory, so infants could no longer use phonological cues to separate the two streams. We then asked if the addition of correlated speaker information would play a more substantial role in dictating infants' ability to discover patterns from these less distinct linguistic streams. We predicted that infants would need this additional cue, and only infants in the Consistent Speaker condition, where structured information was paired with speaker, would discover regularities. An alternative possibility was that infants were relying on the phonological cues and would be unable to segregate the two streams without that demarcation.

Method

Participants.   Experiment 2 included 40 additional monolingual infants (9 females, mean age: 10.2 months, range: 9.1–11.0). Half of the infants were assigned to the Consistent Speaker condition; half were assigned to the Variable Speaker condition. Three additional infants were tested, but excluded for fussiness.

Table 2

Stimuli Used in the Nontarget Stream in Experiment 2

Nontarget stream			
le-we-ji	we-li-ji	de-je-wi	je-wi-li
le-wi-je	we-je-di	de-li-je	je-di-we
li-de-wi	wi-le-de	di-ji-we	ji-de-le
li-ji-de	wi-di-le	di-we-li	ji-le-di

Note. All items were presented twice, in a pseudorandomized order. As in Experiment 1, half of the infants heard these items intermixed with AAB target pattern; the other half heard them intermixed with an ABA pattern.

Stimuli. Familiarization materials again consisted of a Target and Nontarget stream. Infants heard one of the same two Target streams (AAB or ABA) as in Experiment 1. The Nontarget stream still consisted of trisyllabic sequences made up of three unique items, but in Experiment 2, the phonemes used in the Nontarget stream were the same as those in the Target stream (Table 2).

The design of the Consistent and Variable Speaker conditions was identical to Experiment 1. Infants in the Consistent Speaker condition heard the Target stream produced by the female speaker and the Nontarget stream produced by the male speaker. In the Variable Speaker condition, infants heard each speaker produce half of each familiarization stream.

Test materials were identical to Experiment 1, with all test items involving phonemes not found in the familiarization phase, produced by the female speaker.

Procedure. The procedure for Experiment 2 was identical to Experiment 1.

Results and Discussion

In Experiment 2, we performed parallel analyses to those of Experiment 1. Our main (Test Type × Condition) ANOVA revealed that unlike in Experiment 1, there was no main effect of test type ($F(1, 38) = .68$, $p = .42$). The interaction between test type and condition also was not significant ($F(1, 38) = .32$, $p = .58$), and infants did not look significantly longer at Familiar versus Unfamiliar items in either condition (Consistent Speaker: 6.1 s vs. 5.5 s, $t(19) = 1.09$, $p = .29$; Variable Speaker: 6.7 s vs. 6.6 s, $t(19) = .17$, $p = .87$; Figure 3). Only 12 infants in the Consistent Speaker condition and 11 infants in the Variable Speaker condition displayed a familiarity preference. When the Target and Nontarget stream included the same sounds, infants did not demonstrate learning of the Target regularities. Even when speaker information was perfectly correlated with structure in the Consistent Speaker condition, infants failed to exploit this association. Thus, indexical cues alone were not sufficient in highlighting the underlying structure found in the two speech streams.

General Discussion

In two studies, we presented monolingual infants with multidimensional input and assessed how they used different acoustic cues to uncover regularities. Infants encountered two streams of speech, comprised of two sets of sounds, produced by two different speakers. In Experiment 1, when each speech stream consisted of unique sounds, infants detected underlying patterns regardless
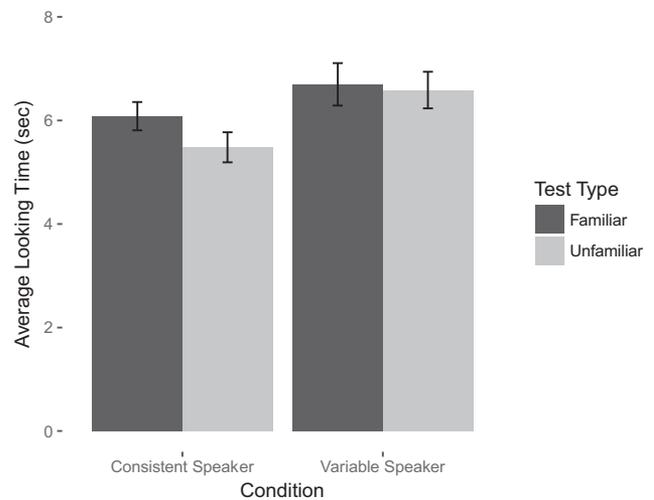

Figure 3. Mean looking times in Experiment 2, where the two familiarization streams consisted of the same syllables. Error bars represent standard errors of the mean.

of whether the patterns were produced by a single speaker or multiple speakers. However, in Experiment 2, when there was no phonological distinction between streams, infants failed to demonstrate learning, even when speakers offered a reliable cue. Thus, sound-based differences appeared to be particularly useful in helping infants to segregate different channels of information, suggesting that infants may prioritize some cues over others in complex environments.

First and foremost, these results highlight infants' robust ability to detect structure. Monolingual 10-month-olds displayed learning of a target speech signal surrounded by inconsistent information. In complex input, learners must discover which regularities to track (e.g., Gerken & Knight, 2015; Mintz, 2002), and even adults rely on concurrent cues, such as sounds or speakers, to track multiple streams of information (e.g., Bulgarelli & Weiss, 2016; Conway & Christiansen, 2006; Gebhart et al., 2009; Karuza et al., 2016; Weiss et al., 2009). In Experiment 1, the phonological division between streams appeared to support infants' discovery of repetition-based regularities in the Target stream. Two potential explanations could account for this successful learning. First, the presence of two different sets of sounds may have helped infants separate the streams. Research on infant bilingualism supports this explanation, as sound-based differences may signal the presence of different languages, structures, or information (e.g., Bosch & Sebastián-Gallés, 2001). Moreover, even infants in single-language environments are exposed to separable speech streams, such as speech directed to the infant versus overheard between adults. This division is marked acoustically (e.g., Fernald & Simon, 1984; Newport, Gleitman, & Gleitman, 1977; Piazza, Iordan, & Lew-Williams, 2017), and infant-directed speech both captures infants' interest and enhances learning, suggesting that acoustic differences between streams may highlight the most relevant information (e.g., Cooper & Aslin, 1990; Weisleder & Fernald, 2013). Another (not mutually exclusive) possibility is that the unique sounds in the Nontarget stream increased the overall variability in the input, and surrounding variability can draw infants'

attention to consistent structure (e.g., Gómez, 2002; Toscano & McMurray, 2010). In either case, infants' successful learning illustrates their keen sensitivity to patterns in their input and suggests that sound-related cues—a pervasive feature of their natural environments—may support learning.

However, not all cues were equally useful in highlighting relevant structure. Contrary to our predictions, dividing information streams by speaker did not facilitate infants' learning in either experiment. Though indexical information may initially be salient for infants (e.g., Jusczyk et al., 1992; Singh et al., 2004; Quam, Knight, & Gerken, 2017), 10-month-old infants can generalize across speakers, suggesting that they learn that differences between speakers are not always meaningful (Houston & Jusczyk, 2000). Indeed, when listening to unfamiliar speech, infants are less sensitive to individual voices, presumably because differences between languages are more salient than differences between speakers (e.g., Johnson, Westrek, Nazzi, & Cutler, 2011; Nazzi et al., 2000). In Experiment 2, infants failed to exploit the usefulness of speakers to help them differentiate the two streams, even when there were no other acoustic distinctions between the streams. Given that listeners use experience to adjust their use of particular dimensions in speech (e.g., Idemaru & Holt, 2011; Potter & Saffran, 2015, 2017; Rost & McMurray, 2009), infants in our experiments may have downweighted the potential value of speakers. Infants do not track all cues equally (e.g., Johnson & Tyler, 2010), and they rely on their everyday experience with language to shape their learning strategies (Thiessen & Saffran, 2007). While speakers may be critical *socially*, they are unlikely to be a reliable cue to important linguistic variation in natural environments, and learners overlook correlated cues that do not typically co-occur with language regularities, such as changes in background color (e.g., Mitchel & Weiss, 2010). On the contrary, sounds can be used to support infants' differentiation of languages (e.g., Bosch & Sebastián-Gallés, 2001; Nazzi et al., 2000), potentially making them a better candidate cue to segregate different information streams. Thus, by 10 months, infants may have discovered it is not advantageous to segregate their learning by speaker, but that different sounds can mark relevant and valuable structural differences.

These studies included only monolingual infants, whose prior experience may have led them to expect that all speakers use a single language. Bilingual infants, whose experience includes regular exposure to multiple, independent structures, might not display similar learning. Prior studies have suggested that bilingual infants may be better able to learn two structures simultaneously (e.g., Antovich & Graf Estes, 2018; Kovács & Mehler, 2009), and bilingual infants may attend to distinctions that monolingual infants ignore, such as nonnative sound contrasts (e.g., Petitto et al., 2012; Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012). In bilingual environments, reliable correlations exist between speakers and languages; monolingual speakers use one language exclusively, and bilinguals tend to preferentially use one language (Grosjean, 2001). Bilingual infants may be sensitive to these relations, and in fact, it has often been explicitly recommended that parents employ a "one parent-one language" strategy, where each parent speaks a separate language to the child (e.g., Barron-Hauwaert, 2004; Ronjat, 1913). It could be that bilingual infants divide their learning by speaker and might benefit from correlated speaker information to learn regularities. Specifically,

one could predict that bilingual infants would show successful learning in the Consistent Speaker condition of Experiment 2. Alternatively, it may be that bilingual infants, like monolinguals, would ignore speaker cues, as all infants must learn to generalize across speakers. Future studies will explore these questions and can provide insight into how infants use experience to inform their use of different cues.

Infants' performance in these experiments demonstrates their ability to focus on the distinctions that are most likely to be useful (e.g., Kuhl et al., 2006; Werker & Tees, 1984). For monolingual infants, a lack of attention to individual speakers may reflect an adaptive strategy. In complex auditory and social environments, infants encounter many cues that vary substantially in their reliability, both in aggregate and in individual contexts of processing. Infants' ability to segregate different speech streams, possibly by recognizing phonological differences, suggests that infants detect and take advantage of reliable cues, an ability that may also support their ability to distinguish one natural language from another. Thus, these studies demonstrate that infants selectively exploit the cues available in their input to find reliable structure in a noisy environment.

References

Antovich, D. M., & Graf Estes, K. (2018). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science, 21,* e12548.

Barron-Hauwaert, S. (2004). Language strategies for bilingual families: The one-parent-one-language approach. *Parents' and teachers' guides; no. 7* (Vol. 1, p. xv, 220). Clevedon, IK: Multilingual Matter. http://dx.doi.org/10.1017/S0272263105250283

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science, 20,* 539–542. http://dx.doi.org/10.1111/j.1467-9280.2009.02327.x

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy, 2,* 29–49. http://dx.doi.org/10.1207/S15327078IN0201_3

Bulgarelli, F., Benitez, V., Saffran, J., Byers-Heinlein, K., & Weiss, D. (2017). Statistical learning of multiple structures by 8-month-old infants. In M. LaMendola and J. Scott (Eds.), *Proceedings of the 41st Annual Boston University Conference on Language Development* (pp. 128–139). Somerville, MA: Cascadilla Press.

Bulgarelli, F., & Weiss, D. J. (2016). Anchors aweigh: The impact of overlearning on entrenchment effects in statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 1621–1631. http://dx.doi.org/10.1037/xlm0000263

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. http://dx.doi.org/10.1006/cogp.1998.0681

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17,* 905–912. http://dx.doi.org/10.1111/j.1467-9280.2006.01801.x

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584–1595. http://dx.doi.org/10.2307/1130766

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* 303–306. http://dx.doi.org/10.1126/science.171.3968.303

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and

biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports, 6,* 25434. http://dx.doi.org/10.1038/srep25434

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113. http://dx.doi.org/10.1037/0012-1649.20.1.104

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 15822–15826. http://dx.doi.org/10.1073/pnas.232472899

Floccia, C., Nazzi, T., & Bertoncini, J. (2000). Unfamiliar voice discrimination for short stimuli in newborns. *Developmental Science, 3,* 333–343. http://dx.doi.org/10.1111/1467-7687.00128

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & Psychophysics, 54,* 287–295. http://dx.doi.org/10.3758/BF03205263

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. http://dx.doi.org/10.1111/j.1551-6709.2009.01041.x

Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition, 98,* B67–B74. http://dx.doi.org/10.1016/j.cognition.2005.03.003

Gerken, L., & Knight, S. (2015). Infants generalize from just (the right) four words. *Cognition, 143,* 187–192. http://dx.doi.org/10.1016/j.cognition.2015.04.018

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications, 4,* 1490–1496. http://dx.doi.org/10.1038/ncomms2430

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431–436. http://dx.doi.org/10.1111/1467-9280.00476

Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7,* 567–580. http://dx.doi.org/10.1111/j.1467-7687.2004.00381.x

Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7,* 183–206. http://dx.doi.org/10.1207/s15327078in0702_4

Gonzales, K., Gerken, L., & Gómez, R. L. (2015). Does hearing two dialects at different times help infants learn dialect-specific rules? *Cognition, 140,* 60–71. http://dx.doi.org/10.1016/j.cognition.2015.03.015

Graf Estes, K., Edwards, J., & Saffran, J. R. (2011). Phonotactic constraints on infant word learning. *Infancy, 16,* 180–197. http://dx.doi.org/10.1111/j.1532-7078.2010.00046.x

Graf Estes, K., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology, 51,* 1517–1528. http://dx.doi.org/10.1037/a0039725

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nichol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Cambridge, MA: Blackwell.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1570–1582. http://dx.doi.org/10.1037/0096-1523.26.5.1570

Idemaru, K., & Holt, L. L. (2011). Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1939–1956.

Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science, 13,* 339–345. http://dx.doi.org/10.1111/j.1467-7687.2009.00886.x

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14,* 1002–1011. http://dx.doi.org/10.1111/j.1467-7687.2011.01052.x

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to sound patterns of native language words. *Journal of Memory and Language, 32,* 402–420. http://dx.doi.org/10.1006/jmla.1993.1022

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition, 43,* 253–291. http://dx.doi.org/10.1016/0010-0277(92)90014-9

Karuza, 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,* 1484–1500. http://dx.doi.org/10.1162/jocn_a_00990

Kovács, A. M., & Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science, 325,* 611–612. http://dx.doi.org/10.1126/science.1173947

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior & Development, 6,* 263–285. http://dx.doi.org/10.1016/S0163-6383(83)80036-8

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9,* F13–F21. http://dx.doi.org/10.1111/j.1467-7687.2006.00468.x

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition, 122,* 241–246. http://dx.doi.org/10.1016/j.cognition.2011.10.007

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283,* 77–80. http://dx.doi.org/10.1126/science.283.5398.77

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38,* 465–494. http://dx.doi.org/10.1006/cogp.1999.0721

May, L., & Werker, J. F. (2014). Can a click be a word?: Infants' learning of non-native words. *Infancy, 19,* 281–300. http://dx.doi.org/10.1111/infa.12048

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11,* 122–134. http://dx.doi.org/10.1111/j.1467-7687.2007.00653.x

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82,* B101–B111. http://dx.doi.org/10.1016/S0010-0277(01)00157-3

Mazuka, R., Hasegawa, M., & Tsuji, S. (2014). Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology, 56,* 192–209. http://dx.doi.org/10.1002/dev.21193

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition, 30,* 678–686. http://dx.doi.org/10.3758/BF03196424

Mitchel, A. D., & Weiss, D. J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes, 25,* 456–482. http://dx.doi.org/10.1080/01690960903209888

Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy, 19,* 326–337. http://dx.doi.org/10.1111/infa.12041

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43,* 1–19. http://dx.doi.org/10.1006/jmla.2000.2698

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: Task-

dependent and dialect-specific aspects. *Journal of Child Language, 41,* 600–633. http://dx.doi.org/10.1017/S0305000913000111

Newport, E., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109–149). New York, NY: Cambridge University Press.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language, 53,* 225–237. http://dx.doi.org/10.1016/j.jml.2005.02 .011

Ota, M., & Skarabela, B. (2016). Reduplicated words are easier to learn. *Language Learning and Development, 12,* 380–397. http://dx.doi.org/ 10.1080/15475441.2016.1165100

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 80–96. http://dx.doi.org/10.1037/0278-7393.34.1.80

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development, 80,* 674–685. http://dx.doi.org/10.1111/j.1467-8624.2009.01290.x

Perruchet, P., Poulin-Charronnat, B., Tillmann, B., & Peereman, R. (2014). New evidence for chunk-based models in word segmentation. *Acta Psychologica, 149,* 1–8. http://dx.doi.org/10.1016/j.actpsy.2014.01.015

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The "perceptual wedge hypothesis" as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language, 121,* 130–143. http://dx.doi .org/10.1016/j.bandl.2011.05.003

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology, 27,* 3162–3167.e3. http://dx.doi.org/10.1016/j .cub.2017.08.074

Potter, C. E., & Saffran, J. R. (2015). The role of experience in children's discrimination of unfamiliar languages. *Frontiers in Psychology, 6,* 1587. http://dx.doi.org/10.3389/fpsyg.2015.01587

Potter, C. E., & Saffran, J. R. (2017). Exposure to multiple accents supports infants' understanding of novel accents. *Cognition, 166,* 67–72. http:// dx.doi.org/10.1016/j.cognition.2017.05.031

Quam, C., Knight, S., & Gerken, L. (2017). The distribution of talker variability impacts infants' word learning. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8,* 1–27. http://dx.doi .org/10.5334/labphon.25

Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2018). The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Developmental Science.* Advance online publication. http://dx.doi .org/10.1111/desc.12704

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General, 133,* 46–62. http://dx.doi.org/10.1037/0096-3445.133.1.46

Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue.* Paris, France: Champion.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12,* 339–349. http://dx.doi.org/10.1111/j.1467-7687.2008.00786.x

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928. http://dx.doi.org/10 .1126/science.274.5294.1926

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology, 69,* 2.1–2.23. http://dx.doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105,* 669–680. http://dx.doi.org/10.1016/j.cognition.2006.11.004

Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development, 81,* 727–736. http://dx.doi.org/10.1111/j.1467-8624.2010 .01430.x

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,* 994–999. http://dx.doi.org/10.1177/ 0956797612436817

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language, 51,* 173–189. http://dx.doi.org/10.1016/j.jml .2004.04.004

Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development, 3,* 73–100.

Toscano, J. C., & McMurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science, 34,* 434–464. http://dx.doi .org/10.1111/j.1551-6709.2009.01077.x

Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science, 16,* 760–771. http://dx.doi.org/10.1111/desc.12064

Van den Bos, E., Christiansen, M. H., & Misyak, J. B. (2012). Statistical learning of probabilistic nonadjacent dependencies by multiple-cue integration. *Journal of Memory and Language, 67,* 507–520. http://dx.doi .org/10.1016/j.jml.2012.07.008

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24,* 2143–2152. http://dx.doi.org/10.1177/ 0956797613488145

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. http://dx.doi.org/10 .1080/15475440802340101

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24,* 672. http://dx.doi.org/10.1037/0012-1649.24.5.672

Werker, J. F., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 7,* 49–63. http://dx.doi.org/10.1016/S0163-6383(84)80022-3