THE OFFICIAL JOURNAL OF THE
INTERNATIONAL CONGRESS
OF INFANT STUDIES

Infancy, **23**(6), 795–819, 2018

ISSN: 1525-0008 print / 1532-7078 online
DOI: 10.1111/infa.12257

WILEY Blackwell

Intonation Plays a Role in Language Discrimination by Infants

Adam J. Chong
Department of Linguistics
Queen Mary University of London and
Department of Linguistics
University of California, Los Angeles

Chad Vicenik and Megha Sundara
Department of Linguistics
University of California, Los Angeles

Previous research on infant language discrimination has focused primarily on the role of prosody, specifically rhythmic timing cues. This, however, ignores the potentially useful role that intonation, another aspect of prosody, might play in aiding discrimination. In this article, we investigated how and when American English-learning infants discriminate between prosodically similar languages, specifically American English and German, focusing on the role of intonation in infant language discrimination. We found that the ability to distinguish American English and German develops between 5 and 7 months. However, 7-month-olds failed to discriminate the two languages when the natural pitch variation was replaced by a monotone. Thus, intonation is necessary for infants' discrimination of American English and German. Based on these results, we argue for a greater role of intonation in supporting language discrimination by infants.

A number of researchers have explored infants' ability to discriminate languages. Early research on language discrimination supported the hypothesis that newborns' familiarity with and recognition of their native language allowed them to distinguish their native language from other non-native languages or dialects (Bahrick & Pickens, 1988; Bosch & Sebastián-Gallés, 1997; Dehaene-Lambertz & Houston, 1997; Mehler et al., 1988; Moon, Cooper, & Fifer, 1993). Subsequent research has attributed successful language discrimination to infants' sensitivity to prosody, specifically the differences in "the rhythmic, timing properties" (Nazzi, Bertoncini, & Mehler, 1998, p. 757) between languages.

Correspondence should be sent to Adam J. Chong, Department of Linguistics, Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom. E-mail: a.chong@qmul.ac.uk

In the following sections, we critically evaluate the notion of rhythm as it applies to language discrimination by infants, arguing that previous studies have largely ignored the role of another aspect of prosody, namely, intonation. Then, we present four experiments investigating American English infants' ability to discriminate between American English and German. In all experiments, rhythmic timing cues were kept intact, yet infant discrimination varied as a function of the presence of intonation cues. Based on these results, we argue that intonation is a necessary cue that infants exploit to discriminate between languages. We suggest that the extent to which infants use intonational cues likely depends on the role of intonation in a language's prosodic system.

Rhythm and rhythmic timing

Languages have frequently been classified in terms of their rhythm, since Pike (1945) and Abercrombie (1967), as either "stress-timed" or "syllable-timed" (or more recently "mora-timed"). Despite documented evidence that native speakers of languages from different rhythm classes process speech in systematically different ways (for, e.g., Bradley, Sánchez-Casas, & Garciá-Albea, 1993; Cutler, Mehler, Norris, & Segui, 1986; Cutler & Norris, 1988; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Otake, Hatano, Cutler, & Mehler, 1993; Pallier, Sebastián-Gallés, Felguera, Christophe, & Mehler, 1993; Sebastián-Gallés, Dupoux, Segui, & Mehler, 1992), there is little consensus about what constitutes rhythm.

The initial idea of rhythm classes centered around isochrony, an idea which rests on the assumption that a language's rhythm is the result of regularly timed units of speech (e.g., syllables in syllable-timed languages, or stressed feet in stressed-time languages). However, research seeking to demonstrate isochrony in speech production has had limited success (see Arvaniti, 2009; Beckman, 1992; Kohler, 2009; Prieto, Vanrell, Astruc, Payne, & Post, 2012 for a review).

Over the last few decades, the view that linguistic rhythm originates primarily from the phonological properties of a language such as the phonotactic permissiveness of consonant clusters, the presence or absence of contrastive vowel length, and vowel reduction, has gained popularity (Dauer, 1983). Thus, a language that is stress-timed is likely to allow more complex consonant clusters, to have lengthened vowels in stressed syllables and reduced vowels in unstressed syllables. In contrast, a syllable-timed language is more likely to restrict consonant clusters and show comparable vowel length over different syllables.

As a consequence, efforts to quantify linguistic rhythm have focused on the distribution of segmental durations. This line of research has led to the development of a variety of metrics aimed at categorizing languages into classes using duration measures of segmental intervals (e.g., proportion of vocalic intervals; Frota & Vigário, 2001; Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999; Wagner & Dellwo, 2004; White & Mattys, 2007). These metrics have been successful in classifying languages which are often considered clear examples of particular rhythm classes (e.g., syllable-timed Spanish, or stress-timed English) with controlled speech material. They have proven less successful, however, when a wider range of materials, speakers, and a larger set of languages are considered (e.g., Arvaniti, 2009, 2012; Loukina, Kochanski, Rosner, Keane, & Shih, 2011; White & Mattys, 2007; Wiget, White, Schuppler, Grenon, Rauch & Mattys, 2010). Further, perception studies have failed to find consistent evidence for such classes (e.g., Arvaniti, 2012; Arvaniti & Rodriquez, 2013; White, Delle Luche, & Floccia, 2016).

Developmental research shows that even newborns are able to discriminate between languages traditionally classified as being from very different rhythm classes (e.g., stress-timed English/Dutch vs. mora-timed Japanese: Nazzi, Bertoncini, and Mehler, 1998; Ramus, 2002). Similar results have been reported for cotton-top Tamarin monkeys (Hauser, Newport, & Aslin, 2001; Ramus, Hauser, Miller, Morris, & Mehler, 2000), rats (Toro, Trobalon, & Sebastián-Gallés, 2003, 2005), and Java sparrows (Watanabe, Yamamoto, & Uozumi, 2006). To the extent that these studies operationalized rhythm as relative *segmental* duration, these results suggest that sensitivity to distributions of segmental duration is at least partially innate, and species-general. In contrast, the ability to discriminate languages within rhythm classes seems to only develop later in the first year of life and, even then, requires familiarity with at least one of the two languages (Bosch & Sebastián-Gallés, 1997, 2001; Christophe & Morton, 1998; Molnar, Gervain, & Carreira, 2014; Nazzi, Jusczyk, & Johnson, 2000; Nazzi, Bertoncini, et al., 1998).

Some recent results, however, suggest that infants' discrimination of languages is not based on categorical rhythm classes per se, but rather on sensitivity to gradient durational differences at the edges of utterances. For instance, reanalyzing data from Butler, Floccia, Goslin, and Panneton (2011), White, Floccia, Goslin and Butler (2014) found that infants were sensitive to local timing differences, specifically, degree of utterance-final lengthening when discriminating between dialects of British English. Similarly, adults' discrimination of languages within the same rhythm class (dialects of the same language) has also been attributed their sensitivity to the degree of phrase-final lengthening (White, Mattys & Wiget, 2012).

More recently, White et al. (2016) examined 5-month-old British English infants' ability to discriminate between French, Spanish, and Finnish. These three languages are all generally considered "syllable-timed." Yet infants were able to discriminate French from Spanish, but not Finnish from either Spanish, or French. Further, White et al. were unable to find any consistent differences in durational measures to account for their results. So even gradient differences in rhythmic timing fail to fully explain differences in infants' ability to discriminate pairs of languages.

An inadvertent consequence of moving toward a definition of rhythm based on the distribution of phonological, particularly segmental, properties has been a dissociation of another crucial aspect of prosody—intonation—from rhythm. For example, in a number of studies, including language discrimination experiments with infants, linguistic rhythm is primarily equated with segmental duration and timing, the target of these rhythm metrics (e.g., Byers-Heinlein, Burns, & Werker, 2010; Molnar et al., 2014; Nazzi, Bertoncini, and Mehler, 1998; Nazzi et al., 2000; Ramus & Mehler, 1999). Yet infants are not only sensitive to durational differences between segments in the speech signal. In the next section, we present evidence that infants are also sensitive to variation in pitch, that is, intonation, within the first year of life.

Intonation as an intrinsic part of prosody

All known languages use pitch to mark the edges of large phrases and sentences (see Jun, 2005, 2014 for an overview). Additionally, in some languages like Korean or French, pitch is used to mark word edges (e.g., Jun & Fougeron, 2000; Kim & Cho, 2009); and in others like English and German, it is used to mark specific syllables within a word as a function of phrasal prominence (e.g., Grice, Baumann, & Benzmuller, 2005; Pierrehumbert, 1980). Together, these regularities in tonal alternations within an utterance can also contribute to the percept of rhythm (e.g., Barry, 1981; Barry, Andreeva, & Koreman, 2009; Dilley & Shattuck-Hufnagel, 1999; Jun, 2005, 2014; Kohler, 2008; Lerdahl & Jackendoff, 1983; Niebuhr, 2009; Thomassen, 1982). For example, tonal alternations, even nonlocal ones, have been shown to have a stronger effect than syllable duration alternations on segmentation of lexically ambiguous words (Dilley & McAuley, 2008). That is, tonal alternations can affect perceived grouping of words. Thus, perceived rhythm cannot solely be about durational and segmental timing properties.

Indeed, several studies have shown that adults can discriminate languages using only pitch cues or at least find pitch cues necessary for successful discrimination (Komatsu, Arai, & Suguwara, 2004; de Pijper, 1983; Ramus & Mehler, 1999; Szakay, 2008; Vicenik & Sundara, 2013; Willems, 1982). For example, European Portuguese listeners could only discriminate Brazilian and European Portuguese when intonation cues were preserved in the stimuli (Frota, Vigário & Martins, 2002; see also Arvaniti & Rodriquez, 2013 for other language comparisons). This is despite the fact that these two dialects of Portuguese are considered to be from different rhythmic classes. Similarly, adult Swiss German listeners were only able to distinguish between two unfamiliar languages, English and French, when intonational cues were present in addition to rhythmic timing cues (Hagmann & Dellwo, 2014).

Like adults, infants are also sensitive to pitch; in fact, their ability to hear and perceive pitch becomes adult-like within the first year of life (Clarkson & Clifton, 1985; Montgomery & Clarkson, 1997; Schneider, Morrongiello, & Trehub, 1990; Spetner & Olsho, 1990). In the linguistic domain, infants demonstrate a fine-grained sensitivity to pitch, at times involving differences as small as 5 or 10 Hz (Bull, Eilers, & Oller, 1985; Frota, Butler, & Vigário, 2014; Karzon & Nicholas, 1989; Nazzi, Floccia, and Bertoncini, 1998). Furthermore, there is evidence that infants use pitch cues preferentially to process speech (Fernald & Kuhl, 1987; Hirsh-Pasek et al., 1987; Mandel, Jusczyk, & Kemler Nelson, 1994; Nazzi et al., 2000; Schmitz, Höhle, & Weissenborn, 2003; Seidl, 2007; Seidl & Cristià, 2008; Shukla, White, & Aslin, 2011). For example, infants show a preference for paying attention to infant-directed speech (IDS) in which the variability in pitch is greater (Cooper, Abraham, Berman, & Staska, 1997; Cooper & Aslin, 1990; Fernald, 1985), and there is evidence that the exaggerated intonation in IDS can aid in various aspects of early learning (Adriaans & Swingley, 2017; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011; Song, Demuth, & Morgan, 2010; Thiessen, Hill & Saffran, 2005; Trainor & Desjardins, 2002; Werker et al., 2007).

While varying pitch does seem to make speech more salient to infants, there is some evidence that the ability to rely on pitch as a cue for other linguistic purposes (e.g., lexical stress; Quam & Swingley, 2014) and for making pragmatic or paralinguistic associations with emotion, for example, might be more protracted in development (Quam & Swingley, 2012). Consistent with this idea, infants at 12 months but not 7 are able to discriminate pitch timing differences that contrast between narrow and broad focus in European Portuguese (Butler, Vigário, & Frota, 2016). This later development of intonational sensitivity is in contrast to the earlier development of lexical tone perception in Chinese infants between the ages of 6 and 9 months (Mattock & Burnham, 2006; Yeung, Chen, & Werker, 2013). Clearly, the development of sensitivity to pitch depends on the linguistic role pitch plays in a particular language. Regardless, although neonates are able to distinguish languages in separate rhythm classes based on just segmental timing information, their ability to do so improves when pitch differences are preserved rather than degraded (Ramus, 2002).

In this article, we tested whether American English-learning infants require intonational cues to discriminate their native language, American English, from a non-native language that is rhythmically very similar, German. In Experiment 1, we show that 7- but not 5-month-olds can discriminate the two languages. Next, in Experiment 2, we examined if prosodic differences were sufficient for language discrimination by 7-month-olds. Infants were tested on their ability to discriminate between American English and German sentences that were low-pass-filtered. Low-pass filtering attenuates segmental information, while preserving most segmental rhythmic timing and intonation properties. Infants successfully discriminated English and German low-pass-filtered speech. In Experiment 3, we resynthesized the American English and German sentences to eliminate pitch cues while preserving segmental and rhythmic timing cues (monotone; see Seidl, 2007; Seidl & Cristià, 2008). In this case, 7-month-olds failed to discriminate the two languages. In Experiment 4, we controlled more carefully for pitch differences across both language stimuli and tested another set of 5- and 7-month-old infants. We replicated the overall results of Experiment 1, showing that 7- but not 5-month-olds successfully discriminated between the two languages. The direction of infants' listening preference, however, was reversed when pitch differences were more controlled. In all experiments, the rhythmic timing differences were intact. Thus, our results show that intonation differences are necessary for English-learning infants to discriminate their native language from German, a non-native language that is prosodically similar.

EXPERIMENT 1: DISCRIMINATION OF AMERICAN ENGLISH AND GERMAN BY 5- AND 7-MONTH-OLDS

A number of studies have shown that neonates are able to distinguish between languages that are from different rhythmic classes (Christophe & Morton, 1998; Mehler & Christophe, 1995; Mehler et al., 1988; Nazzi, Bertoncini, et al., 1998; Nazzi, Floccia, et al., 1998; Nazzi et al., 2000; Ramus, 2002). The ability to discriminate languages from the same rhythm class, however, only develops later in the first year of life and requires familiarity with at least one of the languages (Nazzi et al., 2000). For example, although 2-month-old English-learning infants cannot discriminate between British English and Dutch, two stress-timed languages (Christophe & Morton, 1998; see also Nazzi, Bertoncini, et al., 1998), 5-month-olds can (Nazzi et al., 2000). Similarly, Spanish- and Catalan-learning 4-month-olds can discriminate between Spanish and Catalan, two syllable-timed languages (Bosch & Sebastián-Gallés, 1997, 2001). By 5 months, American English-learning infants can even distinguish between two dialects of their native language, American English and British English, but not two unfamiliar languages in the same rhythm class, Dutch and German (Nazzi et al., 2000). This ability seems to require at least familiarity with one dialect. Thus, 5-month-olds learning Southwestern British English can only discriminate between Southwestern British English and Welsh English, maintaining this ability at 7 months. But they fail to distinguish between Welsh English and Scottish English, two unfamiliar varieties (Butler et al., 2011).

In Experiment 1, we sought to replicate and extend these previous findings. We were interested in examining if 5-month-old infants learning American English can distinguish their native language from a prosodically similar, non-native one, German, and whether this ability is further maintained at 7 months. Both English and German belong to the Germanic family and share a number of similarities in the segmental as well as prosodic domain. They have very similar consonantal inventories, with some differences in the affricates and fricatives (Kohler, 1999; Ladefoged, 1999; Wiese, 1996). American English has interdental fricatives (/θ/ and /ð/) and postalveolar affricates (/ʧ/ and /ʤ/) that German lacks, whereas German has palatal (/ç/) and dorsal fricatives (/x~χ/or/ʁ/) that are absent in English, as well as bilabial (/pf/) and alveolar affricates (/ts/; note however that the phonemic status of these is disputed; Wiese, 1996). In terms of the vowel inventory, a major difference is the fact that German possesses front rounded vowels, which are absent in English.

As for prosody, German and American English are both traditionally considered 'stress-timed' languages (e.g., Kohler, 1983; Pike, 1945) with very similar intonational systems (see Grice et al., 2005 for an overview of German intonational phonology). First, the default intonation contour in both languages involves a high-fall pitch movement at the end of declarative utterances (Grice et al., 2005; for German; Pierrehumbert, 1980; for American English). Second, prominent words in an utterance as well as at prosodic boundaries in both languages are marked using similar tonal events (pitch accents and boundary tones, respectively), the most common of which is a shallow rise on phrasally prominent stressed syllables (see Jun, 2014 for a summary). Finally, both languages have two levels of prosodic structure above the word: the intermediate and intonational phrase. Despite overall similarities, there are subtle differences in the phonetic realization of intonational categories in American English and German. German speakers tend to align tonal rises on stressed syllables later than English speakers (Atterer & Ladd, 2004). They also use more pitch accents and select pitch accents with a steeper rise more often than in English (Vicenik & Sundara, 2013).

In addition to intonational properties, American English and German also differ on the traditional segmental duration based rhythmic measures (see Vicenik & Sundara, 2013). Compared to German, American English has a higher proportion of sonorant (vowels, nasals, and approximants) durations accompanied by smaller standard deviations in obstruent (stops, fricatives, and affricates) duration. Crucially, perception experiments confirm that American English adults can use either rhythmic timing or pitch differences alone to discriminate English and German sentences (Vicenik & Sundara, 2013), although their low discrimination scores suggest that this is not an easy task.

To examine infants' abilities to discriminate between American English and German, in this and all following experiments, we tested infants using the head-turn preference procedure modified for a familiarization-preference task, as in Nazzi et al. (2000; see also Bosch, 1998). In this task, infants are familiarized with passages produced by multiple speakers of one language and then tested on new passages produced by new speakers from the familiar as well as the novel language. Previous research has found both novelty (e.g., Nazzi et al., 2000) and familiarity effects in discrimination (e.g., Butler et al., 2011) using similar paradigms. Given this, we take any significant difference in listening time to be evidence that infants were able to successfully discriminate between the two languages (Houston-Price & Nakai, 2004).

Methods

Participants

Twenty-two 5-month-olds (ages throughout are reported in months; days format. mean age: 5;02; range: 4;18–5;15; 12 males) and twenty-two 7-month-olds (mean age: 7;02; range: 6;16–7;28; 10 males) from monolingual American English-speaking homes participated in this experiment. On average, infants had 97% of their input in English (range = 85–100) as determined by a detailed language questionnaire administered to the parents (Bosch & Sebastián-Gallés, 1997; Sundara & Scutellaro, 2011). This same questionnaire was also used for all subsequent experiments. None had any exposure to German. Twelve additional infants were tested, but excluded because they failed to complete testing due to fussiness (n = 11) and caretaker interference (n = 1).[1]

Stimuli

The stimuli were modeled on those used in Nazzi et al. (2000) and consisted of eight American English and eight German passages. Each passage was made up of five sentences recorded by the same speaker (the sentences were those used in Nazzi, Bertoncini, et al., 1998). Four female native speakers of each language were recorded in a sound-attenuated booth. Each speaker recorded 10 sentences (two passages). To minimize voice quality differences within and between languages, we chose speakers who we perceived to have similar voice qualities. Utterances were all recorded as adult-directed speech with standard declarative intonation, with a falling tonal contour sentence-finally. Example pitch tracks from a sentence in American English and German are shown in Figure 1.

These sentences are a subset of the sentences acoustically analyzed in Vicenik and Sundara (2013); adult English listeners' perceptual discrimination data for these sentences are also reported in that paper. Passages were normalized for intensity in Praat (Boersma & Weenink, 2012) to 80 dB. The acoustic properties of the stimuli are given in Table 1, including duration, mean f0 and f0 range, and number of syllables. Stimuli in both languages showed similar prosodic phrasing, with sentences usually produced in one intonational phrase, containing between one or two smaller intermediate phrases.

Procedure

The procedure and design were identical to that used by Nazzi et al. (2000). The experiment was conducted using the head-turn preference paradigm (HPP; Kemler-Nelson et al., 1995). Infants were tested individually while seated on their parent's lap

[1]This study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the North General Institutional Review Board at the University of California, Los Angeles.

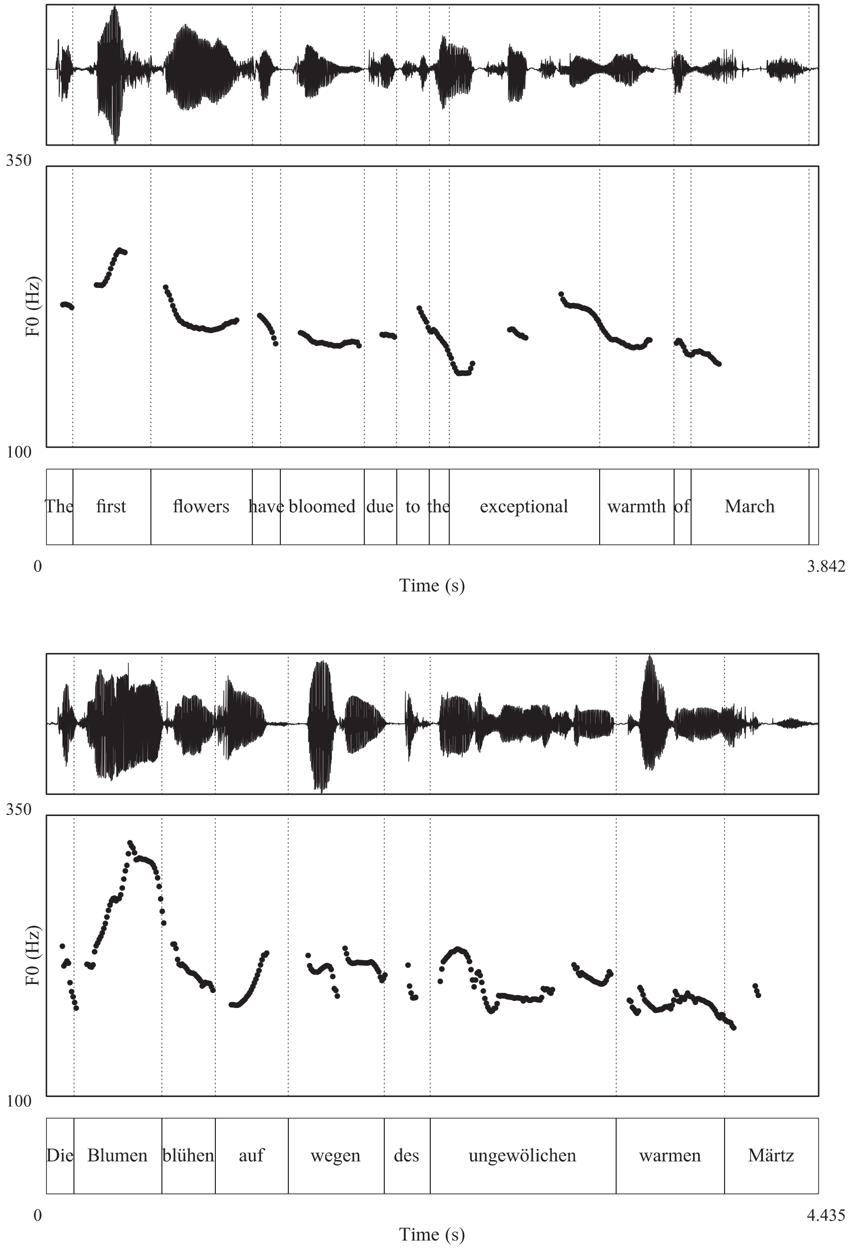

Figure 1 Pitch contours of an example English (top) and German sentence (bottom).

in a three-sided pegboard booth. A red light was mounted on either side panel at the infants' eye level. A Soundworks loudspeaker was hidden behind both side panels. On the center panel, there was one green light and a camera used to record each session. A Sony camera was connected to a monitor, and the lights and speakers were controlled by a computer, both located outside the booth.

TABLE 1
The Acoustic Properties of the Speech Stimuli Used in Experiment 1

	English	*German*
Average number of syllables per passage	89 (6)	91 (6)
Average duration of passages	17.2 sec (1.0)	17.1 sec (2.9)
Mean f0 of passages	215 Hz (18)	185 Hz (21)[a]
Mean f0 range	171 Hz (29)	196 Hz (45)

Notes. Note that the mean f0 (and standard deviations in parentheses) of the English and German passages are different (see text for details).
[a] The mean f0 of the passages was significantly different by unpaired *t*-test ($p = .04$); no other differences were significantly different.

When the experimenter initiated a trial, the green light on the center panel began to blink. Once the infant oriented toward the center light, the center light was extinguished and one of the red sidelights, chosen at random by the program, began to blink. When the infant turned toward the red light (30° head-turn), the auditory stimulus for that trial began to play and continued until the end of the sound file (~17 sec), or until the infant failed to maintain orientation toward the light for two consecutive seconds. A researcher seated at the computer terminal recorded the duration of the infant's head turns. If the infant looked away for <2 sec, but then turned back again, the look away time was not included in the listening time. When the trial ended, or if the infant looked away for more than 2 sec, the sidelight was extinguished and the center light began to blink until the infant reoriented toward the center. At that point, one of the sidelights was randomly chosen by the program to start blinking, initiating another trial. To prevent any influence over the infant's looking time, both the researcher and the infant's caregiver wore sound-attenuating 3M Peltor headphones and listened to music so that they were unaware of the stimuli played during trials.

Design

Each experiment was in two phases: a familiarization phase and a test phase. The familiarization phase consisted of four passages spoken by two speakers from one of the two languages. Half the infants were familiarized with English sentences and the other half with German sentences. To move onto the test phase, infants had to listen to each passage for a total of at least 20 sec (cumulative listening time), for a total minimum familiarization time of 80 sec.

The test phase consisted of eight test trials—four unheard passages of each language spoken by two new speakers per language. The order of presentation of the eight test trials was randomized for each infant. The average listening times to the familiarized and novel languages in the test phase was calculated for each infant and compared statistically.

Results and discussion

Listening times to the familiarized and novel language trials by Age (5- vs. 7-month-olds) are shown in Figure 2. A three-factor repeated-measures ANOVA with Age (5- vs. 7-month-old), Familiarization Condition (English vs. German) as between-subjects

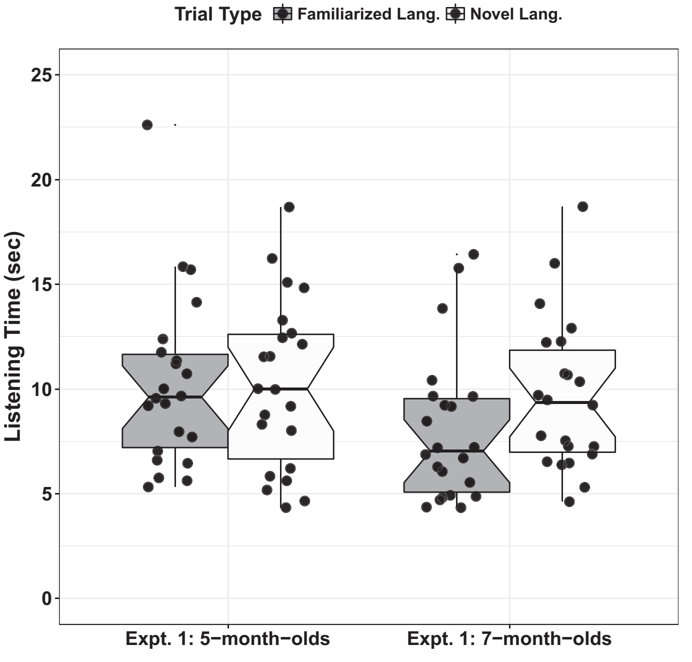

Figure 2 Average listening times by trial type in Experiment 1: full cue with additional supportive pitch cues.

variables, and Test Language (novel vs. familiarized) as the within-subjects variable was used to analyze the results. There was no significant main effect of Age (F(1, 40) = 1.70; p = .20; η_p^2 = 0.04),[2] Familiarization Condition (F(1, 40) = 0.14; p = .71; η_p^2 = 0.003), or Test Language (F(1, 40) = 2.73; p = .11; η_p^2 = 0.06). Neither the interaction of Test Language × Familiarization Condition (F(1, 40) = 0.09; p = .77; η_p^2 = 0.002) nor Familiarization Condition × Age was significant (F(1, 40) = 0.50; p = .48; η_p^2 = 0.01). There was also no significant three-way interaction of Test Language × Familiarization Condition × Age (F(1, 40) = 0.03; p = .87, η_p^2 = 0.001). The interaction of Test Language × Age, however, was marginally significant (F(1, 40) = 3.18; p = .08; η_p^2 = 0.074).

Although we did not find a significant interaction between Age and Test Language, the number of infants who showed a preference for one test language over the other differed across age groups. Only half of 5-month-olds (11 out of 22) showed a longer listening time to the novel language, while a large majority of the 7-month-olds (20 out of 22) did. Moreover, 7-month-olds showed longer average listening times to the novel language (9.65 sec, SD = 3.6) than to the familiarized language (8.02 sec, SD = 3.5), compared to the 5-month-olds, who did not show such a difference (novel: 10.20 sec, SD = 4.0; familiarized: 10.27 sec, SD = 4.1). Given these observations, it is likely that we did not have enough power to detect a robust interaction.

[2]Effect sizes for all experiments are reported as partial-eta-squared values produced by SPSS (IBM Corp., 2013).

To further investigate developmental differences (if any), we examined infants' performance at each age separately. Another repeated-measures ANOVA with Familiarization Condition (English vs. German) as a between-subjects variable and Test Language (novel vs. familiarized) as the within-subjects variable was used to analyze listening times for both age groups separately. For the 5-month-olds, there was no significant main effect of Test Language ($F(1, 20) = 0.005$; $p = .94$; $\eta_p^2 < 0.001$) or Familiarization Condition ($F(1, 20) = 0.05$; $p = .82$; $\eta_p^2 = 0.003$), and no significant interaction between Familiarization Condition and Test Language ($F(1, 20) = 0.005$; $p = .94$; $\eta_p^2 < 0.001$). Contrastively, the 7-month-olds showed a significant main effect of Test Language ($F(1, 20) = 16.04$; $p = .001$; $\eta_p^2 = 0.45$). As with the 5-month-olds, there was no effect of Familiarization Condition ($F(1, 20) = 0.61$; $p = .44$; $\eta_p^2 = 0.03$), and no interaction between Test Language and Familiarization Condition ($F(1, 20) = 0.29$; $p = .60$; $\eta_p^2 = 0.01$). Our results, therefore, indicate that 7- but not 5-month-old American English infants can discriminate their native language from German, a prosodically similar one.

Recall that American English-learning 5-month-olds have been previously shown to discriminate between British English and Dutch, and between British English and their native dialect (Nazzi et al., 2000). In this context, the failure of 5-month-olds to discriminate American English and German suggests that they (or, at least these stimuli) might be more similar than British English and Dutch, at least on the acoustic dimensions attended to by infants. In that case, infants might require more language experience before they are able to discriminate American English and German, explaining why only the 7-month-olds succeeded. We discuss this possibility below in the General Discussion.

EXPERIMENT 2: DISCRIMINATION OF LOW-PASS-FILTERED STIMULI BY 7-MONTH-OLDS

In Experiment 1, we showed that American English-learning infants' ability to discriminate between their native language and German develops between 5- and 7-months of age. What information in the speech signal are 7-month-olds attending to, and what have they learned about language that allows them to discriminate languages that they were previously unable to distinguish?

In Experiment 2, we examined whether 7-month-old infants are able to discriminate between languages with reduced access to segmental information. For this, we low-pass-filtered the American English and German sentences used in Experiment 1. Low-pass filtering attenuates segmental information from the speech signal—which is mostly in the higher frequencies, although some low-frequency information such as the first format of vowels might still be discernible if it is under the cutoff frequency. Infants' success at discriminating languages when the stimuli are low-pass-filtered has traditionally been used as evidence that they are relying on rhythmic timing information (e.g., Bosch & Sebastián-Gallés, 1997; Molnar et al., 2014; Nazzi, Bertoncini, and Mehler, 1998). However, low-pass filtering preserves intonational information in addition to rhythmic timing information. Thus, if infants succeed in discriminating between German and American English low-pass-filtered speech, this would be further evidence that prosodic information alone—both rhythm and intonation- is sufficient for discrimination.

Methods

Participants

Twenty-two American English 7-month-olds (mean age: 7;0; range: 6;19–7;15; eight males) from monolingual English-speaking homes were recruited for this experiment. Infants, on average, had 97% of their input in English (range: 85–100). None had any exposure to German. Four additional infants were tested but their data were excluded from the analysis due to fussiness and completing the experiment (n = 3), and having lower than 80% exposure to American English (n = 1).

Stimuli

To create the stimuli for Experiment 2, the English and German stimuli in both familiarization and test phases in Experiment 1 were modified using a Praat script. The stimuli were low-pass-filtered at a cutoff frequency of 400 Hz following Nazzi, Bertoncini, et al. (1998), Byers-Heinlein et al. (2010), Molnar et al. (2014), among others. This means that any higher frequency segmental cues in the signal are attenuated, leaving intact only rhythmic timing, pitch, and some vocalic cues. Infants in this study, therefore, would not be able to primarily rely on segmental, in particular consonantal, information to distinguish between the two languages.

Procedure and design

The procedure and design used in this experiment were identical to Experiment 1.

Results and discussion

Mean listening times to the familiarized (6.85 sec; SD = 2.14) and novel (8.72 sec; SD = 2.98) language trials in the test phase from Experiment 3 are presented in Figure 3 below. Seventeen out of twenty-two 7-month-olds listened longer to the novel language. Listening times were analyzed using a repeated-measures ANOVA with Familiarization Condition (English vs. German) as a between-subjects variable and Test Language (novel vs. familiarized) as the within-subjects variable. There was a significant main effect of Test Language ($F(1, 20) = 10.68$; $p = .004$; $\eta_p^2 = 0.35$). There was, however, no significant effect of Familiarization Condition ($F(1, 20) = 1.66$; $p = .21$; $\eta_p^2 = 0.08$) and no significant interaction between Test Language and Familiarization Condition ($F(1, 20) = 1.46$; $p = .24$; $\eta_p^2 = 0.07$). Thus, 7-month-olds successfully discriminated American English and German, even with highly attenuated segmental cues. Importantly, infants
were able to discriminate between both languages relying primarily on prosodic cues.

EXPERIMENT 3: DISCRIMINATION OF MONOTONE STIMULI BY 7-MONTH-OLDS

Seven-month-olds' success in discriminating between American English and German in Experiment 3 indicates that they can do so relying on prosodic cues, even with

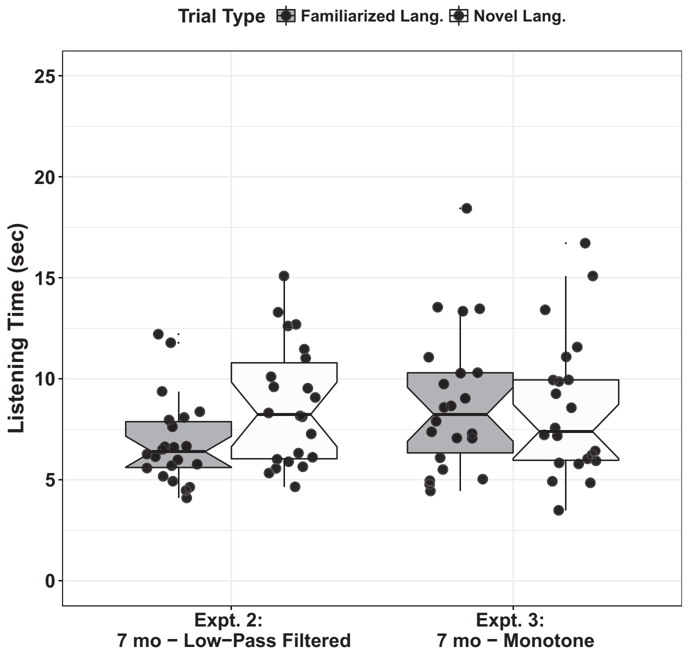

Figure 3 Average listening times by trial type in Experiments 2 (low-pass-filtered) and 3 (monotone stimuli).

attenuated segmental cues. In Experiment 3, we followed previous work by Seidl (2007) and Seidl and Cristià (2008) in neutralizing pitch cues to further examine the importance of these cues in infant speech processing abilities. We resynthesized American English and German stimuli such that the original pitch contours were replaced by a monotone contour. This manipulation removed all intonation information, while preserving both segmental identity and rhythmic differences. If segmental and rhythmic cues are sufficient for language discrimination, 7-month-old English-learning infants should successfully discriminate American English and German monotone stimuli. However, if intonation is necessary for language discrimination, infants were expected to fail to discriminate American English and German monotone stimuli. Therefore, in Experiment 3, we tested whether intonation differences between the two languages are *necessary* for language discrimination.

Methods

Participants

Twenty-two American 7-month-olds (mean age: 7;04 months; range: 6;14–8;0 months; seven males) from monolingual English-speaking homes participated. On average, infants had 98% of their input in English (range = 90–100). None had any exposure to German. An additional four infants were tested, but excluded because they failed to complete testing due to fussiness (n = 3), or equipment problems (n = 1).

Stimuli

The English and German stimuli used in familiarization and test phases from Experiment 1 were modified using a Praat script. The original pitch contours of the passages were extracted and removed. These were then replaced with an artificially generated monotone pitch contour of 220 Hz to approximate the pitch of the American English sentences used previously. Thus, in Experiment 3, any intonational information was eliminated, while preserving rhythmic and segmental differences between the two languages.

Procedure and design

The procedure and design used in this experiment were identical to Experiments 1–2.

Results and discussion

Mean listening times to the familiarized (8.82 sec; $SD = 3.5$) and novel (8.50 sec; $SD = 3.5$) language trials in the test phase from Experiment 3 are also presented in Figure 3 above. Eleven of twenty-two 7-month-olds had a longer listening time to the novel language. A repeated-measures ANOVA with Familiarization Condition (English vs. German) as a between-subjects variable and Test Language (novel vs. familiarized) as the within-subjects variable showed no significant main effects (Test Language: $F(1, 20) = 0.16$; $p = .69$; $\eta_p^2 = 0.008$; Familiarization Condition: $F(1, 20) = 1.03$; $p = .32$; $\eta_p^2 = 0.05$) or interaction ($F(1, 20) = 0.18$; $p = .68$; $\eta_p^2 = 0.009$). Seven-month-olds, therefore, listened comparably to the novel and familiarized language indicating that they could not tell American English and German monotone sentences apart. This shows that 7-month-olds are unable to just use segmental and durational timing information for the purposes of language discrimination. One could argue that the resynthesized speech was unnatural or introduced artifacts making language discrimination difficult for infants in this experiment. Given comparable listening times for 7-month-olds in the experiment using resynthesized speech (Experiment 3) and natural non-manipulated stimuli (Experiment 1), we think this is unlikely.

To confirm that 7-month-olds behaved differently with and without intonation cues, we used a repeated-measures ANOVA with Experiment (pitch, durational timing and segmental cues, i.e., Experiment 1, vs. durational timing and segmental cues only—monotone—i.e., Experiment 3) and Familiarization Condition (English vs. German) as between-subjects variables and Test Language (novel vs. familiarized) as a within-subjects variable. Only the interaction between Experiment and Test Language was significant, ($F(1, 40) = 4.83$; $p = .03$; $\eta_p^2 = 0.11$), driven by the fact that infants could discriminate between the full cue stimuli (Experiment 1), but not the monotone stimuli (Experiment 3). No other effects were significant ($p > .1$). Thus, when intonational cues were absent in the signal, infants were not able to discriminate between American English and German. Given that 7-month-olds were unable to discriminate American English and German monotone stimuli, where rhythmic timing but not intonational cues were present, their success in the previous experiments is likely to have been based on the intonational cues.

EXPERIMENT 4: DISCRIMINATION OF AMERICAN ENGLISH AND GERMAN BY 5- AND 7-MONTH-OLDS—F0 CONTROLLED

Results of Experiments 1–3 suggest that 7-month-olds successfully discriminated between American English and German based on intonational differences between the two languages. An acoustic comparison (Table 1) of the American English and German passages used in Experiment 1, however, shows an average f0 difference of about 30 Hz between speech stimuli in the two languages. Infants might have discriminated between the languages based on this global pitch difference, instead of any intonational differences. Such global pitch differences have been previously proposed to reflect intrinsic differences between the two languages (Mennen, Schaeffler, & Docherty, 2012). Mennen et al. found that female British English speakers in their sample had larger f0 ranges than German female speakers. Regardless of whether the pitch differences between English and German stimuli in our experiment were a result of stimulus selection or stemmed from inherent language-specific differences, Experiment 4 was designed to remove this facilitative f0 difference between the two languages. To do so, we replaced some of the previously used sentences with new ones, also from the same corpus (Vicenik & Sundara, 2013) such that the average f0 was equalized across the passages in the two languages and tested a new group of 5- and 7-month-olds.

Methods

Participants

Another twenty-two monolingual English-learning 5-month-olds (mean age: 4;25; range: 4;13–5;13; 15 males) and twenty-two 7-month-olds olds (mean age: 7;02; range: 6;22–7;16; 10 males) participated in the experiment. Overall, infants had 98% (range = 90–100) input to English. None had any exposure to German. Fifteen additional infants were tested, but excluded because they failed to complete testing due to fussiness (n = 11), caretaker interference (n = 1), because they never looked at the lights (n = 1), or because the looking time difference between the familiar and novel language was more than three standard deviations away from the group mean (n = 2; one positive and the other negative). Note that no other infants in Experiments 1–3 had listening times that were more than three standard deviations away from the mean.

Stimuli

To control for average f0 across both passages, we replaced certain sentences from the stimuli set in Experiment 1 with new sentences to control for average f0 across the passages. The replacement sentences were also drawn from Vicenik and Sundara's original corpus (2013). Passages were normalized for intensity in Praat (Boersma & Weenink, 2012) to 80 dB. The acoustic properties of the new stimuli are given in Table 2, including duration, mean f0 and f0 range, and number of syllables. There were no significant differences between these languages on these measures.

TABLE 2
The Acoustic Properties of the Speech Stimuli Used in Experiment 4

	English	*German*
Average number of syllables per passage	90 (3.7)	91 (5.2)
Average duration of passages	19.6 sec (1.2)	19.6 sec (0.9)
Mean f0 of passages	213 Hz (8.4)	206 Hz (8.5)
Mean f0 range	200 Hz (14.5)	212 Hz (7.3)

Note Standard deviations are presented in parentheses.

Design and procedure

The procedure and design used in this experiment were identical to Experiments 1–3.

Results and discussion

Listening times to the familiarized and novel language trials by Age Group (5- vs. 7-month-olds) are shown in Figure 4. A three-factor repeated-measures ANOVA with Familiarization Condition (English vs. German) and Age (5 vs. 7-month-olds) as between-subjects variables and Test Language (novel vs. familiarized) as the within-subjects variable was used to analyze the overall results. There was a significant main effect of Age ($F(1, 40) = 6.62$; $p = .01$; $\eta_p^2 = 0.14$), with 5 month-olds having longer listening times overall than 7-month-olds, and there was a trend for Test Language ($F(1, 40) = 2.88$; $p = .10$; $\eta_p^2 = 0.07$). The effect of Familiarization Condition was not significant ($F(1, 40) = 0.27$; $p = .61$; $\eta_p^2 = 0.007$). None of the two-way or three-way interactions were significant either (Test Language × Familiarization Condition: $F(1, 40) = 2.44$; $p = .13$; $\eta_p^2 = 0.06$; Familiarization Condition × Age: $F(1, 40) = 0.08$; $p = .78$; $\eta_p^2 = 0.002$; Test Language × Age: $F(1, 40) = 0.14$; $p = .71$, $\eta_p^2 = 0.003$; Test Language × Familiarization Condition × Age: $F(1, 40) = 1.90$; $p = .18$, $\eta_p^2 = 0.05$).

A further inspection of listening times by age confirmed that infants at both ages listened longer to the familiarized language (5-month-olds: 9.87 sec, $SD = 4.0$; 7-month-olds: 7.75 sec, $SD = 2.81$) than the novel language (5-month-olds: 9.39 sec, $SD = 3.2$; 7-month-olds: 6.99 sec, $SD = 2.30$). More 7-month-old infants, however, than 5-month-olds showed this pattern: 14 vs. 12 out of 22. To further examine whether infants at each age discriminated between the familiarized and novel languages, we conducted two-factor repeated-measures ANOVAs for each age group separately, with Familiarization Condition (English vs. German) as a between-subjects variable and Test Language (novel vs. familiarized) as the within-subjects variable was used to analyze the results. For 5-month-olds, there was no significant main effect of Familiarization Condition ($F(1, 20) = 0.02$; $p = .88$; $\eta_p^2 = 0.001$) or Test Language ($F(1, 20) = 0.55$; $p = .47$; $\eta_p^2 = 0.03$), and no significant interaction between Familiarization Condition and Test Language ($F(1, 20) = 2.72$; $p = .12$; $\eta_p^2 = 0.12$). Contrastively, there was a significant main effect of Test Language ($F(1, 20) = 5.21$; $p = .03$; $\eta_p^2 = 0.21$) for 7-month-olds. No other effects or interactions were significant (Familiarization Condition ($F(1, 20) = 0.43$; $p = .52$; $\eta_p^2 = 0.02$; interaction of Test Language and Familiarization Condition ($F(1, 20) = 0.04$; $p = .84$; $\eta_p^2 = 0.002$). Thus, consistent with the results of Experiment 1, 7- but not 5-month-olds successfully discriminated American English and German, even when the stimuli had comparable average f0.

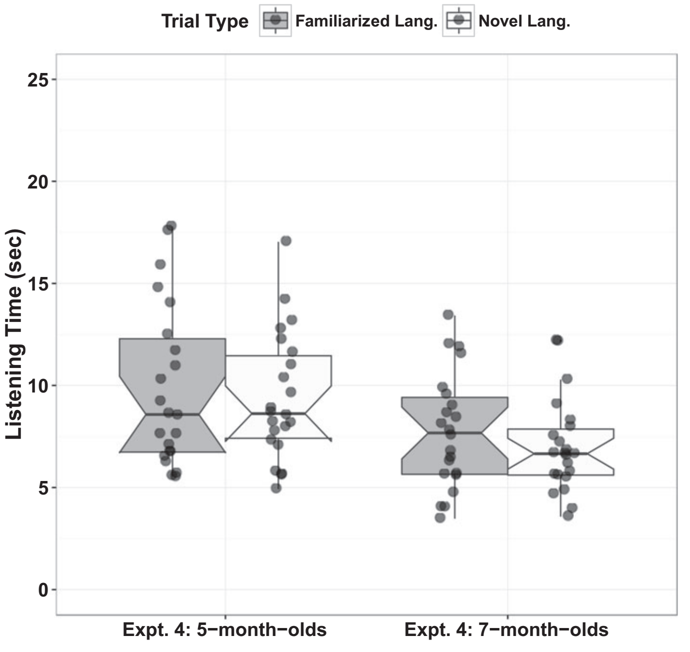

Figure 4 Average listening times by trial type in Experiment 4: full cue without additional supportive pitch cues.

The lack of discrimination by 5-month-olds here is not surprising, given that they were not able to discriminate between both languages even with a 30 Hz supportive pitch difference between American and German stimuli in Experiment 1. While the 7-month-olds showed successful discrimination of familiarized and novel languages here, unlike in Nazzi et al. (2000) and in Experiment 1, infants listened significantly longer to the familiarized language, not the novel language. To confirm that 7-month-olds behaved differently with and without the supportive pitch differences, we analyzed 7-month-olds' listening times in both experiments using a repeated-measures ANOVA with Experiment (1 vs. 4) and Familiarization Condition (English vs. German) as between-subjects variables and Test Language (novel vs. familiarized) as a within-subjects variable. As expected, the interaction of Experiment and Test Language was significant ($F(1, 40) = 20.67$; $p < .001$; $\eta_p^2 = 0.34$). None of the other main effects or interactions were significant ($p > .1$). Thus, 7-month-olds behaved differently in Experiments 1 and 4.

The fact that infants in this experiment showed a familiarity preference is at odds with the novelty preference observed in Nazzi et al.'s (2000) original study. Some recent work, however, examining discrimination of different British English dialects in a similar paradigm (Butler et al., 2011) has also documented preference for the familiar dialect, albeit at 5 months. Interestingly, this preference switched to a preference for the novel accent at 7 months. Factors that have been reported to affect the direction of preference include length of familiarization, age and individual differences in encoding and the salience or complexity of the stimuli (Bornstein, 1985; Cohen, 1969; Houston-Price & Nakai, 2004; Hunter & Ames, 1988). Given that the age of the infants, and the familiarization time was similar for infants tested in Experiment 1 and 4, the direction of preference is likely to be driven by the extent to which the novel and familiar language differed in the test phase. It has been previously shown that infants display a familiarity preference when the familiar choice at test is similar, but not quite identical to the previously experienced stimuli (Gibson & Walker, 1984). In both experiments, talker variation, between familiarization and test phase, is likely to have made the familiarized stimuli similar but not identical. However, the presence of the 30 Hz pitch difference in Experiment 1, but not 4, is likely to have made the distinction between the familiarized and novel test language more salient in Experiment 1, potentially accounting for the novelty preference observed. Regardless, as Houston-Price and Nakai (2004) point out, if the goal is to assess discrimination, any deviation from chance, regardless of direction, is sufficient evidence that infants are discriminating between two types of stimuli.

GENERAL DISCUSSION

In four experiments, using the head-turn preference procedure, we examined American English-learning 5- and 7-month-olds' abilities to discriminate between two languages that are prosodically very similar—American English and German. We were primarily interested in the role of intonation in language discrimination by infants, especially between two languages that are traditionally considered to be in the same rhythm class.

Our results indicate that 7- but not 5-month-olds were able to discriminate between American English and German (see Table 3 for a summary of experiments). They were able to do so with or without supportive pitch cues. Removing the bulk of the segmental information, via low-pass filtering, did not hinder 7-month-olds' ability to discriminate between American English and German passages. Only when speech stimuli were resynthesized to generate monotone sentences, thereby eliminating any pitch cues, did 7-month-olds fail to discriminate between the two languages.

Note that infants' failure to discriminate American English and German when the stimuli were resynthesized to eliminate pitch cues cannot be due to the unnaturalness of the stimuli themselves. Low-pass-filtered stimuli used in Experiment 2 were at least as unnatural as the resynthesized monotone stimuli (also see Bosch & Sebastián-Gallés, 1997; Molnar et al., 2014; Nazzi, Bertoncini, et al., 1998). Yet infants succeeded with low-pass-filtered speech but not monotone speech. Further, 7-month-olds had comparable listening times to natural unmodified stimuli in Experiment 1 and monotone speech in Experiment 3. Instead, we argue that infants' failure to discriminate

TABLE 3
Summary of Experiments and Cues Available in Experimental Stimuli and Discrimination Results

Expt.	*Rhythmic cues*	*Segmental cues*	*Intonation cues*	*Supportive pitch cues*	*Discrimination?*
1	Yes	Yes	Yes	Yes	Yes: 7-month-olds No: 5-month-olds
2	Yes	No	Yes	Yes	Yes
3	Yes	Yes	No	No	No
4	Yes	Yes	Yes	No	Yes: 7-month-olds No: 5-month-olds

American English and German when pitch cues are neutralized is predicted precisely because intonation is perceptually relevant (see also Seidl, 2007; Seidl & Cristià, 2008).

In sum, successful language discrimination in our experiments could not have been based on rhythmic timing information alone because rhythmic timing information was intact in all the testing conditions. Yet, infants' discrimination behavior was variable. In fact, 7-month-olds' failure to use rhythmic timing information alone is in contrast with the performance of adult Americans (Vicenik & Sundara, 2013). Rather we show that intonation is a *necessary* cue for American English-learning infants to discriminate American English and German.

The fact that American English-learning 7-, but not 5-month-olds were able to successfully discriminate their native language from a prosodically similar non-native language, German, indicates that this discrimination ability develops between 5 and 7 months. The failure of 5-month-olds to discriminate American English and German in Experiments 1 and 4 stands in contrast to Nazzi et al.'s (2000) finding that American English 5-month-olds successfully discriminate British English from Dutch. Given the similarity of the methodology used in these experiments and in Nazzi et al., we can only assume that American English and German are perceptually more similar for American English-learning infants than British English and Dutch.

Both pairs of languages (American English and German, as well as British English and Dutch) are considered to be within the same rhythm class; thus, we cannot appeal to categorical differences in rhythm classes to compare the two. There were also no differences in terms of speech rate, average pitch or pitch range between the two pairs. A more detailed examination of the rhythmic and durational characteristics of the stimuli used in the two experiments (Appendix A) showed that both language pairs showed significant differences on some rhythmic features while no differences on others, although British English and Dutch seem to differ on more measures than American English and German. It is possible, then, that American English and German are harder to discriminate because they differ on fewer rhythm measures than British English and Dutch.

Additionally, we propose that American English and German are also intonationally more similar than British English and Dutch (at least in the stimuli compared). A comparison of the British English and Dutch stimuli from Nazzi et al. (2000) found that Dutch had significantly more instances of pitch rises per passage than British English (see Appendix A; see also Jun, 2005, 2014). In contrast, there was no significant difference in the number of pitch rises in American English and German in our stimuli. Instead, as described in the literature (Atterer & Ladd, 2004), the differences between the intonation of American English and German were subtle and typically restricted to the alignment of the high pitch peak on prominent syllables; this peak had a slightly steeper rise in German than in English (Vicenik & Sundara, 2013). It is possible then that the 5-month-olds in Nazzi et al.'s (2000) studies succeeded because of salient differences in pitch modulations in Dutch compared to British English. However, only 7-month-olds were able to use fine-grained phonetic differences of phonologically similar intonation targets necessary to distinguish American English and German. This would be consistent with Butler et al.'s (2016) findings that infants' ability to discriminate distinctions based on pitch timing only develops later in the first year of life.

In summary, we have shown that infants need intonational cues to distinguish rhythmically similar language pairs like American English and German and that this ability develops between 5 and 7 months of age. Thus, a full account of language discrimination needs to go beyond a sole reliance on rhythmic timing cues, and future research should more closely examine how both phonological and phonetic differences in intonation can impact language discrimination.

ACKNOWLEDGMENTS

We would like to thank Volker Dellwo and Constanze Weise for help in recording experimental stimuli and Robyn Orfitelli, Anya Mancillas, and Victora Mateu for recruiting and testing infants. We would also like to thank Thierry Nazzi for sharing stimuli from a previous study with us. Experiments 1 and 2 were part of CV's doctoral thesis. This research was supported by UCLA COR Faculty Research Grant, and NSF BCS-0951639 to MS, and a UCLA Summer Research Mentorship Award to CV. Parts of this research were presented at the 158th meeting of the Acoustical Society of America, San Antonio; the 2010 International Conference on Infant Studies, Baltimore, Maryland; and at the 162nd meeting of the Acoustical Society of America, San Diego. We are grateful to Associate Editor, Suzanne Curtin, and three anonymous reviewers for their helpful feedback, and to Nichol Castro for editorial assistance. The authors have no conflicts of interest, financial or nonfinancial, in the subject matter and materials of this article.

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, UK: Edinburgh University Press.

Adriaans, F., & Swingley, D. (2017). Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. *The Journal of the Acoustical Society of America*, *141*, 3070–3078.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*, 46–63.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, *40*, 351–373.

Arvaniti, A., & Rodriquez, T. (2013). The role of rhythm class, speaking rate, and F0 in language discrimination. *Laboratory Phonology*, *4*, 7–38.

Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics*, *32*, 177–197.

Bahrick, L. E., & Pickens, J. N. (1988). Classification of bimodal English and Spanish language passages by infants. *Infant Behavior and Development*, *11*, 277–296.

Barry, W. J. (1981). Prosodic functions revisited again!. *Phonetica*, *38*, 320–340.

Barry, W. J., Andreeva, B., & Koreman, J. (2009). Do rhythm measures reflect perceived rhythm? *Phonetica*, *66*, 78–94.

Beckman, M. E. (1992). Evidence for speech rhythm across languages. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure* (pp. 457–463). Amsterdam, The Netherlands: IOS Press.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer (version 5.3.04)* [software]. Retrieved from http://www.praat.org/

Bornstein, M. H. (1985). Habituation of attention as a measure of visual information processing in human infants: Summary, systematization, and synthesis. In G. Gottlieb, & N. A. Krasnegor (Eds.), *Measurement of vision and audition in the first year of life* (pp. 253–301). Norwood, NJ: Ablex.

Bosch, L. (1998). Bilingual exposure and some consequences on native language recognition processes at four months. In P. K. Kuhl, & L. Crum (Eds.), *Proceedings of the joint meeting of the Acoustical Society of America and the international conference on acoustics* (pp. 1599–1600). Seattle, WA: Acoustical Society of America.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*, 33–69.
Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*, 29–49.
Bradley, D. C., Sánchez-Casas, R. M., & Garciá-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, *8*, 197–223.
Bull, D., Eilers, R. E., & Oller, D. K. (1985). Infants' discrimination of final syllable fundamental frequency in multisyllabic stimuli. *Journal of Acoustical Society of America*, *77*, 289–295.
Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, *16*, 392–417.
Butler, J., Vigário, M., & Frota, S. (2016). Infants' perception of the intonation of broad and narrow Focus. *Language, Learning and Development*, *12*, 1–13.
Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, *21*, 343–348.
Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, *1*, 215–219.
Clarkson, M. G., & Clifton, R. K. (1985). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *Journal of Acoustical Society of America*, *77*, 1521–1528.
Cohen, L. B. (1969). Uses and misuses of habituation and related preference paradigms. *Infant and Child Development*, *13*, 349–352.
Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for Motherese. *Infant Behavior and Development*, *20*, 477–488.
Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*, 1584–1595.
Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385–400.
Cutler, A., & Norris, D. G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 113–121.
Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*, 51–62.
de Pijper, J. R. (1983). *Modeling British English intonation*. Dordrecht, The Netherlands: Foris.
Dehaene-Lambertz, G., & Houston, D. (1997). Faster orientation latencies toward native language in two-month-old infants. *Language and Speech*, *41*, 21–43.
Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, *59*, 294–311.
Dilley, L. C., & Shattuck-Hufnagel, S. (1999). Effects of repeated intonation patterns on perceived word-level organization. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1487–1490.
Fernald, A. (1985). Four-month-old infants prefer to listen to Motherese. *Infant Behavior and Development*, *8*, 181–195.
Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*, 279–293.
Frota, S., Butler, J., & Vigário, M. (2014). Infants' perception of intonation: Is It a statement or a question?. *Infancy*, *19*, 194–213.
Frota, S., & Vigário, M., & Martins, F. (2002). Language discrimination and rhythm class: Evidence from Portuguese. In *Proceedings of Speech Prosody 2002, Aix-en-Provence*, 315–318.
Frota, S., & Vigário, M. (2001). On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus*, *13*, 247–275.
Gibson, E. J., & Walker, A. S. (1984). Development of knowledge of visual-tactual affordances of substance. *Child Development*, *55*, 453–460.
Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven, & N. Warner (Eds.), *Papers in laboratory phonology 7* (pp. 515–546). Berlin, Germany: Mouton de Gruyter.
Grice, M., Baumann, S., & Benzmuller, R. (2005). German intonation in Autosegmental-Metrical phonology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 55–83). Oxford, UK: Oxford University Press.
Hagmann, L., & Dellwo, V. (2014). Listeners may rely on intonation to distinguish languages of different rhythm classes. *Loquens*, *1*, e008.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, B53–B64.

Hirsh-Pasek, K., Kemler-Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*, 341–348.

Hunter, M., & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier, & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 5, pp. 69–95). Stamford, CT: Ablex.

IBM Corp (2013). *IBM SPSS statistics for windows*, Version 22.0. Armonk, NY: IBM Corp.

Jun, S.-A. (2005). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430–458). Oxford, UK: Oxford University Press.

Jun, S.-A. (2014). Prosodic typology: By prominence type, word prosody, and macro-rhythm. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 520–539). Oxford, UK: Oxford University Press.

Jun, S., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modeling and technology* (pp. 209–242). Dordrecht, The Netherlands: Kluwer Academic.

Karzon, R., & Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception and Psychophysics*, *45*, 10–14.

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

Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of Acoustical Society of America*, *125*, 3373–3386.

Kohler, K. (1983). Stress-timing and speech rate in German. In *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK) No. 20*, 55–97.

Kohler, K. (1999). German. In *Handbook of the International Phonetic Association* (pp. 86–89). Cambridge: Cambridge University Press.

Kohler, K. (2008). The perception of prominence pattern. *Phonetica*, *65*, 257–269.

Kohler, K. (2009). Rhythm in speech and language. *Phonetica*, *66*, 29–45.

Komatsu, M., Arai, T., & Suguwara, T. (2004). Perceptual discrimination of prosodic types and their preliminary acoustic analysis. In *Proceedings of interspeech 2004*, 3045–3048.

Ladefoged, P. (1999). American English. In *Handbook of the International Phonetic Association* (pp. 41–44). Cambridge: Cambridge University Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *Journal of Acoustical Society of America*, *129*, 3258–3270.

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, *7*, 185–201.

Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. (1994). Does sentential prosody help infants organize and remember speech information?. *Cognition*, *53*, 155–180.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*, 241–265.

Mehler, J., & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 943–954). Cambridge, MA: Bradford Books.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, *20*, 298–305.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halstead, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178.

Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *The Journal of the Acoustical Society of America*, *131*, 2249–2260.

Molnar, M., Gervain, J., & Carreira, 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.

Montgomery, C. R., & Clarkson, M. G. (1997). Infants' pitch perception: Masking by low- and high-frequency noises. *Journal of Acoustical Society of America*, *102*, 3665–3672.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day olds prefer their native language. *Infant Behavior and Development*, *16*, 495–500.

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.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, *21*, 779–784.

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.

Niebuhr, O. (2009). F0-based rhythm effects on the perception of local syllable prominence. *Phonetica*, *66*, 95–112.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, *32*, 258–278.

Pallier, C., Sebastián-Gallés, N., Felguera, T., Christophe, A., & Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language*, *32*, 373–389.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English Intonation* (Ph.D. Dissertation). MIT, Cambridge, MA. Distributed 1988, Indiana University Linguistics Club.

Pike, K. (1945). *The intonation of American English*. Ann Arbor, NY: University of Michigan Press.

Prieto, P., Vanrell, M. M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication*, *54*, 681–702.

Quam, C., & Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development*, *83*, 236–250.

Quam, C., & Swingley, D. (2014). Processing of lexical stress cues by young children. *Journal of Experimental Child Psychology*, *123*, 73–89.

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, *2*, 1–14.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and cotton-top tamarin monkeys. *Science*, *288*, 349–351.

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

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.

Schmitz, M., Höhle, B., & Weissenborn, J. (2003). How pause length influences the perception of major syntactic boundaries in 6-month-old German infants. *Paper presented at the conference of Generative Approaches to Language Acquisition (GALA)*, Groningen.

Schneider, B., Morrongiello, B. A., & Trehub, S. (1990). The size of the critical band in infants, children, and adults. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 642–652.

Sebastián-Gallés, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, *31*, 18–32.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, *57*, 24–48.

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*, 596–606.

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of National Academy of Sciences of the United States of America*, *108*, 6038–6043.

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, *128*, 389–400.

Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, *61*, 632–652.

Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, *39*, 505–513.

Szakay, A. (2008). *Ethnic dialect identification in New Zealand: The role of prosodic cues*. Saarbrücken, Germany: VDM Verlag.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant directed speech facilitates word segmentation. *Infancy*, *7*, 49–67.

Thomassen, J. M. (1982). Melodic accent: Experiments and a tentative model. *The Journal of the Acoustical Society of America*, *71*, 1596–1605.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition*, *6*, 131–136.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 95–100.

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, *9*, 335–340.

Vicenik, C., & Sundara, M. (2013). The role of intonation in language and dialect discrimination by adults. *Journal of Phonetics*, *41*, 297–306.

Wagner, P., & Dellwo, V. (2004). Introducing YARD (Yet Another Rhythm Determination) and reintroducing isochrony to rhythm research. In *Proceedings of Speech Prosody 2004, Nara, Japan*, 227–230.

Watanabe, S., Yamamoto, E., & Uozumi, M. (2006). Language discrimination by Java sparrows. *Behavioural Processes*, *73*, 114–116.

Werker, J., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*, 147–162.

White, L., Delle Luche, C., & Floccia, C. (2016). Five-month-old infants' discrimination of unfamiliar languages does not accord with "rhythm class". In *Proceedings of Speech Prosody 2016*, Boston, USA, 567–571.

White, L., Floccia, C., Goslin, J., & Butler, J. (2014). Utterance-final lengthening is predictive of infants' discrimination of English accents. *Language Learning*, *64*, 27–44.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*, 665–679.

Wiese, R. (1996). *The phonology of German*. Oxford, UK: Clarendon Press.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of Acoustical Society of America*, *127*, 1559–1569.

Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht, The Netherlands: Floris.

Yeung, H., Chen, L., & Werker, J. (2013). Referential labeling can facilitate phonetic learning in infancy. *Child Development*, *85*, 1036–1049.

APPENDIX A:COUSTIC COMPARISONS OF NAZZI ET AL.'S (2000) STIMULI USING RHYTHMIC MEASURES FROM VICENIK AND SUNDARA (2013)

TABLE A1
Mean and Standard Deviations (in Parentheses) of Acoustic Measures of British English and Dutch Stimuli (Nazzi et al., 2000) and American English and German Stimuli (Experiments 1–2), Using Measures from Vicenik and Sundara (2013)

	American English	*German*	*British English*	*Dutch*
Raw PVI obst	59.38 (20.92)	63.64 (16.75)	56.06 (23.23)	60.8 (15.93)
Norm PVI obst	0.66 (0.17)	0.66 (0.15)	0.51 (0.11)	0.65 (0.17)[a]
Raw PVI son	100.02 (39.49)	90.2 (47.87)	108.86 (36.95)	90.37 (42.41)[a]
Norm PVI son	0.73 (0.17)	0.66 (0.15)[a]	0.68 (0.18)	0.73 (0.15)
% Son	0.58 (0.06)	0.55 (0.07)[a]	0.59 (0.07)	0.54 (0.07)[a]
% Obst	0.42 (0.06)	0.45 (0.07)[a]	0.41 (0.07)	0.46 (0.07)[a]
ΔSon	88.69 (28.28)	81.81 (44.14)	161.51 (49.81)	131.47 (63.21)[a]
ΔObst	51.82 (18.55)	56.08 (14.73)	97.27 (34.58)	89.93 (31.58)
Mean O	95.05 (18.06)	101.83 (19.77)	108.49 (22.21)	99.36 (11.57)[a]
Mean S	134.67 (29.17)	131.36 (52.57)	157.1 (32.48)	119.67 (35.19)[a]
Varco O	0.54 (0.13)	0.56 (0.14)	0.93 (0.39)	0.92 (0.32)

Table (Continued)

	American English	*German*	*British English*	*Dutch*
Varco S	0.65 (0.14)	0.6 (0.11)	1.05 (0.33)	1.1 (0.42)
Mean f0 (Hz)	215 (17.9)	185 (21.1)[a]	224 (14.6)	217 (21.5)
F0 range (Hz)	196 (45.4)	171 (29.3)	204.5 (41.6)	200 (40.6)
No. of f0 rises per passage	13.1 (4.26)	16.3 (3.34)	16.4 (2.45)	19.6 (3.02)[a]
No. of f0 rises per second	0.8 (0.26)	1 (0.22)	1 (0.16)	1.2 (0.2)[a]
Speech rate (syll./sec)	5.34 (0.7)	5.69 (1.14)	5.61 (0.42)	5.29 (0.33)
nFinalCV	1.32 (0.32)	1.11 (0.35)[a]	1.52 (0.44)	1.34 (0.47)

Notes. Vicenik and Sundara (2013) use sonorant (S) and obstruent (O) intervals instead of consonant and vowel intervals.
nFinalCV = duration of the final consonant + vowel interval divided by the mean consonant + vowel interval duration for each utterance (see White et al., 2012; White et al., 2014); no. of f0 rises per second = no. of f0 rises/total passage duration.
[a]Indicates differences between languages are significant by unpaired t-test at $p < .05$.