Journal of Phonetics 39 (2011) 505–513

Contents lists available at ScienceDirect

Journal of Phonetics

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

Rhythmic distance between languages affects the development of speech perception in bilingual infants

Megha Sundara *, Adrienne Scutellaro [1]

Department of Linguistics, University of California at Los Angeles, 3125 Campbell Hall, Los Angeles, CA 90095-1543, USA

ARTICLE INFO

Available online 18 September 2010

ABSTRACT

The time course and trajectory of development of phonetic perception in Spanish–Catalan bilingual and monolingual infants is different (Bosch & Sebastián-Gallés, 2003a, 2003b, 2005; Sebastián-Gallés & Bosch, 2009). Bosch and Sebastián-Gallés argue that, at least initially, bilingual infants track statistical regularities across the two languages, leading to their temporary inability to discriminate acoustically similar phonetic categories. In this paper, we test bilingual Spanish–English 4- and 8-month-olds' discrimination of vowels. Results indicate that, when the two languages being learned are rhythmically dissimilar, bilingual infants are able to discriminate acoustically similar vowel contrasts that are phonemic in one, but not the other language, at an earlier age. These results substantiate a mechanism of language tagging or sorting; such a mechanism is likely to help bilingual infants calculate statistics separately for the two languages.

1. Introduction

From research in the last two decades we know much about how language experience affects the perception of speech sound categories – consonants, vowels and tones – in monolingual infants. Language experience serves to facilitate the discrimination of native language sound distinctions to bring infants' discrimination abilities up to the high levels at which native adult listeners perform (Kuhl et al., 2006; Narayan, Werker, & Beddor, 2010; Polka, Colantonio, & Sundara, 2001; Sundara, Polka, & Genesee, 2006; Tsao, Liu, & Kuhl, 2006; see Aslin & Pisoni, 1980 for a description of the facilitation pattern). In the absence of this language experience, discrimination of some (Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1983, 1984a), but not all sound distinctions, declines (see Best & McRoberts, 2003 for a review).

At present there is a consensus in the field that this change from language-general to language-specific abilities is not because children are exposed to minimal pairs in speech input. Rather, this change is likely to involve a domain-general, perceptual learning mechanism that exploits infants' sensitivity to the statistical properties of the speech input (Anderson, Morgan, & White, 2003; Kuhl & Meltzoff, 1996; Maye, Werker, & Gerken, 2002; Maye, Weiss, & Aslin, 2008).

In comparison with research on monolingual infants, research on how language experience affects the perception of speech sound categories in bilingual infants is scarce. This gap in the literature is of particular concern given some estimates that there are as many, if not more, children growing up bilingual than monolingual (Tucker, 1998). In this study, we tested Spanish–English learning bilingual infants and their English-learning peers on the discrimination of a vowel distinction (/e - ɛ/) that contrasts meaning in English, but not Spanish.

1.1. Language input to bilingual infants

In acquiring the speech sound categories of their native language, infants face a very challenging task. They must learn to treat some physically non-identical instances of sounds equivalently, for example /a/ produced by men, women and children. This ability is essential for infants learning any language, and can be seen as early as 6 months (Kuhl, 1979, 1983). At the same time, depending on the ambient language, infants must also learn to treat other sound pairs which may be acoustically as equally distinct as belonging to different categories. For example, infants learning English learn to treat stop consonants with a 0 and 60 ms VOT as different (i.e., voiced and voiceless). In contrast, infants learning French learn to treat stop consonants with 0 ms VOT as voiced (e.g., /b/) but stop consonants with 30 and 60 ms VOT as voiceless (e.g., /p/). From research on monolingual infants

* Corresponding author. Tel.: +1 310 825 0041; fax: +1 310 206 5743.
E-mail addresses: megha.sundara@humnet.ucla.edu (M. Sundara), adscut@sas.upenn.edu (A. Scutellaro).

[1] Current address: Institute for Research in Cognitive Science, Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Suite 400A, Philadelphia, PA 19104-6228, USA.

0095-4470/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2010.08.006

we know that infants show attunement to native language sound contrasts sometime in the second half of the first year of life (for a summary see Polka, Rvachew, & Mattock, 2007).

Bilingual infants exposed to two languages from birth face the same challenges of dealing with variability across different talkers as monolingual infants do. They also have to contend with the fact that consonant and vowel inventories in the two languages are seldom the same. The consonant and vowel inventories of languages can differ in quantity, quality or both. First, vowel or consonant distinctions may contrast meaning in one language, but not the other. For example, for a Catalan and Spanish-learning bilingual infant, the difference between either the vowels /e – ε/ or the consonants /s – z/ alters meaning in Catalan, but not in Spanish.

Second, phonemes in different languages may not be phonetically instantiated in exactly the same way. For example, both Spanish and Catalan have the vowels /e/, /o/ and /u/; however, they are produced with systematically different first (F1) and second formant (F2) values in the two languages. Similarly, both English and French have voiced and voiceless stops; however, in English, voiced stops have VOT values between 0 and 30 ms, whereas voiceless stops have VOT values greater than 30 ms. In contrast, in French, voiced stops have negative VOT values, whereas voiceless stops have VOT values between 0 and 30 ms (Lisker & Abramson, 1964; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Sundara, 2005; Sundara, Polka, & Baum, 2006).

Thus, bilingual infants are getting input in two languages such that the distributions of segments in those two languages are sometimes overlapping. This overlapping distribution of segments provides a challenge to the developing bilingual infant.

1.2. Comparison of bilingual and monolingual infants' perception of native language sound categories

The existing research on the development of speech perception in bilingual infants shows two distinct patterns. When tested on the discrimination of phonetic categories that are well separated, bilingual infants show the same time course of development as monolingual infants. Thus, Catalan and Spanish-learning bilingual and monolingual 4- and 8-month-olds are able to discriminate the /e – u/ distinction (Bosch & Sebastián-Gallés, 2005; Sebastián-Gallés & Bosch, 2009). Note that /e/ and /u/ are present in both Spanish and Catalan, although they are not phonetically instantiated in exactly the same way, causing the distributions in the two languages to be overlapping. Critically, /e/ and /u/ are well separated in the F1–F2 space. So, misalignment of phonetic categories alone does not hinder the development of speech perception in bilingual infants. If categories are well separated, then bilingual and monolingual infants show similar developmental patterns.

In contrast, when tested on the discrimination of phonetic categories that are not well-separated, the developmental trajectory of bilingual and monolingual infants diverges. Specifically, bilingual infants, unlike monolingual infants, show a U-shaped developmental pattern when tested with such phonetic contrasts. This has been observed for the Catalan /e – ε/ contrast (Bosch & Sebastián-Gallés, 2003a). Bosch and Sebastián-Gallés report that bilingual 4- and 12-month-olds were able to discriminate the Catalan /e – ε/ distinction whereas bilingual 8-month-olds failed to do so. This was in contrast to the performance of monolingual infants. As expected, monolingual Catalan-learning 4- and 8-month-olds were successful at discriminating Catalan /e – ε/. Further, monolingual Spanish learning 4- but not 8-month-olds succeeded in discriminating Catalan /e – ε/, showing the classic decline-in-the-absence-of-experience pattern.

Such a U-shaped developmental trajectory has also been reported for bilingual Catalan and Spanish-learning infants' discrimination of the Catalan /s – z/ contrast (Bosch & Sebastián-Gallés, 2003b). In Spanish [s] and [z] are in allophonic distribution and thus, do not contrast meaning (Hualde, 2005; Stockwell & Kiddle, 1956). The U-shaped pattern has also been observed for the Spanish and Catalan /o – u/ contrast (Bosch & Sebastián-Gallés, 2005; Sebastián-Gallés & Bosch, 2009). Note, unlike the Catalan /e – ε/ and /s – z/ contrasts, both /o/ and /u/ contrast meaning in Spanish as well as Catalan. Rather, due to differences in phonetic instantiation, bilingual infants find the overlapping distribution of the two vowels in Spanish and Catalan to be challenging.

1.3. Strategies for dealing with overlapping distributions

Bilingual infants could utilize two potential strategies for resolving overlapping distributions of phonetic categories that are not well-separated. These strategies are *not* mutually exclusive. The first strategy is a brute force approach. Specifically, bilingual infants may need to accumulate more input to determine multiple shallow peaks in the overlapping input distribution of speech sounds. The second strategy is one of tagging or sorting input by language. If infants are able to tag sections of the incoming speech stream, even rudimentarily either as Language A, or not Language A, then as far as the bilingual infant is concerned there is no problem of overlapping distributions. In other words, tagging or sorting input allows infants to segregate the two languages, enabling them to calculate statistics for each of the languages separately. In the following section we lay out the existing support for the two strategies leading up to the motivation for the present study.

1.3.1. Strategy I: the brute force approach

There is evidence that bilingual infants may indeed be able to exploit a brute force approach. Recall that 12- but not 8-month-old Catalan and Spanish-learning bilingual infants are able to discriminate the Catalan /e – ε/ (Bosch & Sebastián-Gallés, 2003a) and the Spanish & Catalan /o – u/ contrasts (Bosch & Sebastián-Gallés, 2005). Based on the U-shaped developmental curve observed in bilingual infants it is evident that older bilingual infants are able to successfully discriminate phonetic differences with overlapping distributions in the input. One proposal for how older bilingual infants accomplish this task appeals to the fact that older infants have accumulated more input (Bosch & Sebastián-Gallés, 2003a). Although by no means conclusive, the U-shaped developmental curve is consistent with the use of a brute force approach by bilingual infants.

More compelling evidence that a brute force approach may be used by infants to successfully discriminate phonetic categories that have overlapping distributions comes from a study investigating the discrimination of English and French[2] /d/ by English and French-learning bilingual and monolingual infants (Sundara, Polka, & Molnar, 2008). English /d/ is phonetically described as alveolar whereas the French /d/ is described as dental. Acoustically, English and French /d/ differ in burst spectra and amplitude measures, although there is a substantial overlap on these measures in the productions of monolingual (Sundara, 2005) as well as bilingual speakers (Sundara, Polka, & Baum, 2006; Sundara, Polka, & Genesee, 2006). Furthermore, the difference between English and French /d/ is not used to contrast meaning in

[2] Throughout this paper French refers to Canadian French.

either language. Dental-alveolar differences for coronal stops are thought to contrast meaning in a few Australian and American languages (Busby, 1980; Dart, 1991; Dixon, 1980).

In sum, differences in place of articulation for coronal stops are subtle, highly marked, and perceptible to human listeners. Crucially, in both English and French, coronal stops are some of the most frequent segments occurring about 12–15% in spoken English (Denes, 1963; French, Carter, & Koenig, 1930; Tobias, 1959) and 9% in spoken French (Malécot, 1974). Sundara et al., (2008) report that when tested on the discrimination of English and French /d/, bilingual infants do not show a U-shaped developmental curve (similar results have been reported for English and French /p-b/ by Burns, Yoshida, Hill, & Werker, 2007). In fact, bilingual infants keep pace with their monolingual peers.

In summary, French vs. English /d/ are very similar phones; given Bosch and Sebastián-Gallés's (2003a) account that statistics are calculated over the two languages together, bilingual infants should show a U-shaped developmental curve for the discrimination of the dental-alveolar stop contrast. However, they do not do so. In Sundara et al., (2008) we have argued that this is due to the high frequency of input of these phones. The absence of a U-shaped developmental curve for discrimination of highly frequent phones by bilingual infants supports the idea for a "brute force" approach. Together, data from Catalan and Spanish, and French and English-learning infants supports a strategy that bilingual infants are able to overcome the challenge of overlapping distributions when they accumulate enough input.

1.3.2. *Strategy II: tagging or sorting input*

The other strategy bilingual infants could use is one of tagging or sorting input in order to separate the two languages. Once separated, statistical learning mechanisms can be applied to each of the two languages, allowing bilingual infants to circumvent the problem of overlapping distributions altogether. There is indirect evidence that it is feasible for infants to tag or sort the languages in their input.

To be able to tag or sort languages, it is essential that infants discriminate between languages. Numerous studies indicate that infants' ability to discriminate languages is mediated by the typological distance between languages (e.g., Nazzi, Bertoncini, & Mehler, 1998). One way of classifying languages based on typological distance is by using a metric of speech rhythm. Based on the acoustic-phonetic characteristics of languages, more specifically, the duration and the variability in the duration of consonant and vowels attested in a language, there is evidence that languages cluster into three groups (Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999). Without being completely categorical, these three groups correspond to what has been traditionally delineated as rhythm classes (Abercrombie, 1967; Pike, 1946). Languages have been classified into one of three rhythm classes: stress-timed (e.g., English, German, Dutch); syllable-timed (e.g., Spanish, French, Italian); and mora-timed (e.g., Japanese, Telugu).

There is plenty of evidence demonstrating that infants can discriminate between languages from different rhythm classes from birth. Regardless of experience with specific languages, newborns are able to discriminate between languages from different rhythm classes (Byers-Heinlein, Burns, & Werker, 2010; Mehler et al., 1988; Nazzi et al., 1998; Ramus, Hauser, Miller, Morris, & Mehler, 2000). For example, French as well as English-learning newborns can discriminate between English and Spanish (Moon, Cooper, & Fifer, 1993; Nazzi et al., 1998). Besides human newborns, cotton-top Tamarin monkeys (Ramus et al., 2000) as well as rats (Toro, Trobalón, & Sebastián-Gallés, 2003) are also able to distinguish between languages from different rhythm classes. Thus, infants' ability to distinguish languages from different rhythm classes seems to be based on an innate, species-general auditory sensitivity.

However, in order to distinguish two languages within the same rhythm class, infants require experience with at least one of the languages. Further, this ability only emerges by 4–5-months of age. For example, 4–5-month olds, but not newborns, distinguish languages like Catalan vs. Spanish or English vs. Dutch, but only if one is their native language (Bosch & Sebastián-Gallés, 1997; Nazzi, Jusczyk, & Johnson, 2000).

In sum, infants can discriminate two languages from different rhythm classes from birth, whereas learning to discriminate two languages that are from the same rhythm class requires at least 4–5 months of listening experience. Thus, bilingual infants may be expected to be able to tag or sort two languages from different rhythm classes earlier in acquisition when compared to two languages from the same rhythm class. This may allow bilingual infants to circumvent the dip in discrimination observed for infants learning two rhythmically similar languages like Spanish and Catalan.

1.3.3. *The present study*

In this study, we compared discrimination of the English /e – ɛ/ contrast by bilingual Spanish and English-learning infants with monolingual English-learning infants. We tested infants at 4- and 8-months of age. Recall that previous research by Bosch and Sebastián-Gallés (2003a) tested bilingual Catalan and Spanish-learning infants and monolingual Spanish and monolingual Catalan-learning infants on their discrimination of the Catalan /e – ɛ/ contrast. Compared to Spanish that has a 5-vowel system, Catalan and English both have larger vowel inventories. Not counting either schwa or diphthongs, Catalan is reported to have 7 vowels (Recasens & Espinosa, 2006). American English has been described as having between 9–11 vowels (excluding /ə/, r-colored vowels and diphthongs; Ladefoged, 2006; Tobias, 1959).

More specifically, the phonetic instantiation and distribution of the /e – ɛ/ distinction in Catalan and English have several similarities. In both languages the mid-front /e – ɛ/ vowel distinction primarily involves a difference in F1, with additional differences in F2. Further, the frequency of occurrence of this contrast in Catalan and English is also very comparable. In Catalan, [e] and [ɛ] together account for less than 9% of all the vowels (Rafel, 1980). In English, they account for about 11.9% of all the vowels (Tobias, 1959). More specifically, of all consonants and vowels in English [e] occurs 1.94% and [ɛ] occurs 2.92% (Denes, 1963; Tobias, 1959). Thus, in both Catalan and English the /e – ɛ/ distinction has a low frequency of occurrence.

Despite similarities, the /e – ɛ/ distinction is not identical in Catalan and English. This is particularly the case for the vowel /e/: in Catalan it is a monophthong whereas in most dialects of American English it is a diphthong, typically transcribed as [eɪ].

Crucially for the purpose of this study, Catalan and Spanish are rhythmically very similar; whereas Spanish and English belong to different rhythm classes (Ramus et al., 2000). In fact, there is evidence that newborns (Moon et al., 1993) as well as 5-month olds (Bahrick & Pickens, 1988) are able to distinguish between Spanish and English. In contrast, 4-month olds are able to discriminate between Catalan and Spanish but only if they have experience with one of the languages (Bosch & Sebastián-Gallés, 1997).

In this study we tested whether infants learning two rhythmically different languages like Spanish and English would be better able to deal with the overlapping distribution of vowels in the two languages. If the rhythmic differences between Spanish and English do not help bilingual infants, we expect the same results as reported by Bosch and Sebastián-Gallés (2003a). Like bilingual Catalan and Spanish-learning 8-month olds, bilingual

Table 1
Subject sample.

Measures	4-month olds		8-month olds	
	Monolingual	Bilingual	Monolingual	Bilingual
Number of subjects (n)	20	20	20	20
Girls, boys	10, 10	10, 10	9, 11	7, 13
Additional subjects tested	9	9	10	5
Attrition (%)	31	31	33	20
Mean age (days)	128	126	238	244
Range (days)	108–138	109–142	225–259	227–271

Spanish and English-learning 8-month olds should differ from their monolingual peers. Specifically, they should fail to distinguish English /e – ɛ/. However, if being able to discriminate between the rhythmically different Spanish and English from birth helps the bilingual infant, there should be no difference between Spanish and English-learning bilingual 8-month olds and their monolingual peers. Specifically, both groups should successfully distinguish English /e – ɛ/.

2. Materials and methods

2.1. Subjects

Eighty infants were included in the final sample (20 infants/group). An additional 33 infants were tested, but excluded from analyses for fussiness (8), failure to habituate (5), failure to dishabituate to the post-test trial (16, 14 of these infants were 4-month olds)[3] or experimenter error (4). Table 1 shows information on the subject sample for each of these age and language groups.

All infants were full-term with no known health problems. Parental reports of infant language exposure were collected using a detailed questionnaire (previously used in Bosch & Sebastián-Gallés, 2001; Sundara et al., 2008). In this language questionnaire, parents were asked to identify (a) the people who regularly interact with the infant (e.g., parents, grandparents, child care workers); (b) the number of hours each of these people spent with the infant during the week and during the weekends; (c) and to describe the language(s) each of these people addressed to the infant. Based on these numbers and parent assessment, weekly percentages were calculated to sort infants into appropriate groups. Infants with more than 10% of their language input coming from a language other than English or Spanish were excluded.

To be classified as monolingual, subjects had to have at least 90% of their total language input in English; in practice, all infants classified as monolingual had 100% of their input as English-only. To be classified as bilingual, infants' daily exposure to Spanish and English could range from 50–50% to 80–20%. In Sundara et al. (2008), infants with 70–30% or more were classified as bilingual.[4] We reran all our analyses excluding the 3 bilingual infants that had 80% and 20% exposure to Spanish and English respectively (one 4-month old and two 8-month olds). The pattern of results was identical to that reported here with the 80–20% criterion. To increase the number of subjects in the bilingual group, we opted for the more liberal 80–20% criterion. The details of language exposure for infants in the bilingual group are presented in Table 2. For the majority of infants, Spanish was the maternal native language, and thus, the exposure to Spanish was at home.

2.2. Stimuli

The stimuli were comprised of multiple tokens of the English phonemes /e/ and /ɛ/ from 8 female English speakers and were a subset of the Hillenbrand vowel database (Hillenbrand, Getty, Clark, & Wheeler, 1995), which contains productions of the vowels /i, ɪ, e, ɛ, æ, a, ɔ, o, ʊ, u, ʌ, ɝ/ embedded in /hVd/ syllables from 45 men, 48 women and 46 children. Only the stimuli from the Hillenbrand corpus that were correctly identified 100% of the time by 20 undergraduate and graduate students who had not participated in the recording were included in the final set.

Acoustic analyses of the stimuli are presented in Table 3. The tokens used in the present study were chosen to minimize differences in vowel duration between /e/ and /ɛ/. This was done to match our stimuli as closely as possible to the Catalan vowels without compromising the F1 and F2 differences in American English. Further, we minimized the excursion of F1 for the /e/ stimuli by selecting tokens with the smallest difference in F1 values at 20% and 80% of vowel duration. The resulting set had an F1 excursion of 13.6% (approximately 74 Hz).

Fig. 1 shows the F1–F2 values at vowel steady state for the individual tokens of the /e/ and /ɛ/ stimuli in Hz. From Table 3 and Fig. 1 it is evident that the English /e/ and /ɛ/ tokens have non-overlapping distributions of F1, and also for the most part for F2. Although there was an overlap in the duration of English /e/ and /ɛ/, there was a significant difference between the two. Thus, the acoustic analyses confirmed that the English /e/ and /ɛ/ tokens in this study could be distinguished based on F1, F2 or duration or some combination of these cues.

In Table 4 we compare the /e/ and /ɛ/ tokens produced by 8 female English speakers used in this study to the /e/ and /ɛ/ tokens produced by 5 female Catalan speakers used by Bosch and Sebastián-Gallés (2003a). Note that in this study we used monosyllabic words (CVC) because most English spoken words are monosyllabic (Cutler & Carter, 1987). In contrast, in keeping with the typical length of Catalan words, the Catalan stimuli were disyllables (CVCV). Overall, the English stimuli have higher F1 and lower F2 values indicating that they are produced lower and further back than their Catalan counterparts. Further, the durations of the English vowels in this study appear to be shorter than those used by Bosch and Sebastián-Gallés (2003a).

2.3. Procedure

Infants were tested using the infant-controlled visual habituation procedure (Werker et al., 1998) presented using Habit X

[3] The number of infants who did not habituate in 25 trials is quite high in this experiment. We think this is because we used stimuli produced by four different talkers in the habituation phase. This variability in talkers is likely to be of great interest, particularly to the younger infants, resulting in their failure to dishabituate.

[4] In previous investigations of the development of phonetic perception, word learning and vocabulary acquisition in bilingual infants and toddlers, the classification criteria for bilinguals ranges from 79 to 21% (Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009), through 75–25% (Fennell, Byers-Heinlein, & Werker, 2007), up to 65–35% (Bosch and Sebastián-Gallés, 2001, 2003a, 2003b, 2005; Sebastián-Gallés & Bosch, 2009). The only study that systematically looks at the effect of percent input on vocabulary development in 8–30-month-old bilingual infants shows that the vocabulary learning in both languages seems to be correlated with proportion input in that language (Pearson, Fernández, Lewedeg, & Oller, 1997). Thus, the demarcation between monolingual and bilingual language input appears not to be categorical.

Table 2
Summary of language exposure profiles for bilingual 4- and 8-month olds.

Measures	4-month olds ($n=20$)		8-month olds ($n=20$)	
	English	Spanish	English	Spanish
Percentage exposure {mean (range)}	56.75 (20–77)	43.25 (23–80)	48.75 (20–80)	51.25 (20–80)
Maternal language {count (%)}[a]	18 (90%)	14 (70%)	18 (90%)	15 (75%)
Location of Spanish exposure {count (%)}[a]				
Home		18 (90%)		18 (90%)
Daycare		4 (20%)		4 (20%)

[a] Counts may reflect subjects reporting membership in more than one group (e.g. maternal language is both English and Spanish or child hears Spanish both at home and at daycare).

Table 3
Stimulus characteristics.

Cues	Mean (range)		Cue status	Difference between means
	/e/	/ɛ/		
Correct identification (%)	100	100	ns	
Duration (ms)	314 (231–399)	271 (177–386)	+	43.375
F0 (Hz)	217 (203–234)	209 (173–223)	ns	8.375
Formants (Hz)				
F1				
Steady state	541 (442–645)	720 (680–794)	*	−178.75
20% of vowel duration	541 (442–643)	738 (673–813)	*	−196.75
50% of vowel duration	473 (435–573)	719 (683–801)	*	−245.125
80% of vowel duration	462 (394–501)	657 (620–711)	*	−195.125
F2				
Steady state	2457 (2237–2598)	1996 (1851–2256)	+	461.25
20% of vowel duration	2460 (2232–2603)	1995 (1888–2103)	*	465.125
50% of vowel duration	2541 (2293–2773)	1993 (1823–2216)	*	548.75
80% of vowel duration	2627 (2423–2813)	2034 (1866–2255)	*	593

An asterisk (*) denotes that the two distributions are completely non-overlapping. A plus (+) indicates overlapping distributions, but still significantly different based on paired t-tests ($p<0.001$). Measures with "ns" are not significantly different.

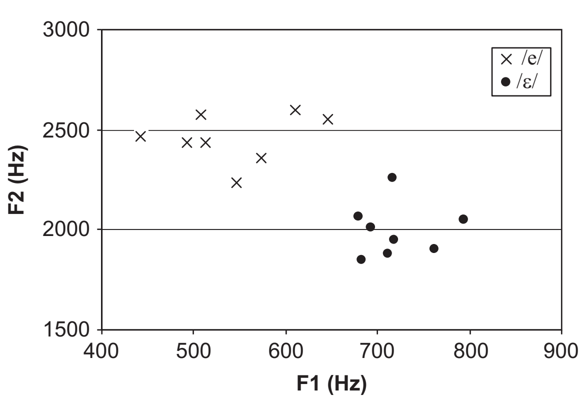

Fig. 1. Steady state first (F1) and second (F2) formant values for the English /e-ɛ/ stimuli.

Table 4
A comparison of the acoustic characteristics of English and Catalan /e-ɛ/ tokens.

Cues	/e/		/ɛ/	
	English	Catalan	English	Catalan
Average F1 at vowel midpoint (Hz)	541	442	719	675
Average F2 at vowel midpoint (Hz)	2457	2714	1996	2524
Average duration[a] (ms)	314	477	270	441

[a] The duration measures for the Catalan stimuli are for the initial syllable (CV), not just the vowel. Also note that the Catalan stimuli were disyllabic.

(Cohen, Atkinson, & Chaput, 2004). The infant sat on a caregiver's lap at a distance of 3.5 feet facing a 46″ television screen. Part of the screen was occluded by black curtains so that the visual stimuli were only presented on a 24.5′ × 21′ display. Curtains covered all of the walls and the other equipment, including JBL speakers and a Sony digital video camera below the screen. An experimenter observed the infant from outside the testing room on a screen connected to the video camera. The experimenter and the caregiver both wore Peltor headphones and listened to music to avoid influencing the infant's behavior. The infant's attention was drawn to the screen at the beginning of each trial with a flashing red light. When the infant oriented to the screen, a checkerboard image was shown and the accompanying audio stimulus was presented over the speakers. The experimenter recorded the infant's visual fixations to the screen as a measure of infant listening time. Trial durations were infant-controlled and stopped when the infant looked away for more than 1 s or at the end of the trial (16 s). A minimum look time of 2 s was required for a trial to be counted. At the end of each trial, the red flashing light appeared to re-orient the infant to the screen to begin a new trial.

2.4. Design

The experiment consisted of four phases: pre-test, habituation, test, and post-test. The pre- and post-tests were each only one trial in length and were included to monitor infant engagement with the task. These trials were composed of an animated female voice repeating the syllable 'pok'. In the habituation phase, half of the infants in each age group heard repeated tokens of the syllables with /e/ from four talkers and the other half heard syllables with /ɛ/ from four talkers. The habituation phase continued until the looking time for the last three consecutive trials dropped below 50% of the looking time for the longest three consecutive trials. Only infants who were habituated in 25 or fewer trials were included in the analysis.

After the habituation phase, two trials were presented in the test phase in counterbalanced order. Half the subjects heard the familiar vowel (vowel presented during habituation) first and the other half heard the novel vowel first. Test stimuli were each produced by two new talkers (total 4 new talkers). Infants' listening time to the familiar and novel vowel in the test phase were compared. To demonstrate discrimination infants had to ignore the novelty of the new talkers in the test phase and listen significantly longer to the novel vowel when compared to the familiar vowel. Data from infants that did not listen longer to the post-test trial compared to the average of the last three habituation trials were excluded from the analysis. This was done to ensure that when children failed the discrimination task, this was not due to an overall disengagement with the task.

3. Results

All four groups of infants were comparable on pre- and post-test listening times (Table 5). A two-way ANOVA, with age (2 levels: 4- and 8-months) and language group (2 levels: monolingual and bilingual) as the between-subjects independent variables, and pre-test listening time as the dependent variable showed no significant differences. There was no significant main effect of age, $F(1, 76)=0.58$, $p=0.45$, or language group, $F(1, 76)=1.4$, $p=0.24$. There was also no significant interaction of age and language, $F(1, 76)=0.07$, $p=0.8$. Similarly, there were no significant differences when post-test listening time was the dependent variable. The main effects of age, $F(1, 76)=0.1$, $p=0.76$, and language group, $F(1, 76)=0.60$, $p=0.44$, were not significant. And there was no significant interaction of age and language, $F(1, 76)=0.06$, $p=0.81$.

To determine whether infants in the four groups differed in their interest in the stimuli, total habituation times were compared using a two-way ANOVA, with age (2 levels: 4- and 8-months) and language group (2 levels: monolingual and bilingual) as the independent variables. There was no significant main effect of age, $F(1, 76)=0.33$, $p=0.57$, or language group, $F(1, 76)=1.14$, $p=0.29$. There was also no significant interaction of age and language group, $F(1, 76)=2.07$, $p=0.16$. A two-way ANOVA, with number of trials to habituation as the dependent variable also showed the same pattern. Finally, there was no significant main effect of age, $F(1, 76)=0.001$, $p=0.98$, or language group, $F(1, 76)=3.27$, $p=0.07$, nor was the interaction of age and language group significant, $F(1, 76)=0.1$, $p=0.75$, when percent habituation levels were used as the dependent variable. Thus, the four groups were comparable on their engagement with the task at the beginning and at the end of testing, as well as their interest and engagement with the experimental stimuli.

Average listening time to familiar and novel stimuli for each of the four groups is presented in Fig. 2. A three-way ANOVA with age (2 levels: 4- and 8-months) and language group (2 levels: monolingual and bilingual) as the between-subjects variables and trial-type (familiar and novel vowel) as the within-subjects variable, and listening time as the dependent variable was used to compare the four groups.[5] There was no significant main effect of age, $F(1, 76)=1.35$, $p=0.25$, partial $\eta^2=0.02$, or language group, $F(1, 76)=0.22$, $p=0.64$, partial $\eta^2=0.004$. There was no significant three-way interaction of age, language group and trial type, $F(1, 76)=0.01$, $p=0.92$, partial $\eta^2=0$. None of the two-way interactions were significant either – age × language group, $F(1, 76)=0.001$, $p=0.98$, partial $\eta^2=0$; age X trial-type, $F(1, 76)=0.78$, $p=0.38$, partial $\eta^2=0.01$, or language group × trial-type, $F(1, 76)=0.45$, $p=0.51$, partial $\eta^2=0.006$. The only significant difference was the main effect of trial-type, $F(1, 76)=18.1$, $p<0.001$, partial $\eta^2=0.19$. Thus, ANOVA results showed that overall, infants in all groups listened longer to the novel vowel.

Despite the lack of either a significant main effect of language group or any interaction of trial-type or age with language group, for each group we compared the listening time to novel and familiar vowel using one-tailed paired t-tests. At 4 months, 10 out of 20 monolingual infants and 13 out of 20 bilingual infants listened longer to the novel compared to the familiar vowel. There was no significant difference between the listening time to the novel and familiar vowel for infants in either the monolingual group or bilingual group. However, when results from monolingual and bilingual 4-month olds were analyzed together, infants listened significantly longer to the novel (*Mean*=7.20 s, *SD*=4.65) compared to the familiar (*Mean*=5.63 s, *SD*=3.71) vowel, $t(39)=2.1$, $p=0.02$, Cohen's $d=0.37$.

At 8 months, 15 out of 20 monolingual and 14 out of 20 bilingual infants listened longer to the novel compared to the familiar vowel. For both groups, the difference in listening time to novel and familiar vowels was significant. Monolingual English-learning infants listened significantly longer to the novel compared to the familiar vowel, $t(19)=3.8$, $p<0.001$, Cohen's $d=0.99$. A comparison of listening time to the novel and familiar vowel for bilingual Spanish and English-learning 8-month olds was also significant, $t(19)=2.09$, $p=0.02$, Cohen's $d=0.63$. Based on Cohen's *d*, the effect sizes for the monolingual as well as bilingual 8-month olds are large. In contrast, the effect size for the 4-month olds, even when combined across the monolingual and bilingual groups, is only of medium strength.

Table 5
Average listening times to habituation, pre-, post-test stimuli (seconds), number of trials to habituation and percent habituation for all four age and language groups.

Variables	4-month olds		8-month olds	
	Monolingual	Bilingual	Monolingual	Bilingual
Pre-test listening time (s)	13.2	11.9	13.6	12.8
Post-test listening time (s)	10.3	9.7	10.9	9.8
Total habituation time (s)	101	96	89.7	123
Average trials to habituation	11.8	11.2	10.4	13.1
Percent habituation[a]	59	61	58	61

[a] Percent habituation was defined as 100 × (1 − (sum of looking time for the last 3 trials in the habituation phase/sum of the looking time for the longest three consecutive trials in the habituation phase)).

4. Discussion

In this study, we compared monolingual English and bilingual Spanish and English-learning 4- and 8-month olds' ability to

[5] Polka and Bohn (2011) used the conditioned headturn procedure to test Danish learning 6- to 12-month-olds on their discrimination of two mid-front vowels differing in F1 with somewhat similar quality to the English /e- ɛ/ distinction. They report a directional asymmetry such that discrimination was harder when /e/ was the foreground. This asymmetry was observed in the younger infants, i.e. infants with ages between 6- and 9-months. Keeping this in mind, we replicated the ANOVA with habituation condition (/e/ vs. /ɛ/) as an additional variable. There was no significant main effect or interaction involving the habituation condition. Hence, this variable has been excluded from all other analyses.

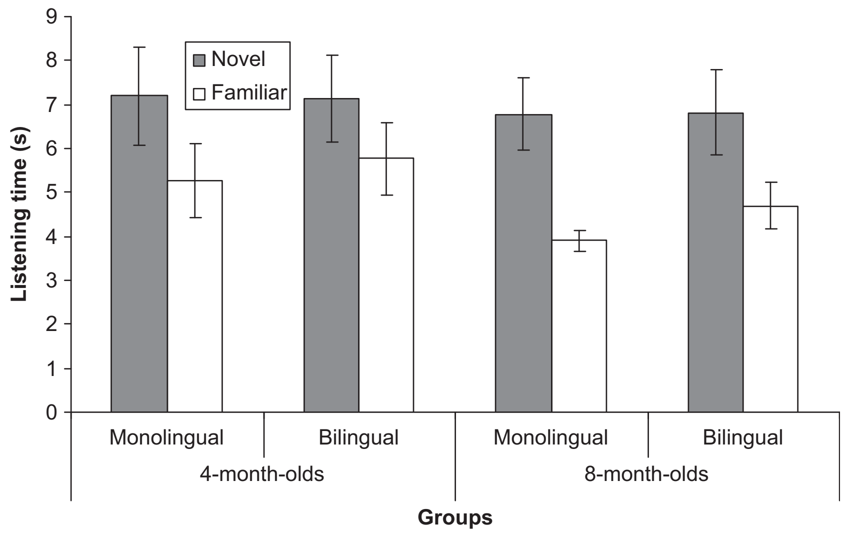

Fig. 2. Average listening time (seconds, ± *SE*) to the familiar and novel vowel in the test phase for monolingual and bilingual groups. The listening times to novel and familiar vowels are significantly different for the monolingual and bilingual 8-month olds.

discriminate English /e – ɛ/. The difference between the vowels /e/ and /ɛ/ is contrastive in English, whereas Spanish has only one mid-front vowel, /e/. Our results indicate that when combined into a single group, monolingual and bilingual 4-month olds were able to discriminate English /e – ɛ/. Of critical importance are the results from 8-month olds. We expected to replicate the results from Bosch and Sebastián-Gallés (2003a), if rhythmic differences between two languages did not help the bilingual infant to track statistics of the two languages separately. Given their findings that Catalan and Spanish-learning bilingual 8-month olds fail to discriminate Catalan /e – ɛ/, we might have expected the bilingual Spanish and English-learning 8-month olds tested in this study to fail as well. In other words, at 8 months, we might have expected bilingual and monolingual infants to differ in their discrimination performance. Instead, we found no difference in the discrimination performance of bilingual and monolingual 8-month olds. Both groups of infants were able to discriminate the English /e – ɛ/ contrast. Thus, these results show that a rhythmic difference between languages affects the trajectory of development of speech perception in bilingual infants.

Interestingly, our results also differ from that of Bosch and Sebastián-Gallés (2003a) for the 4-month olds. Recall that Bosch and Sebastián-Gallés report that 4-month olds learning Catalan, Spanish or both were able to discriminate Catalan /e – ɛ/. In our study, the monolingual and bilingual 4-month-old groups, separately, failed to discriminate English /e – ɛ/. On the surface this seems counterintuitive. Recall that unlike in Catalan where the /e – ɛ/ distinction was a difference primarily in F1 and also F2 to some extent, in English, not only did the vowels differ in F1 and F2, there were additional differences between the vowels in duration. Specifically, English /e/ is typically a diphthong and longer than /ɛ/. So, it may be argued that English /e – ɛ/ tokens are more different than the Catalan /e – ɛ/ and thus, easier to discriminate.

There are two possible reasons for a difference in the results from 4-month olds in the two studies. First, the vowel stimuli used in our experiment were at least a third (if not more) shorter that the vowel stimuli used by Bosch and Sebastián-Gallés (Table 4). Shorter stimuli are likely to be harder to discriminate, perhaps even more so in early stages of development. Second, it is possible that the vowel discrimination task in our study was harder than the one used by Bosch and Sebastián-Gallés. Bosch and Sebastián-Gallés used stimuli produced by 5 talkers, in this study we used stimuli produced by 8 different talkers. Further, Bosch and Sebastián-Gallés used the same talkers in the familiarization and test phase. In contrast, in our study, infants were familiarized with vowel stimuli produced by four talkers, then presented with four new talkers in the test phase. Introducing more talkers into a discrimination task introduces irrelevant variability that needs to be ignored by infants. In fact, in our task infants had to ignore all the four new talkers in the test phase to listen longer to the novel vowel. This is likely to have made the discrimination of vowels harder in our study.

There is some corroboration that vowel contrasts similar in quality to the English /e – ɛ/ distinction are quite difficult, even for older infants. In this issue, Polka and Bohn used the conditioned headturn procedure to test Danish-learning 6- to 12-month olds on their discrimination of two Danish mid-front vowels differing in F1 with somewhat similar quality to the English /e – ɛ/ distinction. The conditioned headturn procedure is a very robust procedure and a version of it is used to test hearing thresholds in infants. Specifically, unlike visual habituation where each infant gets only one (or two) test trial(s), in conditioned headturn each infant can be tested on multiple test trials to determine whether one particular infant can reliably discriminate tokens (Werker, Polka, & Pegg, 1997). Polka and Bohn report that the percent correct score for infants between 6- and 9-months, the ages most comparable to this study, was 55.6%, and for infants between 9- and 12-months was 59.4%. Thus, subtle differences in mid-front vowels are quite difficult for Danish-learning infants as well.

In this study, monolingual and bilingual infants performed similarly. Both monolingual and bilingual 4-month olds failed to discriminate English /e – ɛ/. Recently, Narayan et al. (2010), present another example where 4-month olds ($n=16$) failed to discriminate a subtle phonetic distinction viz., the Tagalog /na – ŋa/ place difference.[6] However, in our study, increasing the sample size (up to 40) by assessing monolingual and bilingual together, showed that 4-month olds are indeed able to discriminate the English /e – ɛ/. Clearly, although difficult, the discrimination

[6] We calculated the effect size for discrimination of the /na – ŋa/ place difference by 4-month-olds tested in the Narayan et al. (2010) study using the mean looking times of 4.44 and 4.17 s and estimated pooled standard deviation (based on the standard error bars in Fig. 4) of 2.56 s. The effect size, Cohen's $d=0.1$, was small. In contrast, in the present study, the effect size for 4-month-olds' discrimination of the English /e – ɛ/ distinction is of medium strength.

of the English /e – ɛ/ distinction is not as difficult as the Tagalog /na – ŋa/ distinction. At 8 months as well, there was no difference between monolingual and bilingual infants. With increasing language experience, the bilingual and monolingual infants' performance was facilitated such that 8-month olds in both groups were able to discriminate English /e – ɛ/.

Of course besides spectral differences, English /e – ɛ/ also differ in duration. It can be argued that this additional cue might have helped 8-month-old bilingual infants. Unfortunately, this is a limitation of using natural language stimuli. We picked the English /e – ɛ/ contrast for its remarkable match in phonetic instantiation and frequency in the input to the Catalan /e – ɛ/ contrast. In both languages, the pair differs primarily in F1 (and also F2), and additionally, is quite infrequent. Although editing the tokens such that the duration differences are neutralized was an option available, we did not do so for two reasons. First, it would have reduced the ecological validity of the stimuli. Second, recall that, as the data from 4-month olds indicates, this is a subtle difference that was hard to discriminate to start with. Neutralizing duration would likely have made the task even harder not just for the bilinguals but also for the monolinguals. For these reasons we chose to retain the duration differences in the natural stimuli. We present the results of this experiment as *proof of concept* that language-tagging can indeed help infants to calculate statistics of units for each of the languages separately.

In light of the results from this study, we can reconsider the evidence presented in Section 1.3.1 above. Recall that French and English learning bilingual infants did not differ from their monolingual peers in discrimination of phonetic contrasts with overlapping distributions in the two languages (Burns et al., 2007; Sundara et al., 2008). In Section 1.3.1, this was attributed to the high frequency of occurrence of the constituent segments. Alternately, perhaps the bilingual infants tested by Burns et al., and Sundara et al., were helped by the fact that French and English are from rhythmically different classes, rather than the high frequency of the individual segments. Thus, it remains to be determined whether the frequency in the input, cumulative over age or otherwise, can modulate the developmental trajectory of bilingual infants in the absence of rhythmic differences between the languages.

One way to completely control the exact phonetic instantiation as well as the frequency of occurrence of segments is to use simulated bilingual environments as has been done by Weiss and colleagues (Mitchel & Weiss, 2010; Weiss, Gerfen, & Mitchel, 2009). In several experiments with adults, Weiss and colleagues exposed adults simultaneously to two artificial languages with different statistical cues to word segmentation, thereby simulating bilingual input. Their results indicate that when indexical cues were provided either by presenting the two languages in a male vs. female voice (Weiss et al., 2009), or by a visual display of two people (Mitchel & Weiss, 2010), adults were able to track two sets of statistics simultaneously. The results from the present study indicate that altering the rhythm of the two artificial languages is another potential cue, at least for infants, and perhaps also for adults, enabling them to track two sets of statistics simultaneously. Additionally, simulated bilingual environments would complement studies using natural language stimuli by allowing a more fine-grained control of individual parameters like acoustic instantiation and frequency of occurrence.

Throughout this paper we have characterized the difference between English and Spanish as primarily being one of rhythm. However, English and Spanish also differ in intonation as well as in their segmental inventory. In fact, adult listeners can and do utilize differences in rhythm, intonation or segmental inventories to distinguish between languages, although they rely on each of these cues to different extents (Vicenik & Sundara, 2009a, 2009b). In principle, bilingual infants could use any cues that help them to discriminate between languages to calculate two sets of statistics simultaneously. However, given that infants are sensitive to prosody – the rhythm and intonation of speech - from birth, prosodic cues are more likely to be useful to bilingual infants early in development.

5. Conclusions

In summary, we used the visual habituation procedure to test bilingual Spanish and English-learning and monolingual English learning 4- and 8-month olds' discrimination of the English /e – ɛ/ contrast. Pooled together, bilingual and monolingual 4-month olds were able to discriminate English /e – ɛ/. By 8 months of age, bilingual as well as monolingual infants were able to discriminate English /e – ɛ/. This result is in contrast with the results reported for Catalan and Spanish-learning bilingual infants. When tested on a similar vowel distinction, Catalan and Spanish-learning bilingual 8-month olds, unlike their monolingual Catalan-learning peers, failed to discriminate the Catalan /e – ɛ/ distinction. We argue that the rhythmic similarity of Catalan and Spanish hinders the ability of bilingual infants to track two sets of statistics simultaneously in Catalan and Spanish. In contrast, bilingual infants are able to exploit the attested rhythmic differences between Spanish and English, to calculate two sets of statistics. This enables bilingual Spanish and English-learning infants to keep pace with their monolingual peers.

Acknowledgement

We would like to thank Kristi Hendrickson, Alvaro Luna, Flor Reyes and Mirian Romero for scheduling and testing infants. Parts of the data reported in this paper were presented at the 2nd ASA workshop on "Cross-language speech perception and variations in linguistic experience" Portland, OR. We would like to thank the audience at the workshop for their feedback, and Daniel Weiss for comments on a previous draft of this paper. This research was funded by the UCLA COR Faculty Research Grant, UCLA International Institute Faculty Research Award, and start-up funds to MS.

References

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

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). The statistical basis for speech sound discrimination. *Language and Speech*, *46*, 155–182.

Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Perception*, Vol. 2 (pp. 67–96). New York: Academic Press.

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

Best, C. T., & McRoberts, G. W. (2003). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. *Language and Speech*, *46*, 289–304.

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 (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29–49.

Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*(2–3), 217–243.

Bosch, L., & Sebastián-Gallés, N. (2003b). Language experience and the perception of a voicing contrast in fricatives: Infant and adult data. In: M. J. Solé, D. Recasens, & J. Romero (Eds.), *International Congress of Phonetic Sciences* (pp. 1987–1990), Barcelona.

Bosch, L., & Sebastián-Gallés, N. (2005). Developmental changes in the discrimination of vowel contrasts in bilingual infants. In: J. Cohen, K. McAlister, K. Roltad, & J. MacSwan (Eds.), *Proceedings of the 4th international symposium on bilingualism* (pp. 354–363). Barcelona.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). Bilingual and monolingual infant phonetic development. *Applied Psycholinguistics*, *28*(3), 455–474.

Busby, P. (1980). The distribution of phonemes in Australian aboriginal languages. *Papers in Australian Linguistics*, *4*, 73–139.
Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, *21*(3), 343–348.
Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of Acoustical Society of America*, *54*(2), 421–428.
Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). *Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0)*. Austin: University of Texas.
Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.
Dart, S. N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics*, *79*, 1–155.
Denes, P. B. (1963). On the statistics of spoken English. *Journal of Acoustical Society of America*, *35*(6), 892.
Dixon, R. M. W. (1980). *The languages of Australia*. Cambridge: Cambridge University Press.
Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child Development*, *78*(5), 1510–1525.
French, N. R., Carter, C. W., & Koenig, W. (1930). The word and sounds of telephone conversations. *Bell Systems Technical Journal*, *9*, 290–324.
Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology*, Vol. 7 (pp. 515–546). Berlin: Mouton de Gruyter.
Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of Acoustical Society of America*, *97*(5), 3099–3111.
Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge: Cambridge University Press.
Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, *66*, 1679–1688.
Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, *6*, 263–285.
Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, *100*, 425–2438.
Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13–F21.
Ladefoged, P. (2006). *A course in phonetics* (fifth ed.). New York: Thomas Wadsworth.
Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384–422.
Malécot, A. (1974). Frequency of occurrence of French phonemes and consonant clusters. *Phonetica*, *29*, 158–170.
Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*(3), 241–265.
Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*, 1367–1381.
Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.
Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*(1), 122–134.
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.
Mitchell, A. D., & Weiss, D. J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, *25*(4), 456–482.
Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day olds prefer their native language. *Infant Behavior and Development*, *16*, 495–500.
Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407–420.
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., 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.
Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*, 41–58.
Pike, K. (1946). *The Intonation of American English* (2nd ed.). Ann Arbor: University of Michigan Press.
Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*, 467–478.
Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/ð/ perception: Evidence for a new developmental pattern. *Journal of Acoustical Society of America*, *5*, 2190–2200.
Polka, L., Rvachew, S., & Mattock, K. (2007). Experiential influences on speech perception and production in infancy. In E. Hoff, & M. Shatz (Eds.), *Handbook of language development*. Blackwell Publishing.
Rafel, J. (1980). Dades sobre la freqüència de les unitats fonològiques del català. *Estudis Universitaris Catalans*, *24*, 473–496.
Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, *59*, 96–121.
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., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.
Recasens, D., & Espinosa, A. (2006). Dispersion and variability of Catalan vowels. *Speech Communication*, *48*, 645–666.
Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it?. *Developmental Science 12*(6), 874–887.
Stockwell, R. P., & Kiddle, L. B. (1956). On phonemes and allophones. *Hispania*, *39*(3), 325–327.
Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *Journal of the Acoustical Society of America*, *118*(2).
Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, *9*(1), 97–114.
Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d- ð/ in monolingual and bilingual acquisition of English. *Cognition*, *100*(2), 369–388.
Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, *108*(1), 232–242.
Tobias, J. V. (1959). Relative occurrence of phonemes in American English. *Journal of the Acoustical Society of America*, *31*, 631.
Toro, J. M., Trobalón, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination by rats. *Animal Cognition*, *6*, 131–136.
Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native fricative-affricate contrasts: Cross language tests on adults and infants. *Journal of the Acoustical Society of America*, *120*, 2285–2294.
Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz, & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education*. Clevedon: Multilingual Matters.
Vicenik, C., & Sundara, M. (2009a). Role of segmental and intonational cues in dialect discrimination. *Journal of the Acoustical Society of America*, *123*(5), 3883.
Vicenik, C., & Sundara, M. (2009b). Role of rhythmic and intonational cues in language discrimination. *Journal of the Acoustical Society of America*, *125*(4), 2764.
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.
Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 349–353.
Werker, J. F., Polka, L., & Pegg, J. E. (1997). The conditioned head turn procedure as a method for testing infant speech perception. *Early Development and Parenting*, *6*, 171–178.
Werker, J. F., Shi, R., Desjardins, R., Pegg, J., Polka, L., & Patterson, M. (1998). Three methods for testing infant speech perception. In A. M. Slater (Ed.), *Perceptual development: Visual, auditory, and speech perception in infancy* (pp. 389–420). London: UCL Press.
Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*, 278–286.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.