ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Cognition xxx (2007) xxx–xxx

COGNITION

www.elsevier.com/locate/COGNIT

Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction

Megha Sundara *, Linda Polka

School of Communication Sciences and Disorders, McGill University, USA

Received 6 September 2006; revised 21 January 2007; accepted 22 January 2007

Abstract

The current study was designed to investigate the timing and nature of interaction between the two languages of bilinguals. For this purpose, we compared discrimination of Canadian French and Canadian English coronal stops by simultaneous bilingual, monolingual and advanced early L2 learners of French and English. French /d/ is phonetically described as dental whereas English /d/ is described as alveolar. Using a categorial AXB task, the performance of all four groups was compared to chance and to the performance of native Hindi listeners. Hindi listeners performed well above chance in discriminating French and English /d/-initial syllables. The discrimination performance of advanced early L2 learners, but not simultaneous bilinguals, was consistent with one merged category for coronal stops in the two languages. The data provide evidence for interaction in L2 learners as well as simultaneous bilinguals; however, the nature of the interaction is different in the two groups.

Keywords: Speech perception; Bilingual; Second language; L2; Simultaneous bilingual; English; French

* Corresponding author. Present address: UCLA, Department of Linguistics, 3125 Campbell Hall, Los Angeles, CA 90095-1543, USA.

E-mail address: megha.sundara@humnet.ucla.edu (M. Sundara).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2007.01.011

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

1. Introduction

There is extensive research on the persistent effects of first or native language (L1) on segmental perception by adults (Best & Strange, 1992; Bosch, Costa, & Sebastián-Gallés, 2000; Flege & Eefting, 1987; Pallier, Bosch, & Sebastián-Gallés, 1997a; Sebastián-Gallés & Soto-Faraco, 1999; see Strange, 1995; and Werker, Gilbert, Humpfrey, & Tees, 1981 for reviews). In recent years, L1 effects on segmental perception by adults have also been demonstrated at the neuro-physiological level (Jacquemot, Pallier, LeBihan, Dehaene, & Dupoux, 2003; Winkler et al., 1999). Researchers have suggested that the difficulty of perceiving a non-native contrast is determined by the phonetic similarity of the non-native contrast to L1 categories (Best, 1993; Flege, 1992). Persistent effects of L1 typically result in problems for adult learners of a second language (L2). Specifically, L2 learners are likely to ignore distinctions in L2 that are not meaningful in their L1. This is illustrated in the well-documented difficulty Japanese listeners have in distinguishing English /r/–/l/ (see Guion, Flege, Akahane-Yamada, & Pruitt, 2001 for a review). Consequently, researchers have described the role of L1 as a 'filter', quite often interfering with the perception of differences in L2 that are not meaningful in L1 (Flege, Schirru, & MacKay, 2003; Iverson et al., 2003; Pallier, Christophe, & Mehler, 1997b; Trubetzkoy, 1939).

Although discrimination of difficult L2 contrasts can improve with training, even highly skilled L2 learners seldom achieve the levels of performance seen in native listeners of their L2 (Flege, 1992; Gottfried, 1984; Tees & Werker, 1984). This is illustrated by research on the perceptual abilities of adults who were early L2 learners with exposure to their second language beginning between three and six years of age. Using several perceptual tasks, Sebastián-Gallés and colleagues (Bosch & Sebastián-Gallés, 2003a; Bosch et al., 2000; Sebastián-Gallés & Soto-Faraco, 1999) tested highly proficient Catalan-Spanish speakers differing in their L1, on their perception of synthetic continua including /e–ɛ /, /o–ɔ/, /s–z/ and /ʃ–ʒ/. Each of these distinctions is used to contrast meaning and is, therefore, phonemic in Catalan but not in Spanish. Sebastián-Gallés and colleagues report that Spanish-dominant listeners, individuals who first learned Spanish at home, and learned Catalan before the age of six, differ from Catalan-dominant listeners. Spanish- and Catalan-dominant listeners have been reported to be different in discrimination functions, in the location of prototypes, in their category boundaries, and in the portion of signal required to correctly identify segments. The residual, but persistent effects of L1 on L2 categories in early L2 learners further supports an L1 interference account.

However, the effects of L1 on perception of L2 categories are not unidirectional. Because Sebastián-Gallés and colleagues (Bosch & Sebastián-Gallés, 2003a; Bosch et al., 2000; Sebastián-Gallés & Soto-Faraco, 1999) did not test monolingual controls in any of the studies, it remains unclear whether early L2 learners discriminate contrasts in their L1 like monolingual listeners. Empirical evidence from investigation of category boundaries using synthetic continua indicates that early L2 acquisition alters boundaries of L1 categories such that they are not aligned with those of either group of monolingual listeners (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Hazan & Boulakia, 1993; Williams, 1979; but see Bohn & Flege, 1992). More

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

importantly, although interference accounts capture the pervasive influences of L1 on perception and production of L2 sounds, they are unable to explain why L2 learners are different from monolingual listeners even in their L1.

Although the effects of L1 on perception of L2 segments have been reported in all second language learners regardless of age of exposure, the effects of L2 on perception of L1 segments have only been reported in early L2 learners. According to Flege et al. (2003) an interaction of L1 and L2 is most likely in early L2 learners because their L1 categories are not yet fully formed. Thus, for L2 phones that are similar to L1 phones, early L2 learners form merged categories with properties of L1 and L2 phones. These merged categories, formed due to interactions between L1 and L2, are the hallmark of early L2 learners' perceptual abilities and cause them to be different from monolinguals learning either language. Thus, early L2 learners are different from native listeners of their L2, because of interference from their L1. They are also different from native listeners of their L1, because of the influence of L2 on L1 categories given their early age of acquisition.

But what about learners exposed to two languages from birth? The purpose of this paper is to compare the discrimination of phonetic contrasts by simultaneous bilingual adults, early L2 learners and monolinguals. Although the term bilingual has been used in the literature to refer to learners exposed to two languages either simultaneously or sequentially, in this paper the term bilingual and simultaneous bilingual are used interchangeably. Learners exposed to two languages sequentially are referred to as L2 learners, with age of exposure to the second language specified. To date, there are no data on speech perception by adults exposed to two languages from birth. However, a few investigations of developing bilinguals exposed to two languages from birth provide a somewhat different view of the consequences of dual language exposure on discrimination of phonetic contrasts (Bosch & Sebastián-Gallés, 2003a, 2003b; Burns, Werker, & McVie, 2003; Sundara, Polka, & Baum, 2006a; Watson, 1995).

Bosch and Sebastián-Gallés (2003a, 2003b) report data on discrimination of the Catalan /e/–/ɛ/ contrast by infants exposed to Spanish, Catalan or both Spanish and Catalan. Spanish has a single mid-front vowel /e/ with first and second formant values intermediate to those of the two mid-front Catalan vowels, /e/ and /ɛ/. In this study, Bosch and Sebastián-Gallés familiarized infants with multiple tokens of Catalan /e/ (or /ɛ/) embedded in natural /dVðI/ syllables produced by five female speakers. Subsequently, infants were presented with new tokens of the vowel presented during familiarization, and tokens of the vowel not previously heard. Bosch and Sebastián-Gallés (2003a) report that all 4-month-olds listened longer to the new vowel when compared to the familiar vowel, confirming young infants' ability to discriminate the vowel contrast, regardless of language experience. However, at 8-months, only Catalan-learning babies retained the ability to discriminate the Catalan vowel contrast. The authors then demonstrated that a group of 12- but not 8-month-old bilingual babies were able to discriminate the Catalan /e/–/ɛ/ contrast. Thus, bilingual infants show a U-shaped developmental pattern, with a decline in discrimination of the Catalan /e/–/ɛ/ contrast at 8-months, and its subsequent recovery at 12-months. A similar pattern has also been demonstrated by Bosch and

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

Sebastián-Gallés (2003b) for discrimination of a fricative voicing distinction (/s/–/z/ contrast) that is phonemic in Catalan but not Spanish. Bilingual babies evidence a decline in their ability to discriminate the /s/–/z/ distinction at 12-months; by 16-months, they are again able to discriminate Catalan /s/–/z/. Bosch and Sebastián-Gallés's data indicate that simultaneous exposure to two languages results in (a) a temporary decline in discrimination of phonetic contrasts (b) such that language-specific phonetic discrimination is observed later in bilingual than monolingual infants.

Preliminary data on the discrimination of voice-onset-time (VOT) differences underlying voicing distinctions also seem to support that bilingual and monolingual infants differ in their rate of development (Burns et al., 2003). Burns et al. (2003) tested 6–8 and 10- to 12-month-old monolingual English-learning and bilingual infants learning English and French on discrimination of a bilabial stop voicing distinction cued by VOT. Both English and French have a two-way voicing distinction, however, they differ in the VOT distinctions used to signal voicing. In English, voiced and voiceless stops contrast in short-lag and long-lag VOT, whereas in French they contrast in lead and short-lag VOT (Caramazza et al., 1973). Burns et al. (2003) habituated infants to tokens with ~~a short-lag~~ VOT, and then presented them with tokens having ~~long-lag and lead~~ VOT. They report that both monolingual English and bilingual English and French-learning 6- to 8-month-olds listened longer to tokens having ~~lead~~, but not ~~long-lag VOT~~. This recovery from habituation indicates that all babies were able to discriminate the French VOT difference (~~lead~~ versus ~~short-lag VOT~~) but not the English VOT difference (~~short~~ versus ~~long-lag VOT~~). By 10- to 12-months of age, as expected, monolingual infants exposed to English were able to discriminate only ~~short~~ and ~~long-lag VOT~~, the contrast that is present in English. However, 10- to 12-month-old bilingual babies were unable to discriminate either of the two differences. The authors then tested three groups of 14- to 21-month-old monolingual English, monolingual French, and bilingual infants. Their results indicate that by this age, monolingual English babies were discriminating only the English contrast, whereas the monolingual French babies discriminated only the French contrast. Overall, the bilingual infants showed a mixed pattern and, as a group, were not comparable to their monolingual peers.

Bosch and Sebastián-Gallés (2003a, 2003b) and Burns et al. (2003) report similar patterns in that in each study bilingual infants did not demonstrate language-specific phonetic perception at the same age as the monolingual infants. Further, across these studies bilingual infants demonstrate a decline in their ability to discriminate a contrast that they had been sensitive to earlier in development, similar to non-native perceivers. These studies also differ in at least two ways. First, because Burns et al. did not test monolingual French 6–8 or 10- to 12-month-olds, they do not present data from a complete cross-language design. Second, as a group, the oldest bilingual infants tested by Burns et al. do not perform like monolingual infants. Consequently, Burns et al.'s data do not provide strong evidence for a U-shaped developmental pattern in bilingual infants.

One possible explanation for Burns et al's results relates to the differential status of VOT as a cue to voicing in English and French. In the production of stop consonants, VOT is just one cue for voicing in English and French (Sundara, 2005;

Sundara et al., 2006a), and quite likely not a primary cue to voicing differences in French (Caramazza et al., 1973). Perception data from a developmental study conducted by Watson (1995) supports the idea that VOT may not play the same role to signal voicing differences in English and French.

Watson (1995) assessed category boundaries for perception of a synthetic continuum in which VOT was varied from −30 to +50 ms in 5 ms steps; he tested 6, 8 and 10-year-old monolingual and bilingual children learning English and French. The continua were embedded in each of two carrier phrases, one in English ("I say…"), and one in French ("Je dis.…") and the children were asked to identify the initial consonant. All children were resident in Paris or London. The bilingual children had been exposed to both English and French at home from birth, and were being educated bilingually. The monolingual children had no knowledge of languages other than their native language. Watson (1995) reports three findings. First, at every age, the bilingual children had a significantly different category boundary in English and French, indicating that their perception of VOT is language-specific. Second, in most cases, boundaries of bilingual children in the two languages were significantly different from those of monolingual children. In Watson's data, the monolingual and bilingual children did not converge on the same developmental pattern even by 10 years of age (although such convergence may emerge in adolescence or adulthood). The difference between monolingual and bilingual children confirms the mutual effects of the two languages in the organization of bilingual perception. Finally, and crucially, for both monolingual and bilingual children, the identification functions were less sharply categorical in French than in English. These differences in the identification functions across the two languages provide evidence that, at least during development, native listeners rely on VOT to a lesser extent in French than in English.

To summarize, research comparing perception by bilingual and monolingual infants indicates that some language-specific categories emerge later in development in infants learning two languages simultaneously from birth. More recently, Sundara, Polka, and Genesee (2006~~b~~) report similar findings in 4-year-olds. They tested monolingual English, monolingual French and bilingual English and French-learning 4-year-olds and bilingual adults on their ability to distinguish /d–ð/, a contrast that is phonemic in English but not in French. English 4-year-olds' discrimination of /d–ð/ was better than English-learning infants' but not as good as English adults, indicating that language experience serves to facilitate discrimination of contrasts. However, bilingual 4-year-olds were no better than monolingual French-learning 4-year-olds, with both groups performing significantly worse than their same-age English-learning peers. Given that bilingual adults were comparable in performance to monolingual English adults, the poor performance of bilingual 4-year-olds indicates that the facilitative effects of language experience are observed later in children learning two languages from birth.

Why do bilingual infants show a decline in discrimination of language-specific native phonetic contrasts? Bosch and Sebastián-Gallés (2003a) attribute the inability of bilingual 8-month-olds to discriminate Catalan /e/–/ɛ/ to their having a single extended mid-front vowel category that includes Catalan /e/, Spanish /e/, and Catalan /ɛ/. The authors suggest that this may be due to the close proximity of the three vowels in acoustic space, and the relatively greater frequency of Spanish /e/ when compared to the two Catalan vowels. They posit that initially, infants track statistical regularities across the two languages, resulting in a shared perceptual space common to the two languages. The subsequent gain in discrimination by bilingual 12-month-olds, they suggest, may result either as a consequence of early word learning, or be due to increased exposure necessary to resolve the extended category into smaller sub-categories.

Although the decline in discrimination of the Catalan /e/ - /ɛ/ contrast by 8-month-old bilingual babies with dual language exposure is indeed consistent with a shared perceptual space common to both languages, Bosch and Sebastián-Gallés' data do not provide any direct evidence for a shared perceptual space. Further, the authors suggest a single shared category for Catalan /e/, Spanish /e/, and Catalan /ɛ/ to be a *temporary* consequence of an immature system that is eventually resolved. However, they are silent as to whether by adulthood, the two languages of the simultaneous bilingual share categories for similar phones, or whether eventually the perceptual space is resolved to form two sets of categories, one for each language. This is the key distinction to predicting perceptual abilities of adults who have been exposed to two languages from birth. The former would predict merged categories for similar phones like that expected for early L2 learners, indicating that merged categories can be routinely expected as a consequence of exposure to two languages. The latter would predict that adult bilinguals form separate categories for similar phones, indicating that merged categories of early L2 learners are a consequence of the sequential acquisition of two language languages. Clearly, data comparing the discrimination of similar phones from monolingual adults, simultaneous bilingual adults, and early L2 learners are needed to address these issues.

2. The present study

The present study was designed to investigate phonetic discrimination by simultaneous bilingual adults to determine whether or not they have shared perceptual categories for similar phones in their two languages. For this purpose, we tested simultaneous bilingual, monolingual and advanced early L2 learners of Canadian French (CF) and English (CE) on their ability to discriminate Canadian French and English /d/- initial CV syllables. CF and CE productions of /d/ differ in voicing as well as place of articulation. With respect to voicing, VOT values for CF and CE /d/ show some overlap. However, in CF, /d/ is typically produced with lead VOT whereas in CE, /d/ is more typically produced with short-lag VOT. With respect to place, in phonetic descriptions, CF /d/ is phonetically described as dental, whereas CE /d/ is described as alveolar (Picard, 1987; Picard, 2001). Acoustically, previous research has shown that differences in dental and alveolar stop place of articulation are cued by relative burst amplitude and burst frequency measures (Jongman, Blumstein, & Lahiri, 1985; Stoel-Gammon, Williams, & Buder, 1994). Compared to coronal stops in CF, in CE /d/ and /t/ produced by monolingual speakers are louder, have a higher mean burst frequency and smaller standard deviation of burst frequency (Sundara, 2005; Sundara et al., 2006a). The dental-alveolar distinction that occurs across CF and CE, is phonemic in some Australian (Busby, 1980; Dixon, 1980; see Tabain & Butcher (1999) for a review) and Amerindian languages (Dart, 1991), but not within CF or CE.

Certain place distinctions are known to be very difficult to perceive for non-native listeners (Polka, 1991; Werker & Logan, 1985; Werker & Tees, 1984b), and resistant to training (Logan, Lively, & Pisoni, 1991). Therefore, a place distinction signalled by spectral cues was selected to be the most sensitive to disambiguate the effects of age of acquisition and dual language input. Similar to the stimuli used by Bosch and Sebastián-Gallés (2003a), we used natural tokens produced by multiple talkers to ensure that the discrimination task used in the experiment reflects what is required of language learners in the process of language acquisition. Use of stimuli produced by multiple talkers in a speech perception task necessitates that listeners treat members of the same category as equivalent, ignoring the not insignificant talker-variability, while discriminating contrastive categories.

To predict the discrimination performance of monolingual listeners, we used Best's Perceptual Assimilation Model (Best, 1993, 1994a, 1994b; Best & McRoberts, 2003) based on how listeners map non-native phones to L1 phoneme categories. According to PAM, if the two sounds in the L2 contrast are assimilated to two phonemes in L1 (TC - Two Category assimilation), high levels of discrimination performance are expected. In contrast, if both L2 sounds are assimilated as equally good instances of one phoneme in L1 (SC – Single Category assimilation), discrimination performance is poor. If the two L2 sounds are assimilated to the same category but differ in the perceived degree of similarity to it (CG – Category Goodness assimilation), PAM predicts intermediate levels of discrimination performance that vary depending on the degree of category-goodness difference.

Given that the dental-alveolar distinction is phonemic in only a few languages around the world, we tested native Hindi listeners to determine the levels of performance consistent with two category assimilation. Although dental-alveolar coronal stops are not contrastive in Hindi, its phonetic inventory includes a distinction between dental and retroflex coronal stops. Anecdotal evidence indicates that when speaking English, native Hindi speakers produce retroflex stops instead of English alveolar stops. There is also experimental evidence that native Hindi speakers identify English alveolar stops as retroflex stops (Ohala, 1978). Ohala (1978) presented dental and retroflex stops produced by four Hindi speakers, and alveolar stops produced by two American-English speakers in /VCV/ context, to native Hindi listeners. The Hindi listeners overwhelmingly (91% of the time) identified English /d/ as a retroflex. Thus, based on two category assimilation documented in previous studies we expected Hindi listeners to discriminate dental and alveolar stops with high levels of accuracy, close to levels observed for similar native contrasts.

Monolingual English listeners are exposed to the dental-alveolar place distinction in fricatives, but not in stops. Specifically, English /⊖/ and /ð/ are produced at dental place while English /s/ and /z/ are produced at alveolar place. Further, monolingual English listeners have some exposure to dental stops, albeit in a highly restricted context (following inter-dental fricatives), as allophones. Finally, English inter-dental fricatives, /⊖/ and /ð/, are acoustically similar to dental stops. Polka (1995) has previously shown that some English speakers identify the Hindi dental stop as the English inter-dental fricative. Thus, based on category goodness or two category assimilation, monolingual English listeners were expected to have intermediate to high levels of accuracy (see Case, Tuller, & Kelso, 2003 for identification, goodness rating and difference rating for a dental–alveolar synthetic continuum by American English listeners). In comparison, monolingual French listeners have exposure to dental obstruents, both stops and fricatives, but not to alveolar obstruents (Picard, 1987). Thus, we expected the monolingual French listeners to assimilate the phones to a single-category. Accordingly, we expected them to show low levels of accuracy and to display poor discrimination compared to native Hindi listeners and monolingual English listeners.

Of particular interest are the predictions for listeners exposed to two languages. From the findings reviewed in the previous section, advanced early L2 learners were expected to form merged categories for similar phones in their two languages. Consequently, advanced early L2 learners were predicted to have low levels of accuracy in distinguishing /d/-initial CV syllables in CF and CE, more similar to monolingual French listeners rather than to the Hindi or English listeners. If merged categories for similar phones result due to dual language exposure alone, a similar performance was expected for simultaneous bilinguals. However, if merged categories for similar phones result as a consequence of sequential acquisition, simultaneous bilinguals were expected to have higher levels of accuracy than early L2 learners, and thus, comparable to native Hindi and monolingual English listeners and better than monolingual French listeners. The latter result would indicate that the nature of the interaction between two languages on the phonetic level is affected by the timing of early dual language experience (simultaneous vs successive).

3. Method

3.1. Subjects

Fifty adults, 10 monolingual English listeners (5 M, 5 F), 10 monolingual French listeners (5 M, 5 F) 10 simultaneous bilingual CE–CF listeners (6 M, 4 F), 10 advanced early L2 learners of French (4 M, 6 F) and 10 native Hindi listeners (5 M, 5 F) participated in the study. Subjects had no history of speech, language or hearing impairment and were compensated for their participation. Their language background was assessed using a detailed language questionnaire and interview.

To be included in the monolingual French or English group, subjects had to meet four criteria. First, the subject's parents were monolinguals. Second, the subjects were schooled in the same language. Third, they rated their ability in their native language with a minimum of 6 on a scale of 1–7 (where 7 represents native-like ability and 1 represents no ability at all). Most people educated in Canada receive formal instruction in both languages at school. Although experiences vary, typically this instruction is in reading and writing, with little emphasis on speaking or listening skills (for a discussion of language pedagogy in Canada, see Lapkin & Swain, 1990; Swain & Johnson, 1997; Swain & Lapkin, 2005). Thus, proficiency of the students in the second language varies vastly. If monolingual subjects had any knowledge of the non-native language, they rated it below three on the same scale. A bilingual interviewer confirmed their lack of proficiency in the non-native language during the interview. Fourth, they had spent no time in a country where a language other than their native language was spoken.

To be included in the simultaneous bilingual group, subjects had to meet the following four criteria. First, subjects had learned both CF and CE simultaneously at home from their parents, each of whom was a native speaker of one language and communicated with the subject primarily in that language. Second, their schooling was completed either in bilingual schools, or at different points in CF and CE. Third, subjects rated their knowledge of both languages with a minimum of 6 on a scale of 1–7. A fluently bilingual interviewer confirmed their self-ratings of level of proficiency in both languages during the interview. Fourth, they lived in bilingual communities such as Montreal or the Ottawa region, and were using both languages consistently within the home and the work context.

To be included in the advanced early L2 learner group, subjects had to meet four criteria. First, subject's parents were monolingual CE speakers. Second, subjects were first exposed to French at 5–6 years of age in immersion programs at school. Third, subjects rated their competence in both languages with a minimum of 6 on a scale of 1–7. Again, a fluently bilingual interviewer confirmed their self-ratings of level of proficiency in both languages during the interview. Fourth, they were using both languages on a daily basis.

Finally, to be included in the native Hindi group, the subjects had to meet the following criteria. First, both parents of the subject were native Hindi speakers. Second, they grew up in India and came to Canada as adults (after 21 years of age). Third, subjects rated their ability in Hindi with a minimum of 6 on a scale of 1–7 (where 7 represents native-like ability and 1 represents no ability at all). The first author, a native speaker of Hindi, confirmed their self-ratings of level of proficiency in conversational Hindi. Fourth, at the time of testing, subjects were in the habit of speaking Hindi regularly with friends and family. Though all subjects spoke English before they arrived in Canada, they had no prior exposure to French. Further, except in movies, they had no previous experience with American or Canadian English before moving to Canada.

3.2. Stimuli

The stimuli consisted of naturally produced /dV/ syllables excised from /d/-initial bisyllabic words in CF and CE. The syllables were excised from the English words - dodo, doper, dagger and dapper, and the French words – doper, dodo, dadais and datcha. Bisyllabic words in CE and CF differ in stress allocation; the stimuli are part of a larger corpus of syllables described acoustically in Sundara (2005). The items selected were produced by three male monolingual speakers of CF and CE selected using the criteria described above for monolingual listeners. In addition, a three-minute spontaneous speech sample produced by each of the monolingual talkers was rated by three native speakers of CF (or CE) with a minimum of 6 on a scale from 1–7, where 7 represents native-like ability and 1 represents no ability.

Stimuli were digitized at 22 kHz and 16-bit quantization. Subsequently, all stimuli were acoustically analyzed and edited using PRAAT (Boersma & Weenink, 2005). Four syllables produced by each talker were selected, two in a front vowel context (/æ/), and two in a back vowel context (/o/). The final set included 12 CF, that is, dental /dV/ syllables (six in a front vowel and six in a back vowel context) and 12 CE, that is, alveolar /dV/ syllables (six in a front vowel, and six in a back vowel context). Within each vowel context, the edited dental and alveolar syllables were matched for duration, intensity, fundamental frequency, and voicing as measured by VOT.

Most voiced stops in CF and some voiced stops in CE are produced with prevoicing evident as a regular low frequency glottal vibration during stop closure and sometimes through the burst. Using a spectrogram and waveform display, prevoicing was edited out, when present, to equate for VOT. This was done so the listeners were not able to distinguish between CF and CF stops based on VOT. Three phonetically trained adults confirmed that all edited tokens, dental as well as alveolar, were perceived to be voiced. Dental tokens had a mean duration of 141 ms (range = 139–143), mean amplitude of 82 dB (range = 80–84) and mean fundamental frequency of 120 Hz (range = 80–140); the alveolar tokens had a mean duration of 143 ms (range = 139–148), mean amplitude of 81 dB (range = 79–84) and mean fundamental frequency of 115 Hz (range = 87–126).

Tokens were also selected to minimize vowel quality differences between dental and alveolar tokens within each vowel context. The vowels /æ/ and /o/ were selected because their articulatory descriptions are identical across the two languages (Picard, 1987; Picard, 2001), and in previous acoustic analyses (Escudero & Polka, 2003; Sundara, 2005; Sundara et al., 2006a), these vowels revealed substantial overlap in the first and second formants (F1 and F2). To obtain a measure of vowel quality difference between dental and alveolar tokens within each vowel context, F1 and F2 derived from LPC analysis with a 15 ms hamming window centred at the midpoint of the vowel steady state were examined.

In front vowel context, dental tokens had a mean F1 of 552 Hz (range = 438–685) and mean F2 of 1787 Hz (range = 1644–1994) whereas the alveolar tokens had a mean F1 of 613 Hz (range = 545–727) and mean F2 of 1677 Hz (range = 1466–1908). In back vowel context, dental tokens had a mean F1 of 500 Hz (412–581) and mean F2 of 1529 Hz (range = 1371–1774) whereas alveolar tokens had mean F1 of 601 Hz (range = 495–673) and mean F2 of 1214 Hz (range = 1098–1329). Comparing the dental and alveolar tokens, these measures show an almost complete overlap in F1 and F2 in the front vowel context and in F1 in the back vowel context. F2 of back vowels produced in dental and alveolar tokens did not overlap. These differences in F2 steady state between dental and alveolar tokens may be related to articulatory set differences in vowel production themselves, to coronal place differences or some combination of the two (see Sundara, 2005 for a more detailed discussion).

Acoustic analyses are reported for the stimuli to identify possible cues that may be used to distinguish between dental and alveolar tokens. Burst intensity and spectral measures as well as formant frequencies at vowel onset are reported in Tables 1 (front vowel) and Table 2 (back vowel context). Based on the difference between the means and the degree of overlap of the distributions of the cue for dental and alveolar stops, the cue status column in this table indicates whether the acoustic cue can be used to identify the distinction. An asterisk (*) in the cue column denotes that the two distributions are completely non-overlapping, providing a very reliable cue for the place distinction. A plus (+) in the cue column indicates overlapping, but well separated means for the two distributions. Although these cues are less reliable, they may be used to distinguish between dental and alveolar tokens. Finally, mea-

Table 1
Acoustic cues distinguishing dental and alveolar tokens in front vowel context

Cue	Mean (range)		Cue status	Difference between means
	CF dental	CE alveolar		
Burst				
Duration of burst (ms)	12.3 (4.8:21.4)	10.4 (8.6:12.8)		
Relative Intensity (dB)	15 (11:22.5)	16.5 (13.7:21)		
Mean frequency of burst (Hz)	2569 (1170:3932)	3872 (3305:4240)	+	−1303
Standard Deviation of burst (Hz)	2471 (2109:3011)	2126 (1878:2664)	+	346
Skewness	1.1 (0.3:2.4)	−0.3 (−0.7:0)	*	1.4
Kurtosis	1.0 (−0.7:5)	−0.4 (−0.7:1.1)		
Formants				
F1 at onset (Hz)	496 (405:566)	515 (492:576)		
F2 at onset (Hz)	1817 (1660:1981)	1775 (1646:1928)		
F3 at onset (Hz)	2838 (2440:3324)	2514 (2397:2611)	+	324

An asterisk (*) denotes that the two distributions are completely non-overlapping, providing a very reliable cue for the place distinction. A plus (+) indicates overlapping, but well separated means for the two distributions. While these cues are less reliable, they may be used to distinguish between dental and alveolar tokens. Measures with neither an asterisk nor a plus have overlapping distributions with very similar means.

Table 2
Acoustic cues distinguishing dental and alveolar tokens in back vowel context

Cue	Mean (range)		Cue status	Difference between means
	CF Dental	CE Alveolar		
Burst				
Duration of burst (ms)	14.2 (7.7:24)	10.1 (8.5:12.2)		
Relative Intensity (dB)	13.1 (9.5:19.9)	18.9 (12.8:25.1)	+	5.8
Mean frequency of burst (Hz)	1817 (1451:2552)	3312 (2380:3895)	+	−1495
Standard Deviation of burst (Hz)	2195 (1884:2590)	1705 (1431:2192)	+	489
Skewness	1.1 (0.7:1.7)	0.3 (−0.5:0.9)	+	0.85
Kurtosis	0.7 (−0.5:2.8)	1.8 (−0.5:5.5)		
Formants at vowel onset				
F1 at onset (Hz)	478 (412:551)	537 (448:608)	+	−59
F2 at onset (Hz)	1543 (1461:1638)	1491 (1381:1734)	+	51
F3 at onset (Hz)	2661 (2516:2836)	2452 (2189:2676)		

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

sures with neither an asterisk nor a plus have overlapping distributions with very similar means. These cues are unlikely to provide cues to distinguish between dental and alveolar stops.

Burst measures were derived using the waveform display supplemented by a wide-band spectrographic display as described in Sundara (2005). Burst intensity and burst spectral measures were calculated over the entire burst beginning at consonantal release. Consistent with previous research (Jongman et al., 1985; Stoel-Gammon et al., 1994: Sundara, 2005; Sundara et al., 2006a), burst intensity was derived as a relative intensity measure where the intensity of the burst (in dB) is subtracted from the maximum intensity of the following vowel (in dB). Thus, a lower relative intensity measure corresponds to a louder burst. Burst spectrum was characterized by the 4 spectral moments – mean, standard deviation, skewness and kurtosis. Bursts were pre-emphasized prior to making spectral measurements. Spectral moments were derived from the power spectra over the entire burst duration for frequencies up to 11025 Hz. Onset F1, F2 and F3 were derived from LPC analysis with a 15 ms hamming window at vowel onset.

For each measure in Tables 1 and 2, there was some overlap in values for dental and alveolar tokens (except skewness of burst in front vowel context). However, despite overlap in their ranges, the burst spectral measures – mean frequency, standard deviation and skewness, were separated for dental and alveolar tokens in both front and back vowel contexts. Additionally, in the back vowel contexts, dental and alveolar tokens could also be distinguished using relative intensity and F1 and F2 at onset. Tokens produced by CF and CE talkers have been previously reported to differ in relative intensity, mean burst frequency, as well as standard deviation of burst frequency (Sundara, 2005; Sundara et al., 2006a).

3.3. Procedure

Perception performance was evaluated in a categorial discrimination task (AXB) using BLISS (Mertus, 1990). In each AXB trial, three syllables were presented. Subjects were told that each syllable was produced by a different talker. The first and the third syllable were always different; one was alveolar and the other dental. The subjects were asked to decide whether the consonant in the second syllable was similar to the consonant in the first or the third syllable by pressing the left or right mouse button, respectively. No feedback was provided. To succeed, subjects were required to ignore speaker differences and attend only to place differences.

Subjects were tested in a sound-treated room. Stimulus presentation was controlled on-line by a Compaq Presario laptop routed through a pair of Harmon/Kardon speakers at about 70 dBA. Even though discrimination of stop consonants is known to be better under headphones, as follow-up experiments with infants were planned, stimuli were presented through loudspeakers to make testing of adults more similar to that of infants. Monolingual groups were instructed in their native language; half the simultaneous bilingual and L2 group was instructed in French, and the other half was instructed in English; native Hindi listeners were instructed in Hindi. Performance accuracy of bilinguals and early L2 learners did not differ as a function of language of instruction. Thus, it was not included as a variable in subsequent analyses.

In the beginning of each testing session, subjects were presented 10 trials in a practice session to familiarize them with the task. Testing was done in two blocks. Order of testing for the blocks was fixed. Subjects completed Block I, the sequence with tokens produced in the front vowel context, followed by Block II, the sequence with tokens produced in the back vowel context. Subjects were given a break in the middle of each block.

Due to differences in the vowel quality of the back vowel as well as greater differences in burst frequency as well as intensity measures, discrimination in the back vowel condition was easier than in the front vowel condition. Thus, stimulus presentation was blocked by vowel context, the back vowel block always presented after the front vowel block. This was done to avoid transfer of learning from the back vowel condition, aided by vowel quality differences, to the front vowel condition.

Each block consisted of 288 trials with an inter-stimulus-interval of 1500 ms and an inter-trial interval of 4000 ms. In each block, AXB sequences included an equal number of 4 trial types with dental (D) and alveolar (A) tokens – DDA, ADD, DAA and AAD. The 12 tokens (6 dental and 6 alveolar) occurred equally often in each trial type. Testing lasted for about 20 min.

3.4. Results

In a categorial AXB task, because a subject may chose either A or B as a response option, the probability of guessing is 1/2 and thus, chance performance is 50%. To determine whether performance of each subject was significantly above chance, the 95% confidence limits were calculated separately for each block based on the normal approximation to the binomial distribution[1] (Boothroyd, 1984; Kishon-Rabin, Haras, & Bergman, 1997). The confidence limits were calculated based on the number of trials ($n = 288$), the probability of guessing (1/2) and the t-value (2, if number of trials is more than 20). The following formula was used $50 \pm 100/\mathrm{sqrt}\,(n)$, to give a confidence interval extending from 44.14 to 55.89. In each of the two blocks, a subject with accuracy greater than 55.89% can be considered to perform significantly above chance with a $p < 0.05$. Accuracy data for individual subjects in each of the 5 groups are presented in Fig. 1.

All native Hindi listeners and simultaneous bilinguals were better than chance at discriminating CF and CE /dV/ in both vowel contexts. In the front vowel context, 1 out of 10 monolingual English listeners, 6 out of 10 monolingual French listeners,

[1] Confidence limits for scores expected from guessing depend on the number of trials (n), and on the probability of guessing (p; in an m-alternative forced-choice, $p = 1/\mathrm{m}$). Based on the normal approximation to the binomial distribution, 95% confidence limits = chance score $\pm 100 * \mathrm{sqrt}(\mathrm{p}(1 - \mathrm{p})/\mathrm{n}) * t$, if scores are in percent. Values for t are taken from the t-tables according to $n - 1$ degrees of freedom. For $n = 20$ or more, $t = 2.0$ for 95% confidence limits. In a 2-interval forced choice, as is the case in this experiment, $m = 2$, and $p = 1/2$. Thus, the formula for confidence limits reduces to 50% $\pm 100/\mathrm{sqrt(n)}$.

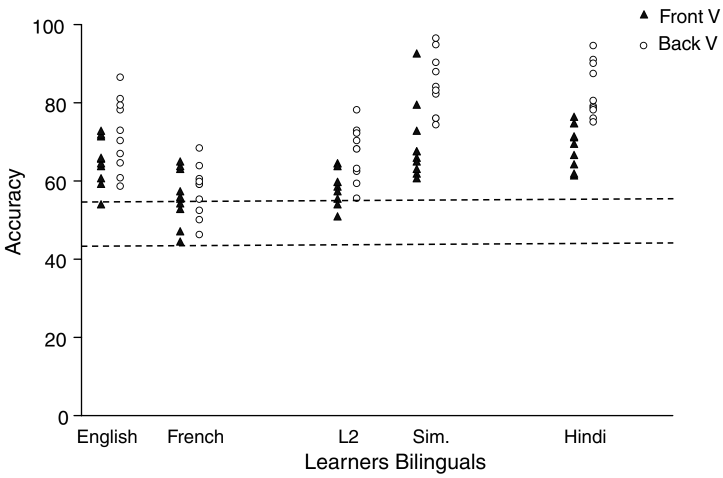

Fig. 1. Scatter plot of accuracy data in front and back vowel context for individual subjects in the each of the 5 groups. monolingual English, monolingual. The dashed lines represent 95% confidence interval for chance performance ranging from 44.14 to 55.89 (see text for details).

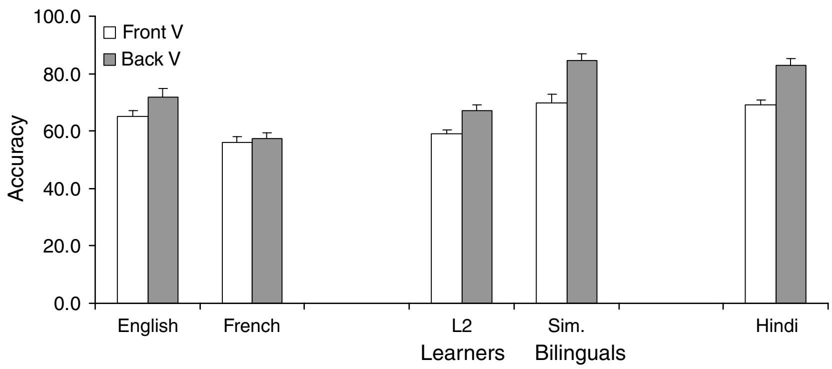

Fig. 2. Mean accuracy data (+SE) for monolingual English, monolingual French, L2 learners, simultaneous French and English learners and native Hindi listeners in the front and back vowel context.

and 3 out of 10 advanced early L2 learners were not able to discriminate CF and CE coronal stops. In the back vowel context, 4 monolingual French listeners and 1 advanced early L2 learner were not able to discriminate CF and CE coronal stops.

The pattern of group differences was identical for analyses using d-prime[2] or percent correct data. Mean accuracy data (percent correct, ±SE) for the four Groups are summarized in Fig. 2. Discrimination performance of the 5 groups was compared

[2] D-prime values were 0.82, 1.01, 1.4, 1.67, 1.77 (front vowel) and 0.89, 1.54, 1.86, 2.68, 2.87 (back vowel) for the French, English, L2 learners, Simultaneous bilinguals and Hindi listeners, respectively. For information on obtaining *d*-prime values in an AXB task, see Macmillan and Creelman (2005).

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

using a General Linear Model (GLM) repeated measures analysis of variance (ANOVA) with Group (monolingual English, monolingual French, advanced early L2 learner, bilingual and native Hindi) as the between-subjects variable and Vowel (front and back) as the within-subjects variable. There was a significant effect of Group, $F(4, 45) = 19.7$, $p < 0.001$, and Vowel, $F(1, 45) = 76.1$, $p < 0.001$. There was also a significant interaction of Group with Vowel, $F(4, 45) = 5.6$, $p = 0.001$.

Overall, performance was poorer in the front vowel than in the back vowel condition; this pattern was evident in all groups. This was expected from the acoustic analyses which showed the alveolar and dental stops to be more acoustically similar in the front vowel context. To assess group differences, paired comparisons were conducted; results are reported in Table 3 for the front vowel and Table 4 for the back vowel context. In each table the mean discrimination score (% correct) and standard error for each group is shown; these groups are arrayed in decreasing order of accuracy. Each cell in the table shows the *t*-test results for the group pair corresponding to the row and column for that cell; the df was 18 for every *t*-test conducted. For significant *t*-values, the group showing the higher discrimination score is indicated along with the *t*-value and *p* level. As there were 10 paired-comparisons, using Bonferroni's correction, only *p*-values of (0.05/10) 0.005 and lower are indicated as significant.

Table 3
Summary of paired-comparisons for discrimination data in the front vowel context

Group Mean (SE)	Bilingual 69.7 (3.1)	Hindi 69.2 (1.7)	English 65.1 (1.9)	Early L2 58.9 (1.4)	French 55.9 (2.1)
Bilingual	–	n.s.	n.s.	Bilingual $t = 3.2$, $p = 0.005$	Bilingual $t = 3.6$, $p = 0.002$
Hindi		–	n.s.	Hindi $t = 4.6$, $p < 0.001$	Hindi $t = 4.8$, $p < 0.001$
English			–	n.s.	English $t = 3.2$, $p = 0.005$
Early L2				–	n.s.
French					–

Table 4
Summary of paired-comparisons for discrimination data in the back vowel context

Group Mean (SE)	Bilingual 84.5 (3.1)	Hindi 83.0 (2.2)	English 71.9 (2.9)	Early L2 67.0 (2.2)	French 57.4 (2.1)
Bilingual	–	n.s.	Bilingual $t = 3.3$, $p = 0.004$	Bilingual $t = 5.3$, $p < 0.001$	Bilingual $t = 8.4$, $p < 0.001$
Hindi		–	n.s.	Hindi $t = 5.2$, $p < 0.001$	Hindi $t = 8.4$, $p < 0.001$
English			–	n.s.	English $t = 4.0$, $p = 0.001$
Early L2				–	n.s.
French					–

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

In the front vowel context, as expected, monolingual English listeners had greater accuracy than the monolingual French listeners; the comparison was marginally significant. The monolingual English listeners, the native Hindi listeners and the simultaneous bilinguals did not differ in their accuracy. As expected native Hindi listeners who were expected to assimilate the CF and CE stops to two-categories, showed high accuracy levels; their discrimination scores were higher than the monolingual French listeners and the L2 learners. Importantly, the simultaneous bilingual group was comparable to the Hindi group. Like the Hindi listeners, discrimination accuracy of simultaneous bilingual group was significantly better than that of the monolingual French group. The accuracy of the advanced early L2 learners was not significantly different from either the monolingual French or the monolingual English listeners but significantly below the accuracy of Hindi listeners and simultaneous bilinguals.

The pattern of group differences in the back vowel context was similar to the findings for the front vowel context. As in the front vowel context, the discrimination accuracy of the monolingual English was significantly higher compared to monolingual French listeners. Further, with the exception of the Hindi listeners, the discrimination accuracy of the simultaneous bilingual listeners was significantly higher than all other groups, including the English monolingual listeners. In the back vowel context as well, the accuracy of the advanced early L2 learners was not significantly different from either the monolingual French or the monolingual English listeners but significantly below the Hindi and simultaneous bilinguals.

To summarize, differences were observed in discrimination accuracy across the five groups when tested with multiple talker stimuli, with a large inter-stimulus-interval and consequently, a high memory load. Overall, in both vowel contexts, the consistently low, near-chance accuracy of the monolingual French listeners is consistent with single-category assimilation and the consistently high, above-chance accuracy of the native Hindi listeners is consistent with two-category assimilation. In both vowel contexts, the simultaneous bilinguals and the advanced L2 learners show divergent levels of discrimination accuracy; the high accuracy of the simultaneous bilinguals is comparable to the Hindi group and the lower accuracy of the advanced L2 learners is comparable to that of the monolingual listeners.

4. Discussion

This study was designed to investigate whether simultaneous bilinguals, like advanced early L2 learners, develop merged categories for similar phones resulting from interactions between their two languages. For this purpose, monolingual English, monolingual French, advanced early L2 learners, and simultaneous bilinguals were tested on discrimination of /dV/ differences in CF and CE. The discrimination performance of all the groups was compared to the discrimination performance of native Hindi listeners who were expected to show discrimination levels consistent with the two categories of coronal stops in their native language. Four main findings emerged.

First, not surprisingly, the monolingual French listeners demonstrated the poorest performance. Recall that monolingual French listeners are exposed to dental but not alveolar obstruents. Second, the monolingual English listeners were significantly better than monolingual French listeners at discriminating the /dV/ in CF and CE in both vowel contexts. Again, this was as expected given the perceptual similarity of the CF /d/ to the CE /ð/, and perhaps because CE listeners are exposed to both dental and alveolar fricatives.

Third, not only was the performance of the advanced early L2 learner group barely above chance, they were significantly poorer than the native Hindi listeners and simultaneous bilinguals in both vowel contexts. Recall that according to Flege and colleagues, interactions between L1 and L2 are most likely to be observed in early L2 learners, because their categories are still immature. Consequently, similar phones in the two languages of early L2 learners tend to merge to produce categories that have characteristics of both phones. The discrimination performance of advanced early L2 learners was consistent with Flege et al.'s (2003) prediction of merged categories in L2 learners. Several researchers have suggested that L1 and L2 systems of L2 learners are not entirely separate (Flege et al., 2003; Grosjean, 1989, 1998; Lambertz & Rawlings, 1969; Paradis, 1978). Discrimination data for /dV/ in CF and CE from advanced early L2 learners in both vowel contexts provides further empirical evidence for interactions between L1 and L2.

The performance of the advanced early L2 learner group in this study cannot be readily explained by L1 differences. Recall that all subjects had English as L1. Unlike simultaneous bilinguals, monolingual English and Hindi listeners, none of the advanced L2 learners had discrimination performance above 70% in the front vowel condition (or above 80% in the back vowel condition). Thus, the discrimination performance of the advanced early L2 learners was consistently poor.

That the advanced early L2 learners were not different from either monolingual group suggests that their perception of /d/ across the languages incorporates dental and alveolar variants. Note that the monolingual English listeners were better at discriminating the CF and CE /dV/ than monolingual French listeners in both vowel contexts. Early extended exposure, as these L2 learners received, appears to have altered their discrimination performance such that unlike the monolingual English listeners, the advanced early L2 learners were not significantly different from monolingual French listeners in either vowel context.

Fourth, unlike the early L2 learners, when tested in a paradigm with high stimulus variability and memory load like natural language acquisition, the simultaneous bilinguals performed consistently above-chance at discriminating /dV/ in CF and CE. Further, as a group, the simultaneous bilinguals did not differ from native Hindi listeners who were expected to categorize /dV/ in CF and CE based on dental and retroflex categories in their native language. These findings indicate that simultaneous bilinguals do not have shared categories for similar phones in their two L1s, either in the front or back vowel context.

This finding is particularly relevant to the discussion of phonological systems of simultaneous bilinguals. Early investigations of children exposed to two languages from birth suggested that while acquiring more than one language, children go through a stage where they are unable to differentiate the languages in their repertoire (reviewed in Genesee, 1989; Vihman, 1985; Volterra & Taeschner, 1978). This undifferentiated language system view, also known as the Unitary Language System or ULS hypothesis, has since been criticized on empirical and methodological grounds (see Genesee, 1989 for a review).

Recent evidence is consistent with an alternative view of bilingual language organization, the Dual Language System or DLS hypothesis (Genesee, 1989). According to this hypothesis, a bilingual has differentiated representation of the two languages at the verbal stage of acquisition. The evidence for DLS comes from investigations of developing pragmatics and syntax; investigations of production phonology of developing bilinguals are equivocal on the issue of unitary or dual language systems.

Although proponents of the ULS and DLS hypotheses investigating phonological and syntactic production abilities of simultaneous bilingual children disagree on the course of development, they unanimously agree that eventually, simultaneous bilinguals develop two systems. There is little argument about the end state of syntactic development, but, to date, there has been no empirical evidence for two phonological systems in adult simultaneous bilinguals – either in perception or production.

To evaluate whether the data presented here are consistent with one or two phonological systems, we need to consider both production and perception data from simultaneous bilinguals. In a previous study, we have shown that simultaneous bilinguals produce language-specific differences in voicing and place of articulation in the two languages, albeit not exactly like the monolinguals (Sundara et al., 2006~~a~~). Acoustic analyses indicate that CF and CE stops produced by monolinguals differ in VOT, burst amplitude, and three measures of burst frequency – the mean frequency, standard deviation and kurtosis of burst spectra. CF and CE stops produced by simultaneous bilinguals differ in VOT and one measure of burst frequency – the standard deviation of burst spectra. Thus, simultaneous bilinguals produce language-specific coronal stops in CF and CE. There is also evidence from acoustic analyses to indicate that simultaneous Quichua-Spanish bilinguals produce language-specific vowels in both languages (Guion, 2003). Language-specific differences in production of target segments in the two languages are usually considered *necessary* evidence for differentiated representation of the two languages. However, production data alone do not provide *sufficient* evidence for differentiation of representation. Consider allophones – they are produced with different phonetic implementation, but they share an underlying representation. Thus, language-specific differences in phonetic implementation are consistent with either one or two phonological systems.

Interestingly, several researchers have previously shown that adults find it difficult to discriminate allophones. They succeed only when tested under sensitive listening conditions – single talker stimuli, with short inter-stimulus-interval minimising memory load in an AX (same-different) task (Carney, Widin, & Viemeister, 1977; Pegg & Werker, 1997; see also Whalen, Best, & Irwin, 1997). The well above-chance performance of the simultaneous bilingual group's discrimination of /dV/ in CF and CE indicates that simultaneous bilinguals do not treat /dV/ in CF and CE as allophones. Rather, simultaneous bilingual listeners have two categories for coronal stops in CF and CE. The evidence that simultaneous bilinguals differentiate between dental and alveolar stops in their production of CF and CE voiced obstruents (Sundara et al.,

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

2006a) and also in their perception ~~(Sundara et al., 2006b)~~ as shown in this study provides the first clear support for the DLS hypothesis.

Recall that even monolingual English listeners were able to discriminate /dV/ in CF and CE in both vowel contexts with accuracy above chance. They were also not significantly different from either simultaneous bilinguals or native Hindi listeners in the front vowel context. Because there was no significant difference between the discrimination accuracy of the simultaneous bilingual and monolingual English group in the front vowel context, it is possible that the two groups use a similar strategy to discriminate /dV/ in CF and CE in the front vowel context. However, if the two groups are using a similar strategy, it is difficult to reconcile the difference in discrimination performance between the simultaneous bilinguals and monolingual English listeners in the back vowel context. Given that the performance of simultaneous bilinguals was virtually identical to that of native Hindi listeners, and even more importantly, significantly different from monolingual English listeners, suggests that the two groups use different strategies to discriminate coronal stops in CF and CE.

One way to determine whether the bilingual and monolingual English speakers use a similar strategy in discriminating CF and CE /dV/ is to obtain identification and rating data from the two groups. However, it is not clear how to obtain interpretable identification and rating data from dual language learners. Several issues complicate this process. Foremost among these is the question of which language mode – monolingual or bilingual – should dual language learners be tested in, in order to be most comparable to monolinguals while giving them the opportunity to exhibit any differences between them. Clearly, determining whether bilinguals and monolingual English speakers use similar perceptual strategies requires further research.

Although it is unclear whether subjects were relying on the burst or vowel cues in the back vowel context, differences between the simultaneous bilinguals and advanced early L2 learners were observed in both vowel contexts. More importantly, unlike the advanced early L2 learners, the discrimination performance of the simultaneous bilinguals in neither vowel context is consistent with merged categories in L1 and L2. Consequently, despite high levels of proficiency, there is a need to investigate and describe the end-state for sequential L2 learners separately from the end state for simultaneous bilinguals.

The difference in discrimination performance by the simultaneous bilinguals and the advanced early L2 learners may be related to differences in the input received by the two groups. Due to social and demographic factors, L2 learners are likely to receive regular input from native speakers of each language less frequently than simultaneous bilinguals. Further, simultaneous bilinguals may also receive input from other simultaneous bilinguals more frequently. It is possible that the tokens produced by simultaneous bilinguals are more perceptually distinct. Evidence to suggest this is indirect. As mentioned previously, simultaneous bilinguals differ from the monolingual groups in the cues they use to signal the dental-alveolar distinction (Sundara et al., 2006~~a~~). Similar to speakers of Australian languages, where dental-alveolar distinctions are phonemic, simultaneous bilinguals produce dental and alveolar tokens differing in standard deviation, but not mean or skewness of burst frequency (M. Tabain, personal communication). Specifically, unpublished moments

Please cite this article in press as: Sundara, M., & Polka, L., Discrimination of coronal stops by ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.01.011

analysis data from Arrernte, Yanyuwa and Yindjibarndi (previously presented in Tabain & Butcher (1999) for analyses of locus equations; also see Tabain, Breen, & Butcher, 2004) indicate that standard deviation but not mean frequency or skewness effectively distinguishes between alveolar and dental stops produced by speakers of all three languages. In languages where dental-alveolar place distinctions contrast meaning, the overlap between these two places of articulation must be minimal, or else these tokens will be perceptually ambiguous. Given the similarity of the acoustic cues used to signal dental-alveolar place differences by simultaneous bilinguals and speakers of Australian languages, it can be argued that compared to coronal stops produced by monolinguals, simultaneous bilinguals produce /dV/ in CF and CE that are more perceptually distinct.

Finally, as simultaneous bilinguals are exposed to the two languages from birth, they hear infant-directed speech in both their languages. Speech addressed to infants is exaggerated at the phonetic level and produced with a large variability in individual items (Burnham, Kitamura, & Vollmer-Conna, 2002; Kuhl et al., 1997). Research indicates that not only do babies prefer infant-directed speech to adult-directed speech (Cooper & Aslin, 1994; Fernald & Kuhl, 1987), but it may also provide several advantages to the infant. Characteristics associated with infant-directed speech are significantly correlated with greater intelligibility (Bradlow, Torretta, & Pisoni, 1996) as well as improved performance on speech discrimination tasks by normal infants (Liu, Kuhl, & Tsao, 2003), children with language delays (Merzenich et al., 1996; Tallal et al., 1996) and adult L2 learners (Lively, Logan, & Pisoni, 1993). There is also evidence to suggest that infant-directed speech facilitates lexical learning in adults exposed to a second language (Golinkoff & Alioto, 1995). Therefore, exposure to two languages from birth may allow a different structuring of perceptual space than that observed even in early L2 learners.

Thus, the differences between simultaneous bilinguals and early L2 learners are most likely to have their roots in infancy. Recall that Bosch and Sebastián-Gallés argue for an initial shared perceptual space in simultaneous bilinguals common to the two languages, that is eventually resolved, either due to continued exposure (input-driven) or due to early word learning (lexico-semantic influences). They attribute it to the frequency of occurrence and distribution of the phones of the two languages. Once the shared categories are resolved, either due to top-down or bottom-up processes, or some combination of the two, a mechanism of *dissimilation* to maintain phonetic contrasts between all categories is likely to result in differences between learners exposed to dual languages and a single language.

Note that postulating perceptual dissimilation as the mechanism of explaining phonetic categories resulting from dual language exposure precludes the two systems of the bilingual being entirely autonomous – at least in infancy, and probably even later in life. Thus, interactions between the two languages are evident in L2 learners as well as simultaneous bilinguals. However, the nature of this interaction is different in L2 learners and simultaneous bilinguals. Specifically, in L2 learners, interactions result in *assimilated* or merged categories, whereas in simultaneous bilinguals, interactions result in *dissimilated* or well-separated categories in production as well as perception.

Flege et al. (2003) have previously posited that interactions between L1 and L2 may be mediated by the mechanisms of category assimilation and dissimilation. To date, early L2 learners have been reported to develop separate categories for production of VOT (Flege & Eefting, 1987) and for perception and production of vowels (Flege, MacKay, & Meador, 1999; Flege et al., 2003; see also Guion, 2003), but not for place differences. Recall from the introduction that certain place differences are more sensitive to early exposure even in the absence of regular language use, are very difficult for non-native listeners, and are very resistant to training. Thus, it is likely that exposure to such place differences in the first few years of life is critical for discriminating them, giving the simultaneous bilingual group an advantage over early L2 learners.

While most of the subjects were better than chance at discriminating CF and CE /dV/, none of the groups had levels of performance that can be characterized as native-like (Polka, 1995). Polka (1995) identifies native-like performance to be above 90%. Of course, it can be argued that the CF and CE /dV/ distinction is not native for any of the groups. Also, unlike previous studies (e.g. Polka, 1995) the subjects were tested using a categorial AXB task, using long inter-stimulus-interval and multiple talkers to make the task demands more like that required in natural language acquisition. Thus, perhaps native-like levels need to be qualified given the specific task demands. Using a 90% correct criterion for native-like performance in the case of high stimulus variability and high memory demands may be too conservative. Likewise, applying a fixed performance level to all types of native contrasts may not be realistic. For example, Jongman (1989) reports findings that show that for some contrasts, native English listeners typically show identification levels that fall short of a 90% accuracy level. Recall that the stimuli used in this experiment were produced by monolingual French and English speakers. Thus, it is possible that the stimuli used in the experiment were not as perceptually distinct as they could have been, had a speaker of a language where the dental-alveolar distinction is phonemic had produced them. Finally, to allow for comparisons to follow-up experiments with infants, the stimuli were presented to adults over loudspeakers. In previous studies of adult perception, the stimuli are typically presented over headphones. When compared to loudspeaker presentation, perception of stop consonants, specifically burst cues, is enhanced under headphone presentation, perhaps explaining why performance was never above 90%.

Two caveats are in order prior to extending the results from this study to simultaneous bilingual phonetic categories in general. These relate to the selection of the languages and the nature of the contrasts tested in this set of experiments. French and English are both official languages in Canada, and in the Montreal and Ottawa regions from where the subjects were selected, both enjoy equal status and are widely used. Clearly most simultaneous bilinguals do not use two languages that fit these criteria. The results of the study are likely to have been affected by these social and demographic factors.

Further, the study reports discrimination of one contrast. And recall that place distinctions were selected precisely because they are most sensitive to age of exposure and language experience. From the pattern of results obtained from previous investigations of effects of language experience, it can be expected that other contrasts, such as vowel or voicing distinctions, will show different patterns. This may be either due to the distributional properties of the phones or differences in acoustic salience. Whereas Bosch and Sebastián-Gallés show evidence for the former, the latter may account for the results from Burns et al.'s study.

In conclusion, advanced early L2 learners, but not simultaneous bilinguals tested in this study, discriminated /d/-initial CV syllables in CF and CE consistent with a merged category. Instead, compared to monolinguals, simultaneous bilinguals were at least as good, if not better, at discriminating /dV/ in CF and CE. Thus, the results of this study provide evidence for differential effects of simultaneous and sequential acquisition of two languages on the organization of phonological categories - specifically, in the nature of the interactions between the two languages. Future research with infants learning two languages will help to determine the developmental roots of these differences.

Acknowledgements

This research forms part of M. Sundara's doctoral thesis. It was supported by an Internal Social Sciences and Humanities Research Council grant administered by McGill University to M. Sundara (202686) and a Discovery Grant from Canada's Natural Science and Engineering Research Council (RGPIN 105397) to L. Polka. We thank Shari Baum, Fred Genesee and two anonymous reviewers for comments on previous versions of this paper; Georgina Hernandez for her help in testing in French; and Marija Tabain, Gavan Breen and Andrew Butcher for sharing the previously unpublished moments analysis data on Yanyuwa, Yindjibarndi and Arrernte.

References

Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: a window one early phonological development. In B. D. Boysson-Bardies, S. D. Schonen, P. W. Jusczyk, P. F. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life*. Dordrecht, The Netherlands: Kluwer.

Best, C. T. (1994a). The emergence of the native-language phonological influence in infants: a perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

Best, C. T. (1994b). Learning to perceive the sound patterns of English. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research*. Hillsdale, NJ: Ablex.

Best, C. T., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech, 46*, 183–216.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*, 305–331.

Boersma, Paul & Weenink, David (2005). Praat: doing phonetics by computer (Version 4.3.19) [Computer program]. Retrieved July 20, 2005, http://www.praat.org/.

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*, 131–158.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech and Hearing Research, 27*, 134–144.

Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12*(2), 189–221.

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.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: global and acousitc fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*(3), 255–272.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new pussycat? On talking to babies and animals. *Science, 296*, 1435.

Burns, T. C., Werker, J. F., & McVie, K. (2003). Development of phonetic categories in infants raised in bilingual and monolingual environments. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th annual Boston University conference on language development*. Cascadilla Press.

Busby, P. (1980). The distribution of phonemes in Australian aboriginal languages. *Papers in Australian Linguistics, 4*, 73–139.

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.

Carney, A. E., Widin, G. P., & Viemeister, N. F. (1977). Non-categorical perception of stop consonants differing in VOT. *Journal of Acoustical Society of America, 62*, 961–970.

Case, P., Tuller, B., & Kelso, J. A. S. (2003). The dynamics of learning to hear new speech sounds. Speech Pathology. Retrieved Nov 17, 2003, from http://www.speechpathology.com/articles/arc_disp. asp?article_id = 50& catid = 560.

Cooper, F. S., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development, 65*, 1663–1677.

Dart, S. N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics, 79*, 1–155.

Dixon, R. M. W. (1980). *The languages of Australia*. Cambridge: Cambridge University Press.

Escudero, P., & Polka, L. (2003). A cross-language study of vowel categorization and vowel acoustics: Canadian English versus Canadian French. In M.J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 861–864).

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279–293.

Flege, J. E. (1992). Speech learning in a second language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research and implications*. Timonium, MD: York Press.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' production and perception of English vowels. *Journal of Acoustical Society of America, 106*, 2973–2987.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*, 467–491.

Genesee, F. (1989). Early bilingual development: one language or two?. *Journal of Child Language 16*, 161–180.

Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: implications for language acquisition. *Journal of Child Language, 22*, 703–726.

Gottfried, T. L. (1984). Effects of consonantal context on the perception of French vowels. *Journal of Phonetics, 12*, 91–114.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3–15.

Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*, 131–149.

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals: age of acquisition effects on the mutual influence of the first and second languages. *Phonetica, 60*, 98–128.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2001). An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *Journal of Acoustical Society of America, 107*(5), 2711–2724.

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech, 1*, 17–38.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., et al. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47–B57.

Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *The Journal of Neuroscience, 23*(29), 9541–9546.

Jongman, A. (1989). Duration of friction noise required for identification of English fricatives. *Journal of Acoustical Society of America, 85*(4), 1718–1725.

Jongman, A., Blumstein, S. E., & Lahiri, A. (1985). Acoustic properties for dental and alveolar stop consonants: a cross-language study. *Journal of Phonetics, 13*, 235–251.

Kishon-Rabin, L., Haras, N., & Bergman, M. (1997). Multisensory speech perception of young children with profound hearing loss. *Journal of Speech, Language and Hearing Research, 40*, 1135–1150.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684–686.

Lambertz, W. E., & Rawlings, C. (1969). Bilingual processing of mixed language associative networks. *Journal of Verbal Learning and Verbal Behavior, 8*, 604–609.

Lapkin, S., & Swain, M. (1990). French immersion research agenda for the 90s: process, product and perspectives. *The Canadian Modern Language Review, 37*(2), 392–428.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*(3), F1–F10.

Lively, S. E., Logan, J., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of Acoustical Society of America, 94*, 1242–1255.

Logan, J., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify /r/ and /l/. *Journal of Acoustical Society of America, 89*, 874–886.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: a user's guide* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum.

Mertus, J. (1990). *BLISS*. Brown University.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*(5245), 77–81.

Ohala, M. (1978). Conflicting expectations for the direction of sound change. *Indian Linguistics, 39*, 25–28.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997a). A limit on behavioral plasticity in speech perception. *Cognition, 64*, B9–B17.

Pallier, C., Christophe, A., & Mehler, J. (1997b). Language-specific listening. *Trends in Cognitive Sciences, 1*(4), 129–132.

Paradis, M. (1978). The stratification of bilingualism. In M. Paradis (Ed.), *Aspects of bilingualism*. Columbia, SC: Hornbeam Press.

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of Acoustical Society of America, 102*(6), 3742–3753.

Picard, M. (1987). *An introduction to the comparative phonetics of English and French in North America*. Amsterdam and Philadelphia: John Benjamin..

Picard, M. (2001). *Phonetics and phonology for ESL and TESL teachers: Comparing Canadian English and French*. Montreal: Concordia University.

Polka, L. (1991). Cross-language speech perception in adults: phonemic, phonetic, and acoustic contributions. *Journal of Acoustical Society of America, 89*(6), 2961–2977.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of Acoustical Society of America, 97*(2), 1286–1296.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72*, 111–123.

Stoel-Gammon, C., Williams, K., & Buder, E. (1994). Cross-language Differences in phonological Acquisition: Swedish and American /t/. *Phonetica, 51*, 146–158.

Strange, W. (1995). *Speech perception and linguistic experience*. Baltimore: York Press.

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), 1026–1037.

Sundara, M., Polka, L., & Baum, S. (2006~~a~~). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition, 9*(1), 97–114.

Sundara, M., Polka, L., & Genesee, F. (2006~~b~~). Language experience facilitates discrimination of /d–ð/ in monolingual and bilingual acquisition of English. *Cognition, 100*, 369–388.

Swain, M., & Johnson, R. K. (1997). Immersion education: a category within bilingual education. In R. K. Johnson & M. Swain (Eds.), *Immersion education: international perspectives* (pp. 1–16). Cambridge: Cambridge University Press.

Swain, M., & Lapkin, S. (2005). The evolving socio-political context of immersion education in Canada: some implications for program development. *International Journal of Applied Linguistics, 15*(2), 169–186.

Tabain, M., Breen, G., & Butcher, A. (2004). CV vs. VC syllables: a comparison of aboriginal languages with English. *Journal of the International Phonetic Association, 34*, 175–200.

Tabain, M., & Butcher, A. (1999). Stop consonants in Yanyuwa and Yindjibarndi: a locus equation perspective. *Journal of Phonetics, 27*, 333–357.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*(5245), 81–84.

Tees, R. C., & Werker, J. F. (1984). Perceptual flexibility: maintenance or recovery of the ability to discriminate non-native sounds. *Canadian Journal of Psychology, 38*, 579–590.

Trubetzkoy, N. (1939). *Principles of Phonology*. Berkeley: University of California Press, Translated by C.A. Baltaxe.

Vihman, M. M. (1985). Language differentiation by the bilingual infant. *Journal of Child Language, 12*, 297–324.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language, 5*, 311–326.

Watson, I. (1995). The effect of bilingualism on the acquisition of perceptual categories underlying the voicing contrast. In K. Elenius & P. Branderud (Eds.), Proceedings of the 13th International Congress of Phonetic Sciences (Vol. 2, pp. 710–713).

Werker, J. F., Gilbert, J. H. V., Humpfrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*, 349–353.

Werker, J. F., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics, 37*, 35–44.

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of Acoustical Society of America, 75*, 1866–1878.

Whalen, D. H., Best, C. T., & Irwin, J. (1997). Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics, 25*, 501–528.

Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception and Psychophysics, 26*, 95–105.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., et al. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology, 36*, 638–642.