# You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments

Christopher Fennell[1] & Krista Byers-Heinlein[2]

University of Ottawa[1] & Concordia University[2]

Abstract

Previous research indicates that monolingual infants have difficulty learning minimal pairs (i.e., words differing by one phoneme) produced by a speaker uncharacteristic of their language environment and that bilinguals might share this difficulty. To clearly reveal infants' underlying phonological representations, we minimized task demands by embedding target words in naming phrases, using a fully crossed, between-subjects experimental design. We tested 17-month-old French-English bilinguals' (N = 30) and English monolinguals' (N = 31) learning of a minimal pair (/kƐm/ - /gƐm/) produced by an adult bilingual or monolingual. Infants learned the minimal pair only when the speaker matched their language environment. This vulnerability to subtle changes in word pronunciation reveals that neither monolingual nor bilingual 17-month-olds possess fully generalizable phonological representations.

Keywords: Infancy, Word Learning, Phonological Development, Bilingualism

# You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments

Infants are surprisingly precocious in their word learning, showing evidence of word knowledge from as young as age 6 months (Bergelson & Swingley, 2013), and this early lexical acquisition appears to proceed in tandem with the development of related perceptual abilities. Each instance of a word is unique (e.g. differing speaker characteristics, emotional valence), leading to different realizations of its constituent sounds. This variability, however, does not affect a word's meaning and should be downplayed in infants' representations of a word's form. Instead, learners need to selectively focus on differences that are *phonemic* in the native-language, for example the distinction between /b/ and /g/ that gives "boat" and "goat" different meanings in English. During the first year of life, infants' ability to discriminate phonetic differences that are meaningful (i.e. phonemic) in the native language is maintained or enhanced (e.g., Kuhl et al., 2006; Polka, Colontonio & Sundara, 2001), while their ability to discriminate phonetic differences that are irrelevant in the native language attenuates (e.g., Werker & Tees, 1984). When can infants apply these *phonetic* sensitivities to word learning in a mature way, by selectively attending to, encoding, and retrieving information that is *phonemic*? The answer to this question is far from straight-forward: infants show complex patterns of success and failure across a variety of ages and tasks (Werker & Curtin, 2005).

One important method used to investigate infants' use of phonetic information in word learning is the Switch task (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). In the classic version of the task, infants are habituated to two novel word-object pairings. At test, infants experience a Same trial where a habituated pairing is presented, and a Switch trial where a pairing violation is presented (e.g. Object A with Word B). Infants demonstrate successful word learning if they look longer to the novel Switch trial than to the familiar Same trial. Around 12-14 months of age, monolingual (Mackenzie, Curtin, & Graham, 2012; Werker et al., 1998) and bilingual (Byers-Heinlein, Fennell, & Werker, 2012) infants successfully learn dissimilar-sounding words (e.g., *lif* - *neem*) in this procedure. When taught minimal pair words that differ by one phoneme (e.g., *bin* - *din*), which directly tests infants' use of detailed phonetic information in word learning, monolingual infants typically do not succeed until 17 months (e.g., Werker, Fennell, Corcoran, & Stager, 2002). Bilingual infants may succeed at the same age as monolinguals (Mattock, Polka, Rvachew, & Krehm, 2010), but other research suggests that this group may have difficulty with minimal pairs up until 20 months of age (Fennell, Byers-Heinlein, & Werker, 2007). This suggests that young infants do not always apply their phonetic sensitivities to word learning.

Significant theoretical work has attempted to reconcile young infants' ability to *perceive* relevant phonetic differences with their apparent inability to *use* this information in some tasks. One proposal is that word learning initially entails high task demands (Werker & Fennell, 2004). Infants must concurrently learn about objects, words, and the connection between them, all while appropriately storing and retrieving the rich phonetic detail of the target words. This might be especially difficult for novice word learners. Thus, task demands and infants' developmental level together influence whether the phonetic detail distinguishing native phonemes will be accessed in a particular task (Curtin, Byers-Heinlein, & Werker, 2011; Werker & Curtin, 2005). Indeed, 14-month-old infants successfully learn minimal pair words when supported by contextual cues that reduce task demands, such as familiarizing infants with the target objects prior to word learning (Fennell, 2012), adding indexical variability while holding the relevant phonetic contrast stable (Rost & McMurray, 2010), and training on lexical contexts that highlight the minimal pair (Thiessen & Yee, 2010).

One task demand inherent to the classic Switch procedure itself is its lack of referential cues. Referential information signals to the infant that the target word should not only be associated with a particular object, but that it should be associated with its abstract conceptual representation (Waxman & Gelman, 2009). In the standard version of the Switch task, to-be-learned words are presented as isolated tokens (e.g., "*Bin!*"), and thus are stripped of the rich referential information characteristic of natural language. One way to enhance the referential nature of the target words is to embed them in naming phrases, which is more typical of parental speech, and provides syntactic information that indicates the target word refers to the object (i.e., it is a noun). When such referential information is provided (e.g. "Look, it's the *bin*!"), 14-month-old infants successfully learn minimal pairs (e.g., Fennell & Waxman, 2010).

In sum, 14-month-old infants only learn minimal pair words when task demands are alleviated, while older infants successfully learn minimal pair words across many situations. What does this developmental pattern indicate about infants' underlying representations? Under the PRIMIR framework of infant speech perception and word learning (Processing Rich Information from Multidimensional Interactive Representations; Curtin et al., 2011; Werker & Curtin, 2005), minimal pair word learning is initially supported by perceptually-based phonetic categories. However, the phonetic information critical to phoneme categories is stored and processed together with other types of acoustic information irrelevant to word meaning (e.g. indexical information like speaker gender). Young word learners do not always attend to the relevant information, and thus can fail in more demanding minimal pair word learning tasks. With experience, infants develop abstract phonemic categories, which summarize across phonetic variation and direct infants' attention to relevant phonological features of a word. PRIMIR predicts that, once infants have developed phonemes, they should robustly succeed in minimal pair word learning tasks as long as the critical phonetic variation falls within the acceptable range of their native phoneme categories.

Surprisingly, infants sometimes fail to learn minimal pairs even at 17 months – an age at which it has been hypothesized that they have developed phonemes (Werker & Curtin, 2005). Using the classic Switch task, Mattock, et al. (2010) tested monolingual (English and French) and bilingual (English-French) infants of 17 months on three realizations of a minimal pair valid in both English and French (/bos/ - /gos/): monolingual English, monolingual French, and bilingual productions. These latter tokens were produced by a bilingual speaker, and included a mixture of English- and French-pronounced tokens. Infants succeeded only when the stimuli matched their language background: bilingual infants succeeded on the bilingual stimuli, French infants on the French stimuli, and English infants on the English stimuli. English and French monolingual infants failed on the bilingual stimuli, and French infants failed on the English stimuli. Mattock, and colleagues hypothesized that bilingual infants could flexibly process the more variable set of bilingual tokens because they experience phonetic variability in their everyday language environments, but that monolinguals could not efficiently process the unfamiliar variation in the targets because they are not normally exposed to such variability. However, Mattock, et al. did not test bilingual infants on monolingual-only stimuli. As such, while it appears that monolingual infants were not using abstracted phonemes to guide their word learning, the claim of enhanced flexibility in bilinguals' phonological representations remains an open question.

Indeed, the results of another Switch study indicate that bilingual infants' word learning might similarly be disrupted when they are presented with phonetic information that does not match their language-learning environment. Using the exact same minimal pair stimuli (/bI/ - /dI/) that English monolinguals learn by age 17 months (Werker, et al., 2002), Fennell, et al. (2007) found that bilingual infants learning English and another language did not accurately learn the minimal pair until 20 months. Fennell and colleagues attributed the difference in results to weaker phoneme representations in bilinguals, perhaps due to their reported difficulties in early phonetic perception (for reviews see Byers-Heinlein & Fennell, 2013; Werker, 2012), or to less overall experience with each language in comparison to monolinguals (Curtin et al., 2011; Werker, Byers-Heinlein, & Fennell, 2009). However, an English monolingual produced these stimuli, which specifically led to monolingual success and bilingual failure at 17 months. Thus, another interpretation of these results is that bilingual infants' difficulty was not related to their bilingualism per se, but rather to an issue common to monolinguals and bilinguals: vulnerability to less familiar pronunciations of native phonemes.

Within a given language, what kinds of pronunciations would be familiar to monolingual as compared to bilingual infants? Monolingual infants are usually raised by parents who are monolingual, while bilingual infants are often raised by at least one parent who is bilingual (Byers-Heinlein, 2013). The majority of bilinguals who acquire a second language in adulthood speak that language with an accent (Piske, MacKay, & Flege, 2001), but even bilingual adults who acquired both languages early in life produce phonemes that differ in small ways from monolingual speakers (e.g., mean voice onset time, greater acoustic variation; Antoniou, Best, Tyler, & Kroos, 2010; Flege & Eefting, 1987; Hazan & Boulakia, 1993). One published study specifically examined productions of bilingual mothers. Bosch & Ramon-Casas (2011) reported that even when the mothers were highly proficient Spanish-Catalan bilinguals, their vowel productions differed as a function of their own early language environments. This slight "accent" in simultaneous and early bilinguals has long been attested in the literature, including for the Canadian French-Canadian English bilingual population involved in the current research (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; MacLeod & Stoel-Gammon, 2009; Sundara, Polka, & Baum, 2006). Together, these studies suggest that there might be systematic differences in the adult pronunciations typical in monolingual and bilingual infants' early language environments, which could in turn affect their ability process phonetic information from different types of speakers.

The aim of the current study was to better understand infants' use of phonetic information during word learning at 17 months. Our study therefore tested monolingual and bilingual infants on both monolingual- and bilingual-produced target words. This fully crossed design has not been undertaken before. Further, our design aims to put monolingual and bilingual infants on equal footing by embedding the tokens in language-specific naming phrases that will clarify language information for the infants (e.g., placing the target in an English sentence highlights that the novel word is English). Past research has shown that naming phrases facilitated monolingual infants' word learning in the Switch task (Fennell & Waxman, 2010). This clarification might be particularly important for bilingual infants, who could need to determine a word's language before appropriately interpreting phonetic information (see Gonzalez & Lotto, 2013, for related adult work). If infants have developed stable, abstract phonemes by this age (Curtin et al., 2011; Werker & Curtin, 2005), then minimal pair word learning should be reasonably robust to non-criterial phonetic variation, such as the slight differences in the phoneme productions of adult bilingual and monolingual speakers. However, if infants' representations are still immature at 17 months, infants might only succeed when tokens match the prevalent phonetic characteristics of their language-learning environment. A final possibility is that our study will reveal overall differences in monolinguals' and bilinguals' ability to learn minimal pair words, with bilinguals either showing greater flexibility (Mattock et al., 2010) or greater difficulty than monolinguals (Fennell, Byers-Heinlein, & Werker, 2007).

## Method

### *Participants*

Infants included in the study were typically developing, without any apparent health or hearing problems, and were at least 37 weeks gestation - all according to parental report. Infants were from the metropolitan area of Ottawa, Canada. Within each language group, infants were randomly assigned to speaker condition (bilingual or monolingual speaker).

*Bilingual Infants:* Thirty bilingual infants successfully completed the study (mean age = 17.48 months; range = 16.13 – 18.62 months). Infants were exposed to English and French from birth and had a maximum of 80% exposure to one language and a minimum of 20% exposure to the other, with 50% mean exposure to English and 50% to French, as assessed by the Language Exposure Questionnaire (Bosch & Sebastián-Gallés, 2001). Sixteen infants (7 females) heard bilingual tokens of the target words (half the infants in English, half in French) and 14 infants (9 females) heard English monolingual tokens. An additional 14 infants were tested but not included in the analyses due to fussiness (10; 3 in the bilingual token condition), parental interference (3; 1 in the bilingual token condition), or distraction during testing (1 infant dropped his toy in the bilingual token condition).

*English Monolingual Infants*: Thirty-one monolingual infants successfully completed the study (mean age = 17.51 months; range = 16.30 – 18.38 months). All monolinguals had greater than 95% exposure to English, as reported by parents. Sixteen infants (7 females) heard English bilingual tokens and 15 infants (8 females) heard English monolingual tokens. An additional 10 infants were tested but not included in the analyses due to fussiness (9; 8 in the bilingual token condition) or being off-camera at test (1 in the bilingual token condition).

*Auditory stimuli*

Auditory stimuli consisted of three novel nonsense words presented in the context of the seven different naming phrases used in Fennell and Waxman (2010) or their French equivalents. Three versions of the stimulus set were created: bilingual tokens recorded from a bilingual speaker in French and English, and monolingual tokens recorded from a monolingual speaker in English. The words used during the habituation and test phases formed a minimal pair differing only in the voicing of the initial velar consonant, /kɛm/ and /gɛm/. These words were chosen because they violate no French or English phonotactic rules and all constituent phonemes are produced in a similar manner across the two languages, although there are differences in voice onset time (VOT) for the velars across the languages. The word /nib/ was used during the pre- and post-test trials, and was chosen because all of its phonemes differ from the target words.

Stimuli were initially recorded in a soundproof booth, and edited using the PRAAT computer program (Boersma & Weenik, 2012). Matched *kem, gem,* and *neeb* versions of each naming phrase were created by excising the /nib/ tokens from the /nib/ sentences (e.g., "Look at the </nib/>"), and splicing /gɛm/ and /kɛm/ tokens from the corresponding sentences into their place. This strategy ensured that infants could not use any slight sound differences in the naming phrases to help them discriminate the target words. Each trial included 9 instances of the novel word: an isolated token at the beginning, seven instances within naming phrases, and another isolated token at the end. See Table 1 for acoustic measurements.

*Monolingual tokens*: A native-English adult female who grew up in the local Ottawa area produced the monolingual tokens. As with the majority of Canadians, she was exposed to French throughout her schooling; however, she had attended only English schools, did not use French on a daily basis, and had always lived in an English monolingual household. The velar consonants conformed to English VOT measures (Lisker & Abramson, 1967). For the /ɛ/ vowel common to both words, the first and second formants were closer to each other than reports of female North American English speakers, but were similar to male values (Hillenbrand, Getty, Clark, & Wheeler, 1995). However, her vowels were perceptible as /ɛ/ and the formant values did not differ across /gɛm/ and /kɛm/ tokens.

*Bilingual tokens*: The bilingual tokens were produced by an adult female raised from birth in a bilingual French-English household in a Canadian city where both French and English are spoken in everyday life. She had both French and English schooling, and used French and English regularly in her daily life. Although she was not originally from the same region as the infants tested, she had spent the previous four years in the Ottawa area and possessed a generic Canadian accent in both languages. This speaker produced two sets of stimuli: one with French naming phrases and one with English naming phrases. Her French productions of /k/ were similar in VOT to previous reports of Canadian French (Ryalls, Cliché, Fortier-Blanc, Coulombe, & Prud'hommeaux, 1997). Her French productions of /g/ had less prevoicing than some reports (Ryalls, et al.), but were similar to others (Caramazza & Yeni-Komshian, 1974). Her English productions of the /k/ and /g/ were similar in VOT to the monolingual's productions (as were the standard deviations), although she had longer lag for both consonants. Her French and English productions of /ɛ/ were typical (Hillenbrand, et al., 1995; Martin, 2002) and she had less variability than the monolingual speaker.

Mattock et al. (2010) had created a bilingual condition by interspersing a bilingual's typical English and typical French productions of target words, whereas we had two separate language sets. Our bilingual condition was solely based on the speaker's background to explore whether infants were sensitive to the subtly different pronunciations of a bilingual versus a monolingual speaker (see Bosch & Ramon-Casas, 2011).

Monolingual and bilingual versions of the stimuli were validated by having three adult monolingual speakers of each language and three adult bilingual speakers identify the target phonemes /g/ and /k/. All adults correctly identified all target consonants.

*Visual stimuli*

See Figure 1 for the three distinctive objects used during testing. The crown and molecule objects moved back and forth across the screen at a slow and constant speed. The water wheel remained in the centre of the screen with its arms rotating.

*Apparatus*

The monolingual and bilingual token conditions were tested in different locations at the same university due to a laboratory move. For the bilingual token condition, testing took place in a 2.38 m by 1.82 m room; dimly lit by a 60W lamp situated 80 cm to the left of the infant. Infants sat on parents' laps facing a 21-inch video monitor approximately 85 cm away, which was surrounded by black cloth. Audio stimuli played at 65 dB, +/- 5 dB, over two loudspeakers, located below the monitor. The lens of the digital video camera peeked out of a hole in the cloth 10 cm below the monitor. A Macbook Pro ran the Habit X computer program, which ordered stimuli presentation and computed looking time data. In a nearby testing room, the experimenter monitored infants' looking time via a closed circuit television system. The researcher coded infant looking by pressing a designated key whenever infants looked at the screen. The testing setup for the monolingual condition was highly similar and only differences are highlighted here. Testing took place in a 2.89 m by 1.30 m quiet room. A NEC Duocom LT280 projector presented the images on to a SmartBoard screen (1.62 m diagonal) 1.5 m in front of infants. The digital video camera was hidden under a table draped with black cloth below the screen, with the lens peeking out of a hole in the cloth 30 cm below the screen.

*Procedure*

After participants arrived, the experimenter explained the procedure, obtained consent, and gave parents the Language Exposure Questionnaire. In the testing room, infants sat on parents' laps facing the screen. Parents listened to female vocal music over headphones to mask the stimuli. From the observation room, the experimenter initiated the first trial by pressing a designated computer key when infants looked at the screen. An animated oval preceded the first trial, and all subsequent trials, to draw infants' attention to the display.

See Figure 1 for the procedure. Infants were tested using either the bilingual or monolingual tokens. For the bilingual token condition, all English monolinguals heard English tokens, whereas half the bilingual infants heard French tokens and half heard English tokens. Habituation trials occurred in a quasi-random order such that each trial type was presented twice within a 4-trial block, with no more than three consecutive trials of the same type. Trial length was fixed at 20 seconds. When average looking time across a four-trial block decreased to 65%, the habituation phase ended. Infants completed a minimum of 8 and a maximum of 24 habituation trials. Immediately after habituation, infants moved to the test phase, where they viewed one Same and one Switch trial in one of eight testing orders that counterbalanced trial order (Same-Switch/Switch-Same) and the particular pairings presented.

*Coding.* Videos of all test trials were coded frame-by-frame (1 frame = 30 ms) by an experienced coder, blind to condition. A second coder then recoded 25% of the videos frame-by-frame, with high reliability ($r = .997$, $p < .001$). These frame-by-frame measurements were used for all analyses, except for one infant for whom online coding values were used because the video was corrupted.

## Results

See Table 2 for infants' looking times to key trials. Preliminary analyses revealed no gender effects in any of the following statistical tests. Also, for bilingual infants in the bilingual token condition, there were no effects related to whether they heard French versus English stimuli. Our research question concerned whether infants' performance differed as a function of whether the stimuli matched (e.g., monolingual infants hearing monolingual-produced tokens) or did not match (e.g., monolingual infants hearing bilingual-produced tokens) their prevalent language experience. Infants in all conditions habituated to the stimuli, showing a significant decrease in attention in the last habituation block compared to the first (*p*s < .001). The number of habituation trials across conditions was statistically equivalent (*p*s >.10), thus any test trial differences cannot be due to differences in perceptual adaptation to auditory stimuli. Further, all groups of infants also showed significant recovery during the post-test as compared to the last habituation block (*p*s < .002); infants were not fatigued or generally disinterested in the task.

A 2 (trial type: Same vs. Switch) X 2 (stimuli match: yes vs. no) X 2 (infant language environment: bilingual vs. monolingual) mixed ANOVA tested whether infants showed different looking times to the Same and Switch trials. See Table 3 for the ANOVA results. We found a significant main effect of trial type, moderated by a significant interaction between trial type and stimuli match. However, there was no interaction between trial type and language environment and no three-way interaction between trial type, language environment, and match. Thus, infants behaved differently across the test trials depending on whether the stimuli matched their learning environment, but this did not differ by language background: monolinguals and bilinguals showed the same pattern. Nevertheless, to explore the significant interaction, we examined infants' test trial performance separately for each possible grouping, using the Bonferroni corrected alpha of .012. Infants who heard stimuli from a speaker that matched their language learning environment looked longer during the Switch than during the Same trial, showing that they detected the phoneme change [Monolinguals: $t(14) = -2.91$, $p = .011$, $d = -0.63$; Bilinguals: $t(15) = -4.02$, $p = .001$, $d = -0.95$]. However, infants who heard stimuli that did not match their language background looked equivalently during both test trials [Monolinguals: $t(15) = -0.80$, $p = .44$; Bilinguals: $t(13) = -1.06$, $p = .31$]. See Figure 2.

For between-subjects effects in the main ANOVA, there was no main effect of match; however, there was a main effect of language, moderated by a language by match interaction.

Monolinguals looked longer to test trials overall, particularly in the condition where they heard bilingual-produced tokens.

Finally, because monolingual-produced stimuli were only in English, we investigated whether bilingual infants with more English exposure outperformed those with less English exposure. We split bilingual infants by language dominance: those with 50% or greater exposure to English (n = 6) and those with less (n = 8). A 2 (trial type: Same vs. Switch) X 2 (English dominance: yes vs. no) mixed ANOVA revealed a significant interaction between trial type and English dominance [$F(1, 12) = 4.98$, $p = .046$, partial $\eta^2 = .29$]. No other effects were significant. English dominant infants looked longer to the Switch trial (*M* = 10.15 s; *SD* = 2.39) than to the Same trial (*M* = 6.44 s; *SD* = 3.11), whereas French dominant infants did not (Switch: *M* = 9.51 s; *SD* = 4.31; Same: *M* = 10.22 s; *SD* = 5.31). However, due to low power engendered by the small subgroup sample sizes, the difference in looking times for the English-dominant subgroup did not achieve statistical significance (p = .096). However, there was a significant positive correlation [$r(12) = .57$, $p = .03$] between amount of English exposure from all sources (e.g., mother, father, grandparents, etc.) and task performance (magnitude of the difference in looking time in favour of Switch), which further supports the above findings. Greater exposure to English facilitated bilinguals' detection of the phoneme change in the monolingual English stimuli.

To investigate whether this pattern might be explained by their nature of the English infants heard from their parents, we classified each parent according to whether or not they grew up as English monolinguals, regardless of their current language usage[1]. We expected that bilingual infants with more exposure to speech from these parents might more readily detect the phoneme changes in words in our English monolingual stimuli. Indeed, there was a significant

[1] We are missing information regarding the linguistic background of one bilingual infant's parents in the monolingual token condition and four in the bilingual token condition.

positive correlation between number of hours per week infants listened to English produced by these parents and their task performance [r(11) = .58, p = .04]. Importantly, there was no relationship between their task performance and the number of hours per week they listened to English produced by parents raised bilingually or raised as French monolinguals [r(11) = -.05, p = .88]. We also examined if exposure to bilingual speakers' productions and to monolingual speakers' productions affected bilingual infants' performance in the bilingual token condition. No significant effects were found (all p values > .4). However, the lack of power due to missing information complicated these analyses. Importantly, our assumption that bilingual infants were hearing bilingual speakers was supported. Twenty-one of 25 bilinguals for whom we had this information were hearing speech from parents raised as bilinguals or from parents raised as monolinguals in the opposite language (e.g., hearing French from a parent raised as an English monolingual). In other words, only four infants had parents adhering to the one parent-one language approach, wherein parents only address their infant in the language in which they themselves were raised as monolinguals.

# Discussion

This study investigated monolingual and bilingual 17-month-old infants' ability to learn minimal pairs produced by an adult representative or unrepresentative of their learning environment. Placing target words in naming phrases should minimize the cognitive load in the task for both monolingual and bilingual infants, thus allowing for a clearer picture of infants' phonological representations. We found that all infants showed vulnerability to stimuli that did not match their language-learning environment: monolingual infants succeeded with a monolingual speaker, bilingual infants with a bilingual speaker, but each group failed with the opposite speaker. Seventeen-month-old infants' overall difficulty with unfamiliar productions of native phonemes in the word-learning task revealed that neither group possessed fully generalizable phonemes. Further, we can reject two hypotheses raised in the introduction: bilinguals this age do not appear to possess more flexible phonological representations than monolinguals (Mattock, et al., 2010) and monolinguals do not appear to have more solid representations than bilinguals (Fennell, et al., 2007).

Our findings are consistent with previous results where monolingual infants failed to learn minimal pair words whose productions did not match their language-learning environment (i.e. French-learning infants failed with English-pronounced stimuli; Mattock, et al., 2010). The current study reveals 17-month-old infants are vulnerable to even more subtle phonetic differences: bilingual and monolingual productions within the same language. As monolingual infants tend to have monolingual parents and bilingual infants tend to have one or more bilingual parents (Byers-Heinlein, 2013), our results are concordant with other evidence suggesting that infants' early discrimination of phonemes is tailored to the specific phonetic properties present in their caregivers' productions (Cristiá, 2012).

The influence of language environment on infants' performance was further highlighted by individual differences amongst the bilingual infants hearing English monolingual tokens: English-dominant bilinguals outperformed French-dominant bilinguals in this condition. Additionally, there was a significant positive correlation between the amount of English infants heard from monolingual English parents and their task performance, yet there was no significant correlation between their task performance and the amount of English they heard from parents who were not raised as English monolinguals. Further evidence of the importance of environment-speaker concordance in infant word learning was found via a re-analysis of 17-month-old bilinguals' performance in Fennell et al. (2007), which also tested bilingual infants on English monolingual stimuli. Splitting infants into English-dominant ($n = 19$) and other-dominant ($n = 18$) groups, we found that the groups differed in their behaviour across Same and Switch trials [interaction: $F(1,35) = 4.22$, $p = .047$, partial $\eta^2 = .11$]. English-dominant infants showed some evidence of learning the minimal pair [Same: $M = 6.2$ s; Switch: $M = 7.8$ s; $t(18) = -1.79$, $p = .09$], while other-dominant infants did not [Same: $M = 6.98$ s; Switch: $M = 5.66$ s; $t(17) = 1.18$, $p = .26$]. Once again, infants whose language environments better matched the stimuli were better able to learn minimal pair words. These findings have important ramifications for studies of bilingual infants' speech perception abilities. While the role of language dominance has been investigated in the context of some speech perception tasks (Sebastián-Gallés & Bosch, 2002), many studies in the bilingual infant literature have not reported analyses with respect to language dominance. Researchers should take both the production of their stimuli and infants' language dominance into account when designing studies and interpreting results. Failing to do so may result in conclusions that bilingual infants generally fail at a perceptual task when, in reality, certain sub-groups of bilinguals can succeed while others are having difficulty.

The current findings recall monolingual infants' reported difficulties in processing phonetic variation in word learning and recognition tasks (Rost & McMuray, 2010), particularly variation associated with accented speech. While monolinguals can recognize familiar words in their own dialect at 6 months (Bergelson & Swingley, 2012), they do not recognize familiar words pronounced in a different dialect until 19 months (Best, Tyler, Gooding, Orlando, & Quann, 2009; Mulak, Best, Tyler, Kitamura, & Irwin, 2013). Further, infants have difficulty learning a novel word produced in a foreign accent up until 30 months of age (Schmale, Hollich & Seidl, 2011). Similar to the idea of generalizable phonological representations (Werker & Curtin, 2005), Best, et al. used the term phonological constancy to describe 19-month-old infants' ability to recognize a word's identity across these natural phonetic variations. Our results add to the accruing data that infants under 19 months do not possess this phonological constancy.

An important question remains with respect to our findings: what specific phonetic information contributed to the difficulties for each infant group? What cues aided monolinguals but impaired bilinguals, and vice-versa? Acoustic measurements of our stimuli did not point to any obvious cause as the bilingual and monolingual speaker differed only slightly, but infants may be tuned to that fine level. Perhaps infants' difficulties stemmed from one or more aspect of our stimuli that we were not able to measure, or a combination of differences. The discovery of which specific phonetic cues aid monolinguals and bilinguals could be a rich area for future study. One further possibility is that infants' perception of speakers' monolingual and bilingual "accents" may have been exacerbated by large amount of phonetic information in the naming phrases. Although we hypothesized that naming phrases would be facilitative, they might have provided infants with further evidence that the speaker was not typical of their language environments. A language clarification manipulation that does not involve naming phrases could test this explanation. For example, one could include a training phase prior to the word-learning phase wherein infants hear a few isolated familiar words in one of their languages, prior to their exposure to the novel minimal pair.

The present study confirmed that both bilingual and monolingual infants have difficulty processing novel minimal-pair words when listening to a speaker of their language who does not match their language-learning environment. Although the optimal stimuli differ for monolingual and bilingual infants, the general pattern of development is the same. Bilingual infants are neither advantaged nor disadvantaged relative to monolinguals, and the inclusion of these two groups illuminates phonological development across all infants. Our results further suggest that using identical stimuli with monolingual and bilingual infants can disadvantage one group or the other, a factor that should be considered in future studies. Mattock, et al. stated that "the first steps to word learning are easier when the phonetic shoes fit" (2010, p. 241). The current study clearly demonstrates just how tightly those shoes need to fit at 17 months of age.

References

Antoniou, M., Best, C. T., Tyler, M . D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics, 39(4),* 558-570. doi: 10.1016/j.wocn.2011.03.001

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA, 109,* 3253-3258. (doi: 10.1073/pnas.1113380109)

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

Boersma, P., & Weenik, D. (2012). Praat: Doing Phonetics by computer: http://www.praat.org

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics, 39*(4), 514-526. doi: 10.1016/j.wocn.2011.02.001

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29–49. doi:10.1207/S15327078IN0201_3

Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition, 16*(1), 32-48. doi: 10.1017/S1366728911000010

Byers-Heinlein, K., & Fennell, C. T. (2013). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychobiology*. Advance online publication. doi:10.1002/dev.21167

Byers-Heinlein, K., Fennell, C. T., & Werker, J. F. (2012). The development of associative word learning in monolingual and bilingual infants. *Bilingualism: Language and Cognition*, *16*(01), 198–205. doi:10.1017/S1366728912000417

Caramazza, A. & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics, 2,* 239-245.

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 the Acoustical Society of America, 54*, 421–428.

Cristià, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America, 129(5)*, 3271–80.

Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*(4), 492–504. doi:10.1016/j.wocn.2010.12.002

Fennell, C. T. (2012). Object Familiarity Enhances Infants' Use of Phonetic Detail in Novel Words. *Infancy*, *17*(3), 339–353. doi:10.1111/j.1532-7078.2011.00080.x

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376–1383.

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. doi:10.1111/j.1467-8624.2007.01080.x

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

Gonzales, K. & Lotto, A. J. (2013). A bafri, un pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science, 24*(11), 2135-2142. DOI: 10.1177/0956797613486485

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French–English bilinguals. *Language and Speech, 36,* 17–38.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97(5)*, 3099-111.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*, F13-F21.

Lisker, L., & Abramson, A.S (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10,* 1-28.

Mackenzie, H., Curtin, S., & Graham, S. A. (2012). 12-Month-Olds' Phonotactic Knowledge Guides Their Word-Object Mappings. *Child Development*, *83*(4), 1129–1136. doi:10.1111/j.1467-8624.2012.01764.x

MacLeod, A.A.N., & Stoel-Gammon, C. (2009). The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics, 30(1),* 53-77. (doi: 10.1017/S0142716408090036).

Martin, P. (2002). La système vocalique du français du Québec de l'acoustique à la phonologie. [The vowel system of Québec French from acoustics to phonology.] *La Linguistique, 38*, 71-88.

Mattock, K., Polka, L., Rvachew, S., & Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: Comparing monolingual and bilingual infants. *Developmental Science*, *13*(1), 229–243. doi:10.1111/j.1467-7687.2009.00891.x

Mulak, K.E., Best, C.T., Tyler, M.D., Kitamura, C., & Irwin, J. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify familiar words spoken in a non-native regional accent. *Child Development.* (doi: 10.1111/cdev.12087)

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191-215. doi: 10.1006/jpho.2001.0134

Polka, L., Colantonio, C., & Sundara, M. (2001). A crosslanguage comparison of /d/–/ð/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109,* 2190–2201. (doi: 10.1121/1.1362689)

Rost, G. C. & McMurray, B. (2010), Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy, 15*, 608–635. doi: 10.1111/j.1532-7078.2010.00033.x

Ryalls, J., Cliche, A., Fortier-Blanc, J., Coulombe, I., & Prud'Hommeaux, A. (1997). Voice-onset time in younger and older French-speaking Canadians. *Clinical Linguistics & Phonetics, 11(3)*, 205-212.

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language, 38,* 1096.1108. (doi: 10.1017/S0305000910000619)

Sebastián-Gallés, N. & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: Role of early exposure. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 974-989. doi: 10.1037/0096-1523.28.4.974

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

Thiessen, E., & Yee, M. (2010). Dogs, Bogs, Labs, and Lads: What Phonemic Generalizations Indicate About the Nature of Children’s Early Word‐Form Representations. *Child*

*Development*, *81*(4), 1287–1303.

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 1–6. doi:10.1016/j.tics.2009.03.006

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234.

Werker, J. F., & Fennell, C. T. (2004). From listening to sounds to listening to words: Early steps in word learning. In D. G. Hall & S. R. Waxman (Eds.), (pp. 79–109). Cambridge, MA: MIT Press.

Werker, J.F. & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 7,* 49-63.

Werker, J. F., Byers-Heinlein, K., & Fennell, C. T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3649–3663. doi:10.1098/rstb.2009.0105

Werker, J. F., Cohen, L., Lloyd, V., Casasola, M., & Stager, C. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289–1309.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1–30.

Werker, J. F. (2012). Perceptual foundations of bilingual acquisition in infancy. *Annals of the New York Academy of Sciences*, *1251*(1), 50–61. doi:10.1111/j.1749-6632.2012.06484.x

Table 1

*Mean acoustic measurements of the target consonant contrast and the subsequent vowels taken from the /gɛm/ and /kɛm/ tokens with standard deviation in parentheses.*

| | **Voice Onset Time (ms)** | | **First Formant at Midpoint (Hz)** | | **Second Formant at Midpoint (Hz)** | |
|---|---|---|---|---|---|---|
| **Speaker** | /gɛm/ | /kɛm/ | /gɛm/ | /kɛm/ | /gɛm/ | /kɛm/ |
| Monolingual English | 14.33 (4.97) | 82.89 (22.52) | 852.33 (86.83) | 970.22 (135.25) | 1526.67 (198.37) | 1575.78 (255.79) |
| Bilingual English | 20.38 (4.78) | 96.38 (20.46) | 739.12 (48.18) | 787.50 (85.26) | 2041.50 (202.20) | 2016.62 (132.66) |
| Bilingual French | -0.14 (29.97) | 74.71 (11.5) | 741.71 (70.43) | 793.71 (65.22) | 2094.57 (186.93) | 2241.57 (69.35) |

Table 2

*Mean looking times in seconds to habituation and test trials by group. Standard deviations are indicated in parentheses.*

| | **Monolingual Infants** | | **Bilingual Infants** | |
|---|---|---|---|---|
| **Habituation and Test Trials** | Monolingual Tokens: Match | Bilingual Tokens: Mismatch | Monolingual Tokens: Mismatch | Bilingual Tokens: Match |
| Pre-test Trial | 18.32 (2.32) | 18.29 (3.48) | 15.71 (3.96) | 16.83 (4.40) |
| First Habituation Block | 17.09 (2.35) | 16.23 (2.10) | 16.02 (3.18) | 15.24 (3.23) |
| Last Habituation Block | 8.94 (2.99) | 9.56 (2.68) | 8.57 (2.94) | 9.01 (3.33) |
| Same Test Trial | 9.04 (4.94) | 13.66 (5.23) | 8.60 (4.76) | 8.13 (3.29) |
| Switch Test Trial | 11.80 (3.71) | 14.75 (3.70) | 9.78 (3.51) | 12.49 (5.58) |
| Post-test Trial | 15.98 (4.06) | 19.33 (1.38) | 14.06 (5.18) | 18.23 (3.11) |

*Note.* Monolingual infants: $n = 31$ (15 in monolingual token condition, 16 in bilingual token condition). Bilingual infants: $n = 30$ (14 in monolingual token condition, 16 in bilingual token condition). All trials were a maximum of 20 seconds.

Table 3

*Main ANOVA.*

| Source | *F* | *p* | *partial* $\eta^2$ |
| --- | --- | --- | --- |
| Between Subjects | | | |
| Stimuli Match (M): Yes vs. No | 1.86 | .18 | .03 |
| Language Environment (L): Mono- vs. Bilingual | 6.86 | .01 | .11 |
| M x L | 6.29 | .02 | ,10 |
| Within Subjects | | | |
| Test Trial (T): Same vs. Switch | 16.65 | .00 | .23 |
| T x M | 4.43 | .04 | .07 |
| T x L | .54 | .47 | .01 |
| T x L x M | .42 | .52 | .01 |

*Note.* $n$ = 61. The *df* for all statistics presented in the table are 1, 57.

*Figure 1:* Outline of the experimental procedure and the audio-visual stimuli. The auditory tokens were presented in naming phrases.

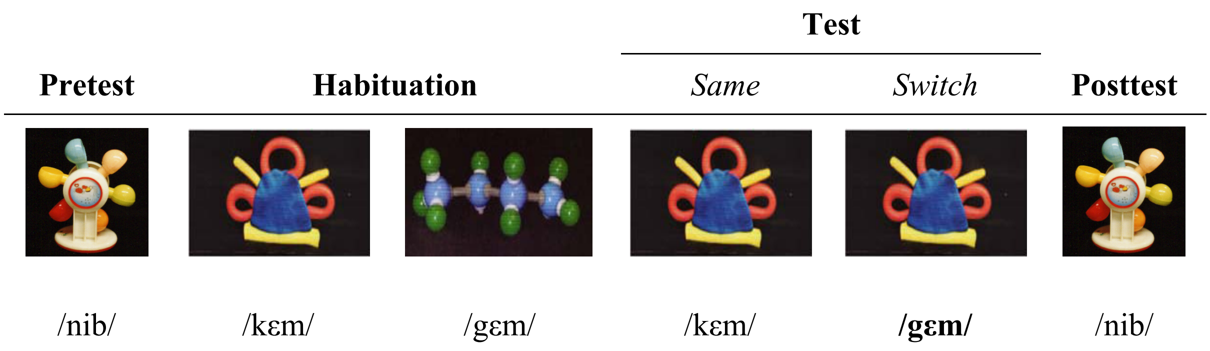
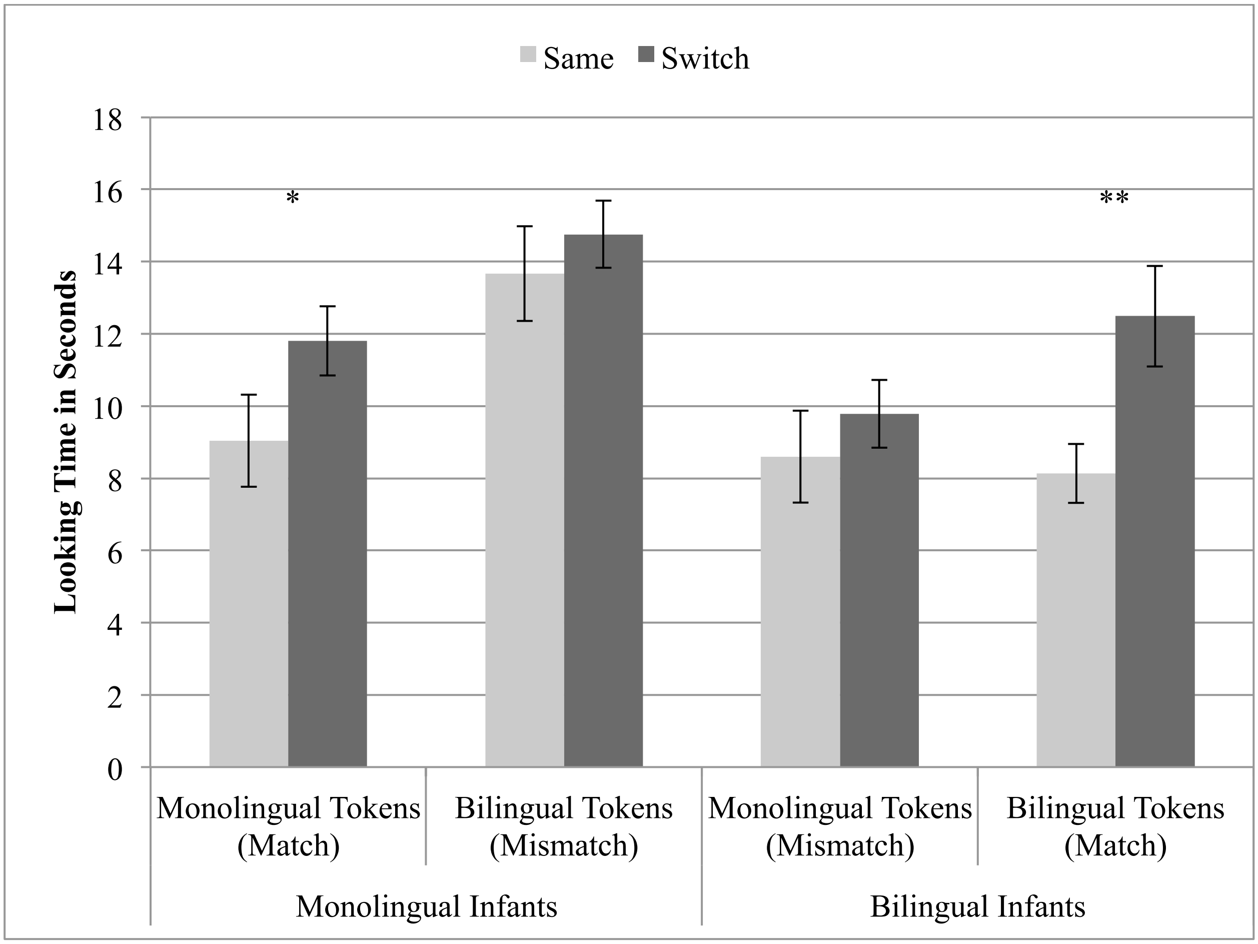


*Figure 2:* Infants' mean looking times in seconds to the test trials across the conditions.




*Note.* Monolingual infants: $n = 31$ (15 in monolingual token condition, 16 in bilingual token condition). Bilingual infants: $n = 30$ (14 in monolingual token condition, 16 in bilingual token condition). All trials were a maximum of 20 seconds. * $p < .05$, ** $p < .01$.