



## PAPER

# The first steps in word learning are easier when the shoes fit: comparing monolingual and bilingual infants

Karen Mattock,<sup>1</sup> Linda Polka,<sup>2</sup> Susan Rvachew<sup>2</sup> and Madelaine Krehm<sup>3</sup>

1. Department of Psychology, Lancaster University, UK

2. School of Communication Sciences & Disorders, McGill University, Canada

3. Department of Psychology, McGill University, Canada

## Abstract

English, French, and bilingual English-French 17-month-old infants were compared for their performance on a word learning task using the Switch task. Object names presented a /b/ vs. /g/ contrast that is phonemic in both English and French, and auditory strings comprised English and French pronunciations by an adult bilingual. Infants were habituated to two novel objects labeled 'bowce' or 'gowce' and were then presented with a switch trial where a familiar word and familiar object were paired in a novel combination, and a same trial with a familiar word–object pairing. Bilingual infants looked significantly longer to switch vs. same trials, but English and French monolinguals did not, suggesting that bilingual infants can learn word–object associations when the phonetic conditions favor their input. Monolingual infants likely failed because the bilingual mode of presentation increased phonetic variability and did not match their real-world input. Experiment 2 tested this hypothesis by presenting monolingual infants with nonce word tokens restricted to native language pronunciations. Monolinguals succeeded in this case. Experiment 3 revealed that the presence of unfamiliar pronunciations in Experiment 2, rather than a reduction in overall phonetic variability was the key factor to success, as French infants failed when tested with English pronunciations of the nonce words. Thus phonetic variability impacts how infants perform in the switch task in ways that contribute to differences in monolingual and bilingual performance. Moreover, both monolinguals and bilinguals are developing adaptive speech processing skills that are specific to the language(s) they are learning.

## Introduction

By their first birthday, monolingual infants display *native* language speech perception skills that seemingly prepare them for word learning. There is a substantial body of literature showing discrimination of native phonetic contrasts to be maintained across the first year, while discrimination of some non-native contrasts declines during this same time frame (Mattock & Burnham, 2006; Polka & Werker, 1994; Werker, Gilbert, Humphrey & Tees, 1981; Werker & Tees, 1984; for a review see Polka, Rvachew & Mattock, 2007). Several studies also show that language experience facilitates discrimination accuracy for native language contrasts with significant change evident in the first year of life for some contrasts (Polka, Colantonio & Sundara, 2001; Sundara, Polka & Genesee, 2006b; Tsao, Liu & Kuhl, 2004; Kuhl, Stevens, Hayashi, Deguchi, Kiritani & Iverson, 2006). These findings have been taken as evidence that infants are honing their perceptual skills to focus on the phonetic variation that is most useful in their language and that which carries meaning.

Despite the early emergence of language-specific perception, infants do not initially make full use of their native language phonetic perception skills when they begin to learn words (Stager & Werker, 1997; Werker & Stager, 2000; Werker, Fennell, Corcoran & Stager, 2002; Pater, Stager & Werker, 1998, 2004). For example, Stager and Werker (1997) found that 14-month-old English-learning infants were unable to learn the association between a novel word and a novel object when the labels for the objects were phonetically similar, /bIh/ and /dIh/. They tested infants using the switch task where infants were habituated to two word–object pairings and were tested on their ability to notice a switch in word–object pairing as indexed by increased looking time to a 'switched' pairing. However, younger 8-month-olds could discriminate /bIh/ and /dIh/ when these phonetic strings were paired with a bounded static visual pattern rather than two moving objects, and 14-month-old infants learned the word–object associations in the switch task when the object labels were phonetically dissimilar, 'lif' vs. 'neem' (Stager & Werker, 1997). Follow-up studies show that 14-month-olds repeatedly fail in the switch task when

Address for correspondence: Karen Mattock, Department of Psychology, Lancaster University, Fylde College, Bailrigg, Lancaster LA1 4YW, UK; e-mail: k.mattock@lancaster.ac.uk

other phonetically similar labels are used, for example when the difference is in manner of articulation, /b/ vs. /p/, or place and manner combined, /d/ vs. /p/ (Pater *et al.*, 2004).

These findings that infants discriminate native consonant contrasts but fail to access native phonetic categories in a word learning task raise questions about the connection between pre-lexical speech perception and lexical acquisition. Most current views of acquisition assume, either implicitly or explicitly, that attunement to native language phonetic categories in the first year prepares the child to acquire linguistic structure. However, direct evidence for such links is limited (Kuhl, Conboy, Padden, Nelson & Pruitt, 2005; Stager & Werker 1997; Tsao *et al.*, 2004; Werker *et al.*, 2002; Werker & Stager, 2000), and some researchers have proposed a discontinuity in the representations underlying pre-lexical speech perception and post-lexical speech processing (Brown, 2000; Brown & Matthews, 1997; Ferguson & Farwell, 1975; Pierrehumbert, 1990).

Werker and colleagues (Werker & Curtin, 2005; Werker & Yeung, 2005; Werker & Fennell, 2004) assume that there is an underlying continuity between pre- and post-lexical representations; they claim that the infant's selective responsiveness to native language categories is maintained across development but is not equally accessible at all points along the way. To explain infants' failure in the switch task, Stager and Werker (1997) proposed the resource limitation (CRL) hypothesis (see also Werker & Fennell, 2004, for a discussion of this issue). Specifically, they argue that the cognitive demands of word learning limits the 14-month-old's ability to form word-object associations in the switch task. In order to successfully form an association between a novel word form and a novel object in the switch task the infant must be able to efficiently access the relevant phonetic units in the speech signal, recognize two visual patterns as distinct whole objects, and hold both types of information in memory long enough to link each phonetic string to the appropriate object without any additional contextual cues to support this mapping. According to the CRL hypothesis, the task of devoting adequate cognitive resources to this mapping process compromises access to fine phonetic differences by the novice word learner. Accordingly, factors that increase processing load can be expected to thwart acquisition of novel word-object associations involving phonetically similar word forms. Likewise, factors that ease processing load are expected to support successful word-object associations involving phonetically similar word forms. Several findings support this perspective. For example, Yoshida, Fennell, Swingley and Werker (2009) found that 14-month-olds *can* learn two similar sounding nonce words, 'bin' and 'din', when a less demanding visual choice paradigm is used, rather than the switch task. In the test phase of the visual choice method both objects are presented side by side on the

screen, a target word is presented auditorily, and to demonstrate word learning the infant is required to look at the correct object (see Swingley & Aslin, 2000). Moreover, using the switch task, Werker and colleagues have shown that more experienced word learners can use phonetic detail to form novel word-object associations. Specifically, 17-month-olds and also 14-month-olds with large receptive vocabularies succeed in the switch task even when phonetically similar non-words are used (Werker *et al.*, 2002). Presumably because these experienced word learners are sufficiently skilled in the word learning task they can allocate more to accessing phonetic detail. In the switch task, 14-month-olds are also more successful with phonetically similar labels when familiar words are paired with familiar objects (Fennell & Werker, 2003) or when novel non-words are paired with familiar objects (Werker & Fennell, 2004). The partial familiarity with the events involved in the mapping presumably reduces the cognitive load. Overall, the CRL hypothesis can provide a coherent and parsimonious account of available data without rejecting the idea that infants' early phonetic representations support their emerging linguistic skills. Nevertheless, more direct evidence showing a clear link between speech perception and word learning is needed to establish continuity between the representations underlying these skills.

To date, infants raised in monolingual English-speaking families have played the starring role in the burgeoning story of speech perception and word learning outlined above. Studies of bilingual infants provide another way to learn more about the relationship between these emerging skills and are essential for the development of a comprehensive model of language acquisition. With respect to speech perception development, several studies have uncovered differences between bilingual and monolingual infants. Bosch and Sebastián-Gallés (2003a) tested monolingual Spanish, monolingual Catalan, and bilingual Spanish/Catalan infants on their perception of the /ε-e/ vowel pair. This contrast is phonemic in Catalan, but not in Spanish, and is difficult for monolingual adult speakers of Spanish to discriminate (Pallier, Bosch & Sebastián-Gallés, 1997). Infants discriminated this contrast at 4 months of age regardless of language background. With increasing age, a decline in discrimination was observed for Spanish infants and maintenance of discrimination was observed for Catalan infants, as was expected on the basis of phonemic status in the input language. The bilingual group showed a U-shaped developmental pattern characterized by a decline in discrimination between 4 and 8 months of age, followed by improved performance at 12 months that was comparable to the performance of monolingual Catalan infants. Similar results emerged when the same language groups were tested across several ages on discrimination of a fricative voicing distinction, /z-/s/, which is phonemic in Catalan but not in Spanish and on discrimination of the /o-/u/

contrast which is phonemic in both Spanish and Catalan (Bosch & Sebastián-Gallés, 2003b, 2005; Sebastián-Gallés & Bosch, in press). Developmental patterns also differed in research on discrimination of the English /d/-/ð/ contrast. Language experience facilitates perception of this contrast but this facilitative effect emerges later in development for children acquiring English and French simultaneously in comparison to children acquiring English only (Polka *et al.*, 2001; Sundara *et al.* 2006b).

Two subsequent studies have failed to line up with these initial findings and instead show that bilingual infants readily resolve subtle phonetic contrasts that are close or overlapping in the phonetic space formed by their dual language input. Sundara, Polka and Molnar (2008) tested monolingual English, monolingual French and bilingual (English/French) infants' discrimination of French /d/ vs. English /d/. These phones differ in place of articulation because /d/ is dental in French and alveolar in English; this place contrast is not phonemic in English or French but is phonemic in other languages. At 6–8 months infants in all three language groups discriminated this place contrast. At 10–12 months the English and bilingual infants continued to discriminate this contrast but the French infants did not, replicating the same language differences that are observed in adult discrimination of this contrast (Sundara & Polka, 2008). Thus at 10–12 months both monolingual and bilingual infants perform similarly to adult perceivers from their respective linguistic communities, and bilingual infants resolve subtle phonetic differences that may or may not be perceived by their monolingual peers. These results show that acquiring two languages does not result in a uniform slowdown in the development of phonetic perception. Findings reported by Burns, Yoshida, Hill and Werker (2007) also support this conclusion. They tested discrimination of bilabial stops differing in VOT by monolingual English and bilingual (English/French) infants. At 6–8 months no effects of language experience were evident; both groups discriminated a VOT difference that crosses the French boundary but not a VOT difference that crosses the English boundary. At 10–12 months and at 14–20 months, bilingual infants discriminated both VOT differences while monolingual English infants discriminated only the VOT difference spanning the English boundary. Thus, both groups shifted to a native language perception pattern within the same age range; in doing so, the bilingual infants learned to perceive two VOT contrasts while the monolingual English infants learned to perceive one.

To sum up, so far speech perception findings show that bilingual infants follow the same developmental path as monolingual infants for some native language contrasts but they need more time and language experience to learn to perceive other native phonetic contrasts. However, data on bilingual infants are still quite limited. Presently, it is unclear when and why developmental differences between monolinguals and

bilinguals occur but they may be due to: (1) bilinguals' reduced exposure to contrasts that occur in one language but not the other; (2) the nature and extent of overlapping acoustic cues associated with certain phonetic categories in the bilingual's input; (3) an adaptive initial focus on phonetic elements that carry the most informational value across both languages.

Although we are moving toward an understanding of speech perception in bilingual infants, their word learning skills have yet to be exhaustively studied. If we adopt a continuity position and assume that the initial tuning of perception to native language phonetic categories is important preparation for subsequent steps in language acquisition, the existing speech perception findings would lead us to expect bilingual infants to be equally or less well prepared than their monolingual peers to begin the task of building a receptive lexicon. A recent study by Fennell, Byers-Heinlein and Werker (2007) supports this view. They tested monolingual English, English-French and English-Cantonese bilinguals in a word learning task that used the [bIh]–[dIh] contrast (spoken by an English speaker). Both groups of bilingual infants failed in the switch task at 17 months, whereas monolingual infants succeeded at this age. However, by 20 months, bilinguals were successful in this task. The critical contrast in this study, /b/ vs. /d/, is phonemic in both languages of the bilinguals. However, there are differences in the phonetic realization of the /bIh/ and /dIh/ word forms across the two languages of these bilinguals that may account for performance differences between the monolinguals and bilinguals.

On the other hand, a different prediction about bilingual word learning emerges from the literature on cognitive and linguistic abilities of bilingual preschoolers. This work points to cognitive advantages in bilinguals compared to monolingual children – most notably in metalinguistic awareness (Bialystok, 1999; Vygotsky, 1962). Metalinguistic abilities include explicit awareness of abstract linguistic representations. For example, bilingual children realize that the relationship between an object and its label is purely arbitrary (Cromdal, 1999), and are more willing to accept a novel or unconventional name for an object than their monolingual peers (Ben Zeev, 1977; Cummins, 1978; Feldman & Shen, 1971; Ianco-Worrall, 1972). Moreover, word-referent problems demand high levels of controlled attention and are solved more proficiently by bilingual children (Bialystok, 2001). The central theoretical assumption is that bilingual children possess greater linguistic flexibility (Cummins, 1978), and greater control of attention when completing linguistic tasks (Bialystok, 2001). These attributes assist the bilingual child to navigate successfully around their language environment, selectively attending to one set of naming labels, while at the same time ignoring equally meaningful labels from the competing language.

Overall, the speech perception abilities of bilingual infants and the cognitive and linguistic abilities observed in bilingual preschoolers lead to contradictory predictions about the early word learning abilities of 17-month-old bilingual infants. On the one hand, if speech perception sets the stage for later word learning, bilingual infants may show comparable word learning skills or have more difficulty learning associations between similar sounding labels and novel objects than monolinguals. The precise outcome will depend on whether the bilinguals possess stable phonetic representations for the specific phonetic categories involved in the task. On the other hand, bilingual toddlers are clearly advantaged on a range of language tasks by 2 to 3 years of age (Bialystok, 2001). Thus older bilingual infants may have better word learning facility than their monolingual peers. This facility may reflect superior cognitive skills, including enhanced cognitive flexibility and greater control of attention (Bialystok, 2001) that make cognitively demanding tasks easier for bilingual infants in comparison to their monolingual peers. Alternatively, a bilingual advantage may be linguistic in nature and more directly related to the specific demands of the word learning task. The added complexity and challenge of word learning when acquiring two languages simultaneously may facilitate the acquisition of this linguistic skill in bilingual children. Thus, the bilingual's real-world experience with two languages may build an early emerging talent for word learning.

To evaluate these conflicting viewpoints we examined the word learning ability of 17-month-old infants who are being raised in bilingual (English-French), monolingual English, and monolingual French families.<sup>1</sup> Specifically, we compared monolingual and bilingual infants' ability to form an association between a novel object and a novel word form in the switch task using phonetically similar object labels that differ only in their initial consonant. Infants of 17 months were chosen because this is the age at which monolingual infants can reliably use fine phonetic detail in a word learning task (Werker *et al.*, 2002). We structured our task to conform to previous studies using the switch paradigm. To construct our task to focus directly on word learning skills, we took several steps to minimize effects of language experience on perception of the object labels. First, we selected a consonant pair that is unlikely to be perceived differently by French and English infants or to pose perceptual difficulty for the bilingual infant, and is also likely to be equally salient and relevant in the emerging lexicon of our monolingual and bilingual infants. To achieve this we chose the consonant pair /b/ vs. /g/ which has converging structural properties in English and French including: (1) it is phonemic in both languages and has a high lexical load (distinguishes many words) in each

language; (2) both consonants have a similar surface phonetic form or pronunciation in English and French; and (3) both consonants have a high frequency of occurrence in each language. These consonants formed the initial phoneme in our object names – /bos/ ('bowce') and /gos/ ('gowce'). Second, the object labels were formed using syllable form and non-target phonemes that also match the phonological structure of both French and English. The non-target phonemes, /o/ and /s/, were selected because they are also frequently occurring and phonetically similar segments in English and French. The phonotactics and the closed syllable form of these nonce words also result in possible word forms in both languages. Finally, an adult bilingual produced French and English pronunciations of each nonce word. Thus, the phonetic variability present in test items included variants that could occur in the input of each monolingual group as well as the bilingual group.

Recall that in previous studies monolingual 17-month-olds have consistently succeeded in the switch task using phonetically similar object labels that differ in a native consonant contrast (Werker *et al.*, 2002). Assuming that speech perception prepares the infant for later word learning, we expected monolingual infants and bilingual 17-month-olds to be equally successful in the switch task implemented in the present study. This outcome is predicted because we selected object labels that we expect bilinguals and monolinguals to perceive with equal proficiency; thus bilinguals should not be disadvantaged in perceiving the word forms. However, if cognitive and linguistic advantages observed in bilingual preschoolers are rooted in their early experience with the processing of two languages, then bilingual 17-month-olds are expected to be more successful word learners compared to their monolingual peers when the speech processing demands of the task are comparable for both language groups.

## Experiment 1

### Method

#### Participants

Forty-eight 17-month-old infants participated in this study. Sixteen infants were English-learning ( $M_{age} = 525$  days,  $SD = 10.6$ ), 16 were French-learning ( $M_{age} = 526$  days,  $SD = 9.3$ ) and 16 were bilingual ( $M_{age} = 526$  days,  $SD = 11.0$ ), (eight boys and eight girls per language group) with exposure to English and French simultaneously from birth. An additional 15 infants were tested but their data were discarded for the reasons of fussiness (11), crying (two), and failure to complete (two). Included infants were born full term and had reportedly normal hearing. Infants were recruited

<sup>1</sup> In this study, French refers to Canadian French and English refers to Canadian English.



from Montreal, Canada where English-French bilingualism is common. Parents answered questions over the phone about their infants' language exposure, and then completed a detailed language exposure questionnaire at the lab visit similar to that used by Bosch and Sebastián-Gallés (2001). The questionnaire is administered via an interview format in which parents are asked to specify the language(s) they use and the language(s) used by family and caregivers who interact regularly with the infant. This information is used to estimate the number of hours the infant hears the language(s) on a daily basis for a typical week and to compute the ratio of regular exposure to English and French. Parents are also asked to estimate the ratio of French/English in their child's input. If substantial discrepancy is noted between our computed ratio and the parent's estimated ratio, the interview questions are reviewed to find the source of the discrepancy. All infants in the bilingual group were estimated to have balanced input in English and French (50–50% or 60–40%). All infants in the monolingual groups were estimated to have 90–100% exposure to either English or French.

### Stimuli

Auditory stimuli were the nonsense words 'bowce' (/bos/) and 'gowce' (/gos/), spoken in an infant-directed style by a male English-French bilingual. The speaker was raised in Montreal by one native Canadian English-speaking and one native Canadian French-speaking parent. He received formal schooling in both languages and rates himself as highly proficient and equivalently proficient in both languages. He reports using both languages on a daily basis.

Tokens of 'bowce' and 'gowce' were recorded onto minidisk in a sound proof booth. The speaker was instructed to produce English and French tokens of 'bowce' and 'gowce'. To induce an English language set, the speaker was given the sentence 'You are showing an infant a new toy, a bowce...', and was instructed to produce carrier sentences containing the target words (e.g. 'Look at the bowce', 'That's a great bowce') followed by 10 isolated productions of the target words in an infant-directed style. This procedure was completed four times – for English bowce and gowce, respectively, and for French bowce and gowce, respectively (e.g. *Vous montrez un nouveau jouet à un enfant – une bowce*, e.g. *'Regarde la bowce', 'C'est une bonne bowce', etc.*). Final stimuli were selected from the pool of individual tokens.

Six tokens of /bos/ and six tokens of /gos/ (three per language) were selected to include items that fall within the same range of values with respect to overall duration, vowel duration, /s/ frication duration, pitch, amplitude, and first (F1) and second (F2) vowel formant frequencies (at vowel mid point) as shown in Table 1. The stops /b/ and /g/ are voiced in English and in French, but there are some differences in voicing cues that specify these segments in each language. For voiced stops, voice-onset

time (VOT) distributions overlap in Canadian French and Canadian English although the modal values found in each language are different. Canadian English voiced stops are typically produced with short-lag VOT but are also produced with prevoicing; Canadian French voiced stops are typically prevoiced (Sundara, 2005). Macleod and Stoel-Gammon (2005) show that the majority of voiced stops produced by monolingual Canadian French speakers are prevoiced. Bilingual Canadian English/Canadian French speakers produce different VOT distributions in each of their languages but they often produce voiced stops in English with prevoicing and thus appear to produce VOTs for voiced stops in a consistent fashion across their two native languages (Caramazza, Yeni-Komshian, Zurif & Carbone, 1973; Gurski, 2006; Macleod & Stoel-Gammon, 2005; Sundara, Polka & Baum, 2006a). Our selected test tokens have VOT values that fall within a similar range across the French and English tokens of each nonce word and also conform to acceptable values for stop production within each language. Formant values for /o/ overlap considerably in English and French (Picard, 1987); this was also observed in our stimuli. Formant measures reveal slight differences across French and English tokens of /bos/ and somewhat larger differences for tokens of /gos/, but the differences do not pattern in a consistent way. For both /bos/ and /gos/ items, the /s/ segments tend to be longer in the English tokens than the French tokens; as a result of this difference the English tokens tend to be longer overall compared to the French tokens, but the duration values overlap nonetheless.

Individual tokens of /bos/ and /gos/ were combined in audio strings with an inter-stimulus interval of 1000 ms. Final /bos/ strings comprised three English and three French tokens of /bos/, and final /gos/ strings comprised three English and three French tokens of /gos/.<sup>2</sup>

Prior to testing infants, perceptual judgments of the stimuli were obtained from adults with no training in phonetics or speech science, including five English monolinguals and five French monolinguals. Adults completed a pen and paper two-alternative forced-choice judgment task. Each token was played twice over headphones and participants were asked to (i) decide if the token was an English or a French pronunciation, (ii) rate their confidence in this judgement on a 5-point scale (1 = not confident at all; 5 = extremely confident), (iii) rate likeness to the native language on a 5-point scale (1 = not native-like at all; 5 = extremely native-like), and (iv) identify the initial consonant as /b/ or /g/. Mean accuracy scores on language judgements was 50.83%, confidence in language judgements was moderate ( $M = 3.5$ ; range = 2.8 to 4.2), and native-ness ratings were similar for native and non-native items. Despite difficulty with correctly identifying

<sup>2</sup> The /bos/ and /gos/ strings are available for listening online at <http://www.medicine.mcgill.ca/srvachew/pubspsresnt.htm>

**Table 1** *Acoustical measurements of mixed Bilingual (Experiment 1), French-only (Experiment 2) and English-only (Experiment 2) tokens. Mean and range values are reported for duration, formant frequencies, and pitch*

	Duration (ms)								Formant frequency (Hz)				Pitch (Hz)	
	Total		Vowel		VOT		/s/		F1 mid		F2 mid		f0	
	bos	gos	bos	gos	bos	gos	bos	gos	bos	gos	bos	gos	bos	gos
Condition														
Mixed														
Max	514	521	255	236	-23	32	228	309	649	818	1772	2109	318	333
Min	445	450	181	170	-9	-21	162	201	356	432	835	904	209	127
Mean	482	477	207	188	-15	7	202	239	491	569	1013	1330	275	257
Range	69	71	74	66	14	53	66	108	293	386	937	1205	109	206
French														
Max	557	483	255	214	-88	-43	293	285	525	504	917	1041	483	229
Min	463	416	157	152	-84	-30	162	172	293	492	720	883	123	112
Mean	497	460	198	187	-85	-34	206	229	441	456	825	961	275	155
Range	94	24	98	62	4	13	131	113	232	12	197	158	360	117
English														
Max	514	521	214	236	12	32	269	309	649	818	1772	2109	257	403
Min	431	423	180	126	8	21	196	201	400	485	772	904	183	111
Mean	469	467	201	176	11	25	231	268	513	576	1098	1337	200	180
Range	83	98	34	110	4	11	73	108	249	333	1000	1205	74	292

the language of the tokens, all participants correctly identified the initial consonant 100% of the time. Thus, all tokens were highly intelligible and acceptable instances of /b/ or /g/ for English and French perceivers. Importantly, stimulus variability – for the whole syllables – was not different enough for tokens to be identified as a French or an English pronunciation by monolingual adults listeners. These data confirm that the stimuli selected were highly intelligible for French and English listeners and none of the items were perceived as accented or non-native.

Visual stimuli were two novel objects of different color and shape developed by Stager and Werker (1997) for a word learning experiment. The objects moved back and forth across the screen asynchronous with their paired auditory stimulus. A multi-colored clockwise spinning water wheel was the visual stimulus for the pre- and post-test trials (Stager & Werker, 1997).

#### Apparatus

Testing took place in a curtained sound proof booth, dimensions 240 × 180 cm. Infants sat on their parent's lap facing a 21-inch Sony television monitor, 1.5 metres from the infant. A Sony DCR TRV27 digital video camera, positioned 10 cm below the monitor, was used to record looking times online. The lens of the camera was positioned through an aperture in the curtain. Auditory stimuli were presented using Audio TRAK, BSI-90 loudspeakers at 65dBA (General Radio USA 1565-B sound-level meter) via a Yamaha RX-396 amplifier, positioned beneath the television monitor. Parents wore headphones playing vocal music to ensure that they did not influence their child's responses during the experiment.

Outside the test booth, a computer (Mac OS X version 10.2.8) running Habit 2000 software (Cohen, Atkinson &

Chaput, 2000) presented the auditory and visual stimuli to the monitor and speakers in the test room, and stored the looking time data. The experimenter viewed the infant on a Panasonic television monitor via the digital video camera in the test booth. Onset and offset of looks to the television monitor during trial presentation were recorded by key-press, and Habit software calculated total looking time per trial.

#### Procedure

Infants were tested using the Switch paradigm. Infants were seated on their caregiver's lap. A flashing red light appeared on a television monitor to center the infant's attention. When the infant looked at the flashing light, a trial commenced and a visual object and auditory stimulus were presented. The word-object pair was presented for as long as the infant fixated the visual display; when the infant looked away for more than 2 seconds, presentation of the word-object pairing ceased. Each new trial began with the flashing red light. The dependent variable was looking time to the word-object pairings. Looking time was calculated online and recorded by key-press.

The habituation phase involved presenting infants with two word-object combinations, for example, 'bowce' – object 1, and 'gowce' – object 2. Looking time to the two word-object pairings was averaged over a sliding three-trial window and the maximum average looking time was computed. Maximum trial length was always 20 seconds. Trial presentation continued until looking time (averaged over a three-trial window) decreased to a criterion of 50% below maximum average looking time, indicating habituation. If the infant did not reach the habituation criterion after 25 trials, the session was aborted. Every block of four habituation trials contained two instances of each word-object pairing presented in random order.

Following habituation, infants were presented two test trials, a 'Same' trial where a familiar word–object pairing was presented, and a 'Switch' trial, in which a familiar word and a familiar object from the habituation phase were paired, but in a *new combination*. Recovery of attention by way of an increase in looking time to the Switch trial is interpreted as evidence that a change in the word–object pairing was detected. As a control for decreases in looking time due to fatigue or boredom, a novel word–object pairing was presented in a pre-test phase before habituation, and in a post-test phase following the test trials. The pre-test and post-test stages consisted of one trial where infants heard a female speaker repeating the syllable /pok/ in an infant-directed speech style, that was paired with presentation of a spinning water-wheel object. Minimum pre-test and post-test listening times were required to ensure that the infant was engaged at the beginning and end of the task.

### Results

Mean and standard deviation of looking time (in seconds) to Control (pre- and post-test), Habituation, and Test trials are reported in Table 2. A preliminary analysis was conducted to determine whether infants' interest was maintained throughout the task by comparing looking times to pre- vs. post-test control trials. A 3 (Group: English, French, Bilingual)  $\times$  2 (Control: Pre, Post) mixed ANOVA with looking time to the Control trials as the dependent variable yielded no effects of language group [ $F(2, 45) = .802, p = .455, \eta^2 = .034$ ] or control [ $F(1, 45) = .448, p = .507, \eta^2 = .010$ ]

**Table 2** Mean and standard deviation of looking time in the Control, Habituation, and Test phases for English, French and bilingual infants in Experiment 1

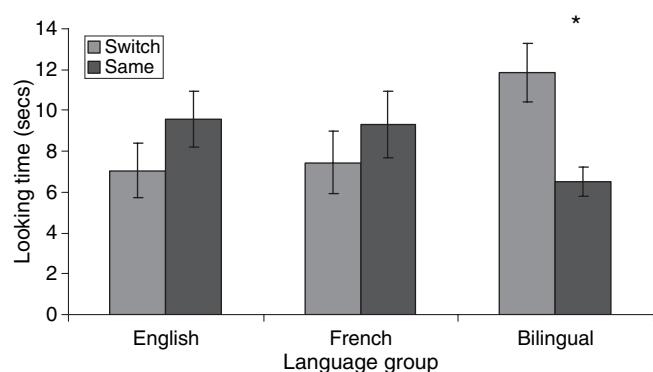
	English	French	Bilingual
Control			
Pre-test			
Mean	16.475	17.750	18.406
SD	5.510	3.039	3.157
Post-test			
Mean	17.725	18.226	18.080
SD	3.424	3.157	2.731
Habituation			
First 3			
Mean	43.237	43.025	49.725
SD	11.528	9.810	8.314
Last 3			
Mean	16.238	18.225	19.563
SD	4.527	4.440	6.394
Recovery			
Last habit.			
Mean	4.106	5.160	5.406
SD	2.346	3.238	4.245
Test			
Switch			
Mean	7.0623	7.450	11.837
SD	5.388	6.206	5.797
Same			
Mean	9.569	9.325	6.518
SD	5.440	6.501	2.812

and no interaction [ $F(2, 45) = .427, p = .655, \eta^2 = .019$ ], confirming that infants' attention was maintained throughout the experiment. A 3 (Group)  $\times$  2 (Recovery: Last, Post) ANOVA comparing looking time to the last habituation trial vs. the post-test trial revealed that infants increased their looking time to the post-test trial significantly [ $F(1, 45) = 366.363, p = .001, \eta^2 = .893$ ]. No group [ $F(2, 45) = .692, p = .506, \eta^2 = .031$ ] or interaction effects [ $F(2, 45) = .163, p = .850, \eta^2 = .007$ ] were found, confirming that infants were equally attentive across the experiment. Thus any difference between language groups is not a function of difference in compliance or attention. A 3 (Group)  $\times$  2 (Habituation: first 3, last 3) ANOVA performed on looking times to the first three vs. last three habituation trials revealed an effect of Habituation [ $F(1, 45) = 741.697, p = .001, \eta^2 = .943$ ], where infants looked longer to the first three habituation trials. No effect of group [ $F(2, 45) = 2.138, p = .130, \eta^2 = .087$ ] or interaction [ $F(2, 45) = 2.407, p = .102, \eta^2 = .097$ ] was found, indicating that regardless of group, infants received equivalent habituation to the word–object pairings.

The main set of analyses concerned infants' performance on the test trials. A 3 (Language: English, French, Bilingual)  $\times$  2 (Trial: Switch, Same) mixed ANOVA was performed on looking times to the Switch versus Same trial displayed in Table 2. There were no main effects of trial [ $F(1, 45) = .072, p = .790, \eta^2 = .002$ ] or language [ $F(2, 45) = .265, p = .768, \eta^2 = .012$ ], but a Trial  $\times$  Language interaction was revealed [ $F(2, 45) = 4.633, p = .015, \eta^2 = .171$ ]. Three *t*-tests were performed on mean looking time to Switch versus Same trials *within* each language group. English, French and Bilinguals' mean looking times to the Switch and Same trials are presented in Figure 1. There was no significant difference in looking time to the Switch versus Same trial for either the English [ $t(15) = 1.270, p = .117, d = .317$ ] or French [ $t(15) = .773, p = .226, d = .193$ ] group. In contrast, bilingual infants looked longer to the Switch trial, thereby demonstrating that they recognized a switch in the word–object pairing [ $t(15) = 3.394, p = .002, d = .848$ ].

### Discussion

Experiment 1 was designed to compare the word learning performance of 17-month-old infants being raised monolingually in English or in French with infants who are growing up bilingually with regular exposure to both languages. We structured a switch task to test infants on their ability to form two novel object + novel word associative pairings using the nonce words, 'bowce' or 'gowce'. The nonce words were selected and recorded to minimize differences in perceptual proficiency across the language groups so that we could look more directly at word learning skill. The nonce words selected differed only in their initial consonant, /b/ and /g/. These



**Figure 1** Mean looking times to the Switch versus Same trial for the English, French and Bilingual infants in Experiment 1. Error bars represent standard error of the mean.

consonants form a native contrast in both English and French; and English variants of this contrast have been found to be discriminated equally well by monolingual (French and English) and bilingual infants in previous studies (Polka & Rvachew, 2005; Polka, Rvachew & Molnar, 2008). The /b/ vs. /g/ contrast also distinguishes many lexical items in each language. The /b/, /g/, /o/ and /s/ segments in our nonce words have very similar surface forms in English and French. French and English pronunciations of each nonce word were also used to label each novel object in the switch task.

Using these stimuli we found that bilingual infants display a clear advantage in forming word/object associations in the switch task when compared to both of their monolingual peer groups. Bilinguals looked significantly longer to the switch trial, providing clear evidence that they detected the violation in the word–object pairing. In contrast, monolinguals performed poorly on this word learning task. Monolinguals did not look significantly longer to the switch trial, indicating a failure to associate the word–object pairings. Although monolingual and bilingual infants were habituated to the same criterion level, the results indicate that the monolinguals did not encode the word–object associations well enough during the habituation stage of the task to detect a violation of the pairing in the test stage.

The bilingual advantage observed here is consistent with research with preschool and school-age children showing bilingual children to outperform their monolingual peers in certain cognitive, linguistic, and metalinguistic tasks. Thus, one interpretation of the present finding is that bilingual 17-month-olds are better prepared for the task of word learning compared to their monolingual peers. However, the observed bilingual advantage could also be due to the specific properties of the stimuli implemented in our switch task. This alternative is clearly viable given that the monolingual groups not only performed below the level of bilinguals; they failed altogether. We structured our switch task to conform to prior studies using this paradigm; we used identical objects, test protocol, habituation criterion, and methods for developing auditory labels (infant-directed

speech, multiple tokens). We departed from earlier methodology in only one respect – we included both French and English pronunciations of the object labels. Recall also that in prior work, monolingual 17-month-olds as a group have been uniformly successful in the switch task when phonetically similar native consonant contrasts are used (Werker *et al.*, 2002). Thus, the failure of the monolingual infants in the present study is clearly inconsistent with previous word learning studies. At the same time, the present study shows that when bilingual infants are tested using stimuli constructed to emulate their native language input – when the stimuli include variants from both languages – they pattern just like 17-month-old monolinguals from previous word learning studies (e.g. Werker *et al.*, 2002).

These facts clearly suggest that the bilinguals' success may relate to the infants' familiarity with the bilingual nature of the auditory strings (word labels). Thus, despite our efforts to minimize the influence of language-specific perception in our word learning task, we may have failed to achieve this objective. Specifically, the bilingual advantage in this task could be due to the closer match between the test stimuli and their ambient language input. Likewise, the monolingual failure may be due to the poorer match between the test stimuli and their input resulting from the inclusion of both French and English pronunciations of each word form. Overall, the present findings point to a clear link between infants' prior language experience and their performance in a word learning task at 17 months of age but the nature of this link is unclear. In Experiments 2 and 3 we examine potential factors underlying this link to better understand the bilingual infants' superior performance in the switch task.

## Experiment 2

In Experiment 1, monolingual infants failed when the switch task involved both French and English pronunciations of 'bowce' and 'gowce'. This finding suggests that 17-month-old infants can succeed in the switch task when there is a correspondence between the stimuli presented in the task and their prior language experience. However, it is unclear what aspect of this correspondence is critical. Monolingual infants in Experiment 1 may have had difficulty recognizing the unfamiliar /b/ and /g/ variants occurring in the target portion of the nonce word tokens. As outlined above, the /b/ and /g/ phones are phonetically similar, but not identical, in English and French. The release bursts may differ and the VOT values included in the set of bowce and gowce tokens are acceptable in both languages but included some VOT values that do not line up with the modal values that characterize the language input of the monolinguals.

Experiment 2 was designed to determine whether the presence of unfamiliar pronunciations of the nonce



words in Experiment 1 had a negative impact on the monolingual performance in the switch task. We tested monolingual 17-month-olds (English and French) using English-only and French-only pronunciations of 'bowce' and 'gowce'. New items selected from the original recording of our bilingual talker were added to the tokens from Experiment 1 to create an English-only and a French-only stimulus set. In addition, prevoicing was added to French tokens to make them more French-like; English tokens were made more English-like by modifying /b/ and /g/ phones of the word items. English infants were tested with the English-only stimuli; French infants were tested with the French-only stimuli. We expected both groups of monolingual infants to succeed, as has been shown in earlier studies. This outcome would support the interpretation that monolingual infants' failure with the mixed language stimuli in Experiment 1 was due to their lack of familiarity with the surface form of some of the 'gowce' and 'bowce' items used in that task.

## Method

### Participants

Thirty-two full-term 17-month-old monolingual infants, 16 English-learning (eight boys, eight girls,  $M_{age} = 522$  days,  $SD = 8$  days) and 16 French-learning (eight boys, eight girls,  $M_{age} = 526$  days,  $SD = 6$  days) participated in this study. To qualify for this study infants were estimated to have at least 90% exposure to either English or French from birth, with minimal to no exposure to a second language, as determined by parental interview and questionnaire. Infants were recruited from Montreal, Canada, and had no known hearing impairment. An additional 15 infants were tested but their data were discarded because of fussiness (11), crying (one) and equipment failure (three).

### Stimuli

Monolingual stimulus strings were created by returning to the original stimulus recording and selecting three new French tokens of /bos/ and /gos/, respectively, and three each of English /bos/ and English /gos/. Tokens that provided the best match acoustically to the three French and three English tokens from Experiment 1 were selected. The new tokens were pooled with the tokens from Experiment 1 to make six-token strings of French-only /bos/, French-only /gos/, English-only /bos/ and English-only /gos/.

The French tokens were then slightly modified to best approximate the production of French stops by native monolingual Canadian French speakers (Macleod & Stoel-Gammon, 2005). Canadian French /b/ and /g/ tend to be prevoiced, so using TFR and Praat, we extracted prevoicing (negative VOT) from one of the French tokens, adjusted (lengthened)

prevoicing to the modal value VOT in French using Praat software and then spliced the prevoicing to the beginning of each token so that all French tokens of /bos/ and /gos/ contained prevoicing prior to burst release. Although there was variation in prevoiced VOT (see Table 1), all instances were in the acceptable range for their native category (Macleod & Stoel-Gammon, 2005).

Canadian English stops typically have short lag voice onset time, although prevoicing is sometimes present (Sundara, 2005). English tokens with prevoicing prior to burst release were edited to create a short-lag positive value using TFR and Praat, thus ensuring that all tokens were short lag, as per modal production of English stops (Sundara, 2005). VOT was the only acoustic property that was altered in both French and English tokens.

As per Experiment 1, /bos/ and /gos/ strings were 30 seconds long with a 1000 ms ISI separating tokens. Acoustics of Experiment 2 stimuli are presented in Table 1.

### Adult ratings

As in Experiment 1, French and English tokens were rated for degree of native-likeness by five native speakers of French and English, respectively (different raters from Experiment 1). Mean accuracy scores on language judgments was 66% for English speakers and 54% for French speakers, confidence in language judgments was moderate, and native-ness ratings were similar for English and French listeners. Despite difficulty with correctly identifying the language of the tokens, all participants correctly identified the initial consonant 100% of the time, indicating that the tokens were highly intelligible and acceptable instances of /b/ or /g/ for English and French perceivers and that our modification of VOT did not affect initial phoneme identification.

### Apparatus and procedure

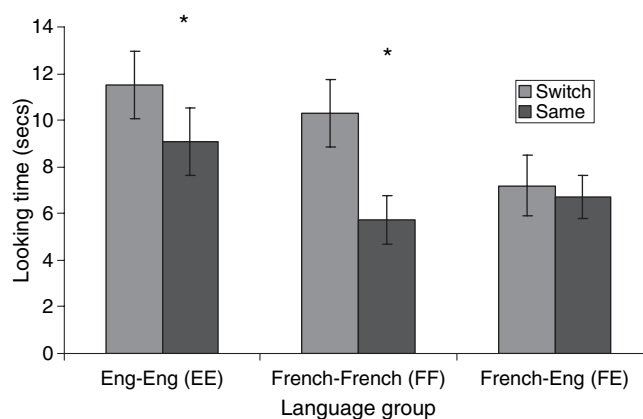
Apparatus and procedure (Switch task) were identical to Experiment 1; only the audio stimuli were changed.

### Results

Mean and standard deviation of looking time (in seconds) to Control (pre- and post-test), Habituation, and Test trials are reported in Table 3. Analysis of Variance (ANOVA) tests were performed on looking time data. Preliminary analyses were conducted to determine whether infants' interest was maintained throughout the task. A 2 (Group: English, French)  $\times$  2 (Control: Pre, Post) mixed ANOVA with looking time as the dependent variable yielded no main effects of group [ $F(1, 30) = 1.082$ ,  $p = .307$ ,  $\eta^2 = .035$ ], control [ $F(1, 30) = .188$ ,  $p = .668$ ,  $\eta^2 = .006$ ] or interaction [ $F(1, 30) = .364$ ,  $p = .551$ ,  $\eta^2 = .012$ ]. A 2 (Group)  $\times$  2

(Recovery: Last, Post) ANOVA revealed longer looking to post-test [ $F(1, 30) = 218.457, p = .001, \eta^2 = .886$ ], and no effect of group [ $F(1, 30) = .058, p = .811, \eta^2 = .002$ ] or interaction [ $F(1, 30) = .721, p = .403, \eta^2 = .025$ ], confirming that infants' attention was maintained throughout the experiment. A 2 (Group)  $\times$  2 (Habituation: first block, last block) revealed an effect of habituation [ $F(1, 30) = 256.155, p = .001, \eta^2 = .895$ ], with longer looking to the first block of three habituation trials, no effect of group [ $F(1, 30) = .277, p = .603, \eta^2 = .009$ ], and a group  $\times$  habituation interaction [ $F(1, 30) = 195.650, p = .016, \eta^2 = .179$ ]. However, two paired samples *t*-tests revealed that both groups looked significantly longer to the first three habituation trials and are therefore equivalent in habituation.

The main analyses concerned infants' performance on the test trials which is plotted in Figure 2. Abbreviation EE denotes English infants tested on English stimuli, and FF denotes French infants tested on French stimuli. Looking times to each trial type were compared in separate paired samples *t*-tests. Looking times to Switch versus Same trials were significant for both the EE [ $t(15) = 1.753, p = .049, d = .438$ ] and FF groups [ $t(15) = 3.090, p = .003, d = .772$ ], confirming that in each monolingual group, infants recognized the switch in word-object pairing when word items contained only the native language pronunciations of the nonce words. To further assess whether the monolingual groups showed comparable performance levels in the switch task we also conducted a 2 (Group: EE, FF)  $\times$  2 (Trial: Switch, Same)



**Figure 2** Mean looking times to the Switch versus Same trial for the English-English (EE, Experiment 2) French-French (FF, Experiment 2) and French-English (FE, Experiment 3) infants. Error bars represent standard error of the mean.

mixed ANOVA on looking times to the Switch versus Same trial. This ANOVA yielded a significant effect of Trial [ $F(1, 30) = 11.920, p = .002, \eta^2 = .284$ ], but no effect of Group [ $F(1, 30) = 1.829, p = .186, \eta^2 = .057$ ], and no Trial  $\times$  Group interaction [ $F(1, 30) = 1.106, p = .301, \eta^2 = .036$ ]. This outcome indicates that the monolingual English and monolingual French infants were equally successful in the switch task when tested with productions of /bos/ and /gos/ in their native language.

### Discussion

Experiment 2 investigated whether monolingual infants succeed on the Switch task when the /bos/ and /gos/ variants match their monolingual input by containing only familiar native language phonetic variants. We found that this was the case for both monolingual groups – French infants tested with French variants, and English infants tested with English variants. This outcome supports the interpretation that infants' failure to learn the word-object associations in Experiment 1 was due to the presence of non-native phonetic variability in the /bos/ and /gos/ items. However, our efforts to create a better match to native language VOT values for each monolingual group also reduced the overall range of VOT values within each monolingual stimulus set compared to the mixed language stimuli. Thus, the better performance in our English-only and French-only conditions could also be due to the narrowing of VOT variation rather than the increased match to native productions *per se*. This issue was addressed in Experiment 3.

### Experiment 3

Monolingual infants' success in Experiment 2 could be due to a general reduction in phonetic variability or to

**Table 3** Mean and standard deviation of looking time in the Control, Habituation, and Test phases for Experiment 2 English-English (EE) and French-French (FF) groups, and Experiment 3 French-English (FE) group

	EE	FF	FE
Control			
Pre-test			
Mean	16.9000	15.065	16.073
SD	3.954	5.479	4.390
Post-test			
Mean	16.713	16.2000	14.713
SD	3.827	4.441	5.687
Habituation			
First 3			
Mean	41.763	36.700	38.838
SD	12.901	10.857	11.970
Last 3			
Mean	16.419	18.350	16.150
SD	5.416	5.428	5.783
Recovery			
Last habit.			
Mean	3.248	4.456	5.025
SD	2.604	2.356	2.902
Test			
Switch			
Mean	11.494	10.294	7.175
SD	5.781	6.132	5.209
Same			
Mean	9.069	5.744	6.700
SD	5.795	4.197	3.732

the closer match with their native language in the English-only and French-only stimulus sets. To tease apart these two accounts we tested another group of monolingual French infants on the English-only stimulus set. If the success of monolingual infants in Experiment 2 was due solely to the general reduction in VOT values then we would expect monolingual French infants to perform equally well when tested with the English-only stimuli or the French-only stimuli. Alternatively, if the closer match to the surface forms of the native language contributed to the monolingual infants' success in Experiment 2, then we expect French infants to do better when tested with the French-only stimuli compared to the English-only stimuli.<sup>3</sup>

### Method

#### Participants

The participants were 16 full-term monolingual French-learning 17-month-old (eight boys, eight girls,  $M_{age} = 526$  days,  $SD = 13$  days) infants recruited from Montreal, Canada who had at least 90% language exposure to French, and reportedly normal hearing. Sixteen infants were tested in the Switch task with the English variants of /bos/ and /gos/ (FE) used in Experiment 2. Data from eight additional infants were discarded due to fussiness/failure to complete the task, and another three infants were discarded due to incorrect language background.

Data from FE infants were compared with the performance of FF infants in Experiment 2.

#### Stimuli

The stimuli were English /bos/, /gos/ strings from Experiment 2, that is, for these French infants the variants presented non-native phonetic variability.

#### Apparatus and procedure

Apparatus and procedure was identical to Experiments 1 and 2.

### Results

Mean and standard deviation of looking time (in seconds) to Control (pre- and post-test), Habituation, and Test trials are reported in Table 3. Analysis of Variance (ANOVA) tests were performed and statistics ( $df$ ,  $F$ ,  $p$ , and  $\eta^2$ ) are reported in Table 3. Abbreviation FE denotes French infants tested on English stimuli, and FF denotes French infants tested on French stimuli. A preliminary analysis was conducted to determine

whether infants' interest was maintained throughout the task. A  $2$  (Group: FF, FE)  $\times$   $2$  (Control: Pre, Post) mixed ANOVA with looking time as the dependent variable yielded no main effect of group [ $F(1, 30) = .006$ ,  $p = .940$ ,  $\eta^2 = .03$ ] or control [ $F(1, 30) = .004$ ,  $p = .984$ ,  $\eta^2 = .02$ ] and no interaction [ $F(1, 30) = 1.113$ ,  $p = .296$ ,  $\eta^2 = .40$ ], confirming that infants' attention was maintained throughout the experiment. A  $2$  (Group)  $\times$   $2$  (Recovery: Last, Post) ANOVA uncovered an effect of recovery only [ $F(1, 30) = 103.869$ ,  $p = .001$ ,  $\eta^2 = 3.72$ ], where looking was longer to the post-test versus the last habituation trial. There was no group effect [ $F(1, 30) = .085$ ,  $p = .771$ ,  $\eta^2 = .110$ ] or interaction [ $F(1, 30) = 1.232$ ,  $p = .270$ ,  $\eta^2 = .410$ ]. A  $2$  (Group)  $\times$   $2$  (Habituation: first block, last block) ANOVA highlighted infants' longer looking to the first three habituation trials compared to the last three habituation trials [ $F(1, 30) = 64.618$ ,  $p = .001$ ,  $\eta^2 = 2.94$ ], but there was no group effect [ $F(1, 30) = .868$ ,  $p = .355$ ,  $\eta^2 = .26$ ] and no interaction [ $F(1, 30) = 3.477$ ,  $p = .067$ ,  $\eta^2 = .68$ ]. Thus infants in the FF and FE groups maintained attention across the experiment, were habituated to an equivalent level, and showed equivalent recovery to post-test.

Looking times for each group are presented in Figure 2. A  $2$  (Group: FE, FF)  $\times$   $2$  (Trial: Switch, Same) mixed ANOVA was performed on looking times to the Switch versus Same trial. There was no effect of group [ $F(1, 30) = .568$ ,  $p = .457$ ,  $\eta^2 = .019$ ]; however, there was a significant effect of Trial indicating longer looking to Switch versus Same trials [ $F(1, 30) = 6.650$ ,  $p = .015$ ,  $\eta^2 = .181$ ], and a Trial  $\times$  Group interaction [ $F(1, 30) = 4.373$ ,  $p = .045$ ,  $\eta^2 = .127$ ]. As reported for Experiment 2, a  $t$ -test revealed that FF infants looked significantly longer to the Switch versus Same trials, indicating successful word learning. However, a  $t$ -test for the FE group revealed that looking times to Switch versus Same trials were not significantly different [ $t(15) = .372$ ,  $p = .357$ ,  $d = .093$ ], indicating that monolingual French infants failed to notice the switch in word-object pairing when word items contained non-native English phonetic variants. The Trial  $\times$  Group interaction confirms that French infants clearly performed better in the switch task when tested with French pronunciations than they did when presented with English pronunciations of the nonce words.

### Discussion

The results of Experiment 3 confirm that French infants' success in learning word-object associations in Experiment 2 is due to familiarity with the phonetic variants and not simply to the reduction in VOT variation in the English-only and French-only stimulus sets. Specifically, French infants learned the word-object associations when the French tokens were presented, as evidenced by longer looking to the novel word-object pairing at test. This position is further supported by the

<sup>3</sup> English infants were not tested on French-only stimuli for practical reasons only; English infants who do not receive regular exposure to French are much more difficult to access in Montreal.

failure of the French infants to learn the word–object associations when the test items failed to provide a good match to French infants’ real-world native language input. Specifically, French infants failed to learn the word–object associations under conditions in which (i) English-only tokens were presented (Experiment 3) and (ii) French and English tokens were combined (Experiment 1). Recall that the differences between the French and English variants are subtle; the non-native productions are not detectable by monolingual English and French adults as accented or foreign-sounding. Thus, the present results show that 17-month-old infants are more sensitive to sub-phonemic variability than adults and this sensitivity impacts their performance in the switch task.

## General discussion

Experiment 1 shows that bilingual 17-month-olds have a processing advantage relative to their monolingual peers in some word learning tasks. Specifically bilingual infants acquiring French and English succeeded in forming an associative link between a novel word form and a novel object when the stimuli used to name the object included both French and English productions of the nonce words /bos/ and /gos/. Bilingual infants appear to have formed a robust word/object pairing indicated by an increase in looking to the switch trial (compared to end of habituation), a sign that they noticed the violation of the word–object pairing. Bilingual infants also showed a further decrease to the same pair (compared to the end of habit), a sign that this pairing is well encoded and they are bored with it. The monolingual infants failed when tested using ‘bilingual’ stimuli as indexed by maintained attention to the same trial (compared to end of habituation phase), and suggesting that they had not yet encoded the object/label pairing.

Experiments 2 and 3 show that monolinguals’ failure in Experiment 1 was due to the presence of the unfamiliar, other-language variants in the stimulus set. When stimuli were restricted to native language productions, monolinguals were able to link word/object events well enough to detect a switch in the word/object pairing. Simply reducing the range of phonetic variation within a set of unfamiliar surface word forms did not lead to success in the switch task for monolingual infants. Thus, at 17 months bilingual infants can process a more diverse set of pronunciations of the same word form when learning new object labels in comparison to their monolingual peers.

We offer several explanations for the bilingual processing advantage demonstrated here which are not mutually exclusive. First, at 17 months of age bilingual infants may be very sensitive to the surface form of words, just like their monolingual peers. This gives them an edge in some word processing tasks because they are

familiar with a larger set of surface forms. This richer phonetic experience supports language processing in their specific bilingual milieu. This specific benefit is clearly evident in the present study. It is also possible that early bilingualism requires and promotes more flexible and robust linguistic skills from a very early age. Although the current study does not provide direct support for such global benefits, this is not incompatible with our findings. Thus, bilinguals’ experience with a more diverse phonetic space may lead to more flexible representations for some phonetic categories that can facilitate processing of non-native speech or foreign accents.

Second, despite apparent differences, bilingual infants, like their monolingual peers, appear to be developing phonetic perception skills that facilitate language acquisition in their natural environment, thus revealing a link between speech perception and emerging language skills. The present findings indicate that the stimulus properties and the prior linguistic experience of the infant are interacting factors that contribute to cognitive load in word learning tasks. We propose that word learning is challenging at 17 months of age, regardless of whether an infant is monolingual or bilingual – but it is challenging in different ways. Bilinguals face more complex phonetic processing demands; this challenge may be beneficial in some respects but not in others. The present study indicates that bilingual infants cope well with the extra variability associated with phonological elements that are shared across their native languages and word learning involving these converging elements is not compromised in the bilingual infant. In fact, this extra variability may lead to more robust and flexible representations for these phonological structures that are, of course, present in both their languages. We do not know how such detailed phonetic representation may advantage the bilingual infant who is building a lexicon as we do not have vocabulary data for these infants. Nevertheless, it is possible that sensitivity to fine phonetic detail assists encoding of lexical items at the word learning stage, but that differences between bilinguals and monolinguals may be short-lived and disappear once the lexicon is better formed.

However, bilinguals will also encounter phonetic structures that are difficult to access efficiently in a mixed language context; this may result in more effortful or slower word learning in bilingual infants compared to their monolingual peers who encounter the same structures in a unilingual context. For example, Fennell *et al.* (2007) found that English-French bilingual infants failed the switch task at 17 months with /bIh/ and /dIh/ nonce words produced in one of the bilinguals’ languages, and when the phones differed in phonetic realization across the bilinguals’ languages. Specifically, both /b/ and /d/ differ in voicing when produced in Canadian French and Canadian English, and French /d/ and English /d/ also differ in place of articulation; French /d/ is produced with a dental place of



articulation, whereas /d/ in English is alveolar. The vowel /I/ is a phonemic category in English but this vowel occurs as an allophonic variant of /i/ in Canadian French. Because of these multiple mismatches between the word forms across their two languages, the bilingual infants may be less certain how to interpret the /bIh/–/dIh/ tokens unless they can clearly recognize the target language, which may be difficult to do when listening to brief isolated segments. In contrast, the success of our bilinguals in Experiment 1 may be specific to the fact that native mixed-language stimuli reflecting bilingual input were used.

Clearly, the bilingual faces an added measure of complexity with respect to phonetic processing, regardless of whether the phonetic representations of each language are independent or interdependent. If representations are independent then it becomes important to sort or recognize speech patterns by language in order to access the appropriate representation. If representations are interdependent, then the infant faces a more complex or crowded phonetic space and must resolve cue conflict or overlap to develop a representation that provides efficient access to both phonological systems. Results on adults who acquired two languages from birth show that simultaneous bilinguals develop interdependent representations (Sundara & Polka, 2008). Interdependent representations for phonological structures that are similar across bilinguals' two languages will likely encompass a wider range of phonetic variability given that languages are rarely identical in their subphonemic details. Thus, more flexibility in linguistic categorization is likely required in bilingual acquisition and it is possible that bilinguals' experience with abundant phonetic variability is a precursor to the metalinguistic advantages seen in bilingual preschoolers. As reviewed earlier, work by Bialystok (1999, 2001) shows that metalinguistic awareness is stimulated by bilingual exposure. Future research should investigate this possibility.

The present study also reveals that considerable effort is required to equate the perceptual demands associated with processing of word forms for bilingual and monolingual infants. Bilinguals have long been compared to the monolingual 'gold standard' for language development. Our findings and also those of Fennell *et al.* (2007) suggest that speech processing demands are typically not equivalent when bilingual and monolingual infants are tested on identical stimulus sets. This means that to fully appreciate the similarities and differences between monolingual and bilingual acquisition we must design studies from both perspectives, and treat monolinguals and bilinguals as typical in their own right.

To summarize, the first steps in word learning are easier when the phonetic shoes fit. At 17 months of age, monolingual infants have difficulty taking steps in word learning when put in the bilingual infant's shoes, but bilingual infants can run in these shoes. Correspondingly,

if the shoe fits, monolinguals also take great strides in word learning. The bilingual advantage observed in the present study is clearly related to the specificity of representations formed during early experience with their native language(s). The challenges of acquiring two languages simultaneously may also require and promote linguistic or cognitive skills that have a global positive impact on infant development. This is an exciting avenue for future exploration.

## Acknowledgements

The authors thank Eugenie Batterby, Lauren Tittley, Eleanor Roff, and Monika Molnar for testing infants, Rachael Carson and Alexandra Giorgi for scheduling participants, and extend appreciation to the parents and infants who participated. This work was supported by a postdoctoral fellowship from the Centre for Research on Language, Mind, & Brain to K. Mattock, a Social Sciences and Humanities Research Council of Canada grant (410-2006-1841) to L. Polka and a Natural Sciences and Engineering Research Council grant (261363-07) to S. Rvachew. We thank Chris Fennell, Krista Byers-Heinlein, and Janet Werker for helpful discussion and two anonymous reviewers for their comments on this manuscript.

## References

- Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development*, **48**, 1009–1018.
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, **70** (3), 636–644.
- Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, **21**, 169–181.
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, **2** (1), 29–49.
- Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, **46**, 217–243.
- Bosch, L., & Sebastián-Gallés, N. (2003b). Language experience and the perception of a voicing contrast in fricatives: infant and adult data. Paper presented at the 15th International Conference of the Phonetic Sciences, Barcelona, Spain.
- Bosch, L., & Sebastián-Gallés, N. (2005). Developmental changes in the discrimination of vowel contrasts in bilingual infants. In J. Cohen, K. McAlister, K. Roltad, & J. MacSwan, (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 354–363). Somerville, MA: Cascadia Press.
- Brown, C. (2000). *Speech perception and phonological acquisition from infant to adult: Second language acquisition and linguistic theory*. Oxford: Blackwell Publishing.

- Brown, C., & Matthews, J. (1997). The role of feature geometry in the development of phonemic contrasts. In S.J. Hannah & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 67–112). Amsterdam: John Benjamins.
- Burns, T.C., Yoshida, K.A., Hill, K., & Werker, J.F. (2007). Bilingual and monolingual infant phonetic development. *Applied Psycholinguistics*, **28** (3), 455–474.
- Caramazza, A., Yeni-Komshian, G., 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** (2), 421–426.
- Cohen, L.B., Atkinson, D.J., & Chaput, H.H. (2000). *Habit 2000: A new program for testing infant perception and cognition* (Version 2.2.5c). Computer software.
- Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: analysis and control in young Swedish-English bilinguals. *Applied Psycholinguistics*, **20**, 1–20.
- Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *Journal of Cross Cultural Psychology*, **9**, 131–149.
- Feldman, C., & Shen, M. (1971). Some language-related cognitive advantages of five-year-olds. *Journal of Genetic Psychology*, **118**, 235–244.
- 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.
- Fennell, C.T., & Werker, J.F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, **46** (2), 245–264.
- Ferguson, C.A., & Farwell, C.B. (1975). Words and sounds in early language acquisition. *Language*, **51**, 419–439.
- Gurski, C. (2006). Voice onset time as a parameter for identification of bilinguals. In *Proceedings of the International Association for Forensic Phonetics and Acoustics*, 23–26 July, Göteborg University, Sweden.
- Ianco-Worrall, A.D. (1972). Bilingualism and cognitive development. *Child Development*, **43** (4), 1390–1400.
- Kuhl, P.K., Conboy, B.T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: implications for the 'critical period'. *Language Learning and Development*, **1**, 237–264.
- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation of native language phonetic perception between 6 and 12 months. *Developmental Science*, **9**, 13–21.
- Macleod, A.N., & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, **3** (2), 118–127.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: evidence for perceptual reorganization. *Infancy*, **10** (3), 241–265.
- Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in vowel acquisition. *Cognition*, **64**, B9–B17.
- Pater, J., Stager, C., & Werker, J. (1998). Additive effects of phonetic distinctions in word learning. In *Proceedings of 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America*, 2049–2050.
- Pater, J., Stager, C., & Werker, J. (2004). The lexical acquisition of phonological contrasts. *Language*, **80**, 361–379.
- Picard, M. (1987). *An introduction to the comparative phonetics of English and French in North America* (Studies in the sciences of language series, vol. 7). Amsterdam: John Benjamins.
- Pierrehumbert, J.B. (1990). Phonological and phonetic representation. *Journal of Phonetics*, **18**, 375–394.
- Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/–/ð/ perception: evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, **109** (5), 2190–2201.
- Polka, L., & Rvachew, S. (2005). The impact of otitis media with effusion on infant phonetic perception. *Infancy*, **8**, 101–117.
- Polka, L., Rvachew, S., & Mattock, K. (2007). Experiential influences on speech perception and speech production in infancy. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 153–172). Oxford: Blackwell.
- Polka, L., Rvachew, S., & Molnar, M. (2008). The effect of a nonspeech distractor on infant speech perception. *Cognition*, **13** (5), 421–439.
- Polka, L., & Werker, J.F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 421–435.
- Sebastián-Gallés, N., & Bosch, L. (in press). Developmental shift in the discrimination of vowel contrasts in bilingual infants: is the distributional account all there is to it? *Developmental Science*.
- Stager, C.L., & Werker, J. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, **388**, 381–382.
- 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., & Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition*, **108**, 232–242.
- Sundara, M., & Polka, L. (2008). Discrimination of coronal stops by bilingual adults: the timing and nature of language interaction. *Cognition*, **106** (1), 234–258.
- Sundara, M., Polka, L., & Baum, S. (2006a). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, **9** (1), 97–114.
- Sundara, M., Polka, L., & Genesee, F. (2006b). Language experience facilitates discrimination of /d/–/ð/ in monolingual and bilingual acquisition of English. *Cognition*, **100** (2), 369–388.
- Swingle, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, **76**, 147–166.
- Tsao, F., Liu, H., & Kuhl, P.K. (2004). Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Development*, **75** (4), 1067–1084.
- Vygotsky, L.S. (1962). *Thought and language*. New York: Wiley.
- 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.), *Weaving a lexicon* (pp. 79–109). Cambridge, MA: MIT Press.

- 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., Gilbert, J.H.V., Humphrey, K., & Tees, R.C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, **52**, 349–355.
- Werker, J., & Stager, C. (2000). Developmental changes in infant speech perception and early word learning: is there a link? In J. Pierrehumbert & M. Broe (Eds.), *Papers in laboratory phonology V: Acquisition and the lexicon* (pp. 181–193). Cambridge: Cambridge University Press.
- Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization in the first year of life. *Infant Behavior and Development*, **7**, 49–63.
- Werker, J.F., & Yeung, H.H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, **9** (11), 519–527.
- Yoshida, K.A., Fennell, C.T., Swingle, D., & Werker, J.F. (2009). Fourteen-month-old infants learn similar sounding words. *Developmental Science*, **12** (3), 412–418.

Received: 15 December 2006

Accepted: 11 December 2008