

Fine-tuning language discrimination: Bilingual and monolingual infants' detection of language switching

Esther Schott^{1,2}  | Meghan Mastroberardino^{1,2}  |
Eva Fourakis³  | Casey Lew-Williams³  | Krista Byers-Heinlein^{1,2} 

¹Concordia University, Montreal, Quebec, Canada

²Centre for Research on Brain, Language and Music, Montreal, Quebec, Canada

³Princeton University, Princeton, New Jersey, USA

Correspondence

Esther Schott, 7141 Sherbrooke St. West, Montréal, QC, Canada H4B 1R6.
Email: esther.schott@mail.concordia.ca

Funding information

National Institute of Child Health and Human Development, Grant/Award Number: R01HD095912 and R03HD079779; Fonds de Recherche du Québec-Société et Culture; American Speech-Language-Hearing Foundation; Concordia University Research Chair; Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: 2018-04390 and 402470-2011

Abstract

The ability to differentiate between two languages sets the stage for bilingual learning. Infants can discriminate languages when hearing long passages, but language switches often occur on short time scales with few cues to language identity. As bilingual infants begin learning sequences of sounds and words, how do they detect the dynamics of two languages? In two studies using the head-turn preference procedure, we investigated whether infants ($n = 44$) can discriminate languages at the level of individual words. In Study 1, bilingual and monolingual 8- to 12-month-olds were tested on their detection of single-word language switching in lists of words (e.g., “dog... *lait* [fr. milk]”). In Study 2, they were tested on language switching within sentences (e.g., “Do you like the *lait*?”). We found that infants were unable to detect language switching in lists of words, but the results were inconclusive about infants' ability to detect language switching within sentences. No differences were observed between bilinguals and monolinguals. Given that bilingual proficiency eventually requires detection of sound sequences across two languages, more research will be needed to conclusively understand when and how this skill emerges. Materials, data, and analysis scripts are available at <https://osf.io/9dtwn/>.

KEYWORDS

bilingualism, infants, language discrimination, language switching, word learning

1 | INTRODUCTION

All bilingual environments involve periodic switching between languages either within or across speakers, and an ability to detect these switches is foundational to bilingual infants' successful language acquisition and later language use. Previous research has tested infants' detection of language switches at longer time scales, such as following narrative-like passages (e.g., Bosch & Sebastian Galles, 2001; Nazzi et al., 1998). However, little is known about infants' ability to detect language switches at shorter time scales. Given that language learning in the first year of life is largely focused on sequences of sounds and words, how do infants process switches at the level of individual words? The ability to detect a transition between languages in a sentence such as "Look at the *chien* [fr. dog]" would presumably support bilingual infants' emerging representations of words in two languages (Byers-Heinlein, 2014; Curtin et al., 2011). The present studies tested whether bilingual and monolingual 8- to 12-month-old infants, who are just beginning to learn words, can detect single-word language switches.

Language switches are frequent in many bilingual infants' language environments. While the amount of language switching varies across families and communities, most bilingual parents switch languages in interactions with their children, with some children hearing language switches in as many as 2/3 of utterances (Bail et al., 2015; Byers-Heinlein, 2013; Kremin et al., 2021). Language switching can take several forms. First, parents may alternate their languages at the point of a sentence boundary (e.g., "*Regarde ici* [fr. look here]! Do you see the dog?"). Second, they may switch languages within a sentence by borrowing a single word from another language ("Do you want your *toutou* [fr. stuffed animal]?"). Finally, they can switch across isolated words, such as when teaching new words (e.g., "Look! Dog! *Chien* [fr. dog]!" Byers-Heinlein, 2013). Some of these types of language switches may be more difficult for infants to detect than others.

Most research on infant language discrimination has focused on cross-sentence switching in long passages of speech. In one study, newborns born to bilingual mothers, who were exposed to both English and Tagalog during pregnancy, were tested in a habituation paradigm (Byers-Heinlein et al., 2010). After hearing sentences in one of their maternal languages until they lost interest, newborns showed renewed interest only when the stimuli switched to the other language, suggesting that they could detect the language change. Thus, following prolonged exposure to a language in a long passage, even newborn bilinguals are able to detect language switches for some language pairs. More evidence for infants' ability to discriminate languages across longer time scales comes from studies of Spanish–Catalan bilingual 4-month-olds (Bosch & Sebastian Galles, 2001) and Spanish–Basque 3.5-month-old bilinguals (Molnar et al., 2014), as well as studies including monolingual infants, from birth to 5 months of age (e.g., Bahrack & Pickens, 1988; Nazzi et al., 1998, 2000). In these studies, language rhythm, an aspect of prosody that is largely related to variation in consonant and vowel duration, appears to play a role in detecting language switches (Gasparini et al., 2020). Rhythmic information is richer in multisyllabic utterances (Ramus et al., 1999); thus, these studies leave open the question of whether infants can detect language switches over shorter time scales where rhythmic cues are less available.

Some recent evidence suggests that at least older infants might be able to detect language switches across shorter time scales, for example when switches involve a single word (e.g.,

“Do you see the *chien* [fr. dog] over there?”). In an eye-tracking study (Byers-Heinlein et al., 2017), 20-month-olds looked less at a target image when the target word was language-switched (“Look at the *chien*”) compared to when the target word was of the same language (“Look at the dog”). Similar results were found for English–Spanish bilingual 18- to 30-month-olds (Morini & Newman, 2019; Potter et al., 2018, 2019). Additional evidence for detection of language-switched words comes from a study of English–Welsh bilingual 2- to 3-year-olds’ event-related potentials (Kuipers & Thierry, 2012). Together, these studies suggest that by their second year of life, bilingual infants have some ability to detect a language switch at the level of individual words.

Even during the first year of life, infants possess prerequisite abilities that could help them to detect single-word language switches. While direct evidence for this is sparse, several studies indicate that infants’ budding knowledge of sounds and words could enable them to detect such variation. First of all, monolingual infants learn about the sound patterns in their native language prior to the onset of word production (Werker, 2018), and 6- to 9-month-olds show precursors of word comprehension for frequent words by looking at a labeled referent (Bergelson & Swingley, 2012, 2015; Kartushina & Mayor, 2019). While such studies have yet to be done with bilingual infants, ample evidence suggests that bilingual infants make rapid gains in learning about the sounds (i.e., phonetic inventory) and allowable combinations of sounds (phonotactics) of their two languages. Within the first year, both monolinguals (Kuhl et al., 2006; Werker & Tees, 1984) and bilinguals (Albareda-Castellot et al., 2011; Burns et al., 2007; Sundara et al., 2008) become perceptually specialized to the sounds of their native language, and further, bilingual infants are sensitive to patterns of sounds that are typical in their dominant language (Sebastian Galles & Bosch, 2002). Moreover, bilingual infants start to recognize the sound patterns of frequently-heard familiar words by 11 months (Vihman et al., 2007), though as mentioned above more studies are needed on bilinguals’ early word comprehension. Thus, it is possible that even young bilinguals could exploit their developing knowledge of sounds and sound patterns to detect single-word language switches, at least for certain highly frequent and familiar words.

Here, in two studies, we investigated whether infants are able to detect single-word language switches. We tested 8- to 12-month-old infants, because at this age they have already acquired knowledge about the sounds and sound patterns in their native language(s) and have started to learn about the words in their language(s) as well. Infants listened to single-language and switched-language speech, and their listening times (operationalized as looking times) were measured using the head-turn preference procedure. Words were presented in a word list in Study 1 or following a naturally produced sentence frame in Study 2, which allowed us to test whether the local context of a language switch matters for infants’ detection of single-word language switches. On single-language trials, infants heard only one language, for example, “dog... milk... dog... milk...” in Study 1 or “Do you like the dog? I want the milk!” in Study 2. On switched-language trials, infants heard words from two languages, for example, “dog... *lait* [fr. milk]... dog... *lait*...” (Study 1) or “Do you like the *chien* [fr. dog]? I want the *lait*!” (Study 2). The same infants participated in both studies during the same laboratory visit.

We predicted that infants would succeed in detecting language switches both in word lists (Study 1) and in sentences (Study 2), indexed by significantly different looking times to single-language and switched-language trials, although we did not predict a direction of the difference given the challenge of doing so for a new paradigm (Aslin, 2007). It was also possible that infants would only succeed at detecting language switches when embedded in a sentence. Multi-word utterances are more common than single-word utterances in input to children (Brent & Siskind, 2001), which may facilitate infants using their day-to-day language processing abilities more fully for sentence-embedded (Study 2) compared to single-word (Study 1) language switches.

Furthermore, most previous studies of early language discrimination have used stimuli that contained long passages of speech, and from this research, we know that rhythmic cues are important to language discrimination (e.g., Ramus, 2002). In a similar vein, bilingual toddlers were better at recognizing familiar words when they were embedded in a sentence than when they were heard in isolation (Morini & Newman, 2019). Hearing “*Do you like the...*” provides infants with more exposure to sequences of sounds and words in the language being spoken prior to a language switch.

We tested both bilingual and monolingual infants. While we expected all infants to detect the switches, we expected that bilinguals might show a larger effect (i.e., a larger looking time difference) than monolinguals. This would indicate that everyday exposure to two languages enhances this ability. Another possible outcome would be that bilingual and monolingual infants show equivalent detection of language switches, suggesting that this ability is not related to the experience of hearing both languages. Testing both bilingual and monolingual infants allowed us to investigate how language exposure interacts with infants’ ability to detect language switches.

2 | STUDY 1: LANGUAGE SWITCHES IN WORD LISTS

2.1 | Method

The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. This study was approved by the Concordia University Human Research Ethics Board (certificate #10000439) and the Princeton University Behavioural Research Ethics Board. All parents provided informed written consent prior to their infants’ participation in the study. Materials and methods are available at <https://osf.io/9dtwn/>.

2.1.1 | Participants

Participants were 21 English–French or English–Spanish bilingual infants, as well as 20 monolingual infants. Instead of a power analysis, we conducted a sensitivity analysis, which is useful when there are constraints on time and participant recruitment (e.g., for hard-to-recruit populations like bilingual infants, Schott et al., 2019; Lakens et al., 2018). Using a sensitivity analysis in G*Power (Faul et al., 2007), we found that a mixed ANOVA with a sample size of 20 infants per group would yield 80% power to detect an effect of Cohen’s $d = .45$. This value was used as our minimum effect size of interest for equivalence tests (presented in the Results section), and is comparable to the average effect size found in a meta-analysis of familiar word recognition studies using similar stimuli and testing mostly monolingual samples of the same age (Cohen’s $d = 0.54$, retrieved from MetaLab; see also Carbajal et al., 2021). Our original research plan was to test bilingual infants only, but we were able to further explore the effect of language background when an opportunity arose to test monolingual children in Montreal, Canada. Bilingual children were recruited in both Montreal, Canada (English–French bilinguals, $n = 14$) and in New Jersey, USA (English–Spanish bilinguals, $n = 7$). The language environment in these two communities is quite different: In Montreal, both English and French are widely spoken in the community, while in New Jersey, English is the majority language. The original research plan involved testing a sufficient number of bilingual infants in both locations to enable a comparison, but due

TABLE 1 Information about infant participants in Studies 1 and 2, presented separately for bilingual and monolingual groups

Language background	N	Age			% Male	% English dom.
		Min	Mean	Max		
Study 1						
Bilingual	21	7 months 29 days	9 months 19 days	12 months 16 days	33	48
Monolingual	20	7 months 29 days	8 months 24 days	10 months 2 days	50	30
Study 2						
Bilingual	20	7 months 29 days	9 months 6 days	11 months 17 days	40	55
Monolingual	17	7 months 29 days	8 months 25 days	10 months 2 days	47	29

All infants were tested in both studies in the same laboratory visit, but in a few cases infants only successfully completed one study. For monolinguals, the column % English-dominant denotes the percent of children who were English monolinguals. As monolingual children were only tested in Montreal, children who were not English-dominant were French-dominant.

to recruitment difficulties, only a limited number of infants in New Jersey met the language inclusion criteria articulated below and could be tested. Thus, the data from both locations were combined. Further information on age and gender is displayed in Table 1.

Infants' language exposure was assessed using the Multilingual Approach to Parent Language Estimates (MAPLE; Byers-Heinlein et al., 2020) which assesses infants' month-by-month exposure to languages from birth to the test date. Infants were considered monolingual if they were exposed to 90% or more of English or French, and bilingual if they were exposed to 25–75% of English and 25–75% of French or Spanish. The language that infants were exposed to the most was considered their dominant language, and the percentage of infants dominant in English is reported in Table 1. On average, bilingual infants tested in Montreal heard their dominant language 63% (range: 42%–75%) and their non-dominant language 36% (range: 26%–50%) of the time. Bilinguals tested in New Jersey heard their dominant language 57% (range: 51%–73%) and their non-dominant language 43% (range: 27%–49%) of the time. Monolinguals tested in Montreal heard their dominant (native) language on average 98% (range 91–100%) of the time. Two bilinguals had exposure to a third language (13% and 20% exposure, respectively). The Language Mixing Scale (Byers-Heinlein, 2013) was used to assess exposure to parental language mixing, and bilingual infants heard more mixing than monolingual infants (bilingual infants: 13.05, $SD = 7.39$, range = [0–25]; monolingual infants: 5.55, $SD = 5.01$, range = [0–16]; $t[37] = 3.73$, $p = .001$, Cohen's $d = 1.16$). All participants were reported to have normal vision and hearing. Infants and their families were recruited through government-supplied birth lists, as well as in daycare centers, playgroups, and other child-focused community activities. Infants in New Jersey were tested between August 2016 and November 2017 and infants in Montreal were tested between February 2017 and November 2018.

To achieve the final sample, 68 infants and their families participated in the study. Two participants were tested during a pilot phase to verify that the study procedure was feasible and were not included in data analysis. Of the remaining infants, one was born prematurely (<37 weeks) and thus did not meet our health criteria, and 16 did not meet the pre-specified language criteria. Four infants were excluded for technical difficulties or experimenter error. Additionally, four infants were excluded for contributing fewer than 8 trials with at least 2.5 s of looking time. Trials shorter than 2.5 s were excluded because infants needed to listen at least that long in order to encounter a switched-language word. Some infants completed fewer than 16 trials because they

were fussing and their parent ended the study early, but in these cases, infants were not automatically excluded from analysis and all usable trials were analyzed. There were no other exclusion criteria.

2.1.2 | Stimuli

In the head-turn preference procedure, visual stimuli were presented on both center and side screens to attract infants' attention. The visual stimuli differed between the Montreal and New Jersey test locations due to different laboratory conventions. Infants tested in Montreal saw an animation of a spinning rainbow-colored citrus; infants tested in New Jersey saw a video recording of an orange flashing light. This was not expected to affect the pattern of results, as the visual stimulus only served to attract infants' attention, and was constant across auditory stimuli.

Speech stimuli for infants tested in Montreal were recorded by a native English–French bilingual, and for infants tested in New Jersey by a native English–Spanish bilingual. All words were non-cognates, and were chosen because they are, on average, acquired early in language development. Word pairs presented together in a trial were always of the same grammatical gender and were thematically dissimilar (Willits et al., 2013). To the extent possible, we avoided overlapping word onsets or codas, and we matched word pairs on number of phonemes, word frequency, and stress patterns (which was not possible for English–French disyllabic words due to differences in typical stress patterns across languages). Due to these constraints, the word pairs in the English–French and the English–Spanish versions of the study are partially but not fully overlapping. For the English–French version of the study, word pairs were dog–milk, kitty–book, mouth–door, cookie–foot and their French translations; for the English–Spanish version, word pairs were doggy–balloon, kitty–foot, mouth–milk, cookie–door and their Spanish translations. Examples of the stimuli used in the study can be seen in Table 2, and all original stimuli can be seen in Table S1 and downloaded at <https://osf.io/9dtwn/>.

Each word was recorded separately, in a friendly, infant-directed manner, and then combined to form single-language (e.g., dog...milk) and switched-language trials (e.g., dog...*lait*). Speakers were asked to produce a consistent, hill-shaped prosodic contour across items (Nencheva et al., 2021), and we selected tokens that sounded similar in their prosody across languages. Individual words were presented in alternation with a 500 ms pause between each word. Trials lasted 20.4–22.6 s depending on the length of the audio file. All looking times were capped at the shortest trial length to avoid introducing a difference in looking time based on the length of the audio file.

2.1.3 | Procedure

Infants sat on their caregiver's lap in a sound-attenuated room with three monitors, one centered, one to the left, and one to the right of the infant. The caregiver listened to music through a set of headphones to avoid influencing the infant's reactions. The experimenter controlled the study via custom MATLAB software (Olson, 2017). The experimenter was in a different room for infants tested in Montreal or in the same room but listening to masking music through headphones for infants tested in New Jersey. In both cases, the experimenter was unaware of the experimental condition on each trial. At the start of each trial, the visual stimulus appeared on the center screen. Once the experimenter observed that the infant looked at the center screen, the visual stimulus appeared on either the left or the right screen. When the infant turned their head to look

at the side screen, the experimenter pressed a button and the auditory stimulus started playing. The trial ended when the infant looked away from the side screen for two consecutive seconds, or once the entire trial was complete (~ 21 s). If the infant looked to the side for <2 s total on any trial, that trial was skipped and then automatically repeated after the final trial.

Infants first completed two practice trials, which presented a non-language sound (Montreal: whistle sound; New Jersey: beep tones). Next, infants encountered 16 test trials. Half of the test trials were single-language trials, where all target words were in a single language. The other half were switched-language trials, where the language alternated between English and French (Montreal) or English and Spanish (New Jersey). Single-language trials were presented in a consistent language for each infant, to limit the occurrence of language switches between trials, such that language switches primarily occurred within switched-language trials. Monolingual infants heard the single-language trials in their native language, and bilingual infants were randomly assigned to hear single-language trials in their dominant or their non-dominant language.

The order of trials was pseudo-randomized, with the constraint that the same word pairs could not appear on consecutive trials, and no more than three trials of one type (single-language, switched-language) appeared consecutively. Each trial type appeared equally on the left and the right side, and no more than three consecutive trials appeared on the same side. The side of presentation was independent of the stimulus type.

Parents were asked about their child’s language background using the LEQ structured interview and MAPLE approach (Byers-Heinlein et al., 2020), which asks about the child’s lifetime exposure to different languages. Furthermore, parents filled out the Language Mixing Scale (Byers-Heinlein, 2013), which asks how often parents switch between their languages when speaking to their infant. For one infant tested in Montreal and one tested in New Jersey, scores from the Language Mixing Scale were not available. As part of standard laboratory protocols, Montreal parents also completed a questionnaire with general demographic information, as well as the MacArthur-Bates Communicative Development Inventories in American English (Fenson

TABLE 2 Examples of single- and switched-language trials for English–French and English–Spanish versions of the studies

Language pair	Version	Single-language	Switched-language
Study 1			
English–French	English	Dog... Milk...	Dog... <i>Lait</i> ...
	French	<i>Chien</i> ... <i>Lait</i> ...	<i>Chien</i> ... Milk...
English–Spanish	English	Doggy... Balloon...	Doggy... <i>Globo</i> ...
	Spanish	<i>Perro</i> ... <i>Globo</i> ...	<i>Perro</i> ... Balloon...
Study 2			
English–French	English	Do you like the milk? I want the dog!	Do you like the <i>lait</i> ? I want the <i>chien</i> !
	French	<i>Aimes-tu le lait</i> ? <i>Je veux le chien</i> !	<i>Aimes-tu le milk</i> ? <i>Je veux le dog</i> !
English–Spanish	English	Do you like the balloon? I want the doggy!	Do you like the <i>globo</i> ? I want the <i>perro</i> !
	Spanish	<i>¿Te gusta el globo</i> ? <i>Quiero el perro</i> .	<i>¿Te gusta el balloon</i> ? <i>Quiero el doggy</i> .

Note: Each participant heard only one version (e.g., an English–French bilingual was randomly assigned to hear either the English or the French version of the study). Non-English words are in italics.

et al., 2007) and Quebec French (Trudeau et al., 1999). Questionnaires were completed either prior to or following the experimental portion of the study. Infants received a certificate and/or a small gift for their participation.

2.1.4 | Coding

During the study, a trained experimenter blind to the auditory stimuli watched the infant through a live video feed and pressed buttons corresponding to a look to the left, right, or center monitor. A trained research assistant later re-coded the videos offline, frame-by-frame, with the sound off. For five participants, no offline coding was conducted, due to technical difficulties with the video recording ($n = 4$) and due to a procedural omission because the participant did not complete the minimum looking criterion ($n = 1$), and so the online coded data was analyzed. In these cases, we excluded trials that the experimenter flagged as coding errors, but retained the rest of the trials, as correlations between online and offline coding were high. We performed two checks to compare the experimenter's online coding to the offline coding (following Ferguson & Lew-Williams, 2016). First, the correlation between the offline and online coders' assessment of total looking times for each trial indicated high agreement for infants tested in Montreal ($r = .98$, 95% CI [.98,.99], $t[430] = 110.03$, $p < .001$) as well as those tested in New Jersey ($r = .97$, 95% CI [.95,.98], $t[62] = 31.77$, $p < .001$). Of particular importance for the head-turn preference procedure is whether the trial ended correctly during online coding, as the dependent measure is the total looking time for a trial. Recall that the experimental program ended a trial 2 s after an infant had looked away. A buffer of $\pm .5$ s was allowed to account for the time the online coder needed to react during the study. We thus examined the proportion of trials where the end time was either less than 1.5 s after the infant had looked away (indicating a trial that ended earlier than intended) or more than 2.5 s after the infant had looked away (indicating a trial that ended later than intended), according to offline coding. Overall, 10.89% of trials ended earlier than intended and 3.42% of trials ended later than intended. We did not exclude trials with these errors, but used offline-coded looking times in statistical analyses for all infants (when available).

2.2 | Results

2.2.1 | Main Analyses

We assessed infants' detection of language switching by measuring their looking times (i.e., listening times) to single- vs. switched-language word lists. Looking times were log-transformed for all statistical analyses (Csibra et al., 2016), and figures show looking times prior to log transformation. After excluding trials with low looking time, the final sample of participants contributed on average 15.49 trials (range: 8–16). Data were analyzed using R version 4.0.3 (R Core Team, 2020) and the reproducible manuscript was created using papaja and citr (Aust, 2019; Aust & Barth, 2020). Looking times are shown in Table 3 and Figure 1 for each group (bilingual, monolingual) on each trial type (single-language, switched-language). Figure 2 shows the same data but using difference scores of looking time across trial types (calculated as $M_{switched} - M_{single}$) in order to highlight individual infants' performance. The average looking time for single-language trials was 8.55 s ($SD = 2.61$ s) and for switched-language trials was 8.69 s ($SD = 1.95$ s).

We conducted ANOVAs to investigate whether infants' looking patterns were affected by trial type (single-language, switched-language) and infants' language background (bilingual, monolingual). Due to difficulties in recruitment in New Jersey, we were unable to add test location as a between-subjects variable as we only had seven participants in New Jersey. In the Supporting informations, we reported a separate ANOVA with the same design on data from infants tested in Montreal only, as well as descriptive data on the infants tested in New Jersey, which showed patterns consistent with the combined analysis presented here. In a 2×2 mixed ANOVA with trial type (single-language, switched-language) as a within-subjects factor and language background (bilingual, monolingual) as a between-subjects factor, we found no statistically significant main effects (trial type: $F[1, 39] = 0.21$, $MSE = 0.04$, $p = .652$, language background: $F[1, 39] = 1.13$, $MSE = 0.09$, $p = .295$) or interaction ($F[1, 39] = 0.01$, $MSE = 0.04$, $p = .930$). Thus, we did not find any evidence of greater attention to one trial type over the other, and bilinguals and monolinguals performed similarly.

Given the lack of a statistically significant difference, we then aimed to understand if our results reflected a true null effect or if we had insufficient power to detect a significant effect. To do so, we conducted equivalence tests (Lakens et al., 2018) to examine whether the effect size is likely smaller than our smallest effect size of interest, and therefore test for the absence of an effect. We set the smallest effect size of interest to be $d = 0.45$, based on the minimum detectable effect size we found in our sensitivity analysis (Lakens, 2017). The equivalence test comparing the observed effect size in Study 1 to the smallest effect size of interest (i.e., of $d = 0.45$) was significant ($t[40] = 2.42$, $p = .010$), meaning that the observed effect size ($d = 0.06$) was significantly within the interval from $d = -0.45$ to $d = 0.45$. This result indicates that we can reject a true effect size larger than $d = 0.45$ (or smaller than $d = -0.45$). In other words, the equivalence test suggests a true null result for infants' looking times on single- and switched-language trials in lists of words.

2.2.2 | Exploratory analyses

We conducted additional analyses to examine individual differences in infants' detection of language switching. These analyses should be interpreted with caution, as they were exploratory. First, we calculated a Pearson correlation to examine whether infants who heard more language mixing at home performed differently in the experimental task than those who heard less language mixing at home. However, there was no evidence of a correlation between infants' language mixing score and their difference scores comparing looking time to single-language and switched-language trials ($r = .15$, 95% CI $[-0.17, 0.44]$, $t[37] = 0.93$, $p = .360$).

TABLE 3 Mean looking times (standard deviations) and effect sizes in Studies 1 and 2

	Mean looking time (<i>SD</i>)		
Language background	Single-Language	Switched-Language	Cohen's <i>d</i>
Study 1			
Bilingual	8.78 (2.88)	8.95 (2.23)	0.07
Monolingual	8.30 (2.34)	8.42 (1.62)	0.06
Study 2			
Bilingual	8.01 (3.25)	8.73 (2.66)	0.24
Monolingual	7.54 (2.29)	8.13 (2.20)	0.26

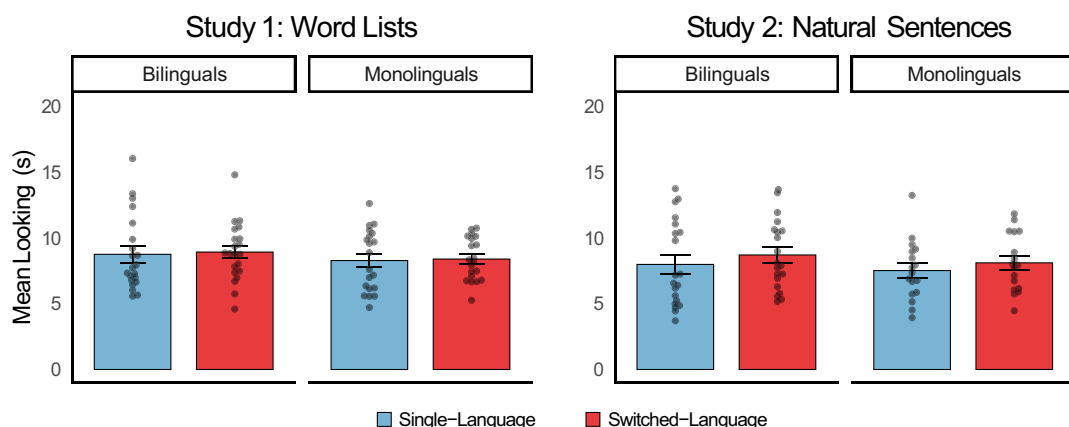


FIGURE 1 Mean looking times for single-language and switched-language trials. Looking times averaged across participants are displayed separately for each language group. Left: Study 1 (word lists). Right: Study 2 (natural sentences). Error bars indicate standard error of the mean

Second, we examined whether bilingual infants who heard single-language trials in their dominant language performed differently from those who heard the single-language trials in their non-dominant language. This was motivated by previous findings that infants' processing of language switching in word comprehension is asymmetrical across the two languages (Byers-Heinlein et al., 2017; Potter et al., 2018, 2019). Descriptively, the effect size for the difference between single and switched trials was larger for the subgroup of bilingual infants for whom the single-language trials were in their dominant language (dominant language: Cohen's $d = 0.38$, $n = 8$; non-dominant language: Cohen's $d = -0.09$, $n = 13$). However, the two-sample t -test comparing difference scores of looking to single and switched trials between participants tested in their dominant compared to non-dominant language was not statistically significant ($t[19] = 1.33$, $p = .198$). We note that the subgroups tested in each language were very small, and thus our analysis is both underpowered and subject to non-meaningful fluctuations in observed effect size. However, our results leave open the possibility that infants would detect language switches if the majority of words were in their dominant language, and this could be examined systematically in future investigations.

Third, we also explored whether bilingual infants who are exposed to their two languages in a more balanced manner (close to 50%/50%) respond differently in our task compared to unbalanced bilingual infants (those closer to 25%/75%). To investigate this, we correlated infants' performance in the task with their exposure to the non-dominant language (ranging from 25% or "unbalanced" to 50% or "balanced"). We found no statistically significant correlation ($r = -.17$, 95% CI $[-0.56, 0.28]$, $t[19] = -0.75$, $p = .465$), suggesting that balanced and unbalanced bilinguals performed similarly in Study 1.

2.3 | Discussion

Study 1 indicated that 8- to 12-month-old bilingual and monolingual infants were unable to detect single-word language switches in word lists. There was no difference between infants' looking times on single-language compared to switched-language trials, and the effect size of the

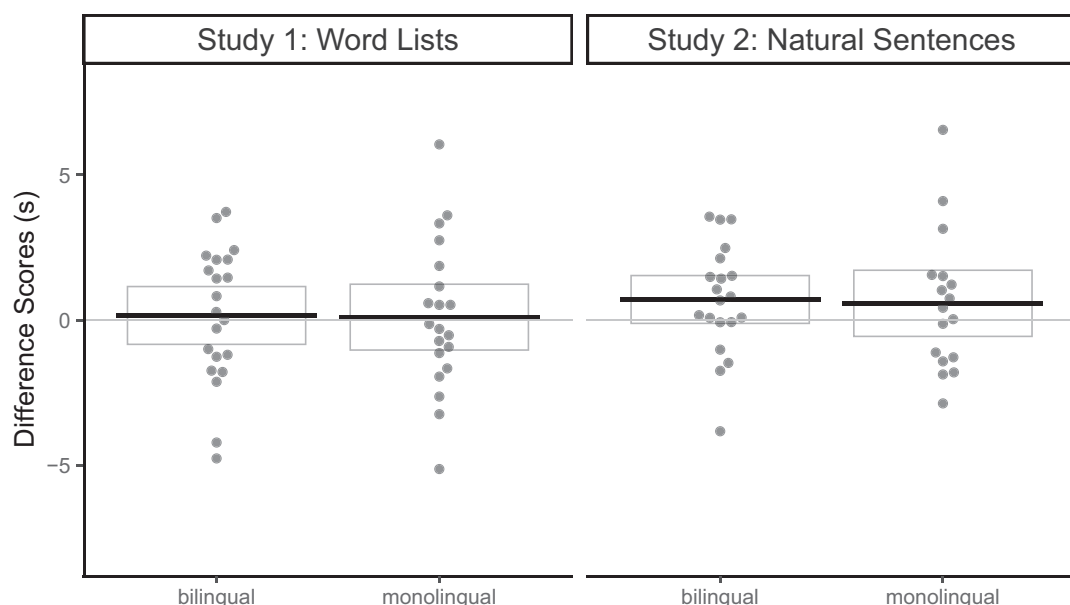


FIGURE 2 Difference scores showing individual participants' looking time to switched-language relative to single-language trials ($M_{\text{switched}} - M_{\text{single}}$). Higher values on the y-axis indicate longer looking to switched-language trials, and lower values indicate longer looking to single-language trials. This figure shows the same data as Figure 1, but displayed to highlight individual participant data

difference was statistically equivalent to zero. It may be that infants were unable to detect the language-switched words because the word lists used in the task are not typical of natural speech. Language input to children largely consists of multi-word utterances and sentences, rather than words in isolation (Brent & Siskind, 2001). Thus, naturally occurring language switches may be more likely to be embedded within a sentence. Study 2 examined whether infants can detect language switches within natural sentences.

3 | STUDY 2: LANGUAGE SWITCHES IN SENTENCES

3.1 | Method

3.1.1 | Participants

Participants in Study 2 were 20 English–French or English–Spanish bilingual infants and 17 monolingual infants, largely overlapping with the infants who participated in Study 1 (see Supporting informations). Infants were tested in Study 1 and then Study 2 during the same visit. Inclusion criteria and number of infants excluded for health and language were the same as those in Study 1. One infant did not participate in Study 2 due to fussiness (their caregiver stopped participation during Study 1). Additionally, infants were excluded for equipment failure or experimenter error ($n = 2$), as well as for contributing fewer than eight trials with at least 2 s of looking time ($n = 9$). The minimum looking time in Study 2 was set to 2 s, which was the first moment infants would be able to hear a language switch.

3.1.2 | Stimuli

The visual stimulus was a video of a colorful animated pinwheel (Montreal) or an orange flashing light (New Jersey), distinct from but comparable in attractiveness and salience to the visual stimuli used in Study 1. Study 2 used the same nouns as Study 1, but they were presented in a sentence context instead of in isolation (see examples in Table 2 and the full list in Table S1 in the Supporting information). Both single- and switched-language trials began with a phrase in the same language, i.e., in English, French, or Spanish (e.g., English: “Do you like the ...”/ “I want the ...”). On single-language trials, the sentence-final target noun was presented in the same language as the sentence frame, whereas on switched-language trials the sentence-final noun was presented in the other language. The same bilingual female speakers produced the stimuli in Studies 1 and 2. The recordings used to create the stimuli for Study 2 were not spliced, but were recorded in a single session to retain the naturalistic articulatory features of the language switch. Care was taken to ensure that prosody, intonation, and phonetic realization were appropriate for the intended language, especially for words adjacent to the language switch. Words embedded in a statement sentence had a hill-shaped prosodic contour, and words embedded in a question had a rising prosodic contour (Nencheva et al., 2021). Trials lasted 20.4–24.6 s and were capped at the length of the shortest sound file, comparable to Study 1.

3.1.3 | Procedure

The trial structure and experimental procedure were the same in Studies 1 and 2. Infants first completed Study 1, and then the researcher asked the parent to play with their infant while Study 2 was set up. This break was typically 2–3 min long, unless the infant needed to be fed or changed. We chose to present Study 2 after Study 1 because we expected that the naturalistic sentences presented in Study 2 would be more engaging to infants than the word lists presented in Study 1, thus increasing the likelihood of infants remaining attentive throughout both studies.

3.1.4 | Coding

The same coding procedure from Study 1 was used for Study 2. For five participants, offline coding was not available, either due to technical problems with the video recording ($n = 4$), or due to a procedural omission because the participant did not complete the minimum looking criterion ($n = 1$). Data from the participants without video recording were retained because correlations between online and offline coding were high. The correlation between online and offline coding of total looking times for each trial was high for infants tested in Montreal ($r = .94$, 95% CI [0.92, 0.95], $t[409] = 53.76$, $p < .001$) as well as New Jersey ($r = .95$, 95% CI [0.87, 0.98], $t[14] = 11.69$, $p < .001$). Offline coding showed 9.51% of trials ended earlier than intended and 6.28% of trials ended later than intended. As in Study 1, we did not exclude trials with these errors, but used offline-coded looking times when possible.

3.2 | Results

3.2.1 | Main Analysis

The analytic strategy for Study 2 paralleled that of Study 1. After excluding trials with low looking and participants who subsequently did not provide the minimum number of trials to be

included, the final sample on average contributed 15.08 trials (range: 9–16). The average looking time for single-language trials was 7.80 s ($SD = 2.82$ s) and for switched-language trials was 8.45 s ($SD = 2.44$ s). The results for Study 2 can be seen in the right panels of Figures 1 and 2. Log-transformed looking times were entered into a 2×2 ANOVA with trial type (single-language, switched-language) as a within-subjects factor and language background (bilingual, monolingual) as a between-subjects factor. Trial type was not statistically significant (trial type: $F[1, 35] = 1.98$, $MSE = 0.04$, $p = .168$), and neither were the other main effects and interactions (language background: $F[1, 35] = 0.24$, $MSE = 0.18$, $p = .626$; interaction trial type \times language background: $F[1, 35] = 0.04$, $MSE = 0.04$, $p = .840$). This indicates that neither bilingual nor monolingual infants detected differences between single-language and switched-language sentences.

As in Study 1, we used equivalence tests to test whether our observed effect size was smaller than our minimal detectable effect size of Cohen's $d = 0.45$. The equivalence test was not statistically significant ($t[36] = 1.31$, $p = .099$), meaning that the observed effect size ($d = 0.25$) in our study may not fall significantly within the interval of $d = -0.45$ to $d = 0.45$. This indicates that we cannot reject the hypothesis that the true effect size is Cohen's $d = 0.45$ or larger (or smaller than $d = -0.45$). Thus, our data are unsurprising in both cases, whether they were sampled from a null distribution centered around 0 (based on our ANOVA), or whether they were sampled from a distribution centered around $d = 0.45$ (based on the equivalence tests). In this case, we have insufficient data to draw a definite conclusion (Lakens et al., 2018), and future research will need to include a larger sample size.

3.2.2 | Exploratory analyses

We again examined individual differences in performance as a function of exposure to language switching and language dominance. There was no correlation between infants' language mixing score and their difference scores between looking time to single-language and switched-language trials ($r = -.08$, 95% CI $[-0.41, 0.26]$, $t[33] = -0.48$, $p = .636$). In the small subgroup of bilingual infants tested in their dominant language, anecdotally effect sizes were larger (dominant language: Cohen's $d = 0.39$, $n = 9$; non-dominant language: Cohen's $d = 0.15$, $n = 11$), although again, a two-sample t -test comparing difference scores for those tested in their dominant and non-dominant language was not statistically significant ($t[18] = 0.25$, $p = .805$), and subgroups tested in each language were very small. Thus, in Studies 1 and 2, descriptively larger effects were observed when single-language trials were presented in the dominant language, and this should be investigated more systematically. The correlation between balance of exposure for bilinguals and task performance trended toward but did not reach statistical significance ($r = -.43$, 95% CI $[-0.74, 0.01]$, $t[18] = -2.04$, $p = .056$), with less balanced children showing a numerically larger difference in looking to single- compared to switched-language trials. This finding should be replicated in a larger sample size before we can interpret it.

3.2.3 | Comparison of Studies 1 and 2.

We compared infants' looking times in Studies 1 and 2 in a $2 \times 2 \times 2$ ANOVA with trial type (single-language, switched-language) and study (Study 1, Study 2) as within-subject factors, and language background as a between-subject factor. We did not include the three-way interaction

between trial type, study, and language background due to concerns about low statistical power. Only participants who had usable data in both studies were included ($n = 33$). The main effect of study was statistically significant ($F[1, 31] = 8.41$, $MSE = 0.06$, $p = .007$), indicating that regardless of trial type, infants looked longer in Study 1 than Study 2 ($M_{\text{Study 1}} = 8.62$, $M_{\text{Study 2}} = 8.12$). None of the other main effects or interactions were statistically significant (trial type: $F[1, 31] = 0.88$, $MSE = 0.06$, $p = .357$; language background: $F[1, 31] = 1.33$, $MSE = 0.19$, $p = .258$; trial type \times study: $F[1, 32] = 0.88$, $MSE = 0.03$, $p = .356$; study \times language background: $F[1, 31] = 0.01$, $MSE = 0.06$, $p = .905$). This suggests that infants' looking times were not affected by trial type across both studies.

4 | GENERAL DISCUSSION

Bilingual infants regularly encounter switches between their two languages. Often these switches occur following long passages of speech, but they can also consist of a single word borrowed from the other language (e.g., "What a cute *chien* [fr. dog]"). Using the head-turn preference procedure, we tested whether bilingual and monolingual infants could detect single-word language switches by measuring their attention to single-language versus switched-language stimuli. We used two complementary types of statistical tests: ANOVAs in a null hypothesis framework and equivalence tests. ANOVAs indicate whether our observed effect is significantly different from zero and surprising under the null hypothesis, and equivalence tests indicate whether the observed effect is within the bounds of the smallest effect size of interest we set, and therefore practically not meaningful in the context of our study.

The results from Study 1 overall indicate that infants could not detect single-word language switches. However, the results from Study 2 were inconclusive: they do not provide strong evidence either way with regards to whether infants can detect switches that occur in naturalistic sentences. The effect size in Study 2 was small, but non-zero ($d = 0.25$). Such an effect size is not unexpected both in the case that infants cannot perform the task and in the case that they can. Furthermore, we had expected language switches in sentences to be easier to detect, yet we did not find a difference between Studies 1 and 2 in the omnibus ANOVA.

There are two main reasons why we had expected a larger effect in Study 2. First, sentences are more typical of infants' everyday experiences, allowing them to engage mechanisms that support real-time language processing, such as statistical learning and prediction (see Potter & Lew-Williams, 2019). This may enable anticipation of the sounds and sound patterns that match the preceding words and "surprisal" when the actual perceived sounds do not match those. Second, the switched-language sentences were produced naturalistically rather than by splicing, which might have afforded extra cues (e.g., coarticulation) in the speech signal even prior to the actual switch location (Fricke et al., 2016). In contrast, the isolated words were recorded in English-only and French-only contexts and thus did not provide additional coarticulatory cues to the switch. Nevertheless, we believe that future studies should prioritize studying language switches in sentences, since the equivalence tests in Study 2 indicated a possibly meaningful effect size. At the same time, it should be noted that infants participated in our studies in a fixed order, to minimize fussiness and attrition as we predicted that Study 2 would be more interesting to the infants. A study published after the current work was undertaken suggests that infants with more previous experience with the head-turn preference procedure in other laboratory visits showed a larger novelty preference than those with less experience (see Santolin et al., 2021). Although in our case infants were tested twice within

the same laboratory visit, we cannot rule out the possibility that infants' experience in Study 1 boosted their performance in Study 2.

We had predicted that bilinguals would show larger differences in response to language switching than monolinguals, but this was not supported by the data. Instead, neither bilinguals nor monolinguals showed evidence of detecting switches in either study. Within the bilingual group in Study 2 (language switches embedded in sentences), there was a non-significant trend for less balanced bilinguals to show a stronger effect than more balanced bilinguals, a finding which should be tested in a larger sample before it can be interpreted. Based on our results, we cannot say whether routine exposure to two languages affects infants' ability to detect deviations from a single language. Even if our results had shown that both bilinguals and monolinguals can detect language switches at this age, they may do so via different underlying processes. It may be that bilingual infants detect switches by recognizing certain sounds and words in each of their languages, which appear in different combinations in single- and switched-language trials. In contrast, monolingual infants were only familiar with one of the languages in the study and thus could succeed on such a task by simply listening longer during trials that contained unfamiliar sounds and words (i.e., switched-language trials). Given our inconclusive results, our data cannot currently speak to these intriguing possibilities. To better understand the reasons behind bilingual and monolingual infants' performance on this task, future studies could avoid these issues of interpretation by using novel words that conform to phonological patterns in each language, but are unfamiliar to both groups.

Although we did not observe an impact of bilingual vs. monolingual language exposure in this investigation, there may be differential sensitivity to language switching depending on the frequency of switching in the bilingual input. This frequency is known to vary across different language communities and individual families (Bail et al., 2015; Byers-Heinlein, 2013; Kremin et al., 2021). In a recent study on English–French bilingual infants in Montreal, parental language mixing was a relatively rare occurrence (Kremin et al., 2021; Orena et al., 2019), but there may be much more language mixing in other communities, such as certain English–Spanish bilingual communities in the United States (Bail et al., 2015). There appears to be variation in how English–French and English–Spanish bilingual toddlers learn words in switched-language sentences (Byers-Heinlein et al., under review), which raises the possibility that there are early-emerging differences in processing across different bilingual communities. While we had originally planned to compare these two populations in our analyses, the small sample size of the English–Spanish group prevented us from doing so. This is a limitation of our study that we hope will be remedied in future work. Studies that examine language mixing and bilingual input in different communities will be crucial for understanding pathways to bilingual proficiency. It would also be interesting to explore how individual differences in language mixing, language balance, and dominant language status interact with the detection of language switches. We computed separate correlation measures for these factors and found no significant relationships, but we found a trend for enhanced sensitivity when detecting switches embedded in the dominant language. We were not able to directly compare the performance of French–English and Spanish–English bilinguals due to sample size limitations. However, we believe that including data from two bilingual communities is likely to make our results more generalizable. Future investigations exploring these moderators together would illuminate which aspects of the bilingual experience are most important to language switching.

Finally, it is important to consider whether our null/inconclusive results could be due to limitations of our experimental design to tap into infants' underlying ability. We tested infants using the head-turn preference procedure, which has yielded strong effect sizes in meta-analyses

(Bergmann et al., 2018), and has been revealing about many aspects of infants' speech perception in infants of this age (e.g., Carbajal et al., 2021; Gasparini et al., 2020). However, we do note that other tasks have shown that older infants and toddlers are sensitive to single-word language switching. For example, several studies using both behavioral and ERP methods have found evidence that infants and toddlers respond distinctly to single-word language switching when viewing visual referents of the words spoken on the task (Byers-Heinlein et al., 2017; Kuipers & Thierry, 2012; Morini & Newman, 2019; Potter et al., 2018, 2019). Each of these studies tested children who were somewhat older than our participants (1.5–3 years old), who are more likely to have robust lexical representations of the stimuli. For such children, seeing a referent such as a picture of a dog when hearing "I like the ..." could help infants generate predictions about the upcoming word, which then might help them recognize when they hear *chien* instead of dog. We cannot rule out that 8- to 12-month-old infants might succeed in a task that provides them with the visual referent. Nonetheless, our findings raise the possibility that perceptual information without referential context may be insufficient to support the detection of single-word language switches in young infants.

Our overarching research question was how bilingual infants make sense of the dual language input they are exposed to. From the literature, we know that 0- to 4-month-olds can show successful cross-sentence language discrimination (for a review, see Gasparini et al., 2020) and that in the second year of life they can detect single-word language switching (Byers-Heinlein et al., 2017; Kuipers & Thierry, 2012; Morini & Newman, 2019; Potter et al., 2018, 2019). Our study was novel because it tested whether infants can detect language switching during the period in-between these ages, when infants first start to recognize words. We did not find evidence that infants can detect single-word language switching at this age. This could mean that the ability to detect single-word language switching only starts to develop in the second year of life. If this is the case, it is unlikely that bilingual word learning in the context of mixed-language input relies on infants implicitly treating words from their two languages as belonging to different categories (Byers-Heinlein, 2014). It could be that more robust lexical knowledge is required for infants to be able to distinguish single words from two different languages. Since we did not find statistically significant results, we cannot resolve this question. We hope that the studies described here can be a stepping stone toward answering this question. Future research can also disentangle which cues are most relevant to children to detect single-word language switches. For example, we have discussed lexical knowledge and coarticulation, but other factors such as phonology and phonotactics can also play a role. Together, existing research reveals that language discrimination is multifaceted, spans multiple levels of language, combines multiple perceptual domains, and may change over the course of development. Future work will need to determine both the nature of "successful" discrimination and the ultimate learning-related value of discriminating languages in the first place.

The present studies provide two main contributions. First, this work highlights a gap in the research into how bilingual infants make sense of their dual language input. While there is ample evidence that bilingual infants can discriminate some languages from birth when they hear whole sentences (Gasparini et al., 2020), it is only later in development around 20 months that the literature reports conclusive evidence for the ability to detect language switches for individual words (e.g., Byers-Heinlein et al., 2017). On the one hand, it may be that this is a much more difficult task with a protracted developmental time course. On the other hand, more sensitive experimental designs might be able to detect this ability earlier in development. This points to a more nuanced bilingual language development than previously thought. Second, we have reported results from two empirical studies aiming to directly test infants' detection of single-word language switches.

The results from Study 1 were null, and those from Study 2 were inconclusive. Although these findings are somewhat unsatisfying, it is nonetheless important that they have a place in published literature rather than in the “file drawer” (Nelson et al., 2018). We have shared our materials, data, and analysis scripts, so that our research can contribute to future empirical studies, comprehensive reviews, and meta-analyses. We hope that this work will motivate further research into how infants navigate the dynamics of language input in environments where two languages are present, in turn illuminating how bilingual infants acquire proficiency in both their languages.

ACKNOWLEDGEMENTS

We would like to thank the participating infants and their families, as well as the members of the Concordia Infant Research Lab and the Princeton Baby Lab. In particular, we would like to thank Michela Martiniello for her help preparing part of the stimuli for this study. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (402470-2011; 2018-04390) to KBH, grants from the National Institute of Child Health and Human Development (R01HD095912; R03HD079779) to CLW and KBH, a grant from the Speech-Language-Hearing Foundation to CLW, support from the Concordia University Research Chairs program to KBH, and a fellowship from the Fonds de Recherche du Québec—Société et Culture and Concordia University to ES. The authors declare no conflicts of interest with regard to the funding source for this study.

ORCID

Esther Schott  <https://orcid.org/0000-0002-0114-4773>

Meghan Mastroberardino  <https://orcid.org/0000-0001-8818-1720>

Eva Fourakis  <https://orcid.org/0000-0003-3891-7538>

Casey Lew-Williams  <https://orcid.org/0000-0002-8781-4458>

Krista Byers-Heinlein  <https://orcid.org/0000-0002-7040-2510>

REFERENCES

- Albareda-Castellot, B., Pons, F., & Sebastian Galles, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm: Acquisition of phonetic categories in bilingual infants. *Developmental Science*, 14(2), 395–401. <https://doi.org/10.1111/j.1467-7687.2010.00989.x>
- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10(1), 48–53. <https://doi.org/10.1111/j.1467-7687.2007.00563.x>
- Aust, F. (2019). *citr: 'RStudio' add-in to insert markdown citations*. R package version 0.3.2, Retrieved from <https://github.com/crsh/citr>
- Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown*. R package version 0.1.0.9997. Retrieved from <https://github.com/crsh/papaja>
- Bahrick, L. E., & Pickens, J. N. (1988). Classification of bimodal English and Spanish language passages by infants. *Infant Behavior and Development*, 11(3), 277–296. [https://doi.org/10.1016/0163-6383\(88\)90014-8](https://doi.org/10.1016/0163-6383(88)90014-8)
- Bail, A., Morini, G., & Newman, R. S. (2015). Look at the *gato*! Code-switching in speech to toddlers. *Journal of Child Language*, 42(05), 1073–1101. <https://doi.org/10.1017/S0305000914000695>
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3253–3258. <https://doi.org/10.1073/pnas.1113380109>
- Bergelson, E., & Swingle, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, 11(4), 369–380. <https://doi.org/10.1080/15475441.2014.979387>
- Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, 89(6), 1996–2009. <https://doi.org/10.1111/cdev.13079>

- Bosch, L., & Sebastian Galles, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29–49. https://doi.org/10.1207/S15327078IN0201_3
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–B44. [https://doi.org/10.1016/S0010-0277\(01\)00122-6](https://doi.org/10.1016/S0010-0277(01)00122-6)
- Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28(3), 455–474. <https://doi.org/10.1017/S0142716407070257>
- 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(01), 32–48. <https://doi.org/10.1017/S1366728912000120>
- Byers-Heinlein, K. (2014). Languages as categories: Reframing the “one language or two” question in early bilingual development. *Language Learning*, 64(s2), 184–201. <https://doi.org/10.1111/lang.12055>
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343–348. <https://doi.org/10.1177/0956797609360758>
- Byers-Heinlein, K., Jardak, A., Fourakis, E. & Lew-Williams, C. (under review). Effects of language mixing on bilingual children's word learning. Manuscript under review. Retrieved from <https://psyarxiv.com/298cz/>
- Byers-Heinlein, K., Morin-Lessard, E., & Lew-Williams, C. (2017). Bilingual infants control their languages as they listen. *Proceedings of the National Academy of Sciences of the United States of America*, 114(34), 9032–9037. <https://doi.org/10.1073/pnas.1703220114>
- Byers-Heinlein, K., Schott, E., Gonzalez-Barrero, A. M., Brouillard, M., Dubé, D., Jardak, A., Laoun-Rubenstein, A., Mastroberardino, M., Morin-Lessard, E., Pour Iliaei, S., Salama-Siroishka, N., & Tamayo, M. P. (2020). MAPLE: A multilingual approach to parent language estimates. *Bilingualism: Language and Cognition*, 23(5), 951–957. <https://doi.org/10.1017/S1366728919000282>
- Carbajal, M. J., Peperkamp, S., & Tsuji, S. (2021). A meta-analysis of infants' word-form recognition. *Infancy*, 26(3), 369–387. <https://doi.org/10.1111/infa.12391>
- Csibra, G., Hernik, M., Mascaró, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, 52(4), 521–536. <https://doi.org/10.1037/dev0000083>
- 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. <https://doi.org/10.1016/j.wocn.2010.12.002>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Fenson, L., Marchman, V. A., & Thal, D. (2007). *MacArthur-bates communicative development inventories: User's guide and technical manual*. Paul H. Brookes Publishing Company.
- Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports*, 6, 25434. <https://doi.org/10.1038/srep25434>
- Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language*, 89, 110–137. <https://doi.org/10.1016/j.jml.2015.10.001>
- Gasparini, L., Langus, A., Tsuji, S., & Boll-Avetisyan, N. (2020). *Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis*. <https://doi.org/10.31219/osf.io/rmn5x>
- Kartushina, N., & Mayor, J. (2019). Word knowledge in six- to nine-month-old Norwegian infants? Not without additional frequency cues. *Royal Society Open Science*, 6(9), 180711. <https://doi.org/10.1098/rsos.180711>
- Kremin, L. V., Alves, J., Orena, A. J., Polka, L., & Byers-Heinlein, K. (2021). Code-switching in parents' everyday speech to bilingual infants. *Journal of Child Language*. <https://doi.org/10.1017/S0305000921000118>. in press.
- 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(2), F13–F21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x/full>
- Kuipers, J.-R., & Thierry, G. (2012). Event-related potential correlates of language change detection in bilingual toddlers. *Developmental Cognitive Neuroscience*, 2(1), 97–102. <https://doi.org/10.1016/j.dcn.2011.08.002>
- Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>

- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/10.1177/2515245918770963>
- Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, 19(3), 326–337. <https://doi.org/10.1111/inf.12041>
- Morini, G., & Newman, R. S. (2019). *Dónde está la ball?* Examining the effect of code switching on bilingual children's word recognition. *Journal of Child Language*, 46(6), 1238–1248. <https://doi.org/10.1017/S0305000919000400>
- Nazzi, T., Bertoni, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756–766. <https://doi.org/10.1037//0096-1523.24.3.756>
- Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43(1), 1–19. <https://doi.org/10.1006/jmla.2000.2698>
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, 69(1), 511–534. <https://doi.org/10.1146/annurev-psych-122216-011836>
- Nencheva, M. L., Piazza, E. A., & Lew-Williams, C. (2021). The moment-to-moment pitch dynamics of child-directed speech shape toddlers' attention and learning. *Developmental Science*, 24, e12997. <https://doi.org/10.1111/desc.12997>
- Olson, R. (2017). *Wisconsin Infant Studies Program*. Retrieved from <https://bitbucket.org/rholson1/wisp/src/default/>
- Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). Reliability of the language environment analysis recording system in analyzing French-English bilingual speech. *Journal of Speech, Language, and Hearing Research: JSLHR*, 62(7), 2491–2500. https://doi.org/10.1044/2019_JSLHR-L-18-0342
- Potter, C. E., Fourakis, E., Byers-Heinlein, K., & Lew-Williams, C. (2019). Bilingual toddlers' comprehension of mixed sentences is asymmetrical across their two languages. *Developmental Science*, 22(4), e12794. <https://doi.org/10.31234/osf.io/q8bjj>
- Potter, C. E., Fourakis, E., Morin-Lessard, E., Byers-Heinlein, K., & Lew-Williams, C. (2018). Bilingual infants process mixed sentences differently in their two languages. In C. Kalish, M. Rau, T. Rogers, & X. Zhu (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 900–905). Cognitive Science Society.
- Potter, C. E., & Lew-Williams, C. (2019). Infants' selective use of reliable cues in multidimensional language input. *Developmental Psychology*, 55(1), 1–8. <https://doi.org/10.1037/dev0000610>
- R Core Team. (2020). *R: A language and environment for statistical computing*. Manual, R Foundation for Statistical Computing.
- Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, 2(1), 85–115. <https://doi.org/10.1075/arla.2.05ram>
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292. [https://doi.org/10.1016/S0010-0277\(00\)00101-3](https://doi.org/10.1016/S0010-0277(00)00101-3)
- Santolin, C., Garcia-Castro, G., Zettersten, M., Sebastian Galles, N., & Saffran, J. R. (2021). Experience with research paradigms relates to infants' direction of preference. *Infancy*, 26(1), 39–46. <https://doi.org/10.1111/inf.12372>
- Schott, E., Rhemtulla, M., & Byers-Heinlein, K. (2019). Should I test more babies? Solutions for transparent data peeking. *Infant Behavior and Development*, 54, 166–176. <https://doi.org/10.1016/j.infbeh.2018.09.010>
- Sebastian Galles, 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. <https://doi.org/10.1037//0096-1523.28.4.974>
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108(1), 232–242. <https://doi.org/10.1016/j.cognition.2007.12.013>
- Trudeau, N., Frank, I., & Poulin-Dubois, D. (1999). Une adaptation en français québécois du McArthur Communicative Development Inventory. *Canadian Journal of Speech-Language Pathology and Audiology*, 22, 151–163.
- Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh, and English-Welsh bilingual infants. *Applied Psycholinguistics*, 28(3), 475–493. <https://doi.org/10.1017/S0142716407070269>

- Werker, J. F. (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics*, 39(4), 703–728. <https://doi.org/10.1017/S0142716418000152>
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)
- Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, 18(6), 1053–1075. <https://doi.org/10.1111/infa.12026>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Schott, E., Mastroberardino, M., Fourakis, E., Lew-Williams, C., & Byers-Heinlein, K. (2021). Fine-tuning language discrimination: Bilingual and monolingual infants' detection of language switching. *Infancy*, 26, 1037–1056. <https://doi.org/10.1111/infa.12429>