

Can Bilingual Children Turn One Language Off? Evidence from perceptual switching.

Leher Singh

Department of Psychology, National University of Singapore;

and

Carolyn Quam

Departments of Speech, Language, and Hearing Sciences and Psychology, University of
Arizona

RUNNING HEAD: Perceptual switching in bilinguals

Address for correspondence:

Leher Singh Dept. of Psychology,
National University of Singapore,
9 Arts Link, Singapore 117570

Acknowledgements: We are grateful to a Ministry of Education Tier 1 Academic Research Fund grant (FY2013FRC2-009) to LS. We thank Annabel Soon, Darrell Loh and Beatrice Ng for assistance with recruitment and testing.

Abstract

Bilinguals have the sole option of conversing in one language in spite of knowing two languages. The question of how bilinguals alternate between their two languages, activating and de-activating one language, is not well understood. In the current study, we investigate the development of this process by researching bilingual children's abilities to selectively integrate lexical tone based on its relevance in the language being used. In particular, the current study sought to determine the effects of global conversation-level cues versus local (within-word phonotactic) cues on children's tone integration in newly learned words. Words were taught to children via a conversational narrative and word recognition was investigated using the intermodal preferential-looking paradigm. Children were tested on recognition of words with stimuli that were either matched or mismatched in tone in both English and Mandarin conversations. Results demonstrated that 3-to-4-year old children did not adapt their interpretation of lexical tone changes to the language being spoken. In contrast, 4-to-5-year children were able to do so when supported by informative within-word cues. Results suggest that preschool children are capable of selectively activating a single language given word-internal cues to language.

One of the most significant challenges facing bilingual learners is the mastery of two systems that are linguistically distinct, yet conceptually linked. This challenge is potentially complicated when a bilingual learners' native languages maintain conflicting rules. One area of potential conflict lies in the language-specific use of phonetic variation: languages often share sources of phonetic information (e.g. voice-onset time, aspiration, pitch movements) yet vary in how they partition these sources of variation to form native phonetic categories. For example, tone languages incorporate systematic variation in pitch movements to distinguish word meanings. However, non-tone languages also incorporate systematic pitch variation in a way that does not distinguish word meanings, but instead may be used to differentiate emotions, questions and statements, or stress and focus (Gussenhoven, 2004). As a result, a bilingual learner attempting to master a tone and non-tone language has to interpret pitch movements in a language-selective manner based on the language being spoken in a given context. The effectiveness with which bilingual children go back and forth in terms of how they interpret phonological variation across their languages, activating and de-activating sensitivity to common cues as befits the language context, remains unclear. This process is known as 'perceptual switching.' Perceptual switching has been quite extensively investigated in bilingual adults. However, the extent to which children can engage in perceptual switching (and modifiers of their ability to do so) remain unclear, yet this ability represents a crucial step in the journey towards bilingual proficiency. The focus of the current study is to determine whether perceptual switching is observable in children and to identify possible cues that influence perceptual switching in childhood.

There have been several investigations of perceptual switching in adults. For example, many languages use variation along the voice-onset-time (VOT) continuum to contrast consonants. However, the location of the boundary between phonetic categories differs across languages, such as in English and Spanish. Adult bilinguals of English and Spanish therefore

have to impose a different VOT boundary on particular sounds when listening to English versus Spanish. In a study with adult bilinguals, Elman, Diehl, & Buchwald (1977) reported that Spanish-English bilinguals could selectively shift their judgments of the same sounds (/b/ vs. /p/) based on whether they were listening to English and Spanish carrier sentences. Participants' capacities for perceptual switching were mediated by their relative proficiency in each of their languages: balanced bilinguals with similar proficiency levels across their languages were better able to shift their judgments in response to the language context than unbalanced bilinguals. Such effects of language context in perceptual switching have also been reported in speech production (Flege & Eefting, 1987; Hazan & Boulakia, 1993) and in other phonetic categories in Spanish (Garcia-Sierra, Diehl, & Champlin, 2009; Garcia-Sierra, Ramirez-Esparza, Silva-Pereyra, Siard & Champlin, 2012, but see Caramazza, Yeni-Komshian, & Zurif, 1974; Caramazza, Yeni Komshian, Zurif, & Carbone, 1973).

Perceptual switching in adults is not limited to conditions where perception is cued by language context. Rather, switching can be enabled by the availability of informative within-word cues to language identity. In recent study by Gonzales and Lotto (2013), instead of varying the broader language context, the authors manipulated within-word phonetic information to cue a particular language. They found that when the word “[b/p]afri” contained a Spanish-like “r”, Spanish-English bilinguals—but not English monolinguals—interpreted a preceding ambiguous initial [b/p] sound as a Spanish phoneme. When the target word contained an English-like “r”, bilinguals Spanish-English speakers and monolingual English speakers interpreted the same ambiguous sound as an English phoneme. Just as subtle differences between phonetic segments can distinguish phonetic categories in two languages, languages also differ – and potentially conflict - in how they utilize suprasegmental cues. A classic example is the family of languages called tone languages. When processing a tone language, learners must integrate particular pitch contours, lexical

tones, into their representations of words in order to successfully contrast word meanings. By contrast, in a non-tone language like English, pitch information does not determine the lexical identity of a word, though it cues other relevant information like the speaker's emotions, the stressed syllables of words, and phrase boundaries. Bilingual learners of a tone and non-tone language must accurately interpret pitch cues with reference to the target language. In a study with bilingual adults learning a tone and non-tone language (English and Mandarin), Quam and Creel (submitted) taught adults a series of novel words in English and Mandarin. Participants' sensitivity to tone when retrieving these words was investigated. Participants were cued to the language identity of novel words by context (the language of the conversation) or within-word cues. Quam and Creel's findings demonstrated that in adult Chinese English bilinguals, within-word cues enabled a language-specific interpretation of tone, whereas language context cues did not when presented in isolation.

The vast majority of studies on perceptual switching have focused on adult bilingual processing of ambiguous phonemes. There has been a minimal focus on the developmental origins of this ability. Developmental investigations of phonological sensitivity in word learning in bilingual children have focused predominantly on sensitivity to phonetic variation within one of their languages (e.g. Byers-Heinlein, Fennell & Werker, 2013; Fennell & Byers-Heinlein, 2014; Fennell, Byers-Heinlein & Werker, 2007; Mattock, Polka, Rvachew & Krehm, 2010). There has been much less focus on how bilinguals alter their phonological sensitivities as they switch back and forth between their languages. However, the ability to recalibrate phonetic perception in response to language input is a crucial component of acquiring bilingual proficiency. Although adults demonstrate a capacity to recalibrate in response to language context, the extent to which children can do the same remains unknown. Furthermore, it remains unknown whether children harness the same set of cues to the target

language when switching between languages as adults or whether they utilize a different set of cues.

There has been one previous study to investigate perceptual switching in infancy. In a study with infants from 7 to 11 months, Singh and Foong (2012) familiarized bilingual Mandarin/English infants with two words and then measured infants' abilities to recognize the familiarized words in sentences. Infants were tested on their capacity to interpret tone variation as lexical in Mandarin and to interpret pitch variation as non-lexical in English. Each infant was tested in two sessions: an English session and a Mandarin session. In one session, infants were exposed to two English words. They were then presented with passages containing familiarized words as well as unfamiliar passages. One of the words matched in pitch between familiarization and test and the other word was mis-matched in pitch. A second session, conducted in Mandarin, was otherwise identical to the English session except that one familiarized word was presented in a form that matched in pitch across familiarization and test. In contrast, the other word changed in pitch, corresponding to a change in lexical tone, between familiarization and test. At 7.5 months, consistent with results of a similar study with monolingual English-learning infants (Singh, White, & Morgan, 2008), bilingual infants only recognized English words when they were pitch matched with words from the familiarization phase. In Mandarin, they only recognized tone matched words at 7.5 months and did not equate pitch or tone variants in English and Mandarin contexts respectively. However at 11 months, infants displayed language-specific integration of pitch variation, recognizing pitch-matched and mis-matched English words. In addition, they only recognized Mandarin words when they were matched in tone across familiarization and test.

Although Singh and Foong's study hint at perceptual switching in infants, there are two important considerations when interpreting their results. First, there was no evidence to suggest that infants had attached meaning to familiarized words. The degree of phonological

precision attached to word representations has often differed for infant word segmentation tasks versus word learning tasks in toddlers that involve establishing the meanings of words (e.g. Jusczyk & Aslin, 1995 (Experiment 3) vs. Swingley & Aslin, 2000). For example, although Singh and Foong (2012) found that bilingual Mandarin-English learners did not differentiate English words based on pitch characteristics at 11 months, Singh, Tam, Chan & Golinkoff (2014) found that at 18 months, Mandarin-English bilingual infants did differentiate English words by pitch contour when associating word forms with meaning. Examples of task dependence in prior investigations of infant's phonological representations are not uncommon (see also Stager & Werker, 1997). In particular, tasks that require auditory sensitivity to sound change versus those that require word-object mapping often yield different findings. The former often gives the appearance of relative strength on the part of infants and the latter often demonstrates a learning cost introduced by weightier task demands (see Quam & Creel, 2015 for a discussion). The process of mapping phonetic variation onto meaning arguably carries a greater cognitive load than recognition of word forms or discrimination of sounds (Curtin, Byers-Heinlein & Werker, 2011; Werker & Curtin, 2005).

A second difference between Singh and Foong's study and prior investigations of perceptual switching in adults is that traditionally, participants are presented with the same type of variation across language contexts. In Singh and Foong's study, the pitch variation built into the English contexts was acoustically distinct from the tone variation incorporated in the Mandarin contexts. Perceptual switching usually involves a language selective interpretation of the same phonetic information. For these reasons, it remains an open question as to whether language-selective sensitivity to tone when children have to re-interpret similar tone changes across languages in different ways.

In the current study, we sought evidence of perceptual switching in bilingual preschool children by investigating whether lexical tone was processed in a language

selective manner. In particular, we investigated influences of discourse-level and word internal cues on bilingual children's integration of tone during word learning in each language. Across two experiments, we taught children novel words in English and Chinese narrative contexts and tested their recognition of these words when they were matched in tone or mis-matched in tone. Of primary interest was whether children would interpret tone variation as lexical in Chinese and as non-lexical in English. In Experiment 1, target words were phonologically plausible in both languages, such that the only cue to the relevance of lexical tone changes was the language used in the narrative context. In Experiment 2, target words in the English context were phonologically legal in English but not in Mandarin. Target words in the Mandarin sentences were identical to those used in Experiment 1 (i.e. phonologically legal in both languages). As such, children had the benefit of two sources of information (language context and word-internal cues for one of the languages) when interpreting lexical tone changes. In both experiments, children were tested at 3 to 4 years and 4 to 5 years of age.

Experiment 1

Methods

In Experiment 1, children were presented with two videos where they were taught the meanings of new words in English and Mandarin. Videos consisted of conversations in which two puppets discussed two novel objects, one of which was explicitly labelled (target) and one of which was discussed but never labelled (distractor). In both languages, children were tested on their ability to recognize the target when it matched versus mis-matched the tone of the training stimulus via a preferential looking paradigm. Target words were phonotactically legal in both English and Mandarin and were equi-biased towards English and Mandarin phonology (Quam & Creel, 2012).

Participants

The sample of participants comprised 34 bilingual preschool children in two age groups: 17 3-year-olds (range: 3;1 to 3;11) and 17 4-year-olds (range: 4;2 to 5;0). All participants were attending bilingual immersion preschools and were judged to have native proficiency in Mandarin Chinese and English. All participants had daily exposure to Mandarin and English and were judged by bilingual experimenters to be equally proficient in both languages. Following a short five-minute conversation with each child, experimenters rated each child on their proficiency in English and Mandarin on a scale from 1 to 5. Each experimenter was a native speaker of English and Mandarin. A score of 5 on this scale indicated that children could understand conversational speech in English and Mandarin and could accurately respond in full sentences, maintain the conversation successfully and that they spoke with a native language accent. Only children who received 5 on this scale for both English and Mandarin were tested further. An additional set of 7 children were tested but not included due to failure to complete testing in both languages (3), equipment failure (2) and zero attention during all the test trials for a trial type (1).

Stimuli

All speech stimuli were recorded by a bilingual female native speaker of English and Mandarin, who was asked to produce the words in infant-directed speech. The speaker was selected on account of having a local accent that would be familiar to the children. She was from the same geographical origin and had the same language background (native speaker of English and Mandarin) as the sample. The novel words, “biufa” and “fipu”, were adopted from Quam & Creel (2012) and were designed to be phonologically equi-biased towards English and Mandarin phonology. Each word contained a Mandarin tone on the first syllable and neutral tone on the second syllable. Tones 2 and 4 were chosen as they are reportedly highly distinguishable to native and non-native speakers of Mandarin (Halle, Chang & Best, 2004). Tone 2 is marked by a rising contour. Tone 4 is marked by a falling pitch contour

(please see Figure 1 for a depiction of pitch contours of syllables assigned Tones 2 and 4).

Across participants, the pairings of words and tones were counterbalanced.

Acoustic analyses were conducted on the first syllable of the target words to determine mean pitch onset and offset (summarized in Table 1). To ensure that the two tones – tone 2 and 4 – used in Experiment 1 were accurately pronounced, 10 adults were native speakers of Mandarin were asked to complete a tone identification task. The first syllable of the target word was excised from carrier sentences and randomly concatenated into an audio file consisting of each target word presented in citation form in each tone. The adults were presented with all tokens in citation form and were instructed to rate them with Mandarin tone numbers, 1, 2, 3 or 4. All stimuli were rated with 100% accuracy. To ensure that there was no language-specific variation in the realization of tones across English and Mandarin on the part of the speaker, target words were exactly the same tokens between the English and Mandarin video. Specifically, they were spliced from the Mandarin context and inserted into the English videos.

Apparatus and Procedure

The experiment was conducted in a quiet, dimly lit room. Stimuli were presented on a Macintosh computer. The experimenter sat on a chair beside the child, with the child facing the centre of the computer screen. Another experimenter sat behind the laptop and video recorded the child's eye movements. Auditory stimuli were presented over speakers at conversational level (65 db). Participants watched two videos (English and Mandarin) each consisting of conversations between two puppets. Each video consisted of a conversational narrative between two puppets, where a word was introduced, followed by a test phase where recognition of the learned words was investigated. Participants watched the entire narrative and were tested on recognition of familiarized words during the test phase in one language before viewing the same sequence of events in the other language.

Each video featured the same bilingual protagonist named Lily, who appeared Chinese (see Figure 2 for the trial sequence). In the Mandarin video, Lily initiated a conversation with another Chinese puppet, Hui Xian, who also had long straight hair and wore Chinese clothing. During the conversation, Lily talked about two objects. One object was explicitly labelled (target) and one was discussed for an equal duration although never labelled (distractor) (for a sample of the conversational narrative, please see Figure 3). Toys were animated by Lily during the conversation and appeared in the lower half of the screen at the centre.

During the English video, Lily introduced two different objects to a Caucasian puppet, Elizabeth. Elizabeth wore western clothing and had blonde hair. The dialogues in the English and Chinese videos were translations of each other, although the background scene changed, as did the target and distractor objects to sustain attention across both videos and reduce interference between word-object mappings. During the conversation, each target label was produced 10 times and each object appeared on screen for 45 seconds, such that participants were equally familiar with the labelled and unlabelled objects prior to entering the test phase. Target words were labelled as a ‘biufa’ or a ‘fipu’. One label was selected to name the target object in the English video and the other was selected to name the target object in the Mandarin video. Children watched an English video and a Mandarin video in succession. After each conversation, a test phase was initiated. The test phase was presented in English following the English conversation and in Mandarin following the Mandarin conversation.

Conversations and test phases were blocked such that participants watched the English conversation and English test block followed by the Mandarin conversation and Mandarin test block or vice versa. For each language, the structure of the test phase was identical except for the language in which auditory stimuli were presented. During the test phase, the labelled object served as the target and the unlabelled object served as the distractor. During

each test trial, the two toys familiarized in the training phase were displayed on screen against a white background side-by-side (side of presentation was counterbalanced across trials). Target words were presented in carrier sentences. Within each test trial, the target word always appeared 1600 milliseconds after the start of the trial. The test block consisted of 4 test trials followed by 2 reminder trials and then by 4 more test trials. During the first block of four test trials, there were two types of trials. In the first trial type, words were matched in tone to the familiarization set. In the second trial type, words were mis-matched in tone to the familiarization set, defined as a shift from the contour of Tone 2 to Tone 4. After the first block of 4 test trials, there were 2 reminder trials to recapitulate the labels assigned to each object. Each reminder trial lasted for 8 seconds. During these trials, participants saw each object on the screen, one at a time. When the target object was presented, it was labelled three times in a carrier sentence. When the distractor object appeared on screen during reminder trials, it was discussed for an equal duration with no labelling. Following this, a second block of test trials was presented that was identical in structure to the first block. The order of presentation of test trials (tone matched versus tone-mismatched trials) was randomized within block. The pairing of labels to objects and to languages was counterbalanced across subjects. The order in which English and Mandarin videos were presented was counterbalanced across subjects.

Eye movements were coded during test trials for fixation to the target and distractor. A trained coder, who was blind to the conditions of the test trials, used the Supercoder software (Hollich, 2005) to code the participants' eye movements using frame-by-frame judgments. Twenty percent of the videos were randomly selected and re-coded by another independent coder. The mean intercoder reliability was high, with a Pearson's correlation coefficient of 0.99 ($p < .001$).

Results

We analyzed the proportion of fixation to the target (PTL) for each object during each test phase from 367 msec to 2000 msec after the onset of the target word, established in prior research as an appropriate window of analyses for spoken word recognition (e.g. Swingley & Aslin, 2002). Proportion of fixation to target is standardly calculated by dividing the fixation to the target by the sum of the total fixation duration to the target and distractor. PTL values are plotted in Figure 4 for each age group and each trial type.

To determine whether performance on the task depended on whether children heard conversation in English or Mandarin first, PTL was included as the dependent variable in a 2 x 2 x 2 x 2 ANOVA with factors Language (Mandarin vs. English), Trial Type (Tone match vs. Tone mismatch), Age (3 years vs. 4 years) and Order (Mandarin conversation first vs. English conversation first). There were no effects or interactions with order ($p > .5$) so PTL values were entered into a 2 x 2 x 2 mixed ANOVA with factors Language (Mandarin vs. English), Trial Type (Tone match vs. Tone mismatch), and Age (3 years vs. 4 years). Results revealed a main effect of Trial Type ($F(1,32) = 5.2, p = .03$): Participants fixated test trials containing tone-matched words for a significantly longer duration than those containing tone mismatched words. However, there was also a significant interaction of Trial Type and Age ($F(1,32) = 10.67, p = .003$). Further analyses were therefore conducted within each age group.

3-4 year-old participants (age range: 3;4 to 3;11 years)

PTL values were subjected to two sets of analyses within each age group. In the first set of analyses, fixation to the target (PTL) was compared to chance fixation to the target (.5) to determine whether there was preferential fixation to the target for each trial type (tone-matched and tone-mismatched trials) in English and Mandarin. This provides an indication of whether auditory labels were associated to visual targets during the test trials. When such an association is made, above-chance fixation to the target is predicted. In a second set of analyses, proportionate fixations to the target (PTL values) were compared across languages

for tone-matched and tone-mismatched words. This analysis was aimed at comparing PTL values for different trial types to one another to determine whether the magnitude of preference varied as a result of language context and trial type.

For the first set of analyses aimed at simply seeking evidence for word recognition within each trial type, PTL during the test phase was compared to fixation predicted by chance (.5, or equal looking to target and distractor pictures) via a series of one-sample t-tests. Fixation times to the target during test trials that are significantly above chance are taken as evidence that participants have mapped the auditory word form onto the target object. Results revealed significantly above-chance looking to target when words matched in tone between the learning and test phases of the experiment both for English and Mandarin videos (English: $t(16) = 2.2, p = .04$; Mandarin: $t(16) = 2.2, p = .04$). For mismatched tones, participants also showed significant above-chance target fixation both for English and Mandarin sentences (English: $t(16) = 2.9, p = .01$; Mandarin: $t(16) = 2.9, p = .01$). This suggests that participants associated tone-mispronounced word-forms with the target object irrespective of whether they were listening to a conversation in Mandarin or English.

For the second set of analyses, a 2 x 2 (Language x Trial type) repeated-measures ANOVA was conducted with PTL as the dependent variable. Results revealed no main effect of language ($F(1,16) = .0001, p = .98$), no effect of trial type ($F(1,16) = .45, p = .51$) and no interaction of trial type and language ($F(1,16) = .69, p = .42$). This suggests that PTL responses did not vary based on whether words were presented in the same tone or a different tone or if they were presented in English and Mandarin videos, nor did they vary as a result of a combination of these conditions.

4-5-year-old participants.

As with 3-year-old participants, a series of planned analyses were computed to compare PTL during the test phase to fixation predicted by chance (.5). Results revealed significantly

above-chance looking to target when words were matched in tone to training both in English and Mandarin videos (English: $t(16) = 3.6, p=.002$; Mandarin: $t(16) = 3.3, p=.004$). However, for mismatched forms, participants did not fixate target pictures at above-chance levels in either language (English: $t(16) = 1.2, p=.25$; Mandarin: $t(16) = .12, p=.9$). As with the younger sample, participants linked tone-matched forms to their referents in both English and Mandarin. However, in contrast to the younger sample, participants did not link tone mispronunciations to the target in either English or Mandarin contexts.

A 2 x 2 (Language x Trial type) repeated-measures ANOVA was conducted with PTL as the dependent variable. Results revealed no main effect of language ($F(1,16) = 1.16, p=.3$) and no interaction of trial type and language ($F(1,16) = .11, p=.74$.) There was however a significant main effect of trial type ($F(1,16) = 16.9, p=.001$). As can be seen in Figure 4, fixation times to tone matched trials were significantly greater than those to tone mismatched trials in both languages.

In the aggregate, the results of Experiment 1 suggest that children at both ages were able to learn novel word-object associations from a conversational context both in English and Mandarin. This was demonstrated by preferential fixation to the target in response to newly learned words in English and Mandarin when words matched in tone across training and test. However, participants' abilities to negotiate language-specific effects of tone variation within newly learned words were far less robust. Younger participants appeared to disregard the phonological relevance of tone in both languages whereas older participants appeared to incorporate tone as a source of lexical contrast in both languages. Both age groups appeared to treat tone variation in a language-nonspecific manner. The goal of Experiment 2 was to determine whether participants would be able to integrate tone in a language selective manner if given an additional cue to the language identity of target words.

Specifically, we investigated whether providing leading within-word phonological cues would facilitate children's interpretation of tone in a language-specific manner.

Experiment 2

The goal of Experiment 2 was to determine whether the availability of phonotactic cues to language identity could facilitate language-specific integration of tone. A new sample of bilingual children were presented with target words as before. However, the phonotactic compatibility of the words with English vs. Mandarin was also manipulated. The pair of target words employed in Experiment 1 was presented in the Mandarin context. A new pair of target words, only phonotactically legal in English, was presented in the English context. Therefore, in addition to language identity, participants had additional phonotactic cues to one language.

Participants

The sample comprised 34 bilingual preschool children in two age groups: 17 3-4-year-olds (range: 3;4 to 3;11) and 17 4-year-olds (range: 4;1 to 4;10). Inclusion criteria matched Experiment 1 (see Experiment 1 Methods). An additional set of 5 children were tested but not included due to failure to complete testing in both languages (4) and equipment failure (1).

Stimulus, apparatus and procedure

Stimuli, apparatus and procedure were identical to that of Experiment 1. The only difference was that the target words in English videos were 'gripu' and 'klafa'. The instantiation of tone on these items, as represented by pitch onset and offset, was similar to that on the 'biufa' and 'fipu' presented in Experiment 1 (see acoustic analyses for all stimuli in Table 1). Tone assignments were evaluated by 10 native speakers of Mandarin Chinese using the same protocol as described in Experiment 1. All tones were identified with 100% accuracy.

Results

As with Experiment 1, to determine effects or interactions with order, proportion of fixation to the target (PTL) was calculated for each object during the test phases. PTL values are plotted in Figure 5 for each age group and each trial type. A 2 x 2 x 2 x 2 (Language: Mandarin/English x Trial Type: Tone match/Tone mismatch x Age: 3-4 years/4-5 years x Order: Mandarin conversation first/English conversation first) ANOVA was computed. There were no effects or interactions with order ($p > .5$) and as such, PTL values were entered into a 2 x 2 x 2 (Language: Mandarin/English x Trial Type: Tone match/Tone mismatch x Age) mixed ANOVA. As in Experiment 1, there was a significant three-way interaction of trial type, language and age ($F(1,31) = 4.2, p = .05$). Further analyses were conducted within each age group.

3-4-year-old participants

A series of planned analyses were computed to compare PTL during the test phase to fixation predicted by chance (.5). Results revealed a significant elevation in fixation to target when words were matched in tone to training both in English and Mandarin videos (English: $t(16) = 2.3, p = .04$; Mandarin: $t(16) = 2.6, p = .02$). However, for mismatched forms, participants showed no difference in fixation to target relative to chance for English or Mandarin videos (English: $t(16) = .02, p = .98$; Mandarin: $t(16) = .63, p = .54$). This suggests that participants consistently associated tone-matched words with the visual target, but consistently did not associate tone-mismatched forms with the visual target in either language.

A 2 x 2 (Language x Trial type) repeated-measures ANOVA was conducted with PTL as the dependent variable to determine whether PTL values differed from one another based on language context. Results revealed no main effect of language ($F(1,16) = .07, p = .8$), an effect of trial type ($F(1,16) = 4.4, p = .05$), but no interaction of trial type and language,

($F(1,16) = 09, p=.77$). PTL values were higher overall for tone-matched versus tone mismatched words although this pattern of results did not depend on the language within which words were taught.

A series of planned analyses were computed to compare PTL during the test phase to fixation predicted by chance (.5). Results revealed significantly above-chance target fixation when tones were matched across learning and test phases both in English and Mandarin videos (English: $t(16) = 3.52, p=.003$; Mandarin: $t(16) = 3.67, p=.002$). When tones were mismatched across learning and test, in the English context, participants also fixated the target significantly above chance ($t(16) = 5.9, p=.0001$). However, in the Mandarin context, their target fixation did not differ from chance ($t(16) = .29, p=.77$). As with the younger sample, participants linked tone-matched forms to their referents in English and Mandarin. However, in contrast to the younger sample, participants interpreted tone mismatches in a language selective manner, treating tone variants as mispronunciations of the target word in the Mandarin context but as correct pronunciations in the English context.

A 2 x 2 (Language x Trial type) repeated-measures ANOVA was conducted with PTL as the dependent variable. Results revealed a main effect of language ($F(1,16) = 13.64, p=.002$) and no main effect of trial type ($F(1,16) = .07, p=.8$). However, there was a significant interaction of trial type and language ($F(1,16) = 14.4, p=.002$). Follow up comparisons revealed a significant decrease in PTL when words were tone-mismatched vs. when they were tone matched ($t(16) = 2.73, p=.02$) in Mandarin, but the opposite pattern in English: a significant increase in PTL when words were tone-mismatched vs. when they were tone matched ($t(16) = 2.1, p=.05$). An increase in attention to tone varying words has been observed in prior studies in English learning preschoolers (e.g. Quam & Swingley, 2010) and is attributed to the high attentional capture of pitch variation.

General Discussion

The purpose of the current study was to determine whether bilingual preschoolers could interpret tone in a language-selective manner based on language context and within-word cues. Children were tested on their ability to recognize newly learned words at two age groups, when words matched or mis-matched in tone in English and Mandarin contexts. Results demonstrated that when children were faced with ambiguous within-word cues to language in combination with leading context cues, they were not able to integrate tone in a language-selective manner between 3 and 5 years of age. In contrast, when children were presented with leading within-word cues in addition to cues based on language context, at the age of 4 to 5 years, children were able to negotiate the functions of tone across each of their languages, integrating tone variation as a source of lexical contrast in Mandarin and disregarding tone variation as a source of lexical contrast in English.

While there have been several studies attesting to bilingual adult's abilities to shift their perception of acoustic-phonetic variation in alignment with properties of their native languages (Bohn & Flege, 1993; Elman, Diehl, & Buchwald, 1977; Flege & Eefting, 1987; Hazan & Boulakia, 1993), there have been few studies thus far investigating the developmental pathway leading up to this ability. In the current study, it appears that the capacity for language selective integration of lexical tones is jointly dependent on maturation and within-word phonological support and appears to mature definitively between 3 and 5 years of age. During this period of maturation, within-word cues seem to be particularly instrumental in enabling perceptual switching for tone. In contrast to a prior study conducted by Singh and Foong (2012) demonstrating language-selective integration of lexical tone in bilingual infants at 11 months, this study points to a relatively prolonged period of maturation of perceptual switching abilities. This difference can presumably be traced to task demands: Singh and Foong's study involved segmenting repetitions of recently heard words from

continuous speech and did not incorporate a referential component. It is likely that the weightier cognitive and linguistic demands associated with the present task contributed to a different profile of ability as compared to that reported by Singh and Foong (2012).

The current set of findings is somewhat surprising given prior evidence suggesting that when presented with words only in one language, bilingual learners can integrate tone in a language-selective manner as toddlers. Bilingual English-Mandarin speaking toddlers have been shown to integrate tone when learning novel words in Mandarin by 2 years of age (Singh et al. 2014). Likewise, bilingual toddlers have been shown to be able to disregard lexical tone when learning novel words in an English context by 2 years of age (Singh et al. 2014). Likewise, when listening to familiar words in Mandarin, Mandarin-English bilingual toddlers successfully integrated tone by 2.5 years of age (Singh, Goh & Wewalaarachchi, 2015). None of these groups required within-word cues to language identity in order to interpret tone in a language-selective manner and interpretation of tone was likely only guided by language context. However, this collection of studies did not require any switching; rather, each participant was tested in a single session spoken in a single language only. It stands to reason that having the added burden of alternating between phonological systems in rapid succession may be much more demanding for bilinguals than single language processing. Language selective integration of tone under these circumstances may place greater demands on children that are only overcome later in development.

A central theoretical question about the nature of bilingual memory – relevant to the current study - is whether bilinguals can selectively activate a single language or whether both languages are co-activated during language processing (Bialystok, Craik, Green, & Gollan, 2009; Genesee, 1989; Grosjean, 1989). This question can be asked in at least two ways. One way is to determine whether both languages are active when bilinguals communicate, even though only one is in use. The second is whether one language can be

selectively de-activated when it is not in use. Both questions can potentially inform our understanding of the structure of early bilingual memory in complementary ways. In terms of whether both languages are active during communication, previous studies have employed cross-language semantic priming paradigms to determine the extent to which a word in one language automatically activates a related word in the other language in bilingual toddlers. One of the major contributions of this arm of research is to provide strong evidence that words in both languages are activated upon hearing a single target language in bilingual toddlers (e.g. Singh, 2014; von Holzen & Mani, 2012). The notion that both languages are jointly activated needs to then be reconciled with the subsequent need to deactivate one language in order to negotiate phonological conflict, the focus of the current study. Our results suggest that early in development, between 3 and 4 years of age, context-cued deactivation proves challenging and children appear to conform to the rules of one language when processing phonological ‘clashes’. However, between 4 and 5 years of age, with phonotactic support, children were better able to engage in selective inhibition and activation, pointing to a language-selective interpretation of tone variation. It is possible that in order to negotiate phonological conflict, the bilingual mind engages advanced executive functions such as inhibition of attention and resistance to potential sources of interference. Inhibitory control and resistance to task-irrelevant information are cornerstones of bilingual proficiency and undergo considerable maturation over the preschool years (Bialystok & Viswanathan, 2009). The involvement of higher order cognitive operations when required to switch rules across native languages in rapid succession may account in part for the late emergence of switching abilities.

Prior studies providing evidence of perceptual switching in adult bilinguals are often recruited as evidence that bilinguals demonstrate separate systems, as a function of being able to activate a single language in response to language context (e.g. Elman et al. 1977; Flege &

Eefting, 1987; Hazan & Boulakia, 1993). It is tempting to interpret the current findings as support for separate systems (or the emergence therefore) by 5 years of age. However, it should be noted, as pointed out by Bohn & Flege (1993), that it is possible to hypothesize perceptual switching in bilinguals without appealing to separate systems. Specifically, languages may be associated with individual acoustic-phonetic ranges (i.e. children could have formed expectations about the range of acceptable lexical pitch variation in English and Mandarin, associating different pitch ranges with words in English versus Mandarin words). Cues to a particular acoustic-phonetic range could rapidly calibrate the perceptual systems of bilinguals in favor of the language with which that range is associated. This argument provides a language-general, integrated mechanism according to which bilinguals may demonstrate perceptual switching simply by sampling acoustic-phonetic ranges associated with particular words, akin to distributional learning accounts of monolingual phonetic category acquisition (e.g. Maye, Werker & Gerken, 2002). Placing our findings within the context of this mechanism, it is possible that language context and/or word internal cues would activate a range of acceptable tone variants within each word and enable perceptual switching, without ever appealing to language-selective activation or to separable linguistic systems. The current study does not easily disambiguate these possibilities and we believe it would be an over-interpretation of the current findings to posit separate linguistic systems in the developing bilingual mind on account of our findings.

The current study provides developmental evidence of a bilingual ability for perceptual switching that is well attested in adults (Bohn & Flege, 1993; Elman, Diehl, & Buchwald, 1977; Flege & Eefting, 1987; Garcia-Sierra, Diehl, & Champlin, 2009; García Sierra, Ramírez-Esparza, SilvaPereyra, Siard, & Champlin, 2012; Gonzales & Lotto, 2013; Hazan & Boulakia, 1993). However, unlike the majority of studies with adults, the current data suggest that bilingual children require more than language context to selectively activate

each language and if it is utilized in older preschoolers, language context may be a secondary cue to word-internal cues. The combination of maturation and within-word phonotactic cues (and optionally, language context) conspire to enable perceptual switching at 4 to 5 years of age. Language context alone was not sufficient to yield perceptual switching at either age group. Even though within-word cues were informative only for Mandarin (the English words remained phonotactically legal in both languages), participants fared better in both languages when one language set incorporated leading phonotactic cues, suggesting that within-word properties do not need to be uniquely specified within each language for perceptual switching to occur.

The current study opens up several lines of future inquiry. In particular, we purposefully selected balanced bilinguals with the expectation of maximizing the potential to elicit perceptual switching. It is possible that children with a greater proficiency differential between their two languages would require greater support in order to engage in perceptual switching. A controlled comparison of switching in balanced versus unbalanced bilinguals would provide evidence as to whether this ability is contingent upon high proficiency in both languages. Secondly, while bilingual children in the current study demonstrated language-selective activation across their languages, we do not know whether their abilities for language-selective activation accord with those observed in monolinguals. A comparison of bilingual and monolingual children's abilities for integration of lexical tone would help to answer another prevailing question in the study of bilingualism, specifically, whether there are limits on bilingual proficiency such that the bilingual phonological percept remains qualitatively distinct (and essentially less 'native') in comparison to the monolingual phonological percept (see Cutler, Mehler, Norris & Segui, 1989; Dupoux, Peperkamp & Sebastián-Gallés, 2010; Navarra, Sebastián-Gallés & Soto-Faroco, 2005 for a discussion of these issues).

In summary, the current study provides developmental evidence that bilingual children demonstrate an impressive facility with alternating between languages, even when the task is complicated by the presence of phonological conflict across languages. Pitch movements, which serve a broad range of communicative functions in human languages, have to be selectively integrated into semantic comprehension and assigned appropriate relevance in every language. The current study suggests that the capacity for assigning lexical relevance to pitch cues across the two languages of a bilingual emerges between 4-5 years of age and relies on phonotactic support. The current study demonstrates that bilingual preschool children are able to negotiate phonological incompatibilities across their language and can rapidly alternate between phonological systems, but that they require different cues to those commonly implicated in adult bilingual processing.

References

- Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition, 112*, 494–500.
- Bialystok, E., Craik, F. I.M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*, 89-129.
- Bohn, O.-S., & Flege, J. E. (1993). Perceptual switching in Spanish/English bilinguals. *Journal of Phonetics, 21*, 267-290.
- Byers-Heinlein, K., Fennell, C.T., & Werker, J.F. (2013). The development of associative word learning in monolingual and bilingual infants. *Bilingualism: Language and Cognition, 16*(1), 198-205.
- Caramazza, A., Yeni-Komshian, G., & Zurif, E.B. (1974). *Bilingual switching: The phonological level. Canadian Journal of Psychology, 28*, 310-318.

- Caramazza, A., Yeni-Komshian, G.H., & Zurif, E.B. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, *54*, 421-428.
- Creel, S.C., & Quam, C. (2015). Apples and Oranges: Developmental Discontinuities in Spoken-Language Processing? *Trends in Cognitive Sciences*, *10/2015*, DOI: 10.1016/j.tics.2015.09.006.
- Curtin, S., Byers-Heinlein, K., & Werker, J.F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*, 492-504.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Letters to Nature*, *340*, 229-230.
- Dupoux, E., Peperkamp, S. & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress "deafness" in simultaneous French-Spanish bilinguals. *Cognition*, *114*, 266-275.
- Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, *62*, 971-974.
- 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*, 1510-1525.
- Fennell, C., & Byers-Heinlein, K. (2014). You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments. *International Journal of Behavioral Development*, *38*, 309-316.
- Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*, 185-202.
- Garcia-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, *51*, 369-378.

- Garcia-Sierra, A., Ramirez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C.A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: an electrophysiological study. *Brain and Language, 121*, 194-205.
- Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language, 16*, 161-179.
- Gonzales, K., & Lotto, A.J. (2013). A *Bafri*, un *Pafri*: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science, 24*, 2135-2142.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3-15.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press, Cambridge, United Kingdom.
- Halle, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*, 395-421.
- Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech, 36*, 17-38.
- Hollich, G. (2005). Supercoder: A program for coding preferential looking (version 1.5) [computer software]. West Lafayette: Purdue University.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1-23.
- Mattock, K., Polka, L., Rvachew, S., & Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: Comparing monolingual and bilingual infants. *Developmental Science, 13*, 229-243.
- Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101-B111.

- Navarra, J., Sebastián-Gallés, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: New evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 912-918.
- Quam, C., & Creel, S. C. (2012). What's in a rise? Effects of language experience on interpretation of lexical tone. *Proceedings of the 36th Annual Boston University Conference on Language Development*, 487-499.
- Quam, C., & Creel, S. C. (submitted). Mandarin-English Bilinguals Process Lexical Tones in Accordance with the Language Context.
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*, 128-142.
- Singh, L., Goh, H. H., & Wewalaarachchi, T. D. (2015). Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation. *Cognition*, *142*, 1-11.
- Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, *17*, 94-109.
- Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, *4*, 157-178.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381-382.
- Swingle, D., and Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.
- von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*, 569-586.

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

Figure 1: Sample pitch contours of Mandarin Tones 2 and 4.

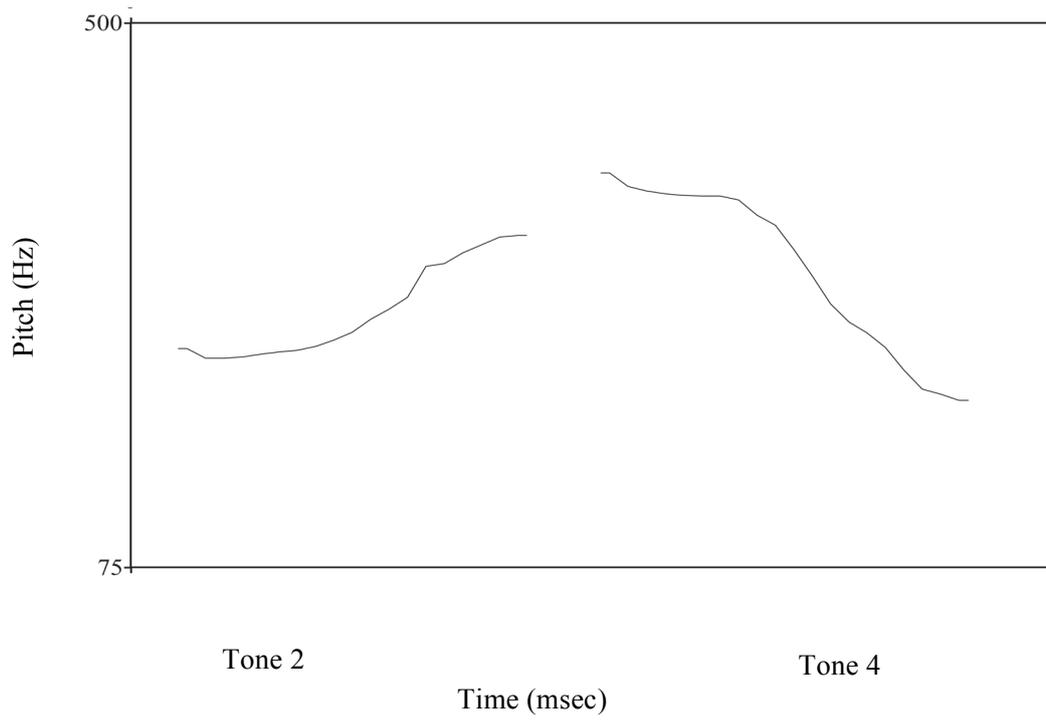


Figure 2: Sequence of events during the experimental session.

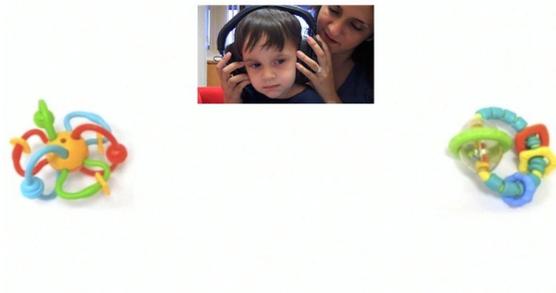
Phase 1: Familiarization with target (biu4fa) and distractor (unlabeled)



Phase 2: Test of Target Recognition (4 trials) (Inset depicts a participant)

“Which one is the *biu2fa*? That’s pretty!”

“Which one is the *biu4fa*? That’s pretty!”



Phase 3: Two Reminder Trials

“Look at the *biu2fa*. *Biu2fa*. See that? The *biu2fa*.”



“Do you see that fun toy? Isn’t it pretty? Would you like to play with it?”



Phase 4: Test of Target Recognition (4 trials)

“Which one is the *biu2fa*? That’s pretty!”

“Which one is the *biu4fa*? That’s pretty!”

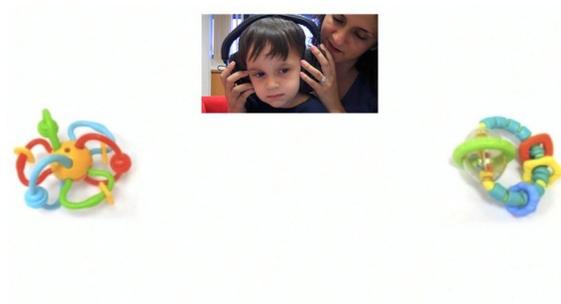


Figure 3: A sample of the conversational narrative in Chinese and English.

	
<p>Mandarin conversation sample</p> <p>Lily: 惠山, 你好!</p> <p>Hui Xuan: 丽丽, 你好! 你在这里做什么</p> <p>Lily: 我在找一个圣诞礼物给我弟弟惠山, 你要帮我选一个礼物吗?</p> <p>Hui Xuan: 圣诞节吗? 我最喜欢圣诞节了。你找到了什么? 我看到了这个玩具, 也找到了这个 <i>biu2fa</i>, 你觉得好吗?</p> <p>Lily: 两个都很可爱啊!</p> <p>Hui Xuan: 你看这个 <i>biu2fa</i>, 这个 <i>biu2fa</i> 看起来很有趣。你喜不喜欢这个 <i>biu2fa</i>?</p> <p>Lily: 是啊。它看起来真的很好玩。</p> <p>Hui Xuan: 那着另一个玩具呢? 这个玩具看起来也很好玩。我想现在就玩它。不知道弟弟会喜欢哪一个玩具。</p> <p>Lily: 不然你先买下这两个玩具。我们可以先试一试, 然后再决定。</p> <p>Hui Xuan: 好主意! 让我先付钱, 然后我们可以去我家玩。</p>	<p>English conversation sample</p> <p>Lily: Hello Elizabeth!</p> <p>Elizabeth: Hello Lily! What are you doing here?</p> <p>Lily: I'm so happy to see you Elizabeth! I'm bored. Would you like to play?</p> <p>Elizabeth: Sure! I just got home from school and I really want to play.</p> <p>Lily: Great! Do you remember my birthday party on Sunday? I got this new toy from my mummy and a <i>biu2fa</i> from my daddy. Shall we play with them?</p> <p>Elizabeth: Ok! That sounds great!</p> <p>Lily: This is my <i>biu2fa</i>. I'll show you how to play with a <i>biu2fa</i>. This <i>biu2fa</i> is really fun.</p> <p>Elizabeth: Wow! I would love to have one too.</p> <p>Lily: Here is my other toy. I'll show you how to play with it too. It's very interesting.</p> <p>Elizabeth: Cool! Do you want to go to the playground to play with them?</p> <p>Lily: Ok! That's a great idea!</p>

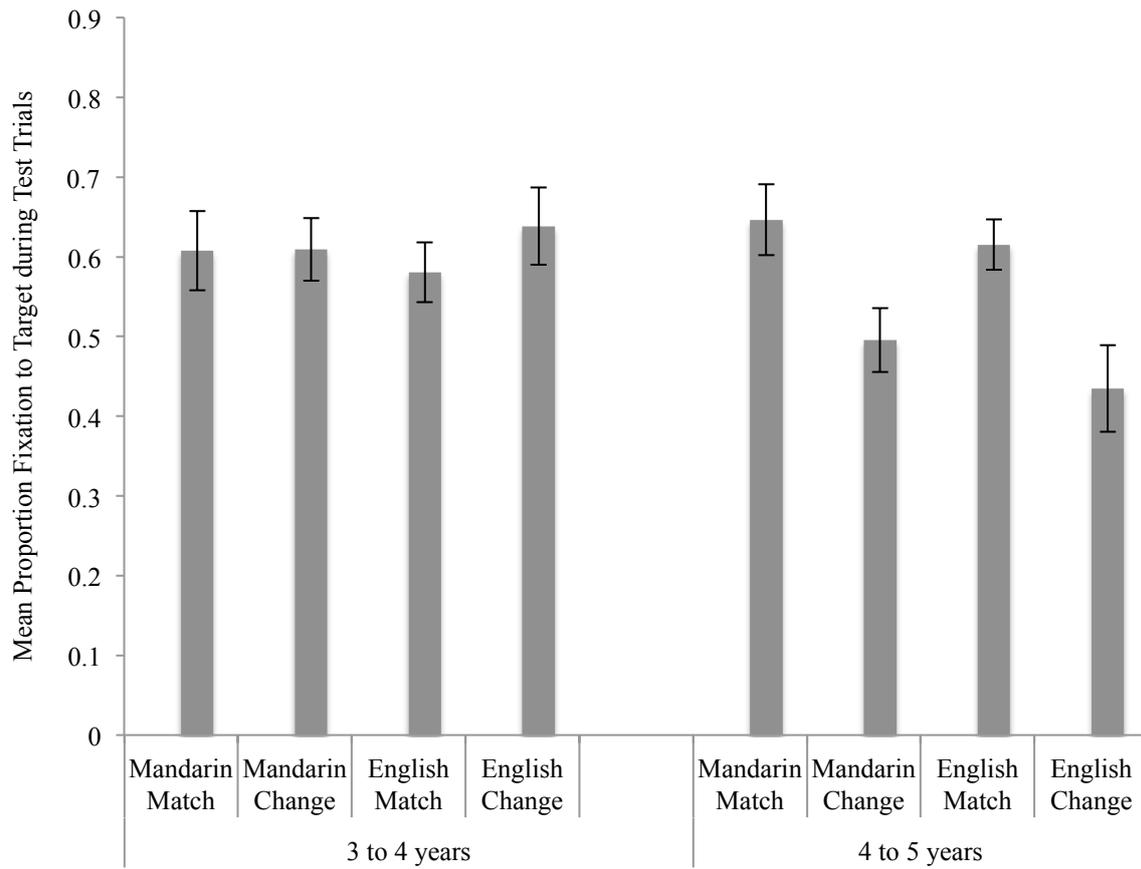


Figure 4: Proportion of Total Looking to Target (Experiment 1). Error bars reflect S.E.M.

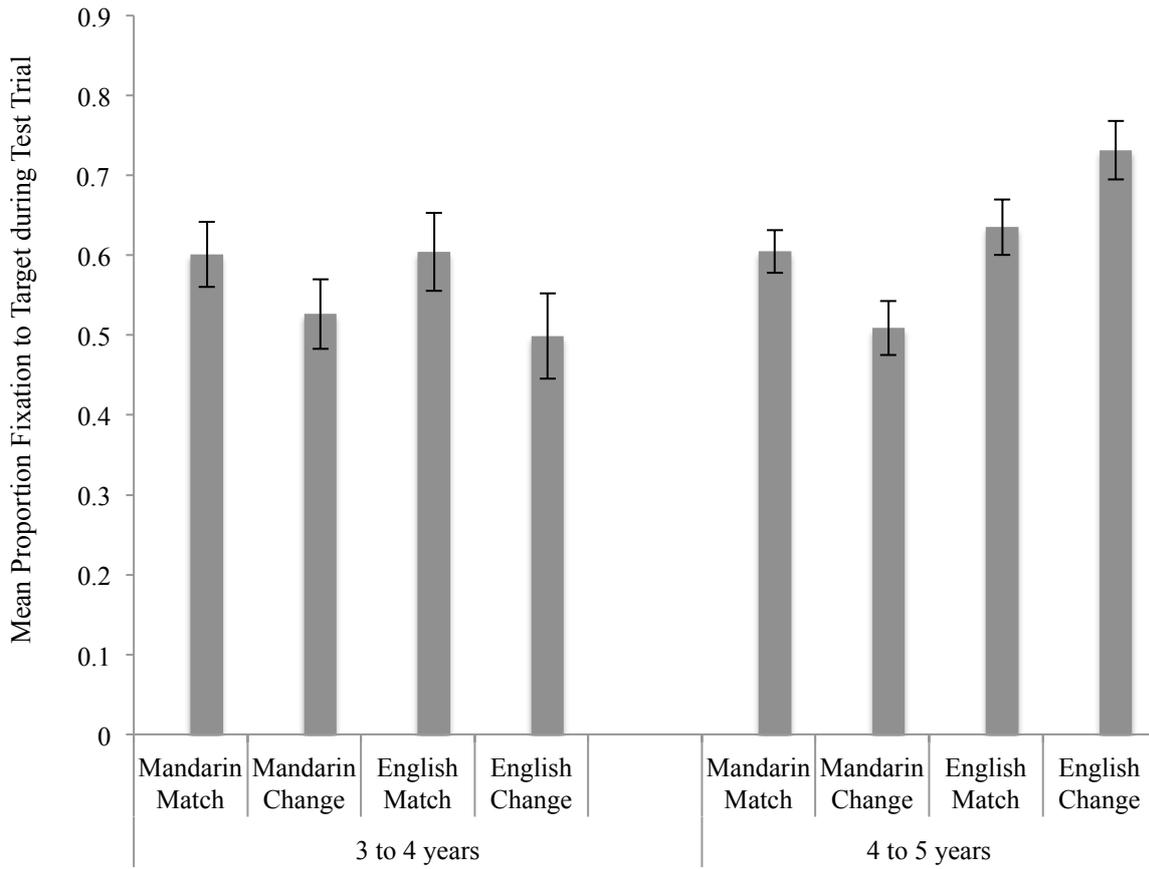


Figure 5: Proportion of Total Looking to Target (Experiment 2). Error bars reflect S.E.M.

Tone	Language	Syllable	Training Trials		Test trials	
			Fundamental frequency (hz)			
			Word onset	Word offset	Word onset	Word offset
2 (rising)	English	“gree”	223.7	350.6	245	330.8
		“kla”	280.5	414	253.6	337.2
	Mandarin	“biu”	232.01	356.56	212.6	298.8
		“fi”	272.8	405.9	252.7	328.1
4 (falling)	English	“gree”	356.87	206.45	378.4	222.9
		“kla”	395.15	237.31	390.3	239
	Mandarin	“biu”	426.12	271.38	378.2	225
		“fi”	429.29	274.93	435.2	277.3

Table 1: Acoustic analyses of target words.