Dónde está la ball? Examining the effect of code switching on bilingual children’s word recognition

Giovanna MORINI1* and Rochelle S. NEWMAN2

1University of Delaware, USA and 2University of Maryland, USA
*Corresponding author: University of Delaware, Communication Sciences and Disorders Program, 100 Discovery Blvd., Room 533, Newark, DE, 19713. E-mail: gmorini@udel.edu

(Received 25 May 2018; revised 23 January 2019; accepted 16 May 2019; first published online 13 August 2019)

Abstract

Hearing words in sentences facilitates word recognition in monolingual children. Many children grow up receiving input in multiple languages – including exposure to sentences that ‘mix’ the languages. We explored Spanish–English bilingual toddlers’ (n = 24) ability to identify familiar words in three conditions: (i) SINGLE WORD (ball!); (ii) SAME-LANGUAGE SENTENCE (Where’s the ball?); or (iii) MIXED-LANGUAGE SENTENCE (Dónde está la ball?). Children successfully identified words across conditions; however, the advantage linked to hearing words in sentences was present only in the same-language condition. This work hence suggests that language mixing plays an important role on bilingual children’s ability to recognize spoken words.

Keywords: bilingualism; word recognition; code switching; sentence processing

Introduction

Parents frequently name objects in sentence frames that may seem to add little substantive information (“This is a ______!”). Nevertheless, work by Fernald and Hurtado (2006) suggests that young children show better word identification when they are presented with familiar words in a sentence, compared to words in isolation. That is, there is an advantage associated with sentence context, presumably linked to having continuous speech that facilitates anticipation of the object name that will come at the end of the frame.

This finding of a sentence-context advantage comes from work with monolingual children, but given that bilingualism is present in most countries around the world and across various ages and levels of society (Grosjean, 2013), many children are growing up receiving input in more than one language. In fact, bilingualism is increasingly common even in primarily monolingual countries like the United States (US Census, 2007; see Shin & Kominski, 2010). Studies examining the characteristics of the input that children in bilingual environments receive suggest that, not only are children exposed to utterances in each language, but there is also frequent exposure to sentences that ‘mix’ the two languages (Bail, Morini, & Newman, 2015;
The mixing of two languages while speaking is referred to as 'code-switching' (CS). It is unclear whether hearing sentences that contain CS influences word recognition, and whether the same sentence context advantage reported in monolingual studies is present with bilingual children. Here we explore this topic.

Acquiring two languages simultaneously may seem challenging or confusing, but infants are extremely skilled at doing this. Studies examining language discrimination in young children raised in bilingual environments suggest that, by 4 months of age, infants are able to tell the two languages apart (Bosch & Sebastián-Gallés, 2001). Before reaching middle childhood, bilingual children are explicitly aware of the fact that they are being exposed to two separate languages, and not a single language that is a mixture of the two (Genesee, Nicoladis, & Paradis, 1995). While on some occasions young bilinguals do produce mixed utterances, this is believed to be the result of not knowing a specific word in the target language rather than confusion (Deuchar & Quay, 1999; Paradis, Nicoladis, & Genesee, 2000). In fact, even highly proficient bilingual adults code-switch (Myers-Scotton, 2006), and in the case of adults, CS is attributed to different contexts (e.g., certain topics or words are always produced in one language) or interlocutors (e.g., if the other person also knows the two languages).

The majority of research on CS has focused on either adult-to-adult speech, or conversations between an adult interlocutor and a school-aged child (Cheng & Butler, 1989; Poplack, 1980); much less is known about CS in speech produced when talking to infants, and the role that this type of input might play on the development of early language skills. Place and Hoff (2011) studied diary data from parents and found that both English and Spanish vocabulary scores at 25 months were linked to the number of English-only and Spanish-only time blocks that a child heard. This relation was not present when taking into account the amount of exposure children received to blocks of time where both languages were heard (i.e., the number of half-hour blocks that contained both English and Spanish input). Similar findings were reported with French–English bilinguals by David and Wei (2008). However, Byers-Heinlein (2012) found that higher rates of CS were linked to smaller receptive vocabularies at 1.5 years, and slightly smaller expressive scores at 2 years compared to children whose parents reported lower rates of CS. The author argues that exposure to CS might make it harder for infants in bilingual environments to rely on cues that facilitate separation of the two language systems, which in turn affects the learning mechanisms that are responsible for early vocabulary growth. However, these findings come from data collected solely through parental report.

In a more recent study, Bail et al. (2015) examined CS in speech productions of Spanish–English bilingual caregivers obtained during a play session with their 18- to 24-month-old children. Over one-third of utterances produced by caregivers contained CS, and these included both inter-sentential (e.g., look at all the toys! tu con cuál quieres jugar?) and intra-sentential (e.g., look at all the juguetes!) switches. Importantly, exposure to CS was not related to delays in vocabulary development in bilingual children, based on Spanish and English MacArthur-Bates Communicative Development Inventory (Fenson et al., 1994) scores. While this finding is encouraging, a vocabulary measure only assesses whether children eventually learn the words from their caregivers; such a measure does not capture any processing difficulties that may take place in the moment when the CS is heard. For example, does hearing a sentence that contains a CS influence bilingual children’s word recognition in ‘real time’?
To our knowledge, only two studies have explored this topic. Byers-Heinlein, Morin-Lessard, and Lew-Williams (2017) tested French–English bilingual 20-month-olds on their ability to recognize familiar words presented as part of either a same-language sentence (e.g., *find the dog!* or a mixed-language sentence (e.g., *Find the chien!*). Analyses of looking patterns suggested that accuracy was lower in mixed-language sentences. Furthermore, the direction of the language switch was important, with CS particularly affecting word recognition when the switch happened from the child’s dominant language (i.e., the one that they were exposed to more frequently) to their non-dominant language. These findings were expanded by Potter, Fourakis, Morin-Lessard, Byers-Heinlein, and Lew-Williams (2019) with Spanish–English bilingual toddlers. Here, participants with varied levels of language dominance were tested in both languages. CS affected word recognition when the sentence frame was heard in the child’s dominant language (as in the earlier study), but not when the sentence frame was in the non-dominant language. Hence, while CS may not lead to global delays in vocabulary learning, it can influence bilingual children’s word recognition in the moment – and this appears to be linked to the direction of the switch. But it is less clear what might be driving this effect, and how it might relate to the sentence-context advantage identified in monolingual children.

One possibility is that sentence frames consistently provide a benefit to young children, but that the benefit is even greater when speech is all in one language. Fernald and Hurtado (2006) found that continuity of speech facilitated word recognition in monolinguals; this might also be the case for code-switched sentences. Yet when the sentence is all in the same language, the sentence frame might also cue the child as to the likely language of the noun, providing an additional benefit. That is, even if mixed-language sentences do not provide as much of an advantage as sentences that do not contain CS, they might still facilitate bilingual word recognition to some extent and be easier to process than hearing words in isolation.

A second possibility is that hearing a sentence that starts in one language and ends in another creates additional cognitive load (Byers-Heinlein et al., 2017) that eliminates the benefits of a sentence frame altogether. Bilingual children might, in essence, treat the sentence as two distinct units: a frame in one language, and a single word in the other, akin to a word presented in isolation. Here, a CS sentence and a single word might be functionally equivalent in some sense.

Logically, a third possibility is that a sentence that contains two languages might actually confuse or mislead young listeners. In this case, children might show poorer recognition for a CS sentence than for a word presented in isolation: by narrowing their predictions to hearing a particular language, children may be led down a lexical garden-path from which it would take time to recover.

Of course, a final possibility is that bilingual children simply do not show a sentence frame advantage at all, even in same-language sentences. Perhaps their frequent experience hearing utterances that contain CS results in them being less likely to use a sentence frame to help predict what might come next.

### Experiment

We examined the role that CS plays on sentence processing and word recognition in Spanish–English bilingual toddlers. Specifically, we examined: (i) whether the sentence-context advantage previously reported with monolingual children also plays...
a role in bilinguals’ word identification; and (ii) whether this advantage is observed with utterances that contain intra-sentential CS.

Methods

Participants

A total of 24 children (12 f, 12 m) between the ages of 18 and 24 months of age (M = 20.4 months; SD = 1.88) participated. Children came from households where both English and Spanish were spoken, and had not been previously diagnosed with any developmental or physiological diagnoses. Participants were exposed to a minimum of 30% and a maximum of 70% of each of the two languages from birth. Language dominance was based on the most-heard and least-heard language (based on parental report), with half of the children hearing Spanish as the dominant language, and the other half English. Children were exposed to different varieties of Spanish and came from both low- and mid-socioeconomic status (SES) homes (based on parental education). Data from an additional 28 participants were excluded due to technical problems (n = 3), caregiver interference (n = 1), not meeting the language exposure requirements (n = 11), and fussiness (n = 13). Fussiness was defined as inattention to the task, and included both children who cried during the study or who refused to sit down and look at the screen.

Stimuli

Participants saw pictures of familiar objects presented in pairs. Simultaneously, they heard speech stimuli of a single word (e.g., ball!) or a sentence instructing them to look at one of the two pictures, with the target word appearing in sentence-final position. Sentences were presented in either the same language condition (e.g., where’s the ball?), or a mixed language condition, where the context was in one language and the target word was in the other language (e.g., dónde está la ball?). The language change always occurred between the determiner and the noun, a common switch location for Spanish–English bilingual speakers (Bail et al., 2015). The auditory stimuli consisted of 6 target words (ball, hand, bed, key, shoe, book) that were familiar to participants, based on caregiver report. Target words were presented in individual children’s dominant language, and therefore stimuli were created in both Spanish and English. All sentences were produced by the same female voice (a Spanish–English bilingual) using child-directed-speech prosody.

The speaker produced speech for all three conditions as natural recordings, rather than via cross-splicing. Plunkett (2006) reported that 17-month-olds have difficulty processing words that include inappropriate coarticulation (e.g., words that were recorded in a sentence and then spliced out). There was some concern that children in this study would similarly not be ‘tolerant’ of absent or misleading coarticulatory cues. Thus, stimuli for the mixed-language condition included natural coarticulation. The speaker started the sentence in Spanish and produced the final word in English (or vice versa). Words for the single word condition were produced in isolation. While these methodological choices meant that the tokens for the target words across the three conditions were different, it was presumed that this was a lesser concern than creating sentences with inappropriate coarticulation. To help reduce any differences across conditions and across language versions, the speaker was
instructed to produce target sentences with similar intonation, and individual tokens that had comparable pitch and intonation were then chosen for the study.

Target objects were chosen such that they did not result in strong semantic competitors, nor in strong phonological overlap. Additionally, we selected pairs of objects that would be of the same grammatical gender in Spanish, so that the same determiner could be legitimately used as part of the sentence frame (e.g., el libro ‘the book’ with el zapato ‘the shoe’, and la cama ‘the bed’ with la llave ‘the key’). This was an important factor given evidence suggesting that bilingual infants can use grammatical gender information to facilitate word recognition (Lew-Williams & Fernald, 2007). The videos used during the study are accessible in a public scientific repository (https://osf.io/k2vxz/).

Between trials a short attention-getter (a bouncing star) appeared on the screen to provide participants with a break from the stimulus and to direct their attention to the center of the screen.

**Apparatus and procedure**

Participants sat on their caregiver’s lap, 4 feet from the stimulus-presentation screen, and completed a version of the Preferential Looking Procedure (Golinkoff, Ma, Song, & Hirsh-Pasek, 2013). Images were presented in pairs, and a video camera positioned above the screen recorded children’s eye-movements. Each trial began with an attention-getter, which continued to play until the experimenter had confirmed via mouse click that the child was looking at the screen. Participants then saw two pictures presented side-by-side on a white background. Speech stimuli were then delivered at approximately 70–75 dB SPL. All trials were the same length (5600ms). A total of 18 test trials (6 in each condition) were included. Additionally, participants saw 3 baseline trials (one per object pair) where there was no audio. Baseline trials were averaged for each child and used to rule out a pre-existing side bias, which was defined as looking to one side for greater than 85% of the time – a value used in other split-screen studies with young children (de Haan, Johnson, Maurer, & Perrett, 2001). As noted above, children who were mostly exposed to English (i.e., English-dominant, n = 12) always heard target words in English, and children who were mostly exposed to Spanish (i.e., Spanish-dominant, n = 12) always heard target words in Spanish. This meant that the carrier sentence in the mixed-language condition was always heard in the non-dominant language. This direction of the language switch is one that was found to be the easiest in the Byers-Heinlein et al. (2017) study with children of the same age. As suggested by Potter et al. (2019), word comprehension can be affected differently when target words are heard in the dominant versus non-dominant language – hence, as a starting point we opted to consistently present target words in the child’s ‘strongest’ language. The position of the objects on the screen and the target noun were counterbalanced across trials. The order in which trials appeared was randomly selected by the computer. Caregivers wore noise-reducing headphones and listened to masking music during testing, to prevent them from possibly influencing the children’s looking behavior.

**Coding**

Participant videos were coded off-line on a frame-by-frame basis using Supercoder (Hollich, 2005). Highly trained coders recorded the durations of participants’
eye-movements to the right or left object on the screen. All videos were coded by two coders blind to the target object location, and any trials where there were discrepancies greater than 15 frames were coded by a third person. Values across coders were compared for reliability and resulted in correlations ranging from $r = .98$ to $r = .99$. Participants’ looking patterns were analyzed for accuracy, with accuracy defined as the proportion of looking relative to the total time spent fixated on either of the two pictures, averaged over a 5000ms window from the onset of the target word. This window of analysis is longer than what has been previously used with other Preferential Looking studies, but given that (i) there is little prior knowledge regarding the effect that CS might have on sentence processing, and (ii) a code-switch might actually slow processing down, we opted to look at a wider window.

**Results**

Examination of the proportions of infants’ fixation patterns revealed that accuracy was highest for the same language condition (63%), followed by the mixed language sentence (54%) (see Figure 1). One-tailed single-sample t-tests indicated that all scores were reliably above chance ($t(23) = 8.21, p < .001$, Cohen’s $d = 1.68$; $t(23) = 3.32, p = .003$, Cohen’s $d = 0.68$; $t(23) = 2.51, p = .02$, Cohen’s $d = 0.51$, respectively). A one-way repeated-measures ANOVA was used to compare mean differences in the fixation proportions across the three conditions. This analysis revealed a significant effect of sentential-context ($F(2,46) = 8.27, p = .001$, $\eta^2_p = .264$). Post-hoc t-tests corrected for multiple comparisons with a Bonferroni correction showed significantly higher accuracy in same-language-sentences than isolated-word ($t(23) = 4.40, p < .001$, Cohen’s $d = 0.89$). There was also a significant difference in accuracy between the same-language and mixed-language sentence trials ($t(23) = 2.28, p = .04$; Cohen’s $d = 0.46$). However, accuracy between the mixed-language-sentence and isolated-word conditions did not differ significantly ($t(23) = 1.33, p = .70$, Cohen’s $d = 0.27$). In other words, children’s word recognition was better when the target word was accompanied by a sentence that did not contain CS, compared to a sentence that contained a language switch or no sentential frame at all.

In order to directly compare the comprehension abilities of children tested with the English versus Spanish target words, a repeated-measures ANOVA with sentential-context (mixed language, same language, single word) as the repeated measure, and language dominance (English, Spanish) as a between-subjects factor, was conducted. This analysis revealed a significant main effect of sentential-context ($F(2,44) = 8.15, p = .001$, $\eta^2_p = .27$), but no main effect of language dominance ($F(1,22) = 0.87, p = .36$, $\eta^2_p = .04$), and no significant interaction ($F(2,44) = 0.62, p = .54$, $\eta^2_p = .02$). This suggests that there was no difference in performance between children tested in English versus in Spanish.

**Discussion**

It is not uncommon for young children being raised in bilingual environments to hear speech that contains code-switching (Byers-Heinlein, 2012). This includes situations where the language change occurs in the middle of a sentence (i.e., intra-sentential CS). Prior findings suggest that the amount of exposure to CS may not necessarily play a critical role on vocabulary development (Bail et al., 2015; David & Wei, 2008;
Less is known about how CS influences sentence processing, and whether young bilingual children experience processing difficulties (and hence poorer word recognition) in the moment when the language switch is heard.

The present study examined what factors may facilitate or negatively affect word recognition in Spanish–English bilingual children. Specifically, we tested the usefulness of having a sentence frame that accompanies the target word, AND what happens when that sentence contains CS. First, we found that accuracy was significantly better when target words were presented as part of a sentence that was all in one language, compared to when the word was heard in isolation. Previous work with monolinguals suggested that, despite being easier to segment, hearing words in isolation leads to less accurate word recognition. The inclusion of sentence-frames, on the other hand, results in a facilitatory effect (Fernald & Hurtado, 2006). Hence, our data extend this previous finding to Spanish–English bilingual toddlers.

Second, when comparing performance across the same language and mixed language sentence conditions, we found that accuracy was significantly better when only a single language was heard (that is, when no CS was present), replicating findings from Byers-Heinlein et al. (2017) with French–English bilinguals. Our work supports the notion that CS influences bilingual children’s word recognition during real-time sentence processing. It also suggests that this is true across different
language combinations. Another difference between the two studies is that in Byers-Heinlein et al., all trials were essentially in the same matrix language, with only an occasional word in the other language inserted; in the present study, the target words were always in a consistent language, while the language of the carrier sentence switched. This meant that participants were exposed to CS both within sentences in a trial, but also across trials (albeit on a somewhat different timescale since an attention-getter video with no speech played in between trials). The fact that the results are comparable across studies speaks to the consistency of the effect across different types of CS sentences.

Third, the comparison between performance in the mixed language sentence and the single word conditions revealed no significant difference. This means that not only is hearing a sentence that contains CS less helpful than hearing a sentence in a single language, but having CS basically removes any advantage provided by the sentence frame, and leads to similar accuracy to when the word is heard in isolation. Hearing a mixture of the two languages within the same sentence may result in greater processing costs for bilingual children compared to when all the information in the sentence is presented in a single language. This implies (i) that children realize that a language switch has taken place, which supports theories suggesting that bilingual children develop differentiated language systems early on (Byers-Heinlein, 2014; Genesee, 1989), and (ii) that adjusting to that switch requires some additional resources, which may be guided by how activation and inhibition of the two languages take place.

When the target word and the sentence frame are in the same language, the sentence provides cues to help the child narrow down what the target word might be. For example, given that bilingual children must acquire two labels for the same referent (one for each language), the presence of a sentence frame might be useful if it cues the child towards the appropriate language. That is, for a monolingual child, look at the indicates that a noun will follow. For a bilingual child, it could also indicate the language of the noun. When the sentence includes CS, it is harder for children to make use of sentence cues. Nevertheless, even if some cues are lost in sentences that contain CS, there is still speech continuity, which Fernald and Hurtado (2006) found to facilitate word recognition with monolinguals. One question we posed was whether that continuity would still lead to some advantages for word identification. Our data suggest that this is not the case, since performance with the mixed-language sentences was the same as with the words in isolation. To process intra-sentential CS, bilingual listeners need to masterfully juggle the two language systems. One possibility is that they start by preferentially activating lexical items in one language, then quickly inhibit that initial language system, switch to access items in the other language (to retrieve the correct word–object mapping), and finally achieve identification. In fact, there is extensive research with adults suggesting that, while there are lexical nodes for each language that are simultaneously activated by a single concept, bilinguals primarily activate the expected language and/or inhibit the non-expected language (Green, 1998; Grosjean, 2001; Spivey & Marian, 1999). Performing multiple steps as CS occurs is likely a highly demanding process, which ends up outweighing the benefits that might be provided by the presence of sentence frames.

Looking at these results another way, the fact that mixed-language sentences were akin to single words, and that accuracy across conditions was above chance, implies
that the CS was not causing active processing difficulties for bilinguals. This is true, at least when target words are heard in the child’s dominant language—a pattern consistent with previous results by Byers-Heinlein et al. (2017) and Potter et al. (2019). Speech addressed to children often includes both multi- and single-word utterances (Brent & Siskind, 2001; Fernald & Morikawa, 1993). It is not the case that words in isolation are considered to be detrimental to children’s sentence processing and language acquisition, but rather that they provide less information to facilitate word identification. Our data suggest that mixed-language sentence frames may not be helpful, but are also not actively preventing children from correctly identifying words that are heard in the dominant language.

Additional research is needed to better understand what specific factors may contribute to the role that CS plays in word recognition. Little is known about maturational differences in bilingual word recognition using mixed-language sentences. Questions also remain regarding how the specific location of the language switch contributes to sentence processing. The Byers-Heinlein et al. (2017) study provides some preliminary findings related to these questions: First, bilingual adults (like children) showed better word recognition when sentences were presented in a single language than when there was CS. This suggests that processing mixed-language sentences comes at a cost, even for adults who have considerable experience using both languages. It is unclear, however, what the degree of that cost is, and whether bilingual adults would similarly show no difference in accuracy between the mixed-language and single-word conditions. Second, when bilinguals were presented with inter-sentential language switches—instead of CS within the same sentence (e.g., That one looks fun! Le chien!)—both infants and adults showed no difference in accuracy for sentences that contained CS and those that did not. This suggests that the location of the language switch does matter but, based on data available to date, it is not clear exactly how particular linguistic structures interact with the processing of CS sentences (e.g., when the determiner and the noun are in the same language but preceded by items in the other language within the same sentence—Look at la pelota!). Last, the present study only tests bilingual children with words presented in their dominant language. Additional research is needed to better understand the role of CS on comprehension when words are heard in the non-dominant language.

To conclude, our findings contribute to the literature on how bilingual children process CS, and the role that hearing this type of input plays on word recognition. While exposure to language mixing does not necessarily lead to delays in bilingual children’s vocabulary learning, it does influence their ability to recognize spoken words. Toddlers can successfully identify words heard as part of a sentence that contains CS. However, the previously reported word-recognition advantage that results from hearing a word in a sentence frame is only present when the sentence and target word are presented in the same language.

**Acknowledgments.** We thank George Hollich for providing the coding software used to analyze the participants’ looking behavior. We also thank Emily Shroads and members of the Language Development Lab for assistance in scheduling and testing of participants. This work was supported in part by a National Science Foundation IGERT Institutional Training Grant (DGE 0801465) awarded to the University of Maryland, by a University of Maryland Center for Comparative and Evolutionary Biology of Hearing Training Grant (NIH T32 DC000046-17), by a fellowship from the Hearing and Speech Sciences Program at the University of Maryland, and by start-up funds from the University of Delaware.
References


Byers-Heinlein, K. (2014). Languages as categories: reframing the ‘one language or two’ question in early bilingual development. Language Learning, 64(s2), 184–201.


---

**Cite this article:** Morini G, Newman RS (2019). Dónde está la ball? Examining the effect of code switching on bilingual children’s word recognition. *Journal of Child Language 46*, 1238–1248. https://doi.org/10.1017/S0305000919000400