

## Bilingual children's comprehension of code-switching at an uninformative adjective

Lena V. Kremin  
Concordia University, CA

Amel Jardak  
Concordia University, CA

Casey Lew-Williams  
Princeton University, USA

Krista Byers-Heinlein  
Concordia University, CA

**Abstract:** Bilingual children regularly hear sentences that contain words from both languages, also known as code-switching. Investigating how bilinguals process code-switching is important for understanding bilingual language acquisition, because young bilinguals have been shown to experience processing costs and reduced comprehension when encountering code-switched nouns. Studies have yet to investigate if processing costs are present when children encounter code-switches at other parts of speech within a sentence. The current study examined how 30 young bilinguals (age range: 37 – 48 months) processed sentences with code-switches at an uninformative determiner-adjective pair before the target noun (e.g., “Can you find *le bon* [the good] duck?) compared to single-language sentences (e.g., “Can you find the good duck?”). Surprisingly, bilingual children accurately identified the target object in both sentence types, contrasting with previous findings that sentences containing code-switching lead to processing difficulties. Indeed, children showed similar (and in some cases, better) comprehension of sentences with a code-switch at an uninformative adjective phrase, relative to single-language sentences. We conclude that functional information conveyed by a code-switch may contribute to bilingual children's sentence processing.

**Keywords:** code-switching; bilingualism; language processing; language acquisition

**Corresponding author(s):** Krista Byers-Heinlein, Department of Psychology, Concordia University, 7141 Sherbrooke St. West, Montreal, QC, H4B 1R6. Email: [k.byers@concordia.ca](mailto:k.byers@concordia.ca).

**ORCID ID(s):** Lena V. Kremin: <https://orcid.org/0000-0002-4010-490X>

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

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

**Citation:** Kremin, L.V., Jardak, A., Lew-Williams, C., & Byers-Heinlein, K. (2023). Bilingual children's comprehension of code-switching at an uninformative adjective. *Language Development Research*, 3(1), 249–276. <https://doi.org/10.34842/zyvj-cv60>

## Introduction

Bilingual children regularly hear both of their languages within a single conversation and even within a single sentence (e.g., *C'est un* [fr. It's a] monkey.). This phenomenon is known as code-switching. Most bilingual children hear code-switching in their daily lives (Kremin et al., 2021), and there is some evidence that over time code-switching may impact a child's vocabulary size (Bail et al., 2015; Byers-Heinlein, 2013) and overall language development (Kaushanskaya & Crespo, 2019). Code-switching can also reduce a child's comprehension in the moment as they process speech (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019). Given the importance of early language processing to early language acquisition (Meylan & Bergelson, 2022), it is important to understand the contexts in which code-switching does or does not affect language processing, as well as underlying mechanisms. To date, research on children's comprehension of code-switching has focused on code-switches at a noun (e.g., "*Dónde está la* [sp. where's the] ball?"), even though everyday code-switching happens at many different parts of speech, such as verbs, prepositions, and adjectives (e.g., "*C'est* [fr. It is] yucky.;" MacSwan, 2012). Here, we extend previous findings with nouns and investigate how code-switching at a mid-sentence determiner-adjective pair affects bilingual children's language comprehension.

A large body of literature has reported that bilingual adults process code-switches more slowly than single-language stimuli (for recent reviews see Beatty-Martínez et al., 2018; Valdés Kroff et al., 2018; van Hell et al., 2018), but researchers have only recently begun to study how young children process code-switches. One eye-tracking study indicated that children process code-switches differently depending on whether the switch happens between sentences or within a single sentence. When hearing between-sentence code-switching (e.g., "That one looks fun! *Le chien* [fr. the dog]!"), 1.5- to 2-year-old children were as accurate at identifying the target object as they were when hearing a single language (e.g., "That one looks fun! The dog!"; Byers-Heinlein et al., 2017). However, when hearing within-sentence code-switching (e.g., "Look! Find the *chien* [fr. dog]!"), children were less accurate at identifying the target object compared to hearing a single language (e.g., "Look! Find the dog!"; Byers-Heinlein et al., 2017; Morini & Newman, 2019). Such studies with young children have focused solely on code-switches at the noun, so they do not address the potential impact of code-switching at other parts of speech. This limitation makes it impossible to draw generalized conclusions about how code-switching may or may not affect comprehension. Children may process code-switching at different parts of speech more readily depending on several factors, such as how often children hear code-switching in that location or what functional information is contained in the code-switched word(s). Evaluating children's comprehension of code-switching at different parts of speech can provide insights into the veracity of two general accounts of what makes

code-switching difficult to process, which we introduce here as the frequency account and the functional account.

### Frequency Account

The frequency account posits that how easily bilinguals process a code-switch depends on how frequently that type of code-switched construction occurs in their everyday life (e.g., Abutalebi et al., 2007; Guzzardo Tamargo et al., 2016; Salig et al., 2023). This account predicts that frequent code-switched constructions will be more easily processed than infrequent code-switched constructions. For example, in one study, Spanish–English bilingual adults more readily processed a common code-switch that included an entire compound verb (e.g., “*los senadores* [sp. the senators] **have requested** the funds”) than an uncommon code-switch that occurred in the middle of the compound verb (e.g., “*los senadores han* [sp. the senators **have**] **requested** the funds”; Valdés Kroff et al., 2018). Similarly, Welsh–English bilingual adults judged code-switching at common parts of speech, such as nouns, to be more acceptable than code-switching at uncommon parts of speech, such as adjectives (Vaughan-Evans et al., 2020). The frequency account could also predict differences in comprehension between bilingual populations if they hear different rates of code-switching in their daily lives (Gosselin & Sabourin, 2021; Valdés Kroff et al., 2018).

If frequency is indeed an important factor in how bilingual adults process code-switching, its importance could also extend to children’s processing. Under the frequency account, children would be expected to understand code-switching at frequently code-switched parts of speech, such as nouns, more easily than at infrequently code-switched parts of speech, such as adjectives. This account could explain existing findings about children’s processing of code-switching. Indeed, when children do hear within-sentence code-switches, they often occur at nouns (Bail et al., 2015). Moreover, children hear more between-sentence code-switches than within-sentence code-switches from their parents (Bail et al., 2015; Kremin et al., 2021), so the frequency account is consistent with the experimental finding that children more easily process between-sentence code-switches compared to within-sentence code-switches (Byers-Heinlein et al., 2017; Morini & Newman, 2019). Thus, if within-sentence code-switches at a relatively common location for code-switching (i.e., the noun) can disrupt children’s processing, then within-sentence code-switches at an uncommon location should be even more disruptive.

### Functional Account

Another account – related to yet different from the frequency account – proposes that bilinguals process code-switches differently based on the *functional* properties of the code-switched word(s), including grammatical properties. While prior research has investigated a variety of functions of code-switching in production – such as adding

emphasis, signaling community identity, and facilitating understanding (Goodz, 1989; Heredia & Altarriba, 2001; Nilep, 2006) – comprehension studies have mainly focused on the functional dimension of grammatical class. One study of German–Russian bilingual adults used event-related potentials (ERPs) to examine the processing of code-switches at open-class words (e.g., nouns) versus closed-class words (e.g., prepositions). While code-switches at both nouns and prepositions elicited a broad late positivity, only code-switches at prepositions elicited a broad early negativity, suggesting that bilinguals process code-switches differently based on their grammatical function (Zeller, 2020). Another ERP study compared how bilinguals processed code-switching at two types of open-class words: nouns and verbs (Ng et al., 2014). When reading a story, Spanish–English bilingual adults processed code-switching at nouns (e.g., “the wind and the *sol* [sp. sun]”) differently than code-switching at verbs (e.g., “they *mira-ron* [sp. saw] a traveler”) as indicated by larger N400 responses and an early Late Positive Component for nouns. The authors proposed that the difference was driven by the effort bilinguals put into integrating and remembering the information contained in each code-switch. That is, nouns are likely to be referenced several times in a story and need to be held in working memory, thus eliciting more cognitive effort compared to verbs that may only be used once. Combined, these results highlight that bilinguals may be sensitive to the functional role of the code-switched words and process them accordingly.

Research has yet to investigate how bilingual children process code-switches with diverse functional or grammatical roles, but evidence from monolinguals shows that children are sensitive to some grammatical classes beginning around 8 months of age (Marino et al., 2020). Moreover, by age 3, children use the meaning of adjectives to predict which noun they refer to (e.g., predicting “heavy” is more likely to be followed by “stone” than “butterfly”; Tribushinina & Mak, 2016). Additionally, monolingual children as young as 2 years old can recognize, but “listen through,” uninformative adjectives to quickly and correctly identify a target noun (Thorpe & Fernald, 2006). For example, when shown a picture of a dog and a bunny, children identified the target object as quickly when it was preceded by an uninformative adjective [e.g., “Where’s the good bunny?”] as when it was not preceded by any adjective (e.g., “Where’s the bunny?”). These results show that young children can attend to the most relevant functional information to efficiently process speech.

Following the functional account, code-switching that occurs at a word that is central to the meaning of the sentence may be particularly challenging for children to process. In many cases, this will be a noun, but in other cases it could be a verb, adjective, or other part of speech, depending on context. This idea is supported by previous research showing that children experience difficulty in understanding functionally-important code-switched nouns (Byers-Heinlein et al., 2017; Morini & Newman, 2019). In contrast, code-switches at parts of speech that play a limited functional role in

comprehension may be relatively easy for children to process, and code-switches that are uninformative in a comprehension task may not elicit any processing difficulties. However, to date, children's comprehension of code-switches at words with limited functional meaning has not yet been investigated; thus there is a lack of empirical evidence for the functional account with children.

### **Current Study**

In the current study, we asked if code-switching within a sentence at an uninformative determiner-adjective pair (which we will hereafter refer to as an uninformative adjective) affects children's comprehension of a target noun that immediately follows it. This allowed us to examine the potential contributions of frequency and/or functional factors in children's processing of code-switching. The frequency account predicts that children will show disrupted processing of a code-switch at an adjective, because it is not a common location for code-switching. This could result in weaker comprehension of the following noun, as processing difficulties earlier in the sentence can negatively affect how children process the end of the same sentence (Truewell et al., 1999). In contrast, the functional account predicts that children may find it relatively easy to process a code-switch at an uninformative adjective as they do not necessarily have to attend to or remember its meaning in the context of the visual scene.

In an eye-tracking experiment, children viewed pairs of pictures of animals, such as a duck and a fish, and heard sentences such as "Can you find *le bon* [fr. the good] duck?" or "Can you see *el buen* [sp. the good] duck?" In trials, both animals were equally consistent with the adjective (e.g., both were depicted as equally "good"). Participants were 30 3-year-old bilinguals, including both French-English bilingual children in Montreal ( $n = 19$ ) and Spanish-English children in New Jersey ( $n = 11$ ). This age group was chosen because, from this age, children can attend to the information in adjectives in real-time sentence comprehension (Tribushinina & Mak, 2016). We included participants from these two testing locations to increase sample size, as bilingual children are a difficult-to-recruit population. This is in line with various sampling strategies in the field of early bilingualism which range from testing homogeneous populations (e.g., all acquiring English and French) to testing heterogeneous populations (e.g., all acquiring English and a variety of other languages; Byers-Heinlein, 2015). Assessing the effects of code-switching at adjectives was appropriate in our sample, because children of this age can generally understand their meaning (Tribushinina & Mak, 2016), and because certain adjectives can occur in the same pre-nominal position across the languages being acquired by our participants (i.e., English, French, and Spanish).

Similar to previous studies on children's processing of code-switching (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019), we expected that code-switching at an uninformative adjective would hinder children's comprehension of the target noun compared to sentences without code-switching. Specifically, we predicted that children would look less towards the target noun after hearing mid-sentence code-switching compared to hearing a sentence entirely in one language. Such a result would be consistent with the frequency account. In contrast, a finding that children's performance was unaffected by an uninformative code-switched adjective would be consistent with the functional account. We also explored whether individual differences such as language dominance, testing location (as a proxy for language pair), SES, or vocabulary size would be related to performance.

### Method

Data collection occurred in two locations: Montreal, Canada and New Jersey, USA. The methods were approved by the Concordia University Human Research Ethics Committee ("Monolingual and Bilingual Language Development"; approval #10000493) and the Princeton University Institutional Review Board ("Language learning and Communication"; approval #7117), and parents provided informed consent prior to their child's participation. Data were collected in Montreal between November 2016 and April 2017 and in New Jersey between March 2017 and January 2018. Final data analysis occurred between May 2020 and June 2021, during the COVID-19 pandemic. As is common in laboratories testing hard-to-recruit populations such as bilingual children, children participated in a second, separate study, either immediately prior to or following participation in this study (the order of the two studies was counterbalanced). The results of that study are reported in a separate manuscript (Byers-Heinlein et al., 2021). All stimuli, data, and analysis scripts for the current study are available via the Open Science Framework at <https://osf.io/ecqwr/>.

### Participants

A total of 30 3-year-old ( $M = 3.57$ , range = 3.10 – 4.05, 14 females) full-term, healthy bilingual children participated in this study. This sample size was sufficiently sensitive to detect an effect size of  $d = 0.46$  at 80% power in a paired-samples  $t$ -test, meaning there were enough participants to detect effect sizes reported in previous related studies (0.56 in Byers-Heinlein et al. 2017; 0.60 in Potter et al. 2019).

Nineteen French–English bilinguals were tested in Montreal, Canada, and 11 Spanish–English bilinguals were tested in New Jersey, USA. In Montreal, children were recruited from a database of families interested in participating in our research, principally identified via government birth lists. In New Jersey, children were primarily recruited from nonprofit organizations. Another 34 children were tested but not

included in the final sample due to not meeting the language criteria ( $n = 15$ ; see details below), fussiness or lack of attention ( $n = 10$ ), technical issues ( $n = 4$ ), health reasons such as low birth weight or gestation period under 37 weeks ( $n = 3$ ), completing an insufficient number of trials ( $n = 1$ ; see below), or having a reported speech delay ( $n = 1$ ). Post-hoc data exclusion resulted in the unbalanced sample between the two locations. Unfortunately, because this discrepancy did not become clear until the time of data analysis, which occurred during the COVID-19 pandemic, we were unable to test additional participants to address this difference. Parents reported their child's ethnicity/race using categories and a free-response option appropriate to each location. Among French-English bilinguals in Montreal, 10 children were European, 2 were Canadian, 2 were Caribbean, 1 was Arab, 1 was Quebecois, and 3 did not report. Among Spanish-English bilinguals in New Jersey, 8 children were Hispanic, 1 was Black, 1 was White, and 1 was from multiple ethnic/racial backgrounds.

### ***Language Background and Proficiency***

Children's language background and proficiency was assessed via a modified version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). Parents were asked about their child's experience with each language, and to rate their child's proficiency in English and French (in Montreal) or in English and Spanish (in New Jersey) compared to monolingual children of the same age. Following a pre-determined inclusion criterion, children had to receive a comprehension score of at least 7/10 for both languages to be eligible for the study, to ensure that children were reasonably proficient in both languages. For each child, their dominant language was established as the language that had the highest comprehension score from the LEAP-Q. Twelve children had equal comprehension scores in both languages, so for these children, the language in which the child had the higher productive vocabulary score (see below) was considered their dominant language. In total, 19 children were dominant in English, 9 were dominant in French, and 2 were dominant in Spanish. Twelve children were regularly exposed to both of their languages from birth, and 18 children were exposed to their second language later in life, between the ages of 2 and 36 months. See Table 1 for details by testing location.

### ***Vocabulary Size***

Children's productive vocabulary size in English was assessed using the Developmental Vocabulary Assessment for Parents (DVAP; Libertus et al., 2015), which consisted of a checklist of words known by children aged 2 to 18 years old based on words used in the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007). We used a parent checklist rather than a direct measure to reduce children's fatigue, as each child participated in two experiments, and we wished to assess their vocabulary in both languages. Moreover, the DVAP has shown strong convergent validity with children's performance on the PPVT ( $\beta = .69$ ; Libertus et al., 2015).

**Table 1. Demographics of participants at each testing location.**

Location	<i>n</i>	Mean age in years (range)	Eng. dom. ( <i>n</i> )	L2 from birth ( <i>n</i> )	Later L2 (age range in months)	Dom. Lang. Vocab. ( <i>SD</i> )	Non-Dom. Lang. Vocab. ( <i>SD</i> )	Parent education ( <i>SD</i> )
Montreal	19	3.47 (3.1 – 3.99)	10	8	6 – 18	76.83 (33.91)	47.83 (30.19)	16.58 (2.17)
New Jersey	11	3.75 (3.19 – 4.05)	9	4	2 – 36	62.36 (26.22)	24.55 (18.34)	12.82 (5.06)

*Note.* Eng. dom. (*n*) lists the number of children at each testing location who were dominant in English; the remainder of children were dominant in either French if tested in Montreal or Spanish if tested in New Jersey. Later L2 (age range in months) only considers participants who were not exposed to both languages from birth.

To assess children's productive vocabulary size in French or Spanish, we adapted a checklist similar to the DVAP, based on words used in the adaptation of the PPVT for Quebec French (Échelle de Vocabulaire en Images Peabody; Dunn et al., 1993) or Spanish (Test de Vocabulario en Imagenes Peabody; Dunn et al., 1986). The words are ordered from easy (e.g., “ball”, “dog”) to hard (e.g., “honing”, “angler”), and parents were asked to indicate which words their child could say. There are 212 items on the English version, 190 items on the French version, and 125 items on the Spanish version. A parent or other adult that was familiar with the child's vocabulary in a particular language filled out the form for that language. In some cases, the forms for each language were completed by different parents who normally interacted with their child in that language, while in other cases it was one parent who filled out both forms if they used both languages with their child. As expected, the number of words children produced in their dominant language ( $M = 71$ ,  $SD = 32$ , range = 24 – 177) was greater than the number of words they produced in their non-dominant language ( $M = 39$ ,  $SD = 28$ , range = 2 – 131),  $t(28) = 7.03$ ,  $p < .001$ ,  $M_d = 32.34$ , 95% CI [22.92, 41.77]. When combining the number of words produced in both languages, on average, children produced 110 total words ( $SD = 55$ , range = 31 – 308). Children in Montreal ( $M = 125$ ,  $SD = 61$ , range = 39 – 308) produced more words than those in New Jersey ( $M = 87$ ,  $SD = 33$ , range = 31 – 138),  $t(26.73) = -2.16$ ,  $p = .040$ ,  $\Delta M = -37.76$ , 95% CI [-73.56, -1.95]), although we note that there were more words on the French version than on the Spanish version of the DVAP.



### **Exposure to Parental Code-Switching**

Children's exposure to parental code switching was measured via the Language Mixing Scale (Byers-Heinlein, 2013), which measures intra-sentential code switching by the primary caregiver. Scores can range from 0 (no switching) to 30 (highest amount of switching). In Montreal, caregivers' average score was 13.50, range = 4 – 28. In New Jersey, caregivers' average score was 14.80, range = 0 – 30. The difference in the amount of switching between the two locations was not significant,  $t(15.88) = -0.39$ ,  $p = .700$ ,  $\Delta M = -1.34$ , 95% CI  $[-8.61, 5.92]$ .

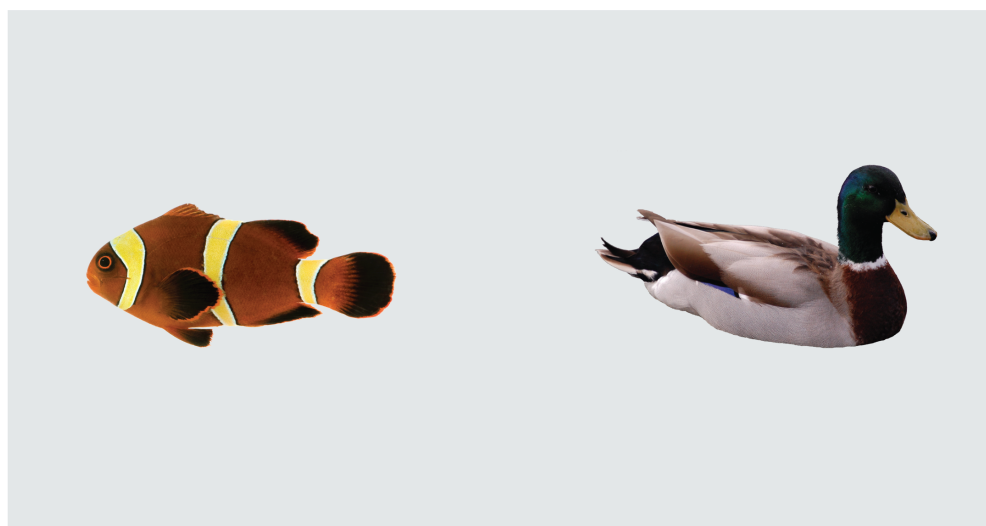
### **Socioeconomic Status**

As a proxy for socioeconomic status (SES), we asked parents to indicate the highest level of education they had attained. As the education systems are somewhat different in the United States and Canada, to be able to compare responses across our two testing locations, we converted these responses to the typical number of years after kindergarten to complete each level of education (e.g., completing a bachelor's degree was equivalent to 16 years of education). For families where both parents' education was provided, the higher level was selected for analysis. On average, parents completed 15.20 ( $SD = 3.89$ ) years of education, which ranged widely from 4 to 21 years. Parents in Montreal reported completing more years of education ( $M = 16.58$ ,  $SD = 2.17$ , range = 13 – 21) than parents in New Jersey ( $M = 12.82$ ,  $SD = 5.06$ , range = 4 – 20),  $t(12.17) = 2.35$ ,  $p = .037$ ,  $\Delta M = 3.76$ , 95% CI  $[0.27, 7.25]$ , suggesting that the participants in Montreal came from a higher SES background than those in New Jersey.

## **Material**

### **Visual Stimuli**

Visual stimuli consisted of 8 pairs of pictures for each language combination (See Table 2 for picture pairs and Figure 1 for an example trial). Each picture in a pair had the same animacy status (i.e., four pairs of animals used in target trials and four pairs of inanimate objects used in filler trials), so that the two pictures had similar visual salience. To ensure that they would be familiar to our 3-year-old participants, we selected pictures whose labels were highly understood by children in American English (Fenson et al., 2007), Quebec French (Boudreault et al., 2007), and Spanish (Jackson-Maldonado et al., 2003). The labels of the picture pairs did not overlap in word onset, had the same grammatical gender in French or Spanish, and are widely used across French and Spanish dialects. Pictures were chosen from free online libraries and digitally edited as necessary.



Look! ... Can you find the good duck?



**Figure 1.** Example and timeline of experimental trial.

### ***Auditory Stimuli***

Auditory stimuli were recorded by a female, native French–English or Spanish–English bilingual with no perceptible accent in either language using infant-directed speech. Each auditory stimulus contained a target word labeling one of the pictures on the screen (e.g., “Look! Can you find the good duck?”). The target noun (e.g., “duck”) was preceded by a determiner (e.g., “the”) and a prenominal adjective (e.g., “good”). Each stimulus sentence was recorded in a single-language version where the determiner and adjective were in the same language as the noun, and a code-switched version where the determiner and adjective were in the other language (e.g., “Look! Can you find *le bon* [fr. the good] duck?” or “Look! Can you see *el buen* [sp. the good] duck?”). Note that the target word (e.g., “duck”) was always in the same language as the initial carrier phrase (e.g., “Look! Can you find...” for French–English and “Look! Can you see...” for Spanish–English). Parallel stimulus sets were created with the carrier sentences in each language (e.g., in French, the previous examples became “*Regarde! Peux-tu trouver le bon canard?*” and “*Regarde! Peux-tu trouver the good canard?*”; in Spanish, the previous examples became “*¡Mira! Puedes ver el buen pato?*” and “*¡Mira! Puedes ver the good pato?*”).

For the animate nouns on target trials, there were a total of four English prenominal adjectives and their French and Spanish translations; similarly, for inanimate nouns in filler trials, there were four prenominal adjectives used (see Table 2). These adjectives were chosen such that they 1) were not cognates across French and English or Spanish and English, 2) did not share phonological overlap with their translation, 3) were not descriptive of one picture more than another, and 4) could precede a noun in French or Spanish. Although both French and Spanish usually place adjectives in a postnominal position, the adjectives we selected can be used prenominally in these grammatical contexts. Each adjective was always used with the same picture pair. All stimuli are available at <https://osf.io/ecqwr/>.

### ***Trial Description***

During each trial, the target and distractor pictures appeared on the screen for 6000ms, and one of the stimulus sentences was played labeling the target picture. The onset of the target noun occurred exactly 3000ms into each trial. The determiner–adjective combinations were of somewhat different lengths, and so occurred between 311 and 1152ms before the noun onset. Trials were combined into four experimental orders of 24 trials: 8 single-language trials (e.g., “Look! Can you find the good duck?”), 8 code-switched trials (e.g., “Look! Can you find le bon [fr. the good] duck?”), and 8 additional single-language filler trials. Filler trials were not analyzed and were mainly used to lower the overall number of trials with code-switching. Target trials (i.e., single-language and code-switched trials) and filler trials were intermixed throughout the study. The language of the carrier phrase was consistent for each child (i.e., always in English, French, or Spanish), but counterbalanced across children at the time of testing. In total, 15 children were tested with carrier phrases in their dominant language (10 French–English and 5 Spanish–English), and 15 children were tested with carrier phrases in their non-dominant language (9 French–English and 6 Spanish–English).

### **Procedure**

In addition to signing a consent form, parents completed questionnaires on their child’s vocabulary (DVAP) and language comprehension (LEAP-Q), on their own language mixing (Byers-Heinlein, 2013), and on basic demographic information. During the study, parents listened to music with headphones, wore darkened glasses, and were instructed not to interfere with the study or provide their child with any instruction. Testing occurred in a darkened room while children sat on their parent’s lap.

**Table 2: Adjective–noun pairs used for French–English and Spanish–English participants. The noun pairs labeled the two pictures shown on screen at the same time. Each noun was used as a target picture in different trials. In single-language trials, the adjective and noun were in the same language. In code-switched trials, the adjective and the noun were in different languages.**

<b>English</b> <i>Look! Can you find ... ?</i>		<b>French</b> <i>Regarde! Peux-tu trouver ... ?</i>	
<b>Adjective</b>	<b>Noun pair</b>	<b>Adjective</b>	<b>Noun pair</b>
<b>Target trials</b>			
the good	duck – fish	le bon	canard – poisson
the little	monkey – sheep	le petit	singe – mouton
the nice	dog – bunny	le gentil	chien – lapin
the pretty	cow – froggy	la jolie	vache – grenouille
<b>Filler trials</b>			
a large	ear – spoon	une grosse	oreille – cuillère
a new	apple – toothbrush	une nouvelle	pomme – brosse à dents
a big	door – hand	une grande	porte – main
an old	coat – pencil	un ancien	manteau – crayon

<b>English</b> <i>Look! Can you see ... ?</i>		<b>Spanish</b> <i>¡Mira! ¿Puedes ver ... ?</i>	
<b>Adjective</b>	<b>Noun pair</b>	<b>Adjective</b>	<b>Noun pair</b>
<b>Target trials</b>			
the good	bear – duck	el buen	oso – pato
the little	butterfly – sheep	la pequeña	mariposa – oveja
the big	bunny – dog	el gran	conejo – perro
the pretty	cow – froggy	la hermosa	vaca – rana
<b>Filler trials</b>			
a beautiful	ear – spoon	una linda	oreja – cuchara
a new	apple – toothbrush	una nueva	manzana – cepillo de dientes
a nice	door – hand	una preciosa	puerta – mano
an old	coat – pencil	un viejo	chamarra – lápiz

Due to differences in lab equipment, the same apparatus was not available at both testing sites. In Montreal, the study was conducted in the lab on a 24-inch Tobii T60XL corneal reflection eye-tracking system using a 5-point calibration, with auditory stimuli played over speakers. In New Jersey, the study was conducted either in the lab (7 children) or at a local community center (4 children), depending on which location was easier for participants to access. In the lab, visual and auditory stimuli were presented using Matlab on a 55" TV monitor. At the community center, visual stimuli were presented in a QuickTime video on a 13" laptop, and auditory stimuli were played through noise-canceling headphones. In both New Jersey setups, a video camera below the screen recorded children's eye movements at a rate of 30 frames per second for later offline coding by trained research assistants.

Before each trial began, a colorful attention-getter was presented to draw the child's attention to the screen. Once the child was looking at the screen, the trial began. An experimenter monitored the status of the study via video camera and controlled the experiment from a computer in another room (Montreal) or within the same room (New Jersey). The total duration of the study was approximately 4 minutes.

### **Coding**

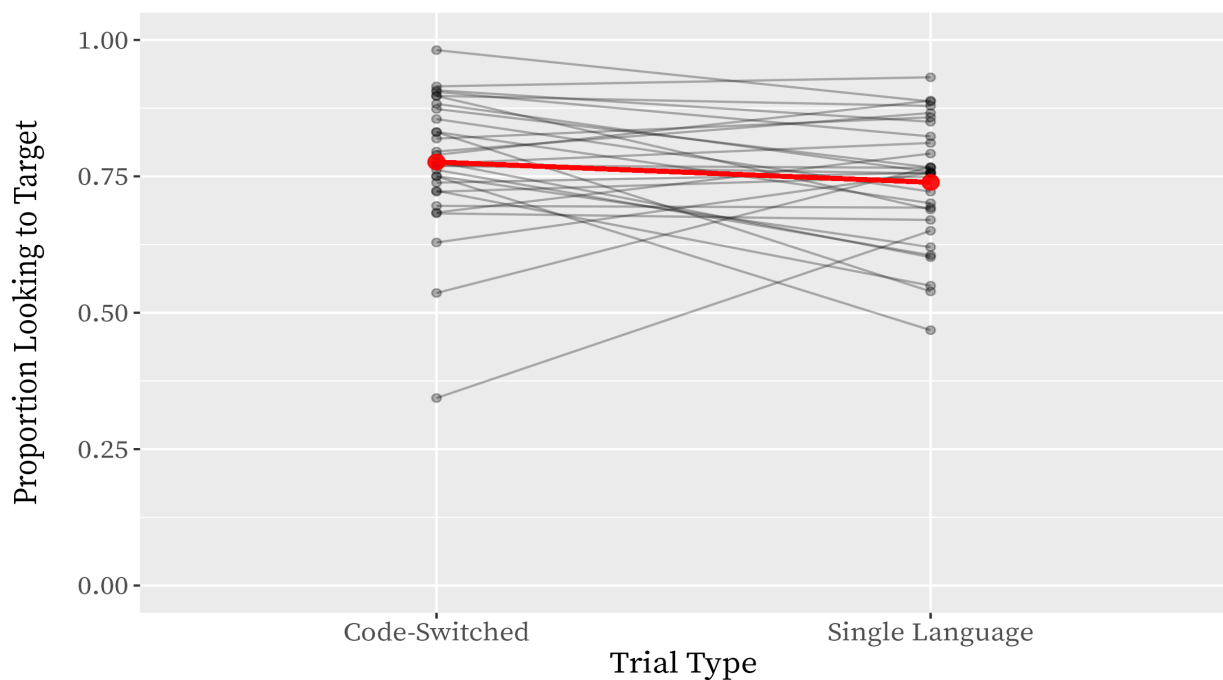
In Montreal, the eye-tracking system collected data on the location of children's eye-gaze and their pupil size at a rate of 60Hz. We defined areas of interest corresponding to a rectangle of 2 cm around each picture presented on the screen. In New Jersey, a trained research assistant used EyeCoder software to code at 33-ms intervals whether the child was looking at the left or right object on the screen, shifting between objects, or inattentive. A second research assistant coded 18% of videos; on the frames surrounding eye movements, inter-coder reliability was 97%. Research suggests that automatic eye tracking and manual gaze coding, although potentially different in their amount of data loss, capture largely similar information (Venker et al., 2020). We did not observe a difference in data loss between the two coding methods. An average of 15.88% (SD = 9.31) of eye tracking data and 15.59% (SD = 8.16) of manually coded data was lost for each participant,  $t(23.37) = 0.09$ ,  $p = .929$ ,  $\Delta M = 0.00$ , 95% CI  $[-0.06, 0.07]$ . Additionally, previous research has combined data across these methods to create a single bilingual sample (Byers-Heinlein et al., 2021), further supporting this approach.

### **Results**

Data for each trial were analyzed between 400 and 2000 ms after the onset of the target noun. While standard approaches typically begin analysis at 367 ms after onset of the target noun (Swingley, 2012), we opted to start our analysis window slightly later in order to create consistent 100-ms time bins to use in a growth curve analysis (see below). Trials where the child was inattentive (i.e., looked at the pictures for less than

750 ms during this window) were excluded from the analyses. Children who did not successfully complete at least 2 single-language and 2 code-switched trials were also removed from the analyses. Out of 8 possible trials of each type, children retained for analysis completed an average of 6.87 single-language trials (range = 3 – 8) and 6.63 code-switched trials (range = 4 – 8). To determine if children demonstrated successful comprehension of the target words, we examined the proportion of time that they looked towards the target picture on each trial. This was calculated by dividing the looking time to the target picture by the total time spent looking at either picture. Analyses were conducted using the R statistical language (R Core Team, 2022).

First, we investigated whether children showed comprehension of the noun on each trial type. One-sample, two-sided  $t$ -tests revealed that children looked significantly above chance ( $\mu_0 = 0.5$ ) to the target picture on both single-language trials,  $t(29) = 11.42$ ,  $p < .001$ ,  $M = 0.74$ , 95% CI [0.70,0.78], and code-switched trials,  $t(29) = 12.03$ ,  $p < .001$ ,  $M = 0.78$ , 95% CI [0.73,0.82], indicating a robust ability to understand the target noun in both trial types (see Figure 2).



**Figure 2. Proportion looking to target picture by trial type for all children. The larger red dots and line represent the grand mean. Smaller gray dots and their connecting lines represent the mean values for individual participants.**

We then compared looking time during the two trial types using a paired-samples  $t$ -test. The effect of trial type was not statistically significant,  $t(29) = 1.49$ ,  $p = .148$ ,

$M_D = 0.04$ , 95% CI  $[-0.01, 0.09]$ , suggesting that children's comprehension of the noun did not differ between single-language and code-switched trials. Contrary to our prediction that children's comprehension of the target noun would be impaired by the code-switching that preceded it, this result indicated that they were potentially unaffected by the code-switched adjective.

### **Growth Curve Analysis**

The previous analyses, which are typical in this area of research, collapsed infants' data across the entire time window and averaged across trial types to yield two data points per child. However, it has long been recognized in the field that time course data can offer revealing information about children's performance (e.g., Fernald et al., 2001). Analytic techniques such as growth curve analysis (Mirman, 2017) offer an approach to quantify differences in time course, and further allow analysis of trial-level data, thus increasing statistical power. We plotted the time course of our data and then conducted an exploratory growth curve analysis, using the same time window of 400 – 2000ms. Looking-time data were binned in 100ms blocks.

Models were built iteratively. We started with a baseline model with only linear and quadratic time terms and by-participant random effects on both time terms. We then added one additional individual difference variable to the model and compared the two nested models with an analysis of variance. Only variables that significantly improved model fit ( $p < .05$ ) according to a chi-squared test were retained. Intermediary models are available in the supplementary materials. The categorical variables of trial type, testing location, and language dominance were coded using a simple contrast coding scheme. SES and vocabulary size were continuous. We estimated parameter estimate degrees of freedom and  $p$ -values using Satterthwaite's method.

To address our main research question of the effect of code-switching on children's comprehension, our first exploratory model added trial type to the baseline model described above. We then conducted additional exploratory growth curve models building from this model looking at the potential individual effects of language dominance, testing location, SES, and vocabulary size.

### **Trial Type**

In the growth curve model investigating the effect of trial type, the fixed effects of the final model included trial type, and linear and quadratic time terms. There was a statistically significant main effect of trial type, indicating that, opposite to our prediction, children were more accurate at gazing toward the target picture when hearing code-switched trials compared to single-language trials  $t(6,100.82) = -3.43$ ,  $p = .001$ ,  $\hat{\beta} = -0.03$ , 95% CI  $[-0.05, -0.01]$  (See Table 3 for full results). This result differs from that of the paired-samples  $t$ -test, which did not find a statistically significant

difference in children's looking between the two trial types.

**Table 3. Growth curve analysis including trial type.**

	Estimate	95% CI	<i>t</i>	<i>df</i>	<i>p</i>
<b>Fixed effects</b>					
Intercept	0.76	[0.72, 0.79]	43.05	29.42	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.86	29.46	.001
Time (Quadratic)	-0.27	[-0.32, -0.23]	-12.36	29.09	< .001
Trial type	-0.03	[-0.05, -0.01]	-3.43	6,100.82	.001
<b>Random effects</b>		<b>Variance</b>			
Participant	Intercept	0.008			
	Time (Linear)	0.154			
	Time (Quadratic)	0.002			

### Individual Differences

As previous studies have found some evidence of individual differences in bilingual children's ability to process code-switching (Byers-Heinlein et al., 2021; Potter et al., 2019), we next investigated how such differences may have affected children's performance on this task. Prior to conducting these individual differences analyses, we first quantified the consistency of children's performance, by estimating the reliability of the looking time to each trial type using an intraclass correlation coefficient (ICC), based on a mean-rating, consistent, 2-way random-effects model (Byers-Heinlein et al., 2022). The estimated consistency was 0.19, 95% CI = [-0.24, 0.51] for single-language trials and 0.39, 95% CI = [0.07, 0.64] for code-switched trials. The magnitude of these ICCs was higher than in many other infant studies (Byers-Heinlein et al., 2022), supporting a cautious investigation of individual differences. However, these ICCs could be considered moderate to low on an absolute scale thus reducing statistical power for detecting correlations with other measures of individual differences.

We investigated four individual difference variables: language dominance, testing location (which was also a proxy for language pair), SES, and vocabulary size. We note that the last three variables were interrelated in our dataset: children from Montreal generally came from higher SES backgrounds,  $t(12.17) = 2.35$ ,  $p = .037$ ,  $\Delta M = 3.76$ , 95% CI [0.27, 7.25], and had a larger vocabulary,  $t(26.73) = -2.16$ ,  $p = .040$ ,  $\Delta M = -37.76$ , 95% CI [-73.56, -1.95], than children from New Jersey. Given our sample size, it was not possible to statistically disentangle these factors. Thus, our approach was to create separate models for each variable to gain some insight into which factor might have the largest explanatory power. We did so by adding each variable to the previous model including trial type as a main effect and in an interaction with trial

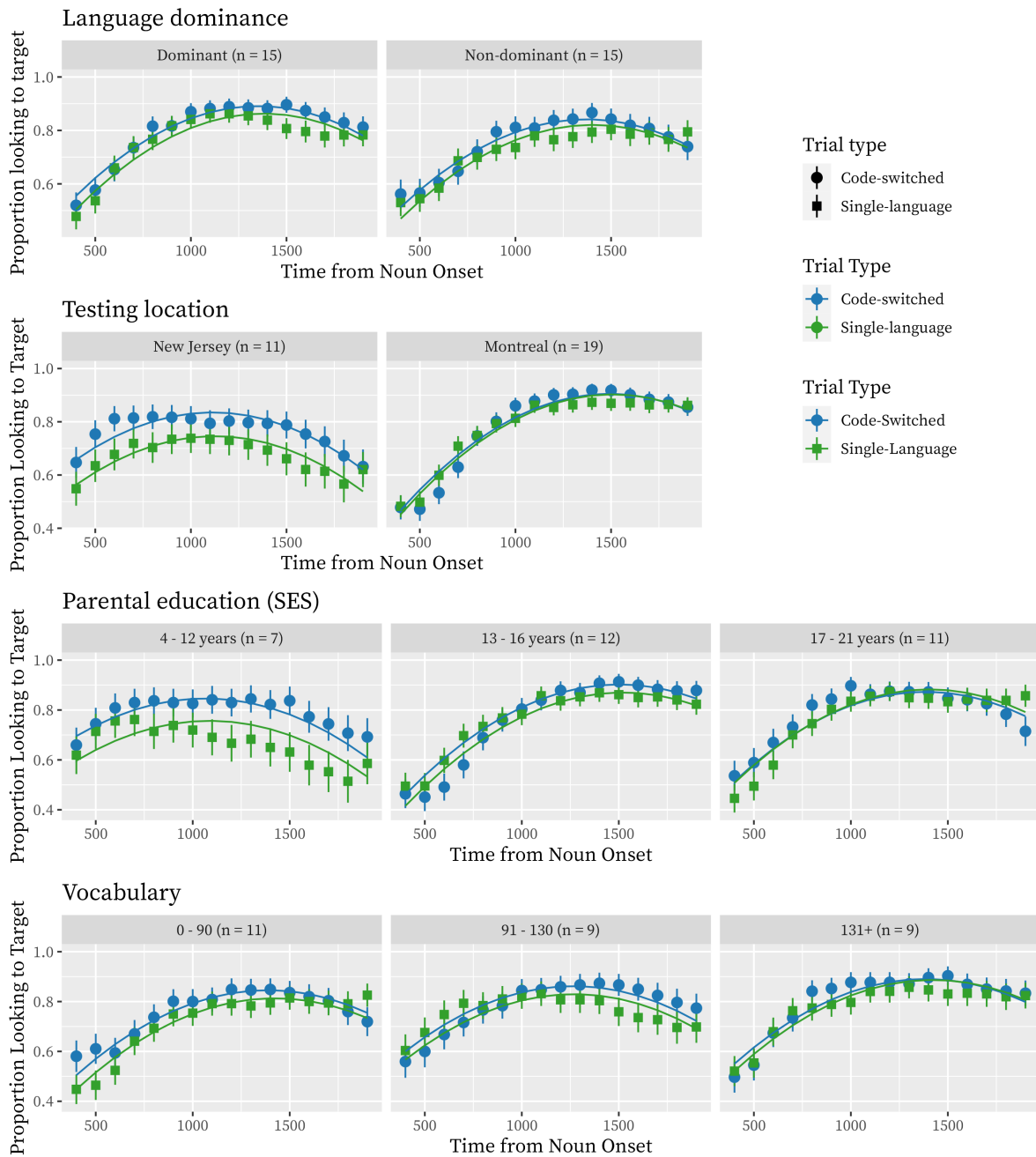


type. Here, we focus on the specific effect of these terms. Full results of these models are reported in the supplementary materials. Note that we also used the same approach to explore the potential impact of a fifth individual difference variable – children’s exposure to parental code-switching as measured by the Language Mixing Scale. This model did not explain significantly more variance than a base model without the code-switching predictor, and thus we did not interpret this model.

In each of the four models there was a statistically significant main effect of trial type, indicating that, opposite to our prediction, children were more accurate at gazing towards the target picture when hearing code-switched trials compared to single-language trials, whether controlling for language dominance,  $t(6,101.58) = -3.39$ ,  $p = .001$ ,  $\hat{\beta} = -0.03$ , 95% CI  $[-0.05, -0.01]$ , testing location,  $t(6,103.15) = -4.67$ ,  $p < .001$ ,  $\hat{\beta} = -0.05$ , 95% CI  $[-0.07, -0.03]$ , SES,  $t(6,106.01) = -4.75$ ,  $p < .001$ ,  $\hat{\beta} = -0.20$ , 95% CI  $[-0.28, -0.12]$ , or vocabulary,  $t(5,899.58) = -2.10$ ,  $p = .035$ ,  $\hat{\beta} = -0.05$ , 95% CI  $[-0.10, 0.00]$ .

We then examined the main effect of each individual difference variable and its interaction with trial type (See Figure 3), and an interesting pattern of results emerged. For language dominance, there was no statistically significant main effect,  $t(29.44) = -1.36$ ,  $p = .183$ ,  $\hat{\beta} = -0.05$ , 95% CI  $[-0.11, 0.02]$ , or interaction with trial type,  $t(6,101.58) = 0.35$ ,  $p = .727$ ,  $\hat{\beta} = 0.01$ , 95% CI  $[-0.03, 0.05]$ , suggesting that both children tested in their dominant language and children tested in their non-dominant language performed similarly across trial types. Effects of testing location, SES, and vocabulary showed similar patterns across models. Analyses of testing location revealed that children from Montreal performed similarly on both trial types, whereas children from New Jersey performed better on code-switched than single-language trials  $t(6,103.14) = -4.16$ ,  $p < .001$ ,  $\hat{\beta} = -0.09$ , 95% CI  $[-0.13, -0.05]$ .

To follow up on the Montreal results, we conducted the pupillometry analyses reported in supplementary materials, which support the main finding that children did not process code-switched and single-language trials differently (these analyses could not be carried out for New Jersey participants, as their data were hand coded from a video recording rather than collected via an eye-tracker). SES analyses showed that children from higher-SES backgrounds performed similarly across trial types whereas children from lower-SES backgrounds performed better on code-switched than single-language trials,  $t(6,103.72) = 4.04$ ,  $p < .001$ ,  $\hat{\beta} = 0.01$ , 95% CI  $[0.01, 0.02]$ . Finally, children with larger vocabularies performed better across trial types (i.e., looked more to the labeled target in general) than children with smaller vocabularies,  $t(28.38) = 2.42$ ,  $p = .022$ ,  $\hat{\beta} = 0.0007$ , 95% CI  $[0.0001, 0.0013]$ , but the effect of vocabulary size did not differ significantly as a function of trial type,  $t(5,896.30) = 0.85$ ,  $p = .396$ ,  $\hat{\beta} = 0.0002$ , 95% CI  $[-0.0002, 0.0005]$ .



**Figure 3.** Proportion looking to target picture throughout the analysis window. Dots represent means averaged over participants, bars represent  $\pm 1$  SEM, and lines represent the growth curve analysis model. SES and vocabulary were included in the model as a continuous variable but have been split into categories for the purposes of visualization. Note that one participant did not have a vocabulary score and was thus excluded from that model.

These results indicate that individual differences in performance across the two trial types were statistically related to testing location and SES, but not to language dominance or vocabulary size. Spanish-English bilingual children from New Jersey, particularly those whose parents had received a high school education or less (i.e., 12 years or fewer; see Figure 3), performed better on code-switched trials compared to single-language trials, whereas French-English bilingual children and those whose parents had more education performed similarly on the two trial types. Together, the findings show the importance of examining individual differences between participants and samples, as bilingual children's comprehension of these code-switched sentences was not uniform.

### Discussion

This study compared bilingual children's comprehension of sentences with code-switching at an uninformative determiner-adjective pair (e.g., "Can you find *le bon* [fr. the good] duck?") to their comprehension of single-language sentences (e.g., "Can you find the good duck?"). We tested 3-year-old bilingual children, including French-English bilinguals in Montreal and Spanish-English bilinguals in New Jersey. We found that bilinguals were, on average, successful at identifying the target noun in both types of sentences, and we did not see evidence that code-switching at an uninformative adjective caused any difficulties in sentence processing. Language dominance did not affect performance, likely because the target noun was always presented in a consistent language, and the switch occurred at the preceding adjective. This finding contrasts with prior reports of dominance effects in studies of children's processing of code-switches (Potter et al., 2019). Surprisingly, we found some evidence that, for certain children, code-switched sentences may have facilitated comprehension relative to single-language sentences. Our experimental design allowed us to evaluate two general accounts of why code-switching impacts speech comprehension. Under the frequency account of code-switch processing, the infrequent nature of code-switching at a determiner-adjective pair should have hindered children's comprehension, perhaps even more so than code-switching at nouns (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019). In contrast, under the functional account, children may have been able to seamlessly process code-switching at an uninformative adjective, because they did not need to integrate the meaning of the adjective to identify the target noun. While these two accounts are not mutually exclusive, our results generally support the functional account as children were able to understand the code-switch sentences as well as the single-language sentences. Below, we further discuss why young children's processing was not disrupted by code-switching at uninformative adjectives. Then, we turn to addressing the observed individual differences between participants and communities.

A key aspect of our experimental design was that the determiner-adjective pair in our sentences was uninformative. Children heard sentences with mid-sentence code-switching, as in “Can you find *le bon* [fr. the good] duck?” Critically, the adjective “*bon*” [fr. good] did not add relevant information for identifying the target object, as there was only one duck on the screen. Children typically process the meaning of adjective-noun phrases incrementally (Fernald et al., 2010; Tribushinina & Mak, 2016), but they can “listen through” the adjective to quickly identify the target object when a prenominal adjective is uninformative and does not disambiguate two objects (Thorpe & Fernald, 2006). Following the functional account, code-switching may not be disruptive when the information it carries does not need to be retrieved or integrated into processing. Children may not have experienced a code-switching cost in the current study, because they did not need to process the meaning of the code-switched adjective to identify the target and were therefore able to ignore it.

Similarly, if code-switching is related to prediction processes during language comprehension (e.g., Yacovone et al., 2021), the unexpected code-switch at the adjective might have led to a brief processing slowdown combined with a simultaneous increase in attention (Reuter et al., 2019), effectively canceling each other out in the context of an uninformative adjective. Thus, derailment in children’s processing of code-switches may be limited to functionally important words or phrases that require them to integrate the information contained in the switch. Importantly – and in contrast to our study design – adjectives often do carry functional importance in a sentence; for example, the word “heavy” can help distinguish between items of different weights, and “yummy” can refer to a food that is more delicious than another.

To further test this possibility, future studies could compare performance on trials like those in the current study and trials with an informative adjective (e.g., by showing a picture of a big and small duck and examining children’s real-time interpretation of the sentence “Do you see *le petit* [fr. the little] duck?”). Under the functional account, sentences with an informative adjective would presumably result in a code-switching cost, because children would no longer be able to “listen through” the code-switched adjective and would potentially need to engage their other language more fully.

While “listening through” could explain why we did not observe a code-switching cost, it does not explain the observed individual differences in children’s performance on code-switched and single-language sentences. Our analyses revealed that testing location and SES accounted for significant individual variation in performance across the single-language and code-switched trials, but language dominance and vocabulary size did not. Specifically, children from higher-SES backgrounds performed similarly across trial types; children from lower-SES backgrounds, particularly whose parents had a high school education or less, performed better on code-switched trials than single-language trials, and were all Spanish-English bilinguals in New Jersey.

In our sample, testing location (a proxy for language pair), SES, and vocabulary size were tightly related: French–English children from Montreal had higher vocabularies and were from higher SES backgrounds on average than Spanish–English children from New Jersey. Because of the correlational nature of this finding and the interrelatedness of these variables, it is not possible to pinpoint the factors driving the individual differences we observed, and thus this is a limitation of our study. However, previous studies have reported similar patterns of individual differences in infants from these same communities; one study suggested that Spanish–English children may have slightly weaker skills in real-time language tasks than French–English children (Byers-Heinlein et al., 2021). Following the functional account, if some children were slower to switch between processing their two languages, or if they were less aware of its meaning, it is possible that they were able to “listen through” the uninformative adjective more easily (or under a prediction-based framework, encountered little to no prediction error). However, note that under this explanation, we would have expected vocabulary size to predict performance, which it did not. Rather, SES was a predictor of performance, a variable which has previously been related to children’s language development (Fernald et al., 2013; Pace et al., 2017; Pungello et al., 2009). We tentatively suggest that experiential factors related to SES might be driving the observed community differences we observed.

There are also other potentially relevant differences between children that we were not able to directly observe that may have affected infants’ performance on our task, which again limits our conclusions. For example, different infants have different experiences with code-switching (Bail et al., 2015; Kremin et al., 2021), which could in turn impact their comprehension of code-switching. The frequency account predicts that bilinguals with frequent exposure to code-switching should experience less disruption in processing compared to bilinguals without frequent exposure to code-switching (Gosselin & Sabourin, 2021; Valdés Kroff et al., 2018). In the context of the current study, experience with code-switching may have been able to build on top of children’s ability to “listen through” the uninformative adjective, supporting aspects of both the frequency and the functional account. Indeed, preliminary evidence from a direct observation study suggests that children’s experiences hearing code-switching may be somewhat different in the two communities from which we sampled, with Spanish-English bilingual caregivers in New Jersey engaging in more code-switching than English-French bilingual caregivers in Montreal (Kosie et al., 2022). It is also possible that production of code-switching varies by SES within the the two communities we studied, although this has not yet been examined directly. Indeed, we speculate that if Spanish-English bilinguals in New Jersey, particularly those from lower-SES backgrounds, were somewhat more accustomed to hearing code-switching than the French-English bilinguals in Montreal, this could result in the observed “boost” in real-time sentence interpretation – at least in the context of sentences with mid-sentence code-switches at uninformative locations. To address this question, additional

research is needed to directly investigate the relationship between the amount and type of code-switching that bilingual children hear and how they process incoming speech input in two languages.

Finally, this work adds important qualifications to the idea that code-switching engenders processing costs in bilingual children. Our study found that code-switching of uninformative adjectives does not hinder children's comprehension of a subsequent noun and indeed it may have facilitated comprehension, at least for some children. Such facilitatory effects have been reported in the adult literature, whereby a code-switch cued participants that a low-frequency word would be heard, allowing listeners to rapidly identify a labeled target (Tomić & Valdés Kroff, 2021). Parents of young bilinguals code-switch at a variety of syntactic locations and for a variety of reasons, and our results support the idea that certain instances of code-switching do not hinder processing but may even support comprehension and learning (Kremin et al., 2021). It is not scientifically sound to tell parents that code-switching is 'good' or 'bad', and future experiments will need to carefully document young bilinguals' everyday experience with code-switching and evaluate how they process instances of typical and atypical switching.

## Conclusion

Code-switching is common in bilingual speech, making it important to understand its effect on children's language comprehension and language learning. Past research has generally found that code-switching leads to processing costs, but in the current study, bilingual children did not show this processing cost. Growth curve analyses revealed that bilinguals showed similar (and in some cases, better) processing of sentences with a code-switch at an uninformative adjective phrase, relative to single-language sentences. These findings demonstrate that linguistic features such as informativeness and location, together with individual-difference variables, may impact how bilingual children process code-switching in natural settings.

## References

- Abutalebi, J., Brambati, S. M., Annoni, J.-M., Moro, A., Cappa, S. F., & Perani, D. (2007). The Neural Cost of the Auditory Perception of Language Switches: An Event-Related Functional Magnetic Resonance Imaging Study in Bilinguals. *Journal of Neuroscience*, 27(50), 13762–13769. <https://doi.org/10.1523/JNEUROSCI.3294-07.2007>
- Bail, A., Morini, G., & Newman, R. S. (2015). Look at the gato! Code-switching in speech to toddlers\*. *Journal of Child Language*, 42(5), 1073–1101. <https://doi.org/10.1017/S0305000914000695>

- Beatty-Martínez, A. L., Valdés Kroff, J. R., & Dussias, P. E. (2018). From the field to the lab: A converging methods approach to the study of codeswitching. *Languages*, 3(2), 19. <https://doi.org/10.3390/languages3020019>
- Boudreault, M.-C., Cabirol, É.-A., Trudeau, N., Poulin-Dubois, D., & Sutton, A. (2007). Les inventaires MacArthur du développement de la communication: Validité et données normatives préliminaires. *Revue Canadienne d'orthophonie Et d'audiologie*, 31(1), 27–37.
- Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism; Cambridge*, 16(1), 32–48. <https://doi.org/10.1017/S1366728912000120>
- Byers-Heinlein, K. (2015). Methods for studying infant bilingualism. In J. W. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing* (pp. 133–154). Cambridge University Press. <https://doi.org/10.1017/CBO9781107447257.005>
- Byers-Heinlein, K., Bergmann, C., & Savalei, V. (2022). Six solutions for more reliable infant research. *Infant and Child Development*, 31(n/a), e2296. <https://doi.org/10.1002/icd.2296>
- Byers-Heinlein, K., Jardak, A., Fourakis, E., & Lew-Williams, C. (2021). Effects of language mixing on bilingual children's word learning. *Bilingualism: Language and Cognition*, 1–15. <https://doi.org/10.1017/S1366728921000699>
- 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*, 114(34), 9032–9037. <https://doi.org/10.1073/pnas.1703220114>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition*. <https://doi.org/10.1037/t15144-000>
- Dunn, L. M., Dunn, L. M., & Thériault-Whalen, C. M. (1993). *Échelle de vocabulaire en images Peabody: EVIP*. Psycan.
- Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *TVIP: Test de vocabulario en imágenes Peabody: Adaptación Hispanoamericana*. American Guidance Service.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories (CDIs)* (Second). Paul H. Brookes Publishing Company Baltimore, MD.

- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16(2), 234–248. <https://doi.org/10.1111/desc.12019>
- Fernald, A., Swingle, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, 72(4), 1003–1015. <https://doi.org/10.1111/1467-8624.00331>
- Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjectivenoun phrases. *Cognitive Psychology*, 60(3), 190–217. <https://doi.org/10.1016/j.cogpsych.2009.12.002>
- Goodz, N. S. (1989). Parental language mixing in bilingual families. *Infant Mental Health Journal*, 10(1), 25–44. [https://doi.org/10.1002/1097-0355\(198921\)10:1<25::AID-IMHJ2280100104>3.0.CO;2-R](https://doi.org/10.1002/1097-0355(198921)10:1<25::AID-IMHJ2280100104>3.0.CO;2-R)
- Gosselin, L., & Sabourin, L. (2021). Lexical-semantic processing costs are not inherent to intra-sentential code-switching: The role of switching habits. *Neuropsychologia*, 107922. <https://doi.org/10.1016/j.neuropsychologia.2021.107922>
- Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, 89, 138–161. <https://doi.org/10.1016/j.jml.2015.12.002>
- Heredia, R. R., & Altarriba, J. (2001). Bilingual Language Mixing: Why Do Bilinguals Code-Switch? *Current Directions in Psychological Science*, 10(5), 164–168. <https://doi.org/10.1111/1467-8721.00140>
- Jackson-Maldonado, D., Thal, D. J., & Fenson, L. (2003). *MacArthur Inventarios del Desarrollo de Habilidades Comunicativas: User's guide and technical manual*. Brookes Pub.
- Kaushanskaya, M., & Crespo, K. (2019). Does Exposure to Code-Switching Influence Language Performance in Bilingual Children? *Child Development*, 90(3), 708–718. <https://doi.org/10.1111/cdev.13235>
- Kosie, J. E., Fibla, L., Tsui, R. K.-Y., Martinez, T., Byers-Heinlein, K., & Lew-Williams, C. (2022). *Infants' exposure to language switching in bilingual homes across two communities*.



- 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*, 1–27. <https://doi.org/10.1017/S0305000921000118>
- Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2015). A Developmental Vocabulary Assessment for Parents (DVAP): Validating Parental Report of Vocabulary Size in 2- to 7-Year-Old Children. *Journal of Cognition and Development*, 16(3), 442–454. <https://doi.org/10.1080/15248372.2013.835312>
- MacSwan, J. (2012). Code-Switching and Grammatical Theory. *The Handbook of Bilingualism and Multilingualism*, 323.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech Language and Hearing Research*, 50(4), 940. [https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067)
- Marino, C., Bernard, C., & Gervain, J. (2020). Word Frequency Is a Cue to Lexical Category for 8-Month-Old Infants. *Current Biology*, 30(8), 1380–1386.e3. <https://doi.org/10.1016/j.cub.2020.01.070>
- Mirman, D. (2017). *Growth curve analysis and visualization using R*. Chapman and Hall/CRC.
- 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*, 1–11. <https://doi.org/10.1017/S0305000919000400>
- Ng, S., Gonzalez, C., & Wicha, N. Y. Y. (2014). The fox and the cabra: An ERP analysis of reading code switched nouns and verbs in bilingual short stories. *Brain Research*, 1557, 127–140. <https://doi.org/10.1016/j.brainres.2014.02.009>
- Nilep, C. (2006). "Code Switching" in Sociocultural Linguistics. *Colorado Research in Linguistics*, 19(1). <https://doi.org/10.25810/hnq4-jv62>
- Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, 3, 285–308. <https://doi.org/10.1146/annurev-linguistics-011516-034226>
- Potter, C. E., Fourakis, E., Morin-Lessard, 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/10.1111/desc.12794>

Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental Psychology*, 45(2), 544. <https://doi.org/10.1037/a0013917>

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*, 55(8), 1656–1665. <https://doi.org/10.1037/dev0000754>

Salig, L. K., Valdés Kroff, J. R., Slevc, L. R., & Novick, J. M. (2023). Linking frequency to bilingual switch costs during real-time sentence comprehension. *Bilingualism: Language and Cognition*, 1–16. <https://doi.org/10.1017/S1366728923000366>

Swingley, D. (2012). The Looking-While-Listening Procedure. In E. Hoff (Ed.), *Research Methods in Child Language* (pp. 29–42). Wiley-Blackwell. <https://doi.org/10.1002/9781444344035.ch3>

Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds “listen through” ambiguous adjectives in fluent speech. *Cognition*, 100(3), 389–433. <https://doi.org/10.1016/j.cognition.2005.04.009>

Tomić, A., & Valdés Kroff, J. R. (2021). Expecting the unexpected: Code-switching as a facilitatory cue in online sentence processing. *Bilingualism: Language and Cognition*, 1–12. <https://doi.org/10.1017/S1366728921000237>

Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: Evidence from eye-tracking. *Journal of Child Language; Cambridge*, 43(2), 425–441. <https://doi.org/http://dx.doi.org.lib-ezproxy.concordia.ca/10.1017/S0305000915000173>

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73(2), 89–134. [https://doi.org/10.1016/S0010-0277\(99\)00032-3](https://doi.org/10.1016/S0010-0277(99)00032-3)

Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Dussias, P. E. (2018). Experimental contributions of eye-tracking to the understanding of comprehension processes while hearing and reading code-switches. *Linguistic Approaches to Bilingualism*, 8(1), 98–133. <https://doi.org/10.1075/lab.16011.val>

van Hell, J. G., Fernandez, C. B., Kootstra, G. J., Litcofsky, K. A., & Ting, C. Y. (2018). Electrophysiological and experimental-behavioral approaches to the study of intra-sentential code-switching. *Linguistic Approaches to Bilingualism*, 8(1), 134–161. <https://doi.org/10.1075/lab.16010.van>

Vaughan-Evans, A., Parafita Couto, M. C., Boutonnet, B., Hoshino, N., Webb-Davies, P., Deuchar, M., & Thierry, G. (2020). Switchmate! An Electrophysiological Attempt to Adjudicate Between Competing Accounts of Adjective-Noun Code-Switching. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.549762>

Venker, C. E., Pomper, R., Mahr, T., Edwards, J., Saffran, J., & Weismer, S. E. (2020). Comparing Automatic Eye Tracking and Manual Gaze Coding Methods in Young Children with Autism Spectrum Disorder. *Autism Research*, 13(2), 271–283. <https://doi.org/10.1002/aur.2225>

Yacovone, A., Moya, E., & Snedeker, J. (2021). Unexpected words or unexpected languages? Two ERP effects of code-switching in naturalistic discourse. *Cognition*, 215, 104814. <https://doi.org/10.1016/j.cognition.2021.104814>

Zeller, J. P. (2020). Code-Switching Does Not Equal Code-Switching. An Event-Related Potentials Study on Switching From L2 German to L1 Russian at Prepositions and Nouns. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01387>

### **Data, code, and materials availability statement**

All stimuli, data, and analysis scripts for the current study are available via the Open Science Framework at <https://doi.org/10.17605/OSF.IO/ECQWR>.

### **Ethics statement**

Ethics approval was obtained from the Concordia University Human Research Ethics Committee (“Monolingual and Bilingual Language Development”; approval #10000493) and the Princeton University Institutional Review Board (“Language learning and Communication”; approval #7117), and parents provided informed consent prior to their child’s participation.

### **Authorship and Contributorship Statement**

The authors made the following contributions. Lena V. Kremin: Formal analysis, Writing – Original draft, Visualization; Amel Jardak: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original Draft; Casey Lew-Williams: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project

administration, Funding acquisition; Krista Byers-Heinlein: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

### **Acknowledgements**

This work was supported by a grant to KBH from the Natural Sciences and Engineering Research Council of Canada (402470-2011), grants to CLW and KBH from the U.S. National Institute of Child Health and Human Development (R01HD095912, R03HD079779), a grant to CLW from the Speech-Language-Hearing Foundation, support for KBH from the Concordia University Research Chairs program, and fellowships to LVK and AJ from the Fonds de Recherche du Québec–Société et Culture. We thank the many members of the Concordia Infant Research Lab and the Princeton Baby Lab who assisted with testing participants, and the children and parents who participated.

### **License**

*Language Development Research* (ISSN 2771-7976) is published by TalkBank and the Carnegie Mellon University Library Publishing Service. Copyright © 2023 The Author(s). This work is distributed under the terms of the Creative Commons Attribution-Non-commercial 4.0 International license (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits any use, reproduction and distribution of the work for non-commercial purposes without further permission provided the original work is attributed as specified under the terms available via the above link to the Creative Commons website.