Adaptive and selective production of syllable duration and fundamental frequency as word segmentation cues by French-English bilinguals

Annie C. Gilbert, Max Wolpert, Haruka Saito, Shanna Kousaie, Inbal Itzhak, and Shari R. Baum

Citation: The Journal of the Acoustical Society of America **146**, 4255 (2019); doi: 10.1121/1.5134781
View online: https://doi.org/10.1121/1.5134781
View Table of Contents: https://asa.scitation.org/toc/jas/146/6
Published by the Acoustical Society of America

Adaptive and selective production of syllable duration and fundamental frequency as word segmentation cues by French-English bilinguals

Annie C. Gilbert,[1,a),b)] Max Wolpert,[2,b)] Haruka Saito,[1,b)] Shanna Kousaie,[3,b)] Inbal Itzhak,[4] and Shari R. Baum[1,b)]

[1]*School of Communication Sciences and Disorders, McGill University, 2001 McGill College Avenue, Montreal, H3A 1G1, Canada*

[2]*Integrated Program in Neuroscience, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, H3A 2B4, Canada*

[3]*Montreal Neurological Institute and Hospital, McGill University, 3801 University Street, Montreal, H3A 2B4, Canada*

[4]*Centre for Research on Brain, Language, and Music, 3640 de la Montagne, Montreal, H3G 2A8, Canada*

(Received 23 June 2019; revised 30 October 2019; accepted 31 October 2019; published online 10 December 2019)

Despite the significant impact of prosody on L2 speakers' intelligibility, few studies have examined the production of prosodic cues associated with word segmentation in non-native or non-dominant languages. Here, 62 French-English bilingual adults, who varied in L1 (French or English) and language dominance, produced sentences built around syllable strings that can be produced either as one bisyllabic word or two monosyllabic words. Each participant produced both English and French utterances, providing both native productions (used as reference) and L2 productions. Acoustic analyses of the mean fundamental frequency (*F*0) and duration of both syllables of the ambiguous string revealed that speakers' relative language dominance affected the speakers' prosodic cue production over and above L1. Speakers also produced different prosodic patterns in English and French, suggesting that the production of prosodic cues associated with word-segmentation is both adaptive (modified by language experience) and selective (specific to each language). © *2019 Acoustical Society of America*. https://doi.org/10.1121/1.5134781

[MS] Pages: 4255–4272

I. INTRODUCTION

In contrast to written words that are separated by spaces, spoken words are embedded in a continuous speech stream without systematic pauses or other explicit word boundary markers. Thus, word segmentation in the acoustic domain is rife with ambiguity. For instance, when hearing a sentence like "She was looking for toucans in the jungle," listeners may initially parse the beginning of the sentence as "She was looking for two cans…," but then, upon hearing "in the jungle," realize that their initial parse is not likely, and that what they initially parsed as two different words (*two cans*) is in fact only one word (*toucans*). Fortunately, native listeners can rely on a wide range of cues to localize word boundaries. Depending on the language, these cues can involve suprasegmental features (Lehiste, 1972; Quené, 1993), the phonotactics of the heard language (McQueen, 1998; Mehler *et al.*, 1981; Skoruppa *et al.*, 2015; Suomi *et al.*, 1997), or even top-down information pertaining to speaker familiarity and lexical knowledge (Davis and Johnsrude, 2007). Nonetheless, questions remain as to how speakers cope with segmenting a non-native language, and how this would affect their production of word boundaries in their non-native language.

An extensive body of research on perceptual word segmentation has demonstrated that native speakers of different languages rely on different cues and strategies to isolate meaningful units from the speech stream. For instance, numerous studies have shown that speakers of stress-timed languages (like English or Dutch; see Pike, 1945) are sensitive to lexical stress and use it to locate word onsets in the speech stream (Cutler *et al.*, 1997; Cutler and Otake, 2002; Jusczyk, 1999; Jusczyk *et al.*, 1999; Mattys *et al.*, 1999). On the other hand, native speakers of syllable-timed languages, like French, have been shown to be sensitive to syllable structure and duration in general (Cutler *et al.*, 1986; Mehler *et al.*, 1981), relying on syllable structure to locate syllable boundaries within words (Cutler *et al.*, 1986, 1992) and on the lengthening of phrase-final syllables to locate word offsets (Christophe *et al.*, 2003; Christophe *et al.*, 2001; Christophe *et al.*, 2008; Christophe *et al.*, 2004; Cutler *et al.*, 1997; Rietveld, 1980). Given the variety of language-specific segmentation strategies, one wonders how second language speakers (L2; or learners) can manage to successfully extract words in their L2, especially if that L2 does not rely on the same segmentation strategy as their L1.

Earlier studies investigating word segmentation in a foreign or second language have suggested that listeners acquire a word segmentation strategy during the early phases of language development and continue to apply it when listening to foreign languages, regardless of its effectiveness in

[a)]Electronic mail: annie.c.gilbert@mail.mcgill.ca

[b)]Also at: Centre for Research on Brain, Language, and Music, 3640 de la Montagne, Montreal, H3G 2A8, Canada.

that language (Cutler *et al.*, 1986; Cutler and Otake, 1994; Cutler *et al.*, 2006; Otake *et al.*, 1993; Weber and Cutler, 2006). For example, monolingual French listeners continue to rely on syllable structure to locate syllable boundaries in English words, whereas monolingual English listeners would not (Cutler *et al.*, 1986). Interestingly, subsequent work has shown that, even though highly proficient simultaneous French-English bilinguals did not show signs of being able to use the appropriate segmentation strategy in their non-dominant language (e.g., French-dominant speakers did not use lexical stress when processing English words and English-dominant speakers did not use syllable structure when processing French), they were at least able to refrain from applying the word segmentation strategy of their dominant language to their non-dominant language (Cutler *et al.*, 1992). Such results led the investigators to suggest that listeners use two types of segmentation procedures: "restricted" procedures that are language specific, learned during infancy and that exploit the rhythmic characteristics of the L1, and "non-restricted" procedures that are available to all languages and rely on "universal rather than language-specific phonological characteristics" (Cutler *et al.*, 1992, p. 408).

The bilingual studies cited above relied on a word or syllable "spotting" paradigm, in which listeners are asked to locate words or syllables within longer words (e.g., locating *bal-* in *balcony*). Although these tasks can be informative concerning what listeners pay close attention to in the signal, they might not reflect how listeners extract words from spontaneous speech per se. For example, even though native English listeners are known to pay close attention to lexical stress and associate it with word onsets, Tyler and Cutler (2009) showed that English listeners also use phrase-final lengthening to locate word offsets in artificial language learning paradigms, like French listeners do (see also White *et al.*, 2015 for a study detailing the independent effects of consonant and rime lengthening in English). These results further demonstrate that listeners rely on both "restricted" and "non-restricted" segmentation strategies when processing speech, even in their L1 (Cutler *et al.*, 1992).

This ability to rely on different segmentation strategies can facilitate the transition between L1 and L2. For example, Tremblay *et al.* (2012) found that as the French-L2 proficiency of English-L1 learners increased, so did their reliance on phrase-final lengthening to locate word boundaries in French. Furthermore, Tremblay *et al.* (2017) showed that English native speakers who have reached a high proficiency level in L2-French used *F*0 as a word *offset* cue in an artificial language learning paradigm, while monolingual English speakers did not. In turn, French native speakers who have reached a high proficiency level in L2-English relied less on *F*0 as a word offset cue than did monolingual French speakers (Tremblay *et al.*, 2017). These results prompted the authors to suggest that "listeners' use of prosodic cues to word boundaries is, at least to some degree, adaptive (i.e., it is modulated by both L1 and L2 experience), and it is not selective (i.e., segmentation strategies cannot be selected as a function of how useful they are for segmenting the unfamiliar language)" (Tremblay *et al.*, 2017, p. 14). These results therefore suggest that listeners have only one speech segmentation strategy at their disposal at any given time, but that the specifics of said strategy are modulated by language experience. Taken together, these studies show that even though listeners' L1 has been found to orient their attention to different aspects of the speech stream (e.g., syllables vs stress patterns), it does not prevent them from using the other acoustic cues present in the signal. These studies also show that L2 learners are able to assign new L2-specific roles to acoustic cues, once they have enough experience with the language. However, the question remains as to L2 learners' ability to adapt their *production* of language-specific acoustic cues to signal word boundaries in their L2.

Adapting one's production of prosodic cues to a second or non-dominant language can be difficult since speakers must manage competition between their first and second languages (De Groot, 2011; Jacobs *et al.*, 2016; Kroll and Bialystok, 2013; Kroll and Gollan, 2014). This may lead to a perceived foreign accent (Magen, 1998; Mareüil and Vieru-Dimulescu, 2006; Piske *et al.*, 2001; Trofimovich and Baker, 2006) or even render speech hard to understand (Anderson-Hsieh and Koehler, 1988; Munro and Derwing, 1995; Trofimovich and Baker, 2006). Despite the importance of prosody in L2 intelligibility, few studies have investigated L2 prosody production. Among those few studies, many have demonstrated that speakers are able to adapt their prosody to the L2 or non-dominant language, at least to a certain degree, both at the sentential level (Kainada and Lengeris, 2015; O'Brien *et al.*, 2014; Rasier and Hiligsmann, 2007; Shen, 1990; Trofimovich and Baker, 2006) and at the word level (Guion *et al.*, 2004; Trofimovich and Baker, 2006). For example, Trofimovich and Baker (2006) investigated the English production of native Korean speakers who had varying language experiences to determine which factors predict more native-like English production. Their results showed that different prosodic variables (stress timing, peak alignment, speech rate, pauses) were affected by different factors relating to an individual's language experience (cumulative amount of L2 exposure or age of L2 acquisition). Namely, stress timing was more native-like among speakers with more cumulative exposure to English, whereas speech rate and pauses were more native-like among speakers who started learning English at a younger age. These results show that given the right language experience, Korean speakers were able to adapt their prosody production to their L2 (i.e., English).

Interestingly, Kim (2019) found what can be seen as a consistent pattern of results while investigating the production and perception of Spanish lexical stress among heritage speakers (i.e., speakers who learned Spanish from birth, but now use it only at home and use English in all other contexts). Comparing their production to that of monolingual Spanish speakers and of English-L1 learners of Spanish, Kim (2019) found that heritage speakers' production of Spanish lexical stress patterns with that of English-L1 learners of Spanish and not with that of monolingual Spanish speakers. On the other hand, Spanish heritage speakers' perception of Spanish lexical stress matched that of monolingual Spanish speakers and not that of English-L1 learners. These results demonstrate that even though heritage speakers

can maintain native-like stress perception, their production of the same stress patterns was influenced by their now dominant L2. Thus, linguistic experience factors have a significant impact on the development and maintenance of stable native-like prosodic productions.

Moreover, one might expect that the transition from any specific L1 to a given L2 may be more or less difficult depending on the prosodic systems of the two languages in question. For instance, transitioning from a syllable-timed language to another syllable-timed language might be easier because both languages require the production of similar prosodic cues. On the other hand, transitioning from a syllable-timed language (like French) to a stress-timed language (like English or Dutch), or vice versa, might present a greater challenge because it would involve learning to produce different prosodic cues (see Tremblay *et al.*, 2016, for alternative predictions according to the Prosodic Learning Interference Hypothesis). For example, Dupoux *et al.* (1997) found that native French listeners seem to be "stress deaf," in that they do not pay attention to the acoustic correlates marking lexical stress when determining if two words are similar or different, even if they are able to process the pertinent acoustic details (Dupoux *et al.*, 1997; Dupoux *et al.*, 2001). Such a pattern of results was later interpreted as a sign that French listeners do not encode contrastive stress, probably due to the fact that French does not use such stress patterns (Dupoux *et al.*, 2008). Therefore, when native French speakers learn English as an L2, they must not only learn to use a new prosodic cue, but they must learn to pay attention to and encode lexical stress to overcome their "stress deafness" (Dupoux *et al.*, 1997). Furthermore, the transition between specific language pairs might be more difficult in one direction than the other. For example, going from a syllable-timed L1 to a stress-timed L2 (i.e., from French-L1 to English-L2) might not be equivalent to going from a stress-timed L1 to a syllable-timed L2 (i.e., from English-L1 to French-L2), although to our knowledge, no study has investigated this issue directly.

The general aim of the present study is thus to investigate French-English bilingual speakers' (FEs) ability to adapt their prosody production to signal word boundaries in both of their languages. Taking advantage of the linguistic diversity of the Montreal area, we recruited FEs with either French or English as their L1, alongside simultaneous bilinguals having learned both languages from birth. Contrary to previous studies, we investigate prosody production in both of the languages spoken by our bilingual participants, namely French and English. Therefore, our dataset includes occurrences of both languages being produced by native speakers, which can be used as references to identify native-like patterns, and by L2 learners with different patterns of language experience. Given that the bilingual experience varies a great deal, which can affect overall L1 and L2 proficiency, we chose to compute a relative language dominance index to investigate possible language dominance effects on the production of prosodic cues irrespective of speakers' L1 (Birdsong, 2015). This dataset allows us to observe the continuous effect of language experience, as reflected in the relative dominance index, independent of L1 effects.

Specifically, the present study aims at observing FEs' use of *F*0 and syllabic duration when producing syllable strings that can represent either two monosyllabic words (*two cans*) or one bisyllabic word (*toucans*). Participants produced such sequences in both their L1 and L2, and their production was analyzed as a function of their L1 and their relative language dominance. These language experience variables are used to address both possibilities put forward in the bilingual word segmentation literature, namely that either speakers acquire a segmentation strategy during early language acquisition and apply it everywhere, regardless of the language to process, or that they can learn L2 appropriate word segmentation processes given appropriate L2 experience (based on perception data from Cutler *et al.*, 1992; and Tremblay *et al.*, 2017). Nonetheless, given the different word segmentation strategies used in French and English (see Tremblay *et al.*, 2017 for a discussion), one might expect to observe more systematic *F*0 variations in French trials than in English trials, while syllabic duration might be used in both languages.

If early L1 experience is the most important factor for word segmentation strategies, then we would expect that a speaker's L1 would predict the use of *F*0 and duration in both languages, regardless of language dominance. Therefore, French-L1 speakers would produce French-like *F*0 and duration patterns in both languages whereas English-L1 speakers would produce English-like prosodic patterns in both languages. On the other hand, simultaneous bilinguals might either be expected to produce native-like prosody in both English and French since they are native speakers of both languages, or to produce native-like prosody only in their most dominant language, replicating previous word segmentation studies (Cutler *et al.*, 1992).

However, if speakers can adapt their word segmentation strategy based on their language experience, then we would expect that language dominance would predict the prosodic cues used in both languages, irrespective of speakers' L1. That is, speakers who are dominant in English (as indexed here by a verbal fluency task) would use more English-like *F*0 and duration patterns in both English and French, while French-dominant speakers would produce more French-like *F*0 and duration patterns in both languages, irrespective of their L1 (replicating the perception results reported in Tremblay *et al.*, 2017, where listeners used the same segmentation strategy in both L1 and L2, strategy that had been modulated by their L2 experience.) In light of previous studies highlighting the role of language experience on L2 prosodic production, we expect a mixed pattern of results, with both L1 and language dominance contributing to the pattern of *F*0 and duration in L2 production.

II. METHOD

A. Participants

Sixty-two English-French bilingual speakers from the greater Montreal area took part in the experiment (18 to 36 years of age, mean = 23.7 years, 41 females). Twenty-one participants reported French as their L1, 22 participants reported English as their L1, and 19 reported having learned

both languages from birth. Information regarding language history and proficiency was collected through a questionnaire adapted from the *Language history questionnaire* (LHQ 2.0, Li *et al*., 2013), and participants did not present any perceptual, speech, or learning impairments, as established during a pre-screening phone interview conducted to determine participants' eligibility. Following completion of the testing session, two participants (one English-L1 and one French-L1) had to be removed from the sample due to missing information in the language history questionnaire. Table I summarizes proficiency and dominance measures (see explanation below) for the 60 participants included in the analyses.

Relative language dominance was estimated using verbal fluency tasks in English and French. During the verbal fluency task, participants had one minute to name as many items as possible corresponding to different semantic or orthographic categories (e.g., "animals" or "words starting with the letter P"). The task included one semantic category and three different orthographic categories per language ("animals" and the letters F, A, and S in English; "fruits" and the letters P, F, and L in French). Participants performed the task in their L1 first, followed by their L2. Simultaneous bilinguals first performed the task in the language in which they felt most comfortable and then in the other. The relative language dominance index was calculated by comparing the total number of English words produced by participants across all conditions to the total number of French words produced across all conditions (English total divided by French total.) A relative language dominance index above one indicates that participants performed better in English than in French, while a dominance index below one indicates that participants performed better in French than in English (Birdsong, 2015; Treffers-Daller and Korybski, 2015).

Of note, using a *relative* language dominance index presents many advantages. Namely, it circumvents task-related effects (i.e., some participants are better at accessing and generating words in general) by comparing each individual's performance across languages, incorporating the two scores into one continuous numerical value. Using a fixed formula where the English score is divided by the French score also means that the relative dominance index can be interpreted independently from the speaker's L1 (Birdsong, 2015).

B. Stimuli

Stimuli consisted of sentence pairs designed around two-syllable strings that could represent either one bisyllabic word or two monosyllabic words (e.g., in English, [tukænz] can be interpreted as "*toucans*" or "*two cans*"; in French, [ɔʁlɔʒ] can be interpreted as "*horloge*" - *Eng. clock* - or "*or loge*" - *Eng. gold is lodged at*). These two conditions allow us to directly compare the same sound sequence produced as a single word (no word boundary between syllables), or as two words (with a word boundary between syllables). Within a pair, sentences were identical until the word(s) representing the two syllables of interest, with the continuation of the sentence providing semantic context to fit the intended interpretation. (Eng.: "*She was looking for toucans in the jungle*" vs "*She was looking for two cans of soup to heat up*" Fr.: "*Le vendeur d'horloges vit à l'hôtel*" – *Eng. The clock salesman lives at the hotel* vs "*Le vendeur d'or loge à l'hôtel*" – *Eng. The gold salesman is lodged at the hotel*). In total, 80 sentence pairs were created (40 per language, English and French), for a total of 160 sentences.

Grammatical content was controlled to the extent possible, with French trials composed of bisyllabic nouns and monosyllabic noun-verb combinations. The majority of English trials were composed of bisyllabic nouns (two bisyllabic verbs) and monosyllabic combinations involving a noun and a word from another grammatical category (see supplementary material, Appendix A, for a complete list of the stimuli).[1] Most of the English bisyllabic words selected bore a trochaic stress pattern (37 out of 40). Given the constraints on the possible combinations of monosyllabic words, the grammatical structure of the sentences was not identical across items of a pair. The majority of monosyllabic combinations were split into two phrasal constituents while the majority of bisyllabic words occurred at the end of a phrasal constituent.

A validation task was performed by 20 monolingual native speakers (10 of each language, who did not participate in the production task). They were asked to rate each sentence on a scale from 1 to 4, where 1 corresponds to "the sentence sounds perfectly natural, it could be heard anywhere (given an appropriate context)" and 4 corresponds to "the sentence sounds quite strange, there is no context where this sentence could be appropriate." The results of the

TABLE I. Age of acquisition and language proficiency measures of participants (self-reported and objective).

	English L1				French L1				Simultaneous			
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
Age of first L2 exposure	5.6	2.8	1	15	6.5	2.9	2	11	0	0	0	0
Self-reported proficiency[a]:	5.4	0.9	3.5	7	6.6	0.5	5.5	7	6.1	1	4	7
French												
English	6.9	0.2	6.5	7	5.9	1.1	4	7	6.6	0.6	5.5	7
Verbal fluency (total):	36.3	8.9	21	53	49.4	11.8	25	70	48.6	11.2	26	69
French												
English	59.7	9.4	39	74	49.1	12.5	28	69	60.4	16	27	103
Relative language dominance index[b]	1.7	0.5	1.1	2.7	1.0	0.3	0.6	1.8	1.3	0.4	0.7	2.7

[a]Out of 7, where 1 = very poor and 7 = native-like.
[b]Total number of English words produced during verbal fluency task divided by total number of French words produced.

validation task confirmed that the majority of sentences were deemed "perfectly" or "somewhat" natural sounding (70 out of 80 in English, 71 out of 80 in French) by at least six out of ten raters. Of the remaining sentences, only two French sentences (from different sentence pairs) and one English sentence were judged "somewhat" or "quite" strange by a majority of raters (more than five out of ten raters). Other sentences received equivocal ratings with no clear positive or negative bias. Removing the trials with lower ratings or with an iambic bisyllabic word from the analysis did not influence the pattern of results observed in either L1 or L2 productions, so all stimuli were included in the final analysis.

Sentences were grouped into two blocks of stimuli per language (four blocks in total). Each block contained only one sentence from each pair and an equal number of sentences of each condition (one bisyllabic word or two monosyllabic words). Sentence order within blocks was fixed and pseudo-randomized to minimize the risk of participants noticing the existence of the two conditions. Participants were offered breaks between blocks. Trial presentation was blocked by language and participants first performed the two blocks of their L1 (block order balanced across speakers) and then the two blocks of their L2 (simultaneous bilinguals were asked to determine in which language they felt most comfortable and performed the task first in that language and then in their other native language). Language order was fixed among sequential bilinguals, taking for granted that most speakers would be dominant in their L1 in an attempt to create comparable testing conditions for sequential and simultaneous bilinguals. Unexpectedly, 9 out of the 20 recruited French-L1 speakers were later found to have performed better on the English version of the verbal fluency task than on the French version, leading to relative language dominance indices above one, suggesting English dominance. Thus, these participants had performed the task in their less dominant language first, instead of their dominant language. A direct comparison of the two groups of French-L1 speakers to determine potential task order effects was not possible given the qualitative difference between groups (one being dominant in French, the other one in English). Nonetheless, removing the English dominant French-L1 speakers from the analyses investigating L1 effects did not affect the overall pattern of results. This suggests that the English dominant French-L1 speakers who first performed the task in their less dominant language did not unduly affect the present results and were therefore kept in the sample.

The stimuli were presented visually on a computer screen using a timed PowerPoint presentation (inter-stimulus interval of 8 s). Participants were instructed to first read the sentences aloud as they appeared on screen, with natural prosody (as if talking to a friend), and then repeat the sentence from memory once the sentence disappeared (after 4 s). Sentence repetition was initially included to compare prosodic production in read and quasi-spontaneous production. Unfortunately, because the repetition task in L2 was too difficult for many speakers, we analyzed only the read sentences.

C. Recording and acoustic analysis

The responses of the majority of participants (39 out of 60) were digitally recorded in a quiet room using a shotgun microphone (Sennheiser ME66) connected to an M-Audio Delta sound card (model 1010LT) controlled by the *Goldwave* software (version 6.10). The other participants were recorded using comparable devices and set-ups in different laboratory settings (13 participants were recorded using a Marantz PMD-670 digital recorder and eight participants were recorded using an Olympus LS-11 digital recorder). All recordings were sampled at 44.1 kHz with a 16-bit resolution (since the focus of our analyses was on duration and fundamental frequency, the dependent variables should not be significantly affected by minor variations in recording equipment). We adopted a relatively conservative approach to trial inclusion. Trials with disfluencies occurring before the syllables of interest or errors directly affecting the production of the syllables of interest were removed from the present analysis. We also removed trials with pauses between the two syllables of interest to focus on the use of more subtle prosodic cues to signal word boundaries. Because pauses are reliable word boundary cues in both English and French, they would not be informative with regard to how bilinguals adapt their prosody to their L2. Finally, removal of any one trial caused the removal of both sentences from the pair, to ensure an equal representation of both conditions in the analysis (see Kim, 2019 for similar trial inclusion criteria in a production task).

Audio files were segmented and manually annotated by trained bilingual coders using Praat (version 5.4.19, Boersma, 2001). Data were extracted using a custom script developed on site and included *F*0 and duration measures for both syllables of the ambiguous region. The duration of the first syllable of the two-syllable string was measured from the onset of the initial consonant (excluding any prior hold for stop consonants) through the end of the final segment of the syllable. The duration of the second syllable of the string was measured from the offset of the first syllable to the offset of the final segment of the syllable. A mean *F*0 was extracted for the entire duration of each syllable using the Pitch function implemented in Praat (based on an autocorrelation method). Of note, in the present design, we compare the effect of conditions (one bisyllabic word vs two monosyllabic words) within the same sentence pair, so the impact of any specific speech sounds or syllable structure on the extracted data should have minimal impact on the overall pattern of results.

To get a sense of how participants control these two parameters across conditions, *F*0 and duration ratios were computed for each trial by comparing the second syllable (S2) to the first syllable (S1; S2 divided by S1 for both *F*0 and duration). Therefore, a duration ratio above one means that the second syllable of the ambiguous region was longer than the first syllable. Conversely, a duration ratio below one means that the second syllable was shorter than the first. (See Kim, 2019, for a similar use of relative prosodic measures in a production task.)

D. Statistical analysis

Results were analysed using a series of linear mixed effects (LME) models evaluating *F*0 and duration ratios in English and French separately. The models were designed to determine the effects of condition (two-level categorical variable: one bisyllabic-word vs two monosyllabic-words) and how individual differences in bilingual experience modulated any condition effects. The condition factor was deviation coded so that both levels were compared to their grand mean (instead of comparing one level to the other), which allowed the testing of main effects in the two-level factor (UCLA: Statistical Consulting Group, 2011). Two different variables related to bilingual experience were tested in these models, namely speakers' L1 (as a three-level dummy coded categorical variable; English-L1, French-L1, and simultaneous bilingual; reference level set to L1 speakers of the language being analyzed, intercept representing the mean of the reference group), thus comparing the two other speaker groups to L1 speakers of that language (UCLA: Statistical Consulting Group, 2011) and relative language dominance (as a scaled continuous variable based on the relative verbal fluency scores). Of note, given the potential collinearity between L1 and language dominance (most speakers being dominant in their L1), the two variables were tested separately (Baayen, 2008; Baayen *et al.*, 2008; Tabachnick and Fidell, 2007). Finally, random slope adjustments were included in the random structure of the models only where it was "justified by the design" (Barr *et al.*, 2013, p. 22; see also Bates *et al.*, 2015, for an argument in favor of "Parsimonious mixed models"). Thus, the random structure of the models took into account participants (intercept only) and sentence pairs produced {intercept and slope adjustments for the effect of conditions; i.e., [*F*0 ratio ~ Condition * scale(Relative language dominance index) + (1 | Participant) + (1 + Condition | Sentence Pair)]}. Random slope adjustments for condition were included to compensate for the fact that we could not completely control sentence structure (presence/absence of a prosodic boundary, word category, position of ambiguous syllable string in the sentence). Including such random slope adjustment allowed the observation of condition effects above and beyond the specifics of each sentence pair. On the other hand, the impact of participant level variables should not vary across speakers, and therefore no slope adjustments were applied to the random effect of participants.

A set of follow-up analyses was also conducted to determine if participants produced similar *F*0 and duration ratios in both languages. These models tested the effect of condition and language of the trial (as a deviation coded two-level categorical variable: French trial vs English trial) on the *F*0 and duration ratios produced by the two groups of sequential bilinguals separately (French-L1 and English-L1 speakers). The random structure of the models took into account participants (intercept and slope adjustments for the language of the trial) and sentence pairs produced (intercept and slope adjustments for the effect of conditions; i.e., {[F0 ratio ~ Condition * Language of trial + (1 + Language of trial | Participant) + (1 + Condition | Sentence Pair)]}. Random slope adjustments for trial language were included to compensate for variations in relative language dominance within L1 speaker groups, which might modulate the effect of the language of the trial. For example, a native speaker of French who is now dominant in English is not going to be affected by the language of the trial in the same way as a native speaker of French who remained dominant in French.

LME models were implemented in RStudio version 3.2.4 (R Development Core Team, 2010), using the lme4 library, version 1.1–7 (Bates *et al.*, 2014) and estimates of *p*-values were obtained using the lmerTest package version 2.0–29 (Kuznetsova *et al.*, 2015). Plots were generated using ggplot2 (version 2.1.0, Wickham, 2009) and Excel as implemented in Office 365.

III. RESULTS

A. Production of English trials

Out of the 2400 sentence pairs produced across participants, 428 sentence pairs were excluded from the analysis because at least one sentence contained an error or a disfluency occurring before the syllables of interest. Sentence pairs where at least one sentence was produced with a pause between the two syllables of interest were also removed from the analysis (387 pairs). Therefore, a total of 1585 sentence pairs were included in the analyses. Due to the variability in language proficiency and the complexity of some sentences, as well as our conservative trial inclusion criteria, there were, on average, 26.4 usable sentence pairs (out of 40) per participant, in which both sentences were produced without hesitations, pauses or other errors [standard deviation (SD) = 6.2, Min = 11, Max = 39]. Of note, although the proportion of removed trials seems high, similar proportions are found in the literature using similar designs (Kim, 2019). Sections III A 1–III A 3 present the analysis of *F*0 and duration ratios separately, and how they are modulated by speakers' L1 and relative language dominance. Only the main statistical results are reported in the text, while complete model outputs are available in the supplementary materials (see supplementary material Appendix B).[1] Of note, the main analyses presented hereafter were also performed on a subset of sentence pairs that were matched in terms of syntactic boundary location and the exact same pattern of interactions was observed. We therefore report the results of the original models including all sentence pairs.

1. Use of fundamental frequency (F0) in English trials

Figure 1 represents the average *F*0 ratios produced during English trials as a function of condition and speakers' L1 (English, French or both). Two main observations can be made based on Fig. 1. First, *F*0 ratios produced in the two-word condition (*two cans*) were in general lower than in the one-word condition (*toucans*) for two out of three speaker groups (French-L1 and simultaneous bilinguals), suggesting an effect of condition in native speakers of French. Second, speakers' L1 seems to have had an impact on the *F*0 ratios produced, with French-L1 speakers producing the largest difference between conditions, followed by simultaneous

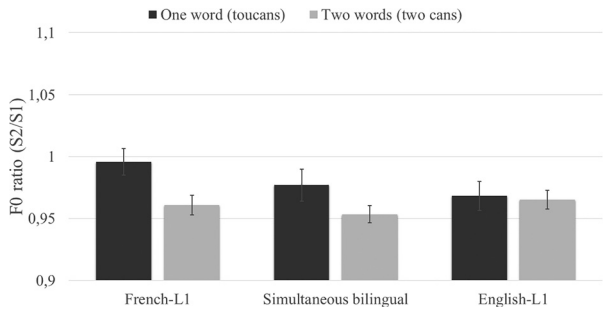

FIG. 1. Average *F*0 ratios produced during English trials as a function of conditions and speakers' L1, error bars represent standard error of the mean (*SE*).

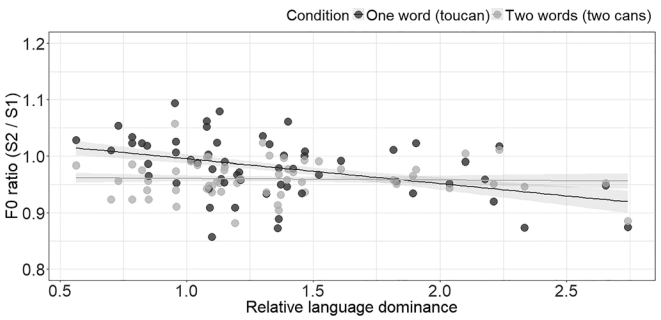

FIG. 2. Average *F*0 ratios produced during English trials as a function of conditions and speakers' relative language dominance index. Shaded areas around linear regressions represent the standard error.

bilinguals, and with English native speakers producing similar *F*0 ratios in both conditions.

To test the significance of this observed data pattern, we used an LME model including condition and speakers' L1 as predictors of *F*0 ratios. In the present analysis, English-L1 served as the reference category; thus, the model tests the significance of the difference between English-L1 and other L1 categories (i.e., difference between English-L1 and French-L1, difference between English-L1 and Simultaneous bilinguals). None of the main effects reached significance ($b < -0.007$, $SE > 0.007$, $t < 0.63$, $p > 0.5$), but, as expected from visual inspection of Fig. 1, the model revealed a significant interaction between condition and L1 for French-L1 speakers ($b = -0.025$, $SE = 0.008$, $t = -2.93$, $p < 0.005$). This finding indicates that the effect of condition differed across French-L1 and English-L1 speakers. The interaction between condition and L1 for simultaneous bilinguals did not reach significance ($b = -0.011$, $SE = 0.008$, $t = -1.31$, $p > 0.1$), indicating that simultaneous bilinguals produced native-like *F*0 ratios in English (i.e., not significantly different from English-L1 speakers, see Table II).

Further investigations focused on the impact of relative language dominance, rather than speakers' L1, on the production of *F*0 ratios in English. Figure 2 represents the average *F*0 ratios produced during English trials as a function of condition and speakers' relative language dominance index. As may be seen in Fig. 2, speakers on the French-dominant end of the spectrum tended to produce larger differences in *F*0 ratios between the one-word (*toucans*) and the two-word (*two cans*) conditions than speakers on the English-dominant end of the spectrum. The significance of the observed data pattern was confirmed using an LME model including the relative language dominance index and condition as predictors. As expected, the model revealed a significant interaction between the relative language dominance index and condition in the expected direction ($b = 0.012$, $SE = 0.004$, $t = 3.29$, $p = 0.001$) (see Table III).

Furthermore, given the similarity of the effects generated by the L1 and relative language dominance analyses of *F*0 productions, follow-up LME models were conducted for the three L1 groups separately, to determine if language dominance affected *F*0 production beyond L1 effects. These

TABLE II. Effect of condition (one-word vs two-words) and native language (English-L1, simultaneous bilingual, and French-L1) on fundamental frequency (*F*0) and duration ratios produced in English trials.

	***F*0 ratios**				**Duration ratios**			
Fixed Effects	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Intercept	0.9739	0.0107	90.961	*<0.00001*	0.9380	0.0769	12.190	*<0.00001*
Condition	−0.0044	0.0070	−0.622	0.5352	−0.1271	0.0480	−2.647	*0.0112*
L1[a]								
Simultaneous bilinguals	−0.0073	0.0118	−0.622	0.5362	0.0097	0.0246	0.396	0.6934
French-L1	0.0039	0.0117	0.331	0.7422	0.0331	0.0246	1.349	0.1831
Condition * L1								
Condition * Simultaneous bilinguals	−0.0110	0.0084	−1.305	0.1920	−0.0016	0.0249	−0.064	0.9487
Condition * French-L1	−0.0253	0.0086	−2.931	*0.0034*	0.0342	0.0256	1.336	0.1816

	Variance		Variance	
Random Effects	Intercept	Slope[b]	Intercept	Slope[b]
Participants	0.0012	—	0.0044	—
Items (sentence pairs)	0.0019	0.0006	0.2248	0.0792
Residual	0.0096		0.0850	

[a]English-L1 used as reference level, thus, the model tests the significance of the difference between English-L1 and other L1 categories (i.e., difference between English-L1 and French-L1, difference between English-L1 and Simultaneous bilinguals).
[b]Random Slope adjustments were done on Condition across items (sentence pairs).

TABLE III. Effect of condition (one-word vs two-words) and relative language dominance on fundamental frequency (*F*0) and duration ratios produced in English trials.

	***F*0 ratios**				**Duration ratios**			
Fixed Effects	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Intercept	0.9724	0.0084	115.325	*<0.00001*	0.9512	0.0757	12.560	*<0.00001*
Condition	−0.0155	0.0053	−2.921	*0.0059*	−0.1175	0.0461	−2.548	*0.0150*
Relative language dominance	−0.0082	0.0046	−1.774	0.0813	−0.0135	0.0099	−1.361	0.1790
Condition * Relative language dominance	0.0116	0.0035	3.288	*0.0010*	−0.0089	0.0104	−0.851	0.3950

	Variance		Variance	
Random Effects	Intercept	Slope[a]	Intercept	Slope[a]
Participants	0.0011	—	0.0043	—
Items (sentence pairs)	0.0019	0.0006	0.2248	0.0793
Residual	0.0096		0.0850	

[a]Random Slope adjustments were done on Condition across items (sentence pairs).

models included condition and the relative language dominance index as predictors and revealed three different patterns of results. First, the model testing production by English-L1 speakers yielded no main effects and no interaction between factors ($b < -0.015$, $SE > 0.007$, $t < -1.84$, $p > 0.082$), showing that language dominance did not affect native productions of English-L1 speakers. Second, the model testing production by French-L1 speakers yielded only a main effect of condition ($b = -0.031$, $SE = 0.007$, $t = -4.70$, $p < 0.0001$), indicating that French-L1 speakers used different *F*0 ratios to differentiate the conditions, but the model revealed no main effect of language dominance or interaction between factors ($b < -0.010$, $SE > 0.006$, $t < 1.44$, $p > 0.1$). Finally, the model examining *F*0 production by simultaneous bilinguals yielded a significant interaction between factors ($b = 0.012$, SE 0.005, $t = 2.257$, $p = 0.024$), with speakers at the French dominant end of the continuum producing different *F*0 ratios than those of speakers at the English dominant end of the continuum, regardless of their shared L1s.

To determine the role of each condition in the interactions affecting *F*0 ratio productions, a separate set of LME models was used to test the impact of L1 and relative language dominance on the two conditions separately. The models revealed no main effects of L1 in either the one-word or the two-word condition ($b < -0.014$, $SE > 0.010$, $t < -1.43$, $p > 0.16$), but revealed a significant main effect of relative language dominance in the one bisyllabic word condition [*toucan* ($b = 0.015$, $SE = 0.006$, $t = -2.40$, $p = 0.020$)], although not in the two monosyllabic words condition [*two cans* ($b = -0.001$, $SE = 0.004$, $t = -0.35$, $p = 0.727$)]. Thus, the interaction in the L1 analyses is not specifically driven by one condition, while the interaction between relative language dominance and condition is likely driven by the language dominance effect in the one-word condition.

2. Use of duration in English trials

Figure 3 represents the average duration ratios produced during English trials as a function of condition and speakers' L1 (English, French or both). Again, two main observations can be made from the average ratios. First, like the *F*0 ratios previously reported, the duration ratios produced in the two-word condition (*two cans*) were lower than in the one-word condition (*toucans*), which suggests an effect of condition. Second, speakers' L1 also seems to have an impact on the duration ratios produced, but here French-L1 speakers seem to produce the smallest difference between conditions, followed by simultaneous bilinguals, and with English native speakers producing the largest difference between conditions.

To determine if speakers' L1 (English, French, or both) significantly modulates the duration ratios produced during English trials, we used an LME model including condition and L1 as predictors, with English-L1 as the reference category. As expected from the grand averages presented in Fig. 3, the model revealed a significant main effect of condition ($b = -0.127$, $SE = 0.048$, $t = -2.65$, $p = 0.011$), with all speaker groups producing different duration ratios in the two conditions (see Table II). However, contrary to expectations, the model revealed no main effect of L1 ($b < 0.033$, $SE > 0.024$, $t < 1.35$, $p > 0.1$) and no interaction between condition and L1 ($b < 0.034$, $SE > 0.025$, $t < 1.34$, $p > 0.1$).

To further investigate the influence of individual differences on prosody production, we investigated the impact of relative language dominance on duration ratios (see Fig. 4). An LME model evaluating the impact of relative language dominance and condition on the use of duration ratios in English trials revealed a pattern of results similar to the one

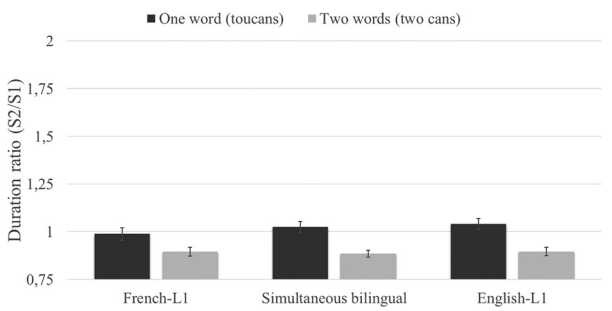

FIG. 3. Average duration ratios produced during English trials as a function of conditions and speakers' L1 (error bars represent standard error of the mean).

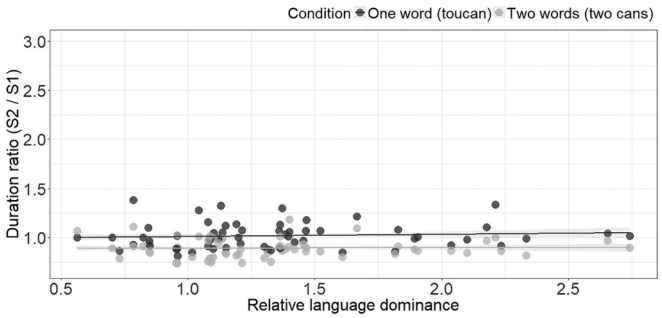

FIG. 4. Average duration ratios produced during English trials as a function of conditions and speakers' relative language dominance index. Shaded areas around linear regressions represent the standard error.

observed for L1, with a main effect of condition ($b = -0.118$, $SE = 0.046$, $t = -2.55$, $p = 0.015$), but no significant main effect of relative language dominance ($b = 0.013$, $SE = 0.010$, $t = -1.36$, $p = 0.179$) and no interaction between factors ($b = -0.009$, $SE = 0.010$, $t = -0.85$, $p = 0.395$) (see Table III). Therefore, bilingual speakers consistently produced different duration ratios across conditions when producing English trials, regardless of their L1 or relative language dominance.

3. Summary of results—English trials

In sum, both speakers' L1 and relative language dominance interacted with condition in predicting the *F*0 ratios produced during English trials, with French-L1/dominant speakers producing greater *F*0 ratios differences between conditions than English-L1/dominant speakers. Although the interaction between condition and L1 did not seem to be driven by one specific condition, the interaction between condition and relative language dominance was found to be driven by the one-word condition, as only this condition was found to be affected by speakers' relative language dominance when analyzing the two conditions separately. The effect of relative language dominance on the production of *F*0 ratios was also observed within the simultaneous bilingual group, with speakers at the French dominant end of the continuum producing more French-like *F*0 ratios than speakers at the English dominant end of the continuum despite the fact that they are native speakers of both languages. A different pattern of results was observed for duration ratios; that is, neither L1 nor relative language dominance significantly affected the production of duration ratios in English, with speakers consistently producing different duration ratios across conditions.

B. Production of French trials

Out of the 2400 sentence pairs produced across participants, 460 sentence pairs were excluded from the analysis because at least one of the sentences contained an error or a disfluency and 826 were removed because of a pause between the two syllables of interest. Thus, a total of 1114 sentence pairs were included in the analyses. Due to the complexity of some sentences and the overall lower French proficiency of our participants (compared to English), participants produced, on average, fewer French sentences compared to English sentences. Specifically, an average of 18.6 sentence pairs (out of 40), in which both sentences were produced without hesitations, pauses or other errors ($SD = 6.9$, $Min = 3$, $Max = 30$) were produced in French. Sections III B 1–III B 3 present the analysis of *F*0 and duration ratios separately, and how they are modulated by speakers' L1 and relative language dominance. Only the main statistical results are reported in the text, while complete model outputs are available in the supplementary materials (see supplementary material Appendix C).[1]

1. Use of fundamental frequency (F0) in French

As with English trials, our first analysis focused on the potential impact of speakers' L1 on the *F*0 ratios produced in the two conditions (see Fig. 5). As may be seen in Fig. 5, all three speaker groups (French-L1, English-L1, and simultaneous bilinguals) produced different *F*0 ratios across conditions, suggesting an effect of condition, with *F*0 ratios above 1 in the one-word condition (*F*0 higher on the second syllable of *horloge*) and below 1, on average, in the two-word condition (*F*0 lower on the second syllable of *or loge*). Speakers' L1 also seems to have an impact on the magnitude of the difference between conditions, with French-L1 speakers producing the largest difference between conditions, English-L1 speakers producing the smallest difference between conditions, and simultaneous bilinguals presenting an intermediate pattern.

To test the significance of the observed differences, we used an LME model including condition and speakers' L1 as predictors. In the present analysis, French-L1 served as the reference category (native speakers), thus the model tests the significance of the difference between French-L1 and other L1 categories (i.e., difference between French-L1 and English-L1, difference between French-L1 and simultaneous bilinguals). As expected from Fig. 5, the model revealed a significant interaction between condition and L1 ($b > 0.044$, $SE < 0.009$, $t > 4.94$, $p < 0.0001$), indicating that the difference between conditions was modulated by speakers' L1, with simultaneous bilinguals and English-L1 speakers producing different *F*0 ratios compared to French-L1 speakers (see Table IV). The model also yielded a significant main effect of condition ($b = -0.124$, $SE = 0.011$, $t = -11.70$, $p < 0.0001$), as well as a significant main effect of L1 for English native speakers [using French-L1 speakers as a

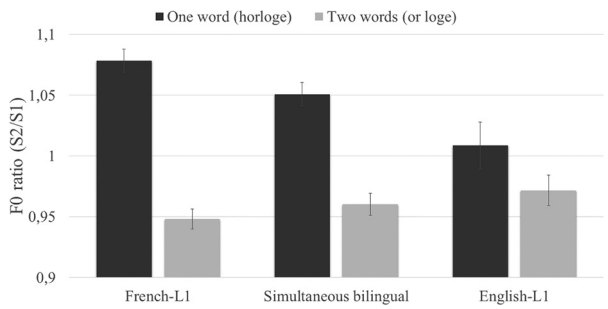

FIG. 5. Average *F*0 ratios produced during French trials as a function of conditions and speakers' L1 (error bars represent standard error of the mean).

TABLE IV. Effect of condition (one-word vs two-words) and native language (English-L1, simultaneous bilingual, and French-L1) on fundamental frequency (*F*0) and duration ratios produced in French trials.

	F0 ratios				Duration ratios			
Fixed Effects	*B*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Intercept	1.0160	0.0101	100.760	*<0.00001*	1.2637	0.0863	14.649	*<0.00001*
Condition	−0.1235	0.0106	−11.697	*<0.00001*	−0.4876	0.0662	−7.361	*<0.00001*
L1[a]								
Simultaneous bilinguals	−0.0123	0.0099	−1.247	0.2189	0.0255	0.0351	0.727	0.4709
English-L1	−0.0239	0.0099	−2.421	*0.0195*	0.1323	0.0353	3.748	*0.0005*
Condition * L1								
Condition * Simultaneous bilinguals	0.0443	0.0090	4.935	*<0.00001*	−0.1678	0.0392	−4.281	*<0.00001*
Condition * English-L1	0.0745	0.0092	8.123	*<0.00001*	−0.1850	0.0400	−4.622	*<0.00001*

	Variance		Variance	
Random Effects	Intercept	Slope[b]	Intercept	Slope[b]
Participants	0.0007	—	0.0080	—
Items (sentence pairs)	0.0018	0.0024	0.2313	0.1220
Residual	0.0077		0.1475	

[a]French-L1 used as reference level, thus, the model tests the significance of the difference between French-L1 and other L1 categories (i.e., difference between French-L1 and English-L1, difference between French-L1 and Simultaneous bilinguals).
[b]Random Slope adjustments were done on Condition across items (sentence pairs).

reference ($b = -0.024$, $SE = 0.010$, $t = -2.421$, $p = 0.02$)]; simultaneous bilinguals did not differ significantly from French-L1 speakers ($b = -0.012$, $SE = 0.010$, $t = -1.25$, $p > 0.2$]).

To determine if simultaneous bilinguals differ significantly from English-L1 speakers in their use of *F*0 ratios in French, we refitted the same LME model using English-L1 speakers as the reference group instead of native French speakers. This model yielded significant interactions between condition and L1 for French-L1 speakers (replicating the previous model using French-L1 as reference) and for simultaneous bilinguals ($b = -0.030$, $SE = 0.010$, $t = -3.13$, $p = 0.002$), indicating that simultaneous bilingual speakers produce different *F*0 ratios in French trials compared to English-L1 speakers. Thus, although all speaker groups produced different *F*0 patterns to differentiate the two production conditions in French, the magnitude of the *F*0 difference between conditions varied as a function of speakers' experience with French, with speakers having the least experience with French producing the smallest *F*0 differences between conditions, and speakers having long term experience with both languages producing larger *F*0 differences than English-L1 speakers, but smaller differences than French-L1 speakers.

Further investigations focussing on the impact of relative language dominance, rather than L1, on the production of *F*0 ratios in French trials also revealed a significant interaction between condition and relative language dominance ($b = 0.030$, $SE = 0.004$, $t = 7.90$, $p < 0.0001$) in keeping with the L1 effect (see Table V). That is, participants at the French-dominant end of the relative language dominance spectrum produced the largest difference between conditions, while participants at the English-dominant end of the spectrum did not produce significantly different *F*0 patterns between conditions (see Fig. 6).

TABLE V. Effect of condition (one-word vs two-words) and relative language dominance on fundamental frequency (*F*0) and duration ratios produced in French trials.

	***F*0 ratios**				**Duration ratios**			
Fixed Effects	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Intercept	1.0050	0.0085	118.875	*<0.00001*	1.3103	0.0841	15.572	*<0.00001*
Condition	−0.0875	0.0095	−9.251	*<0.00001*	−0.5947	0.0630	−9.439	*<0.00001*
Relative language dominance	−0.0081	0.0037	−2.180	*0.0339*	0.0483	0.0138	3.492	*0.0009*
Condition * Relative language dominance	0.0299	0.0038	7.896	*<0.00001*	−0.0742	0.0165	−4.503	*<0.00001*

	Variance		Variance	
Random Effects	Intercept	Slope[a]	Intercept	Slope[a]
Participants	0.0007	—	0.0084	—
Items (sentence pairs)	0.0018	0.0024	0.2313	0.1217
Residual	0.0078		0.1480	

[a]Random Slope adjustments were done on Condition across items (sentence pairs).

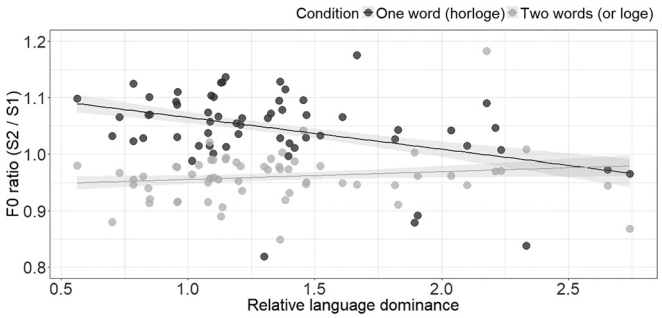

FIG. 6. Average *F*0 ratios produced during French trials as a function of conditions and speakers' relative language dominance index. Shaded areas around linear regressions represent the standard error.

Again, given the similarity of the effects generated by the L1 and relative language dominance analyses, follow-up LME models were conducted for the three L1 groups separately to determine if language dominance affected prosody production above and beyond L1 effects. The new models included condition and the relative language dominance index as predictors and revealed three different patterns of results. First, the models on *F*0 ratios produced by French-L1 speakers yielded only a main effect of condition ($b = -0.122$, $SE = 0.012$, $t = -10.24$, $p < 0.0001$), [no effect of language dominance ($b < -0.001$, $SE = 0.005$, $t = -0.14$, $p = 0.888$) and no interactions between factors ($b = -0.001$, $SE = 0.005$, $t = -0.227$, $p = 0.821$)], suggesting that language dominance did not affect native productions of French-L1 speakers. Second, the model on *F*0 production by simultaneous bilinguals replicated the model examining French-L1 speakers with only a main effect of condition ($b = -0.080$, $SE = 0.012$, $t = -6.64$, $p < 0.0001$). Finally, the model on *F*0 ratios produced by English-L1 speakers revealed a significant interaction between condition and relative language dominance ($b = 0.033$, $SE = 0.008$, $t = 4.354$, $p < 0.0001$), where more balanced speakers produced more native-like prosody in French, while speakers at the English dominant end of the spectrum did not produce different *F*0 patterns to differentiate the conditions. Of note, this pattern of *F*0 ratios differs slightly from the one observed in the previous analysis focusing on speakers' L1, in that not all participants used different *F*0 ratios to differentiate the two speaking conditions (compare Fig. 5 and Fig. 6). This discrepancy might be due to the fact that some English-L1 speakers are very heavily dominant in English and do not behave in a similar fashion to English-L1 speakers who are more balanced bilinguals, a difference that is masked when averaging them all together as a function of their shared L1.

To determine the role of each condition in the interactions affecting the production of *F*0 in French, a separate set of LME models was used to test the impact of L1 and relative language dominance on the two conditions separately. This analysis revealed significant effects of L1 (English) and relative language dominance consistent with the interactions observed in the original analyses in both models on the one-word condition ($b > -0.024$, $SE < 0.019$, $t > -3.39$, $p < 0.0013$). Conversely, the models on the two-word condition did not reach significance ($b < 0.019$, $SE > 0.004$, $t < 1.82$, $p > 0.073$). Thus, the present interactions between language experience factors and condition seem to be driven by the one-word condition.

2. Use of duration in French

Figure 7 represents the average duration ratios produced during French trials as a function of condition and speakers' L1 (English, French, or both). Figure 7 shows that all three speaker groups produced different duration ratios between conditions, with duration ratios above 1 in the one-word condition (second syllable longer than first syllable in *horloge*) and below 1 in the two-word condition (first syllable longer than second in *or loge*). The L1 also affected the magnitude of the difference between conditions, with English-L1 speakers producing the largest difference between conditions, French-L1 speakers producing the smallest difference between conditions, and simultaneous bilinguals presenting an intermediate pattern.

To test the significance of these observations, we analysed the French duration ratios using an LME model including speakers' L1 and condition as predictors. As expected from the grand averages presented in Fig. 7, the model revealed a significant interaction between condition and L1 ($b > 0.168$, $SE < 0.040$, $t > -4.28$, $p < 0.001$) indicating that the difference between conditions was modulated by speakers' L1, with English-L1 and simultaneous bilingual speakers producing significantly larger duration differences between conditions than French-L1 speakers (see Table IV). The model also yielded a significant main effect of condition ($b = -0.488$, $SE = 0.066$, $t = -7.36$, $p < 0.0001$), as well as a significant main effect of L1 for English native speakers [using French-L1 speakers as a reference ($b = 0.132$, $SE = 0.035$, $t = 3.75$, $p < 0.001$), while simultaneous bilinguals did not differ significantly from French-L1 speakers ($b = 0.026$, $SE = 0.035$, $t = 0.73$, $p = 0.471$)].

To determine if simultaneous bilinguals differ from English-L1 speakers in their use of duration ratios in French, we refitted the previous LME model using English-L1 speakers as the reference group instead of native French speakers. The model replicated the significant interactions between condition and L1 for French-L1 speakers, but yielded no significant interaction between condition and L1 for simultaneous bilinguals ($b = 0.017$, $SE = 0.042$, $t = 0.41$, $p = 0.682$), indicating that the difference in simultaneous bilingual speakers' use of duration between the two conditions in

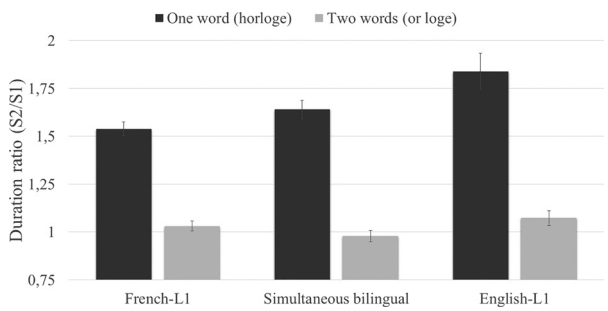

FIG. 7. Average duration ratios produced during French trials as a function of conditions and speakers' L1 (error bars represent standard error of the mean).

French trials did not differ from that of English-L1 speakers. Therefore, all participants produced different duration ratios between conditions in French, regardless of their L1, but speakers having more experience with English (English-L1 speakers and simultaneous bilinguals) produced greater duration ratio differences between conditions than French-L1 speakers.

Furthermore, the investigation of the effect of relative language dominance on the use of duration ratios in French again revealed a pattern of results similar to the one observed for speakers' L1. Namely, we observed a significant interaction between relative language dominance and condition ($b = -0.074$, $SE = 0.016$, $t = -4.50$, $p < 0.0001$) where speakers at the English-dominant end of the spectrum produced a larger difference between conditions than speakers at the French-dominant end of the spectrum (see Fig. 8 and Table V).

As above, follow-up LME models were conducted for the three L1 groups separately. These models revealed patterns of results consistent with the equivalent analyses performed on *F*0 ratios during French trials. Namely, the models on French-L1 speakers and simultaneous bilinguals yielded only a main effect of condition ($b > -0.49$, $SE < 0.073$, $t > -7.41$, $p < 0.0001$), showing that language dominance did not affect the production of native speakers of French, while the model on English-L1 speakers revealed a marginally significant interaction between factors ($b = -0.069$, $SE = 0.036$, $t = -1.92$, $p = 0.056$), where speakers at the English dominant end of the spectrum produced larger duration differences than speakers from the French dominant end of the spectrum.

To determine the role of each condition in the interactions affecting the production of duration ratios in French, follow-up LME models were used to test the impact of L1 and relative language dominance on the two conditions separately. These models revealed significant effects of L1 (English) and relative language dominance consistent with the interactions observed in the original duration ratio analyses in all models on the one-word condition ($b > 0.092$, $SE < 0.062$, $t > 3.86$, $p < 0.001$). On the other hand, none of the models on the two-word condition reached significance ($b < 0.027$, $SE > 0.014$, $t < 0.79$, $p > 0.4$). Thus, as was observed in the *F*0 analyses, the interactions between language experience factors and condition seem to be driven by the one-word condition in French, as observed for the English trials.

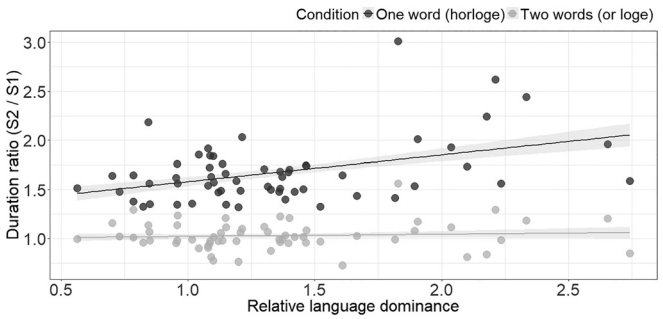

FIG. 8. Average duration ratios produced during French trials as a function of conditions and speakers' relative language dominance index. Shaded areas around linear regressions represent the standard error.

3. Summary of results—French trials

To summarize, both *F*0 and duration were affected by interactions between condition and either speakers' L1 or relative language dominance in French trials, albeit not in the same direction. That is, French-L1/dominant speakers produced larger *F*0 ratio differences between conditions than English-L1/dominant speakers, who instead produced larger duration ratio differences. Both interactions were found to be driven by the one-word condition in subsequent analyses conducted on the two conditions separately. An effect of relative language dominance was also observed within the English-L1 group, with more balanced speakers producing more native-like prosody (*F*0 and duration ratios) in French than speakers at the English-dominant end of the spectrum.

C. Comparing the production of English and French trials

Having observed L1 and language dominance effects on the *F*0 and duration ratios produced in both languages, one might ask whether non-native speakers attempted to adapt their prosody to their L2, albeit imperfectly, or if they simply produced the same prosodic cues in both languages. To determine if speakers produced different *F*0 and duration ratios in their L1 and L2, follow-up LME models were used to test the effects of condition (one bisyllabic word or two monosyllabic words) and language of the trial (English or French) in the two groups of sequential bilinguals separately (English-L1 and French-L1). All models revealed significant interactions between condition and language of the trial, with French trials being produced with greater *F*0 and duration ratio differences between conditions than English trials in both speaker groups [*F*0: ($b > -0.046$, $SE < 0.013$, $t > -3.61$, $p < 0.001$), duration: ($b > -0.409$, $SE < 0.084$, $t > -5.49$, $p < 0.0001$)]. These results indicate that speakers did adapt their prosodic production to their L2, even if their production was not quite native-like.

To determine the role of each condition in these interactions, a separate set of LME models was used to test the impact of language of the trial on the two conditions separately. This analysis revealed that, similar to the interactions involving the condition factor previously observed, these interactions seem to be driven by the one-word condition, which varies the most between languages among both French-L1 speakers [*F*0: ($b = 0.081$, $SE = 0.021$, $t = 3.96$, $p < 0.001$), duration: ($b = 0.497$, $SE = 0.132$, $t = 3.76$, $p < 0.001$)] and English-L1 speakers [F0: ($b = 0.037$, $SE = 0.023$, $t = 1.57$, $p = 0.127$), duration: ($b = 0.785$, $SE = 0.150$, $t = 5.25$, $p < 0.001$)]. On the other hand, the prosodic patterns associated with the two-word condition remained quite stable across languages in both the French-L1 speakers [F0: ($b = -0.010$, $SE = 0.014$, $t = -0.70$, $p = 0.486$), duration: ($b = 0.091$, $SE = 0.110$, $t = 0.83$, $p = 0.408$)] and the English-L1 speakers [*F*0: ($b = 0.0001$, $SE = 0.015$, $t = 0.003$, $p = 0.997$), duration: ($b = 0.168$, $SE = 0.101$, $t = 1.67$, $p = 0.099$)]. See supplementary material Appendix D for complete model outputs.[1]

IV. GENERAL DISCUSSION

The goal of the present paper was to analyze the use of *F*0 and duration ratios as word boundary cues by French-English bilinguals with different language experiences. As expected, our results demonstrate that the speakers' use of prosodic cues is influenced not only by their L1, but also by their relative language dominance. The results therefore support a hybrid model of the production of L2 prosody, incorporating both the restricted segmentation strategies developed during L1 acquisition (Cutler *et al.*, 1992) and the adaptive impact of language experience (Tremblay *et al.*, 2017). That is, even though a speaker's L1 predisposes them to using specific prosodic patterns, they remain able to modulate their prosody production to the specifics of an L2 if provided with sufficient L2 experience. In addition, our results show that the two cues—*F*0 and duration—are affected differently by speakers' L1 and relative language dominance, suggesting that some aspects of the restricted segmentation strategy acquired by L1 speakers might be harder to modulate than others. In Secs. IV A–IV D, we discuss the overall effects of L1 and relative language dominance separately and demonstrate how they support our proposal of a hybrid model of L2 prosody production.

A. L1 effects

Analyzing the production of English trials, we found that speakers' L1 interacted with condition (one bisyllabic word or two monosyllabic words) to predict the *F*0 ratios produced, but that it did not affect speakers' production of duration ratios. Specifically, while speakers from all three language groups used similar duration patterns to differentiate the two conditions, only French-L1 speakers also produced different *F*0 ratios across conditions (simultaneous bilinguals did not differ significantly from the English-L1 group used as a reference, indicating that they produced native-like *F*0 ratios in English). Based on the production of the syllables of interest in the English-L1 group, native-like production of these sentence pairs in English is signalled by different duration ratios, but not different *F*0 ratios. Therefore, the prosodic patterns produced by the French-L1 speakers represent an accented or non-native-like use of *F*0 in English.

Moreover, given that the *F*0 ratios produced by simultaneous bilinguals (who learned both French and English from birth) did not differ from that of English-L1 speakers (who learned only English from birth, and French later in life), one could hypothesize three possible explanations as to why French-L1 speakers failed to produce native-like *F*0 ratios in English. A first explanation relates to the time period of first exposure to English, suggesting that speakers need to be exposed to English from birth (like English-L1 and simultaneous bilinguals) to achieve native-like use of *F*0 in English, so that the later the first exposure to English, the less native-like the *F*0 ratios produced in English. A second explanation relates to French-L1 speakers' limited cumulative exposure to English, given that they learned English only later in life (mean age of first exposure to English = 6.5 years). A similar explanation was previously put forward by Trofimovich and Baker (2006), who found that native Korean speakers produced more native-like stress timing in English as their cumulative exposure to English increased. A third possible explanation would relate instead to relative language dominance, suggesting that speakers need to reach a certain proficiency level in English to use *F*0 in an English-like manner.

To determine which of the three potential interpretations is most likely in the present dataset, we used additional LME models including condition and either age of first exposure to English, cumulative exposure to English (chronological age minus age of first exposure to English), or the number of words produced in the English version of the fluency task (as an index of proficiency) as predictors of *F*0 ratios produced by French-L1 speakers (all as scaled continuous variables). These models revealed different patterns of results. First, we observed a significant interaction between condition and age of first exposure to English, where speakers who were exposed to English at an earlier age produced smaller *F*0 differences between conditions (i.e., more native-like use of *F*0 in English) than speakers whose exposure to English first occurred at a later age. On the other hand, the interaction between condition and cumulative exposure to English failed to reach significance, meaning that the amount of time spent in English did not explain the use of *different* *F*0 ratios in these two conditions. And finally, the interaction between condition and English proficiency was significant ($p = 0.04$), with speakers who named more items during the English verbal fluency task producing smaller *F*0 ratio differences between conditions (i.e., more native-like use of *F*0 in English) than speakers who named fewer items in the same task; see supplementary material Appendix B, Tables 3a, 3 b, and 3c for complete model outputs).[1] Therefore, it seems that the age of a speaker's first exposure to English and their attained English proficiency level were better predictors of their use of different *F*0 ratios to differentiate the two conditions (or not). Interestingly, the age of a French-L1 speaker's first exposure to English did not correlate with their English proficiency as indexed by the verbal fluency task, suggesting that each of these two variables might have a specific impact on *F*0 ratio production in English.

Unlike with English trials, the analysis of French trials revealed a significant interaction between speakers' L1 (English-L1 or simultaneous bilinguals) and condition (one bisyllabic word or two monosyllabic words) in predicting the production of both *F*0 and duration ratios, although the two effects were in opposite directions. Namely, while French-L1 speakers produce larger *F*0 ratio differences between conditions than English-L1 speakers, they also produce smaller duration ratio differences to differentiate the two conditions compared to English-L1 speakers. Using the production of the syllables of interest in the French-L1 group as a native-like reference, the production of these sentence pairs in French relies on both *F*0 and duration ratios. Therefore, the prosodic patterns produced by the English-L1 speakers represent an accented or non-native-like use of prosody in French; English-L1 speakers appear to under-use *F*0 ratios to differentiate the two speaking conditions, and instead over-use duration ratios to compensate.

Given that the *F*0 and duration ratios produced by simultaneous bilinguals (who learned both French and English from birth) differed from native-like productions in French, one might conclude that variables pertaining to L2 exposure (age of first exposure and cumulative exposure to French) had little impact on the prosodic cues produced by native speakers of English (simultaneous bilinguals and English-L1 speakers), otherwise simultaneous bilinguals would be expected to produce native-like prosody during French trials. Interestingly, although simultaneous bilinguals' production of *F*0 ratios did not pattern with those of French-L1 speakers, it also did not pattern with those of English-L1 speakers. Rather, the simultaneous bilinguals presented an intermediate pattern of results, in which they produced significantly smaller *F*0 ratio differences between conditions than French-L1 speakers, but significantly larger *F*0 ratio differences than English-L1 speakers. Therefore, having been exposed to English from birth might have influenced their use of *F*0, somewhat hindering their ability to produce native-like *F*0 ratios during French trials. With regard to duration patterns, simultaneous bilinguals did not differ from English-L1 speakers, meaning that they, too, produced exaggerated duration ratios compared to French-L1 speakers. These complex patterns of results suggest that L1 is not a sufficient predictor of the use of *F*0 and duration ratios in French; it may be, as suggested above, that the simultaneous acquisition of English influenced their prosody production, or it may be important to take relative language dominance into account to help explain the observed results. Although not perfect, the relative language dominance index gives us an idea regarding relative language proficiency levels which are independent from a speaker's L1.

B. Relative language dominance effects

When analyzing the production of English sentences in relation to speakers' relative language dominance (irrespective of L1), we found results largely consistent with those reported for L1. Namely, all speakers produced different duration ratios to differentiate the two conditions, but only speakers at the French dominant end of the continuum also produced different *F*0 ratios for the different conditions. Thus, as bilinguals become more English dominant, their production of *F*0 ratios becomes more English-like (in English), regardless of their L1. Subsequent analyses further supported this interpretation, by demonstrating that the production of *F*0 ratios by simultaneous bilinguals was modulated by their relative language dominance, regardless of the fact that they are native speakers of English. Thus, even if, as a group, simultaneous bilinguals behaved like native speakers of English, the production of each individual speaker was nonetheless affected by their long-term experience with French affecting their relative language dominance. These results cohere with those of Cutler *et al.* (1992), who found an effect of language dominance among simultaneous bilinguals' use of prosodic cues in word/syllable spotting tasks. Moreover, these results demonstrate that the production of native-like *F*0 ratios in English is more variable than previously expected, and that native exposure to English is not sufficient to maintain native-like production of *F*0 in English if exposed to another language with different prosodic patterns (see Kim, 2019 for similar effects on Spanish lexical stress production.) Nonetheless, *F*0 production of the English-L1 speaker group (who learned English from birth and French only later in life) was not affected by their relative language dominance, which might give the impression that *F*0 production in their L1 was somewhat immune to influences from their L2. However, this result is likely due to the limited French proficiency of our English-L1 speakers who were all English dominant according to our relative language dominance index (i.e., they all performed better on the English version of the verbal fluency task than in the French version of the task, leading to index scores above 1). Thus, the English-L1 speakers might not have had enough contact with their L2 for it to have affected their *F*0 production in L1.

Of note, given the overall effects observed in the L1 and relative language dominance analyses, we expected to see an interaction between condition and relative language dominance in the follow-up analyses performed on the English productions of French-L1 speakers. The absence of a significant effect might be due to the limited range of relative language dominance of the French-L1 group (range of 1.2 index points for French-L1 speakers compared to 1.6 index points for English-L1 speakers and 2 index points for simultaneous bilinguals).

Subsequently, when analyzing French trials in terms of relative language dominance, we also found effects in line with those found in the L1 analyses. Speakers from the French dominant end of the spectrum produced different *F*0 and duration ratios between conditions (like French-L1 speakers), whereas speakers from the English dominant end of the scale produced exaggerated duration ratio differences but no *F*0 ratio differences between conditions. Thus, congruent with the results obtained for the English trials, as bilinguals become more French dominant, their prosody production becomes more French-like (in French), regardless of their L1. This interpretation was further supported by follow-up LME models on the different speaker groups separately. These models showed that both the production of *F*0 and duration ratios by English-L1 speakers were modulated by their relative language dominance. That is, more balanced speakers (none were French dominant according to our index) produced more native-like prosody in French, while speakers at the English dominant end of the spectrum did not produce different *F*0 to differentiate conditions, but instead produced larger duration differences between conditions. This effect demonstrates that even though English-L1 speakers might have acquired a restricted use of *F*0 during infancy (Cutler *et al.*, 1992), they were still able to learn to modulate their *F*0 production to the specifics of their L2, at least to a certain degree, after acquiring sufficient language experience. English-L1 speakers whose linguistic experience does not allow them to modulate their *F*0 production in French instead compensate by modulating syllable duration to signal word boundaries. These results partly cohere with those of Tremblay *et al.* (2017), who demonstrated that language experience affected listeners' use of *F*0 in speech

segmentation. Contrary to our results, they found that the use of *F*0 by both French and English native listeners was affected by their L2 proficiency, whereas in our results, this effect seems to be limited to English-L1 speakers, possibly due to the limited range of relative language dominance of the French-L1 group in the present dataset.

Of note, the French productions of French-L1 speakers, like the English productions by English-L1 speakers, seemed somewhat immune to influences from exposure to the L2, even though some French-L1 speakers are in fact English-dominant according to our relative language dominance index. Moreover, simultaneous bilinguals patterned with French-L1 speakers in the follow-up analyses that looked at speaker groups separately, suggesting that in simultaneous bilinguals the different use of prosody in French might be due to their exposure to English from birth and not to their ongoing experience with either language.

C. On the adaptive and selective nature of prosody production in L2

Taken together, the present results confirm our hypothesis that both speakers' L1 and language dominance have a significant impact on the use of *F*0 and duration cues in L2. Also, when comparing the production of English and French trials within sequential bilinguals (French-L1 and English-L1), we observed that speakers did, in fact, produce different *F*0 and duration ratio patterns across languages. Thus, all speakers indeed attempted to adapt their prosody to their L2, albeit imperfectly. Therefore, we can interpret the present results as demonstrating that, even though speakers' L1 predisposes them to use certain *F*0 and duration patterns as word boundary cues (regardless of the language they are speaking), L2 speakers can nonetheless acquire native-like production patterns if they have sufficient L2 experience. Thus, speakers may continue to apply the restricted prosodic patterns associated with their L1 to their L2 during the early phases of language learning, but they learn to adapt their prosody to native-like patterns as their language dominance shifts towards their L2. These results partly cohere with the perception results from Tremblay *et al.* (2017), who suggested that the use of prosodic cues to segment words is adaptive, in that it is modulated by language experience variables, but non-selective (cannot be adjusted as a function of the language to process). As in their study, we found that language experience (as indexed by relative language dominance) affected how speakers use prosodic cues to segment words, but contrary to their results, we observed that our participants were able to produce F0 and duration ratios in a language-selective way. That is, our speakers did not rely on a single intermediate pattern of *F*0 and duration ratios in both languages, but instead used two parallel systems, specifically tailored to French and English, with language experience affecting how native-like their production was in both languages. Thus, language experience appears to have a different impact on the production and the perception of prosodic cues associated with word segmentation (see also Kim, 2019).

D. On the specific challenges of the transition between English and French (and vice-versa)

To address the specific challenges of learning English as a French-L1 speaker, one has to relate the present results to the specifics of the stimuli used, specifically to the one-word condition that is driving all the interactions observed in both *F*0 and duration ratios. First, it is important to remember that, as a stress-timed language, English words carry lexically-coded stress that needs to be produced independently of sentence-level prosody. For example, most of the bisyllabic words used in the one-word condition in English have a trochaic stress pattern (37 sentence pairs out of 40). Therefore, one would expect the initial syllable of these words to be produced with higher *F*0, longer duration, and greater intensity than the second syllable (Beckman, 1986; Lieberman, 1960). Also, in the present task, most of these bisyllabic words were placed at the end of prosodic phrases, meaning that they also bear phrasal prosody cues, including final lengthening as well as a potential pitch accent, depending on the specific construction of the sentence (see Wagner and Watson, 2010, for a review of prosodic cues associated with sentence-level boundaries). Previous studies have shown that, in English, final lengthening not only affects the final syllable of the phrase (Vaissière, 1983; Wightman *et al.*, 1992) but also the preceding stressed syllable (Turk and Shattuck-Hufnagel, 2007). Therefore, the production of trochaic bisyllabic words in phrase final position might present a particular challenge for non-native French-L1 speakers since they are known to have difficulty encoding lexical stress (Dupoux *et al.*, 1997) and they signal phrase boundaries by lengthening only the final syllable of a phrase (Fletcher, 1991).

Furthermore, even though French L1 or dominant speakers are able to vary *F*0 independently from duration in certain contexts, these two acoustic cues often cooccur as phrase boundary markers in casual speech (Fletcher, 1991; Vaissière, 1983). Therefore, French L1/dominant speakers might have difficulty modulating these two cues to signal word-level and sentence-level prosody in a coordinated fashion (i.e., producing a *F*0 rise on the first syllable to indicate lexical stress, but not on the second syllable to signal phrase boundary, while lengthening both syllables to indicate phrase boundary as opposed to lengthening only the second syllable as in French). Evidence of this difficulty comes from the observed production of similar *F*0 and duration values for both syllables of bisyllabic English words (leading to ratios of about 1) by French L1/dominant speakers, which seems to indicate that they used the same acoustic cue combination to signal both the lexical stress and the phrase boundary. Three different patterns of prosodic cue use could explain the patterns of *F*0 and duration ratios that emerged. A first pattern would involve French-L1 speakers failing to produce lexical stress in the first syllable altogether but managing to refrain from using French prosody to mark the phrase boundary (leading to two "short" syllables with low *F*0 values). A second pattern would involve French-L1 speakers managing to produce native-like lexical stress on the first syllable of the bisyllabic word, but not being able to

refrain from using French prosody to mark the phrase boundary on the second syllable (leading to two lengthened syllables with high *F*0 values). A third pattern would present a middle ground between the previous two explanations, with French-L1 speakers modulating their prosody to signal lexical stress, albeit not in a native-like manner, and mitigating their production of phrase boundary markers.

Additional follow-up LME models investigating the effect of language of production on French-L1 speakers' *F0* and duration of each syllable of the English bisyllabic word separately (for example, comparing the raw *F*0 value of the first syllable of the bisyllabic words in French vs English trials, or comparing the raw duration of the second syllable of the bisyllabic word across languages) seem to support the third and more nuanced explanation of the flat *F*0 and duration ratios proposed earlier (see supplementary material Appendix E, Table 1a, for complete model outputs).[1] These analyses on raw *F*0 values showed that the *F*0 of the second syllable was affected by the language of the trial (lower *F*0 in English than in French), but not the *F*0 of the first syllable. On the other hand, the duration of both syllables was affected by the language of the trial, with English trials being produced with a longer first syllable and a shorter second syllable. This pattern of results indicates that French-L1 speakers were able to refrain from marking phrase boundaries in a French-like manner during English trials but were not able to use *F*0 as a lexical stress cue, instead relying only on syllabic duration. Thus, French-L1 speakers show signs of having modulated their prosody to the specifics of their L2, albeit imperfectly. The *F*0 and duration ratios of 1 (indicating that similar *F*0 and duration were produced for both syllables) is therefore the result of two simultaneous prosodic modulations affecting word- and sentence-level prosody in English. However, French-L1 speakers did not use *F*0 variations to signal lexical stress, indicating that they likely do not associate *F*0 modulations with word-level prosody. This suggests that it is easier for French-L1 speakers to modulate syllabic durations to the specifics of English than to produce *F*0 in an English-like manner. These results are consistent with those of Trofimovich and Baker (2006), who found that native Korean speakers did not reach native-like F0 peak alignments in English, no matter their language experiences, even if they were able to produce native-like lexical stress. Taken together, these results suggest that it might be difficult for native speakers of languages without lexical stress to learn to adapt their F0 production to the specifics of a language with lexical stress, like English. Interestingly, our paradigm also allows for the investigation of a potential converse effect that could affect English-L1 speakers use of *F*0 in French.

As a syllable-timed language, French does not encode word-level prosody, so the equivalent bisyllabic words used in the French trials need only be marked for the phrasal boundaries. In French, phrase boundaries are marked by a syllabic lengthening accompanied by an *F*0 rise, both occurring on the final syllable of the phrase (also marking word offsets; Fletcher, 1991; Vaissière, 1983). Therefore, French trials present a different challenge for English native speakers, who more often associate *F*0 rises with lexical stress than with phrase boundaries (Beckman and Ayers Elam, 1997). Native English speakers might then have difficulty using *F*0 as a boundary marker in French, especially when producing a phrase-final bisyllabic word, as indicated by their use of equivalent *F0* values in both syllables of these words (leading to *F*0 ratios of about 1). On the other hand, the use of syllabic lengthening to signal word/phrase boundaries in French should not be problematic for English L1/dominant speakers because duration is also used to signal sentence-level prosody in English to a certain degree (Vaissière, 1983; Wightman *et al.*, 1992). This could, in turn, explain why English L1/dominant speakers produce exaggerated duration ratio differences between conditions in French; they appear to rely on the prosodic cue that is shared across the two languages (duration) to compensate for the limited ability to modulate *F*0 in a French-like manner.

Additional follow-up LME models investigating the effect of language of production on English-L1 speakers' *F*0 and duration of each syllable of the bisyllabic French word separately suggest that, as expected, English-L1 speakers are able to modulate *F*0 independently from syllabic duration, and that they are able to produce *F*0 rises and syllabic lengthening as phrase boundary markers in French, albeit not quite in a native-like manner (consistent with Tremblay *et al.*, 2017; see supplementary material Appendix E, Table 1 b, for complete model outputs).[1] Unlike the analyses on English trials produced by French-L1 speakers, the present analyses revealed that the *F*0 of the second syllable was affected by the language of the trial (higher *F*0 in French than in English), but not the *F*0 of the first syllable. Duration of both syllables was also affected by the language of the trial, with French trials being produced with a shorter first syllable and a longer second syllable. This pattern of results indicates that English-L1 speakers were able to mark phrase boundaries in a French-like manner but were not able refrain completely from producing an English-like stress pattern on bisyllabic French words, producing the *F*0 rise on the first syllable typical of the trochaic stress pattern of English. Thus, English-L1 speakers show signs of having modulated their sentence-level prosody to signal phrase boundaries in a native-like manner (longer final syllable with higher *F*0), but of only partly adapting their word-level prosody, producing shorter first syllables in French, but with similar *F*0 patterns in both languages, regardless of the absence of lexical stress in French. These results suggest that it is easier for English-L1 speakers to adapt their sentence-level prosody to the specifics of French than to refrain from producing word-level prosody, specifically controlling *F*0 in context where a lexical-stress would be produced in English. Taken together, the results from French-L1 and English-L1 speakers' prosodic modulations support a more general principle according to which syllabic duration is easier to adapt to the specifics of an L2 than *F*0 production (at least among French-English bilinguals), and sentence-level prosody is easier to adapt than word level prosody.

Of note, the complex effects observed in the one-word condition are not mirrored in the two monosyllabic words condition in either language. In the two monosyllabic words condition, all sentence-level and word-level (in English)

prosodic cues fall on the same syllable, avoiding the potential confusion of where to produce lexical stress (or not) and where to produce phrase boundary cues. Such contexts are therefore more likely to be produced in a native-like manner in either French or English, even by speakers having difficulty modulating their F0 and duration ratios.

V. CONCLUSION

In sum, the present results demonstrate that bilingual speakers are not bound to using the restricted word segmentation procedure of their L1 in the production of their L2. Specifically, the bilingual speakers included in the present study show signs of being able to adapt their *F*0 and duration production to the specifics of their L2 if they have sufficient L2 experience, albeit duration seems somewhat easier to adapt than *F*0. Furthermore, the fact that our speakers produced different prosodic cues in English and in French demonstrates that language exposure does not simply modify the restricted word segmentation procedure learned through contact with the L1, but instead helps to create a parallel set of word segmentation procedures specifically tailored to the L2. Thus, even if a speaker has learned to use the English word segmentation strategy during infancy, they nonetheless remain able to learn to use the French word segmentation strategy (given that they reach a certain level of French experience) and keep using the English strategy when speaking their L1. The present results also demonstrate that adapting the production of sentence-level prosodic cues (phrase boundary marking) seems easier than adapting word-level prosodic cues (lexical stress or lack thereof). Therefore, the transition between English and French does not seem easier in one direction than the other, but rather involves direction-specific challenges, with French-L1 speakers having difficulty producing trochaic *F*0 patterns in English while English-L1 speakers have difficulty refraining from producing the same pattern in French. Given the discrepancy between the results from the present production study and previous studies on acoustic word segmentation involving similar language-pairs (Cutler *et al.*, 1992; Tremblay *et al.*, 2017), a perception task relying on the same sentences is presently in progress to determine if language experience has a different impact on the perception and production of bilingual speakers, as reported by Kim (2019).

ACKNOWLEDGMENTS

This work was supported by a team grant from the Fonds de Recherche du Québec—Société et Culture, and the Blema and Arnold Steinberg Family Foundation.

[1]See supplementary material at https://doi.org/10.1121/1.5134781 for a complete list of the stimuli and complete model outputs (Appendixes A–E).

Anderson-Hsieh, J., and Koehler, K. (**1988**). "The effect of foreign accent and speaking rate on native speaker comprehension," Lang. Learn. **38**(4), 561–613.

Baayen, R. H. (**2008**). *Analyzing Linguistic Data: A Practical Introduction To Statistics Using R*. (Cambridge University Press, Cambridge, UK).

Baayen, R. H., Davidson, D. J., and Bates, D. M. (**2008**). "Mixed-effects modeling with crossed random effects for subjects and items," J. Mem. Lang. **59**(4), 390–412.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (**2013**). "Random effects structure for confirmatory hypothesis testing: Keep it maximal," J. Mem. Lang. **68**(3), 255–278.

Bates, D., Kliegl, R., Vasishth, S., and Baayen, R. H. (**2015**). "Parsimonious mixed model," arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (**2014**). "lme4: Linear mixed-effects models using Eigen and S4. (Version R package version 1.1-6)," http://CRAN.R-project.org/package=lme4 (Last viewed November 20, 2019).

Beckman, M. E. (**1986**). *Stress and Non-Stress Accent* (Foris, Dordrecht, the Netherlands).

Beckman, M. E., and Ayers Elam, G. (**1997**). *Guidelines for ToBI Labelling*, http://www.speech.cs.cmu.edu/tobi/ (Last viewed November 20, 2019).

Birdsong, D. (**2015**). "Dominance in bilingualism: Foundations of measurement, with insights from the study of handedness," in *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, edited by C. Silva-Corvalán and J. Treffers-Daller (Cambridge University Press, Cambridge, UK), pp. 85–105.

Boersma, P. (**2001**). "Praat, a system for doing phonetics by computer," Glot Int. **5**(9/10), 341–345, available at https://hdl.handle.net/11245/1.200596.

Christophe, A., Gout, A., Peperkamp, S., and Morgan, J. (**2003**). "Discovering words in the continuous speech stream: The role of prosody," J. Phon. **31**, 585–598.

Christophe, A., Mehler, J., and Sebastián-Gallés, N. (**2001**). "Perception of prosodic boundary correlates by newborn infants," Infancy **2**(3), 385–394.

Christophe, A., Millotte, S., Bernal, S., and Lidz, J. (**2008**). "Bootstrapping lexical and syntactic acquisition," Lang. Speech **51**, 61–75.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., and Mehler, J. (**2004**). "Phonological phrase boundaries constrain lexical access. I. Adult data," J. Mem. Lang. **51**, 523–547.

Cutler, A., Dahan, D., and van Donselaar, W. (**1997**). "Prosody in the comprehension of spoken language: A literature review," Lang. Speech **40**(2), 141–201.

Cutler, A., Mehler, J., Norris, D., and Segui, J. (**1986**). "The syllable's differing role in the segmentation of French and English," J. Mem. Lang. **25**(4), 385–400.

Cutler, A., Mehler, J., Norris, D., and Segui, J. (**1992**). "The monolingual nature of speech segmentation by bilinguals," Cogn. Psychol. **24**(3), 381–410.

Cutler, A., and Otake, T. (**1994**). "Mora or Phoneme? Further evidence for language-specific listening," J. Mem. Lang. **33**(6), 824–844.

Cutler, A., and Otake, T. (**2002**). "Rhythmic categories in spoken-word recognition," J. Mem. Lang. **46**, 296–322.

Cutler, A., Weber, A., and Otake, T. (**2006**). "Asymmetric mapping from phonetic to lexical representations in second-language listening," J. Phon. **34**(2), 269–284.

Davis, M. H., and Johnsrude, I. S. (**2007**). "Hearing speech sounds: Top-down influences on the interface between audition and speech perception," Hear. Res. **229**(1), 132–147.

De Groot, A. M. B. (**2011**). *Language and Cognition in Bilinguals and Multilinguals: An Introduction* (Psychology Press, New York).

Dupoux, E., Pallier, C., Sebastian, N., and Mehler, J. (**1997**). "A destressing 'deafness' in French?," J. Mem. Lang. **36**(3), 406–421.

Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. (**2001**). "A robust method to study stress 'deafness'," J. Acoust. Soc. Am. **110**(3), 1606–1618.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., and Peperkamp, S. (**2008**). "Persistent stress 'deafness': The case of French learners of Spanish," Cognition **106**(2), 682–706.

Fletcher, J. (**1991**). "Rhythm and final lengthening in French," J. Phon. **19**, 193–212.

Guion, S. G., Harada, T., and Clark, J. J. (**2004**). "Early and late Spanish–English bilinguals' acquisition of English word stress patterns," Biling. Lang. Congn. **7**(3), 207–226.

Jacobs, A., Fricke, M., and Kroll, J. F. (**2016**). "Cross-language activation begins during speech planning and extends into second language speech," Lang. Learn. **66**(2), 324–353.

Jusczyk, P. W. (**1999**). "How infants begin to extract words from speech," Trends Cogn. Sci. **3**(9), 323–328.

Jusczyk, P. W., Houston, D. M., and Newsome, M. (**1999**). "The beginnings of word segmentation in English-learning infants," Cogn. Psychol. **39**(3–4), 159–207.

Kainada, E., and Lengeris, A. (**2015**). "Native language influences on the production of second-language prosody," J. Int. Phon. Assoc. **45**(3), 269–287.
Kim, J. Y. (**2019**). "Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress," Biling. Lang. Cogn. (published online).
Kroll, J. F., and Bialystok, E. (**2013**). "Understanding the consequences of bilingualism for language processing and cognition," J. Cogn. Psychol. **25**(5), 497–514.
Kroll, J. F., and Gollan, T. H. (**2014**). "Speech planning in two languages: What bilinguals tell us about language production," in *The Oxford Handbook of Language Production*, M. Goldrick, V. S. Ferreira, and M. Miozzo (Oxford University Press, Oxford, UK), pp. 165–181.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2015**). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-29.
Lehiste, I. (**1972**). "The timing of utterances and linguistic boundaries," J. Acoust. Soc. Am. **51**(6B), 2018–2024.
Li, P., Zhang, F., Tsai, E., and Puls, B. (**2013**). "Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool," Biling. Lang. Cogn. **17**(3), 673–680.
Lieberman, P. (**1960**). "Some acoustic correlates of word stress in American English," J. Acoust. Soc. Am. **32**(4), 451–454.
Magen, H. S. (**1998**). "The perception of foreign-accented speech," J. Phon. **26**(4), 381–400.
Mareüil, P., and Vieru-Dimulescu, B. (**2006**). "The contribution of prosody to the perception of foreign accent," Phonetica **63**(4), 247–267.
Mattys, S. L., Jusczyk, P. W., Luce, P. A., and Morgan, J. L. (**1999**). "Phonotactic and prosodic effects on word segmentation in infants," Cogn. Psychol. **38**, 465–494.
McQueen, J. M. (**1998**). "Segmentation of continuous speech using phonotactics," J. Mem. Lang. **39**(1), 21–46.
Mehler, J., Dommergues, J. Y., Frauenfelder, U., and Segui, J. (**1981**). "The syllable's role in speech segmentation," J. Verbal Learn. Verbal Behav. **20**(3), 298–305.
Munro, M. J., and Derwing, T. M. (**1995**). "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners," Lang. Learn. **45**(1), 73–97.
O'Brien, M. G., Jackson, C. N., and Gardner, C. E. (**2014**). "Cross-linguistic differences in prosodic cues to syntactic disambiguation in German and English," Appl. Psycholing. **35**(01), 27–70.
Otake, T., Hatano, G., Cutler, A., and Mehler, J. (**1993**). "Mora or syllable? Speech segmentation in Japanese," J. Mem. Lang. **32**(2), 258–278.
Pike, K. L. (**1945**). *The Intonation of American English* (University of Michigan Press, Ann Arbor, MI).
Piske, T., MacKay, I. R., and Flege, J. E. (**2001**). "Factors affecting degree of foreign accent in an L2: A review," J. Phon. **29**, 191–215.
Quené, H. (**1993**). "Segment durations and accent as cues to word segmentation in Dutch," J. Acoust. Soc. Am. **94**(4), 2027–2035.
R Development Core Team. (**2010**). "R: A language and environment for statistical computing," Vienna, Austria: R Foundation for Statistical Computing, www.R-project.org (Last viewed November 20, 2019).
Rasier, L., and Hiligsmann, P. (**2007**). "Prosodic transfer from L1 to L2. Theoretical and methodological issues," Nouveaux Ling. Française **28**, 41–66, available at http://hdl.handle.net/2078.1/86887.
Rietveld, A. M. C. (**1980**). "Word boundaries in the French language," Lang. Speech **23**(3), 289–296.
Shen, X. S. (**1990**). "Language by speakers of a tonal language: Chinese speakers learning French prosody," Int. Rev. Appl. Ling. **28**, 119–134.
Skoruppa, K., Nevins, A., Gillard, A., and Rosen, S. (**2015**). "The role of vowel phonotactics in native speech segmentation," J. Phon. **49**, 67–76.
Suomi, K., McQueen, J. M., and Cutler, A. (**1997**). "Vowel harmony and speech segmentation in Finnish," J. Mem. Lang. **36**(3), 422–444.
Tabachnick, B. G., and Fidell, L. S. (**2007**). *Using Multivariate Statistics*, 5th ed. (Pearson/Allyn and Bacon, Boston).
Treffers-Daller, J., and Korybski, T. (**2015**). "Using lexical diversity measures to operationalize language dominance in bilinguals," in *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, edited by C. Silva-Corvalán and J. Treffers-Daller (Cambridge University Press, Cambridge, UK), pp. 106–133.
Tremblay, A., Broersma, M., Coughlin, C. E., and Choi, J. (**2016**). "Effects of the native language on the learning of fundamental frequency in second-language speech segmentation," Front. Psychol. **7**, 985.
Tremblay, A., Coughlin, C. E., Bahler, C., and Gaillard, S. (**2012**). "Differential contribution of prosodic cues in the native and non-native segmentation of French speech," Lab. Phonol. **3**, 385–423.
Tremblay, A., Namjoshi, J., Spinelli, E., Broersma, M., Cho, T., Kim, S., Martínez-García, M. T., and Connell, K. (**2017**). "Experience with a second language affects the use of fundamental frequency in speech segmentation," PLoS One **12**(7), e0181709.
Trofimovich, P., and Baker, W. (**2006**). "Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech," Studies Second Lang. Acq. **28**(01), 1–30.
Turk, A. E., and Shattuck-Hufnagel, S. (**2007**). "Multiple targets of phrase-final lengthening in american English words," J. Phon. **35**, 445–472.
Tyler, M. D., and Cutler, A. (**2009**). "Cross-language differences in cue use for speech segmentation," J. Acoust. Soc. Am. **126**(1), 367–376.
UCLA: Statistical Consulting Group. (**2011**). *R Library Contrast Coding Systems for Categorical Variables*, https://stats.idre.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables/ (Last viewed Septebmer 6, 2019).
Vaissière, J. (**1983**). "Language-independent prosodic features," in *Prosody: Models and measurements*, edited by A. Cutler and D. R. Ladd (Springer, New York,) pp. 53–66.
Wagner, M., and Watson, D. G. (**2010**). "Experimental and theoretical advances in prosody: A review," Lang. Cogn. Process. **25**(7–9), 905–945.
Weber, A., and Cutler, A. (**2006**). "First-language phonotactics in second-language listening," J. Acoust. Soc. Am. **119**(1), 597–607.
White, L., Mattys, S. L., Stefansdottir, L., and Jones, V. (**2015**). "Beating the bounds: Localized timing cues to word segmentation," J. Acoust. Soc. Am. **138**(2), 1214–1220.
Wickham, H. (**2009**). *ggplot2: Elegant Graphics for Data Analysis*. (Springer, New York).
Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (**1992**). "Segmental durations in the vicinity of prosodic phrase boundaries," J. Acoust. Soc. Am. **91**(3), 1707–1717.