Prosodic cue weighting in the perception of lexical stress by French-English bilinguals: Preliminary behavioural data

Annie C. Gilbert[1,2], Claire T. Honda[1,2], Shari R. Baum[1,2]

[1]Centre for Research on Brain, Language, and Music, Montreal, Canada
[2]School of Communication Sciences and Disorders, McGill University, Montreal, Canada
annie.c.gilbert@mail.mcgill.ca , claire.honda@mail.mcgill.ca , shari.baum@mcgill.ca

ABSTRACT

Learning to adapt one's prosodic processing to a second or non-dominant language can be difficult, particularly if the second language uses different prosodic constructs than the native language. In the present study, we examine how French-English bilinguals with varied language experience (native language, relative language dominance, proficiency) assign weight to F0 and syllabic duration cues during the perception of English lexical stress. Participants were asked to report, via key press, which stress pattern they perceived when presented with words produced with either their original prosody or with F0 or duration edited. Preliminary analyses of the results from 16 participants suggest that language experience has a significant impact on the weight assigned to F0 cues in perceiving English lexical stress, but not on the weight assigned to duration cues. These results align with previous findings suggesting that French-English bilinguals have an easier time producing L2-specific duration cues than L2-specific F0 cues.

Keywords: bilingualism, speech perception, prosody, lexical stress, acoustic cue weighting

1. INTRODUCTION

Learning a second language (L2) is not a trivial task. Beyond learning to recognize new speech sounds, L2 learners must also learn to interpret prosodic cues appropriately. In many cases, this process requires the learner to assign new meaning to acoustic cues, because the same cue can play different roles during native language (L1) versus L2 perception. For instance, while rises in fundamental frequency (F0) are often associated with lexical stress in English [1], they are often associated with phrase boundaries in French [2], where they co-occur with phrase-final lengthening. Of note, lexical stress and phrase boundaries both provide important information about how to segment the speech stream, with English lexical stress tending to fall on word onsets [3] and French phrase boundaries falling on word offsets. Thus, the misinterpretation of the F0 cue in L2 (as lexical stress instead of phrase boundary, or vice-versa) could lead to segmentation errors, hindering speech comprehension.

Past research with bilingual participants has revealed that successful learning of L2 prosodic processing varies as a function of learners' experience and of the specific characteristics of the languages involved. For example, in terms of language experience, more native-like use of L2 or non-dominant language cues was predicted by greater length of immersion in an L2-speaking environment, higher L2 or non-dominant language proficiency, and relative language dominance [4-7].

Moreover, research has shown that it is easier to learn L2-specific prosody processing when both L1 and L2 make use of the same cue. For instance, one study compared English-French and Dutch-French bilinguals' use of F0 rise as a cue to word offset in French (their L2) [8]. English and Dutch both differ from French in that they typically use F0 rise to signal word onset rather than word offset; however, they also differ from each other in that the F0 rises associated with lexical stress/word onset in English generally co-occur with changes in vowel quality, which is not the case in Dutch. Thus, F0 rises have less functional weight in English than in Dutch. Interestingly, compared to the English-L1 speakers, the Dutch-L1 speakers showed earlier and greater use of the F0 rise cue in French [8]. This finding suggests that it was easier for Dutch-L1 listeners to successfully reallocate F0 rises to a new function in their L2 (i.e., identifying word offset instead of onset), compared to English-L1 listeners who usually encounter F0 rises paired with an additional cue (vowel change). It therefore appears that the more functional weight a cue has in one's L1, the easier it is to use the cue in an L2.

Thus, previous research has established the factors modulating learners' ability to learn to segment an L2 (or any non-dominant language) in a native-like manner, but questions remain as to how learners associate different acoustic cues to prosodic constructs in their L2—particularly when such constructs do not exist in their L1. Various studies have provided evidence that it is difficult to perceive and produce lexical stress when it is not used in one's L1 (e.g., [4, 6, 9]). As such, French-L1 learners of English have the challenging task of learning not only

to assign new meaning to prosodic cues to identify lexical stress instead of phrase boundaries, but also of learning to use lexical stress in the first place. The present study is part of a larger project aimed at clarifying the processes involved in such prosodic learning. We focus here on lexical stress perception, and on how non-native listeners assign weight to F0 and duration cues during lexical stress perception.

In the present study, we examined how French-English bilinguals with varied language backgrounds (native language, relative language dominance, English proficiency) assign weight to F0 and syllabic duration during the perception of English lexical stress. To do so, we created stimuli with conflicting F0 and syllabic duration cues, allowing us to observe on which cue participants rely. Based on previous research on speech segmentation in French-English bilinguals [4], we expected participants' use of F0 and duration cues to be modulated by their language experience.

We also expected that French-L1 participants might rely less heavily on F0 cues than native listeners to assign lexical stress in English because F0 rises often coincide with lengthened syllables in French, and therefore likely have a different weight in French compared to English.

2. METHOD

2.1. Participants

Sixteen French-English bilinguals were recruited from the Montreal area (18 to 41 years old, $M = 25$). Five reported French as their L1, five reported English as their L1, and six reported having been exposed to both languages from birth. Participants' proficiency in each language was assessed using a verbal fluency task. During this task, participants had one minute to name as many words as possible belonging to a specific category (one semantic category like "fruits" or "musical instruments" and three orthographic categories with words starting with specific letters per language.) The total number of words produced across trials in one language was used as a proficiency index for that language, and a relative language dominance index was computed by dividing the total number of English words produced by the total number of French words produced (see [4] and [5] for similar computations). Table 1 presents language proficiency scores of participants from each L1 group separately. Kruskal-Wallis Tests yielded significant differences across L1 groups in relative language dominance (p = 0.024) and English proficiency (p = 0.049). Pairwise comparisons revealed significant differences in relative language dominance only between English-L1 and French-L1

participants (p = 0.023), while English proficiency scores different significantly between English-L1 and French-L1 participants (p = 0.021) and between French-L1 and simultaneous bilinguals (p = 0.049).

	Language proficiency score		Relative language dominance
	English	French	English / French
	M (*SD*)	M (*SD*)	M (*SD*)
English-L1	63.8 (*11.8*)	41.2 (*10.5*)	1.61 (*0.40*)
Simultaneous	64.3 (18.0)	54.3 (18.4)	1.28 (0.46)
French-L1	45.0 (8.2)	46.6 (10.1)	0.99 (0.23)

Table 1: Average language proficiency (English and French) and relative language dominance scores as a function of participant native language.

2.2. Stimuli

Stimuli consisted of five bisyllabic base words that can be produced as either nouns or verbs depending on the prosodic structure used (trochaic nouns vs. iambic verbs; *implant*, *protest*, *refund*, *survey*, *torment*). The five words were recorded as both nouns and verbs by a native speaker of English in a sound-attenuated booth. Recordings were manually annotated by a trained phonetician to extract duration and mean F0 values for each vowel. F0 and duration ratios were computed for each version of the base word (trochaic noun and iambic verb) by dividing the value from the second syllable by the value from the first syllable. For example, a duration ratio above 1 indicates that the vowel of the second syllable was longer than the vowel of the first syllable while a ratio below one indicates that the first vowel was longer than the second (see [4] and [5] for a similar use of F0 and duration ratios).

Original recordings were then edited using *Praat* version 6.1.26 [10] to match the F0 or vowel duration ratio of the other stress pattern, leading to the creation of two edited conditions per stress pattern, as illustrated in Figure 1.

Stimuli were presented within three blocks following a practice block. The first block consisted of five repetitions of the original recordings of each base word produced with both trochaic and iambic stress patterns (as nouns or verbs). Stimulus order was
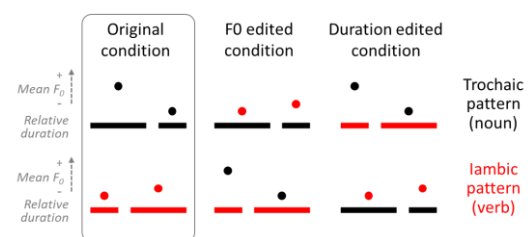

Figure 1: Schematic representation of conditions.

pseudorandomized to avoid consecutive repetitions of the same base word (50 trials in total). Results from this block are used a reference condition to which the other conditions are compared. The other two blocks (order counterbalanced across participants) consisted of five repetitions of each base word in the six conditions represented in Figure 1 (for a total of 150 trials per block). Stimulus order was again pseudorandomized to avoid consecutive repetitions of the same base word and the same condition. Participants had the opportunity to take up to three short breaks during each block as well as a break between blocks.

2.3. Procedure

Participants sat in a quiet room about 60 cm from a computer screen and listened to the stimuli through insert earphones (EARTONE 3A) at 66 dBA. They were instructed to listen to the word presented in isolation and to simply indicate which stress pattern they perceived via button press. Special emphasis was placed on the fact that participants should respond based on their first impression, even if this was not a speeded task. Half the participants signalled perceiving stress on the first syllable (trochaic pattern or noun) with the left hand and half with the right hand. Visual reminders of response key assignment were provided onscreen where stress was represented by capital letters. For instance, when hearing a trial involving the base word "protest", participants would see the written forms "PROtest" and "proTEST" above their assigned response buttons.

2.4. Statistical analyses

Responses were coded as "1" if the participant assigned the original stress pattern to the word regardless of the edited F0 or duration, and as "0" if they assigned the opposite stress pattern. Responses were analysed using a series of generalized linear mixed effects models (GLME) comparing results from the edited conditions to results from the first block of original recordings. Such comparison allowed us to determine if participants relied on the edited F0 or duration ratios to assign stress to the presented words. Each edited condition and original stress pattern was analysed separately.

Moreover, the models examined whether individual differences in language experience impacted reliance on F0 and duration during stress assignment. Based on previous publications, we investigated the potential impact of relative language dominance (log transformed), English proficiency (scaled), and L1 (as reference group = English-L1) [4, 5, 11]. In cases where a significant effect of L1 was observed to affect non-native (French-L1)

participants but not simultaneous bilinguals, a follow-up model was fitted using French-L1 as reference group to compare French-L1 and simultaneous bilinguals. The random structure of the models took into account participants (intercept) and base words (intercept; [Response ~ Condition * L1 + (1 | Participant)] + (1 | Base word)], family = binomial). GLME models were implemented in RStudio version 3.2.4 [12], using the lme4 library, version 1.1-7 [13], and estimates of p-values were obtained using the lmerTest package version 2.0-29 [14].

3. RESULTS

3.1. Trochaic pattern (nouns)

3.1.1. F0 edits vs. originals

The model examining the effect of L1 on the use of F0 in stress assignment revealed a significant interaction between L1 and condition (original vs. edited F0), where French-L1 listeners presented a smaller difference between conditions than native listeners (English-L1 and simultaneous bilinguals; ($|z|$ > 2.972, p < 0.003). Replacing L1 by relative language dominance yielded no significant effect of dominance as a main effect or as part of an interaction ($|z|$ < 1.698, p > 0.089).

On the other hand, the model taking English proficiency into account revealed a significant interaction between condition (original vs. edited F0) and English proficiency (z = 2.962, p = 0.003). Namely, participants who named more English words during the verbal fluency task showed a greater difference between conditions than participants who named fewer words. This pattern of results suggests that higher-proficiency English listeners rely more heavily on F0 when assigning stress than lower-proficiency English listeners.

3.1.2. Duration edits vs. originals

The models yielded no significant effect of individual differences as either main effects or as part of interactions ($|z|$ < 1.647, p > 0.099). Only the main effect of condition remained significant across comparisons ($|z|$ > 4.272, p < 0.001), with participants reporting having heard the original stress pattern less often in the duration edited condition than in the original condition.

3.2. Iambic pattern (verbs)

3.2.1. F0 edits vs. originals

As observed earlier, the model examining the effect of L1 on the use of F0 in stress assignment revealed a

significant interaction between L1 and condition. That is, replacing the original F0 ratio by the ratio from the trochaic condition had a smaller impact on stress assignment for French-L1 than for English-L1 participants ($z = -2.133$, $p = 0.033$). The model investigating the impact of relative language dominance yielded no significant effect of dominance (either as a main effect or part of an interaction; $|z| < 0.461$, $p > 0.645$).

Interestingly, the model investigating the effect of English proficiency on the use of F0 in stress assignment yielded no significant interaction between English proficiency and condition ($z = 1.294$, $p = 0.196$), but revealed a main effect of English proficiency, where less proficient English listeners were less likely to recognize the iambic stress pattern even in the original condition ($z = 1.999$, $p = 0.046$).

3.2.2. Duration edits vs. originals

As observed for the trochaic pattern, the models yielded no significant effect of individual differences as either main effects or part of interactions ($|z| < 1.939$, $p > 0.053$) while the main effect of condition remained significant across comparisons ($|z| > 4.117$, $p < 0.001$), suggesting that listeners' stress assignment was affected by modulation of duration patterns regardless of their linguistic experience.

4. DISCUSSION / CONCLUSION

The goal of the present study was to determine how much weight French-English bilinguals assign to different acoustic cues (F0 and duration) during lexical stress perception in English, and whether individual differences in language experience affect this weighting.

Our preliminary results suggest that experience did affect the use of F0 as a cue to lexical stress. Specifically, French-L1 bilinguals assigned less weight to F0 compared to English-L1 or simultaneous bilinguals. This finding held across both trochaic (noun) and iambic (verb) conditions. The use of duration as a cue, on the other hand, was not affected by experience. All participants assigned similar weight to duration regardless of their L1.

These results are in line with work that compared French-English bilinguals' prosodic cue weighting during the production of English sentences [4], which had found that bilinguals showed difficulty adapting their F0 production to the specifics of their L2, whereas they showed no difficulty adapting their duration production—in fact, some participants even overcompensated for their inability to adapt F0 production by relying more heavily on duration as a cue. It appears that, at least for French-L1 learners of

English, duration is an easier prosodic cue to adapt than F0.

Interestingly, the effect of F0 on stress perception depended not only on language background, but also in part on language proficiency. The more proficient the bilinguals were in English, the more they used F0 to perceive lexical stress in nouns. During verb perception, there was only a main effect of proficiency rather than a proficiency-by-condition interaction as had been found for noun perception. This pattern could have arisen for a couple of reasons. The F0 values for English verbs (F0 rise on the second syllable) are more similar to the values that would be found for French words, so perhaps this made it more complicated for French listeners to use F0 in an English-like way in the verb condition, regardless of proficiency. It could also be that, because the lexically stressed syllable was less marked in the verb condition than in the noun condition (i.e., the noun stimuli had greater F0 and duration differences across syllables compared to the verbs), proficiency provided less of an advantage when recognizing the more subtle stress cues in verbs.

Contrary to the results for F0, the effect of duration on lexical stress perception did not depend on language proficiency or language dominance. This is again in line with the previous finding that duration is an easier prosodic cue to adapt to one's L2 [4].

Our preliminary results also suggest that French-English bilinguals are capable of perceiving lexical stress, even if they are native or proficient speakers of French (which does not use lexical stress). Earlier work had suggested that French speakers have great difficulty with lexical stress perception (e.g., [9]), but more recent work corroborates our finding that French-L1 bilinguals are capable of showing similar lexical stress perception to native listeners [15].

Overall, we found that even though English and French are languages involving different weighting and different use of the same acoustic cues, French-English bilinguals are able to adapt their prosodic cue use to the language being perceived (in this case, English). For example, French-L1 speakers—who would normally use F0 as a cue to phrase boundaries—were generally able to use F0 as a cue to English lexical stress, albeit to a lesser extent than native listeners. This ability to adapt cue use appears to develop to some extent with increased L2 proficiency, in line with other findings [8].

By manipulating two acoustic parameters and studying a group of bilinguals with varied language experience, we helped to elucidate the acoustic and experiential factors involved in stress perception. Future work could further investigate the extent to which language background affects the use and adaptation of various prosodic cues from L1 to L2.

7. REFERENCES

[1] Morton, J., Jassem, W. 1965. Acoustic Correlates of Stress. *Lang. Speech.* 8, 159–181.

[2] Jun, S.-A., Fougeron, C. 2002. Realizations of accentual phrase in French intonation. *Probus* 14, 147-172.

[3] Cutler, A., Norris, D. 1988. The role of strong syllables in segmentation for lexical access. *J. Exper. Psych.: Human Percep. Perform.* 14, 113-121.

[4] Gilbert, A. C., Wolpert, M., Saito, H., Kousaie, S., Itzhak, I., Baum, S. R. 2019. Adaptive and selective production of syllable duration and fundamental frequency as word segmentation cues by French-English bilinguals. *J. Acoust. Soc. Am.* 146, 4255-4272.

[5] Gilbert, A. C. et al. 2021. Spoken Word Segmentation in First and Second Language: When ERP and Behavioral Measures Diverge. *Front. Psychol.* 12, 3974.

[6] Tremblay, A. 2008. Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Appl. Psycholinguist.* 29, 553-584.

[7] Tremblay, A. et al. 2017. Experience with a second language affects the use of fundamental frequency in speech segmentation. *PLOS ONE* 12, e0181709.

[8] Tremblay, A., Broersma, M., Coughlin, C. E. 2018. The functional weight of a prosodic cue in the native language predicts the learning of speech segmentation in a second language. *Bilin.: Lang. Cogn.* 21, 640-652.

[9] Dupoux, E., Pallier, C., Sebastian, N., Mehler, J. 1997. A Destressing "Deafness" in French? *J. Mem. Lang.* 36, 406-421.

[10] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5, 341-345.

[11] Anonymous. Submitted. Phrase parsing in a second language as indexed by the Closure Positive Shift: The impact of language experience and acoustic cue salience.

[12] R Development Core Team. 2010. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

[13] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-6.

[14] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2015. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-29.

[15] Gilbert, A. C., Honda, C. T., Phillips, N. A., Baum, S. R. 2021. Near native-like stress pattern perception in English-French bilinguals as indexed by the mismatch negativity. *Brain. Lang.* 213, 104892.