Finding the Key in Kiwi During L2 Spoken Production:

Low Proficiency Speakers Sound More Native-Like if they Live in Mixed-Language Environments.

Annie C. Gilbert[1,2], Jason Gullifer[2,3], Shanna Kousaie[2,4], Max Wolpert[2,5], Debra Titone[2,3] & Shari R. Baum[1,2]

[1]School of Communication Sciences and Disorders, McGill University, Montreal, Canada

[2]Centre for Research on Brain, Language, and Music, Montreal, Canada

[3]Department of Psychology, McGill University, Montreal, Canada

[4]Montreal Neurological Institute and Hospital, McGill University, Montreal, Canada

[5]Integrated Program in Neuroscience, Montreal Neurological Institute, McGill University,

Author Note

Annie C. Gilbert https://orcid.org/0000-0001-5179-1453

Jason Gullifer is now at Computer Science, Marianopolis College, Montreal, Canada. https://orcid.org/0000-0003-2959-9072

Shanna Kousaie is now at School of Psychology, University of Ottawa, Ontario, Canada. https://orcid.org/0000-0002-8218-2746

Max Wolpert https://orcid.org/0000-0001-5312-710X

Debra Titone is https://orcid.org/0000-0001-9060-9896

Shari R. Baum https://orcid.org/0000-0002-0141-0760

Dr. Debra Titone is chief editor of Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale.

Correspondence concerning this article should be addressed to Annie C. Gilbert, School of Communication Sciences and Disorders, 2001 McGill College Avenue, 8th floor, Montreal, QC, H3A 1G1 Canada. Email: annie.c.gilbert@mail.mcgill.ca

Mixed-language environment increases L2 nativeness.

Abstract

The goal of this paper was to determine if L2 speakers benefit from living in mixed-language environments, and whether said benefit applies across proficiency levels. To this end, we reanalyzed a subset of data from Gilbert et al. (2019) considering language entropy scores as a proxy for linguistic environment predictability. The task involved producing sentences designed around oronyms in French and English. Participants produced sentences in both languages, allowing the comparison of L1 and L2 productions. Their results demonstrated the production of L2-appropriate prosodic cues, albeit after having reached a high level of L2 proficiency. Adding language entropy scores to the original statistical models revealed significant interactions suggesting that participants benefited from living in mixed-languages environment whereby even low-proficiency speakers produced L2-appropriate prosodic cues. However, low-proficiency L2 speakers living in predictable linguistic environments failed to adapt their prosodic production to their L2, as previously observed. These results suggest that, irrespective of proficiency, the language environment has a significant impact on non-native language production. This has implications for language development and models of language acquisition.

Public Significance Statement

The present study suggests that unpredictable linguistic environments might be beneficial in terms of the development of native-like prosody in a second language. The results show that low proficiency second language speakers living in unpredictable linguistic environments sound more native-like in their second language than low proficiency second language speakers who live in predictable language environments.

Keywords: Prosody, Speech Production, Speech Segmentation, Bilingualism

Mixed-language environment increases L2 nativeness.

Finding the Key in Kiwi During L2 Spoken Production: Low Proficiency Speakers Sound More Native-Like if they Live in Mixed-Language Environments.

When learning a second language (L2), learners not only have to learn the sounds, words and structure of the new language, they also need to learn and integrate a new word segmentation strategy. Understanding the factors that favour or hinder the mastery of L2 word segmentation cues and strategies is crucial for people who have to interact in their L2 in potentially high-stake situations (e.g., Rothermich et al., 2023). For example, imagine an L2 speaker producing a sentence beginning with "Have you seen the [ki] [wi]…". Upon hearing this, a first language (L1) listener might erroneously conclude that the speaker is searching for a key last used by a group of people. However, when the rest of the sentence arrives, "in his lunch box?", the L1 listener will then realize that the speaker was actually referring to a fruit ("Have you seen the *kiwi* in his lunch box?") and must then revise their initial interpretation. Of relevance here, this kind of communication challenge often arises when L2 speakers fail to produce appropriate word segmentation cues, thereby leading the L1 listener to treat the two syllables of [kiwi] as two separate words instead of one bisyllabic word. Such ambiguities often occur when a speaker's different languages rely on different word segmentation cues and strategies as is the case when their L1 and L2 belong to different prosodic typology (stress-timed vs. syllable-timed), as will be discussed below.

Previous work on word segmentation has demonstrated that listeners of different languages rely on different prosodic cues to signal word boundaries. For instance, listeners whose first language is English (English-L1) tend to rely on lexical stress (marked by rising F0, longer duration, higher intensity, no vowel centralization; Beckman, 1986; Fry, 1955; Grimson, 1980; Lehiste, 1976; Lieberman, 1960; Mousikou et al., 2024) to locate word *onsets*

Mixed-language environment increases L2 nativeness.

(Cutler et al., 1997; Cutler & Otake, 2002; Jusczyk, 1999; Jusczyk et al., 1999; Mattys et al., 1999) whereas listeners whose first language is French (French-L1) tend to rely on phrase final phenomena (syllabic lengthening with optional F0 rise; Cutler et al., 1986; Mehler et al., 1981) 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). Thus, English speakers learning to speak French (and vice versa) must learn to rely on different prosodic cues to segment words (lexical stress vs phrase-final lengthening) and must also learn to rely on different segmentation strategies altogether (identifying onsets vs offsets). In the previous example, a French-L1 English-L2 speaker might have issues producing lexical stress on the initial syllable of “kiwi” (i.e., producing a longer initial syllable with higher intensity and higher F0), which would interfere with the perception of a word onset by English listeners. Furthermore, they might also not be able to refrain from producing a French-like phrase ending on the second syllable of the word (i.e., producing a longer syllable with a higher F0, two cues usually associated with lexical stress in English), which would hinder the English-L1 listener’s segmentation of the syllable string (perceiving “key we” instead of “kiwi”).

Many studies have investigated how L2 learners adapt their *perception* of word segmentation cues to the specifics of their L2 (Cutler et al., 1986, 1992; Cutler & Otake, 1994; Cutler & Pasveer, 2006; Cutler et al., 2006; Otake et al., 1993; Tremblay et al., 2018; Tremblay et al., 2016; Tremblay et al., 2012; Tremblay et al., 2017; Weber & Cutler, 2006). Generally, these studies show that L2 learners can learn to rely on L2-specific word segmentation cues, albeit imperfectly (see Gilbert et al., 2021, for a discussion). Interestingly, despite the important role of prosody in L2 intelligibility in general, few studies have investigated L2 prosody *production* (Gilbert et al., 2019; Guion et al., 2004; Kainada &

Lengeris, 2015; O'Brien et al., 2014; Rasier & Hiligsmann, 2007; Shen, 1990; Trofimovich & Baker, 2006, 2007), let alone the role of L2 prosody in word segmentation. Among these few studies, Trofimovich and Baker (2006) investigated the production of English lexical stress (a word onset cue) by Korean-L1 speakers with duration of residence in the US ranging from 3 months to 10 years. . Their results showed that English stress timing was more native-like among speakers with more cumulative exposure to English (i.e., longer residence in the US), demonstrating that Korean speakers were able to produce English-specific word segmentation cues in a native-like manner given sufficient cumulative exposure (Trofimovich & Baker, 2006).

Kim (2019) further demonstrated the importance of language experience on prosodic production among Spanish heritage speakers (speakers who grew up in Spanish-speaking families but are now dominant in English). They reported that Spanish heritage speakers' production of Spanish lexical stress more closely resembled that of English-L1 learners of Spanish than that of monolingual Spanish speakers regardless of the fact that the heritage speakers had learned Spanish from birth. Interestingly, their perception of Spanish lexical stress remained native-like, suggesting that prosody production is more sensitive to experience than prosody perception. Further investigating the impact of language experience on prosody production among bilinguals, Gilbert et al. (2019) investigated the production of prosodic word segmentation cues within French-English bilinguals. Their results suggested that participants' relative language dominance (based on their relative proficiency in English versus French) was a better predictor of native-like prosody production across languages than participants' L1. These results concord with findings reported by Trofimovich and Baker (2006) in suggesting that L2 speakers are able to learn to produce L2-appropriate word segmentation cues if they reach a sufficient level of L2 proficiency. Furthermore, in general, the participants from Gilbert et al. (2019) produced more native-like prosody during English trials, which was attributed to the fact that a majority of speakers were dominant in English, irrespective of their L1. These results also cohere with those from Kim (2019), showing the significant impact of language experience even on languages learned from birth, further supporting the impact of proficiency and language dominance on (prosodic) word segmentation cue production.

Conversely, a later study by Gilbert et al. (2020) suggested that increasing L2 exposure might temporarily hinder L2 production. In their study on speech planning, Gilbert et al. (2020) found that current L2 exposure levels had a different impact on participants whether they had been experiencing said exposure levels for a shorter or longer period of time. Their participants either had experienced a *recent increase* in L2 exposure (i.e., started attending university in their L2 after having done their schooling mostly in their L1) or not (i.e., were already attending school in their L2 prior to starting university in their L2). During this task, participants simply had to read short equations composed of two numbers of varying length (in syllables) and a mathematical symbol (e.g., 3 + 18) in their L2. Participants did not have to solve the equation and only the speech onset delay was analyzed. The results suggested that participants who had been living in high L2 exposure environments for an extended period of time (i.e., participants who did not experience a recent increase in L2 exposure but had been living in high L2 exposure environments for a few years already) used a shorter scope of planning in their L2, initiating speech after having planned only the first number of the equation, in effect finishing the planning of the second number during the production of the first number.

Mixed-language environment increases L2 nativeness.

On the other hand, participants who experienced a recent increase in L2 exposure (participants who had been living in high L2 exposure environments only for a short period of time) seemed to use a longer scope of planning, strategically considering the characteristics of the second number before initiating speech. Taken together, these results suggest that even though increased L2 exposure is usually associated with better L2 performance, sudden increases in L2 exposure might momentarily reduce speakers' confidence in their L2 abilities or increase cross-language interference, temporarily hindering L2 production abilities until speakers successfully adapt to their new L2 exposure level (see Gilbert et al., 2020, for a more detailed discussion). Such results further suggest that the linguistic environment in which L2 learners live can have a measurable impact on bilingual speakers' speech production abilities, perhaps even modulating the effects of language proficiency, as predicted by models of bilingual language use in socioecological contexts (e.g., Green & Abutalebi, 2013; Titone & Tiv, 2023).

However, questions remain as to the impact of the specifics of L2 speakers' daily language experience on their ability to produce native-like L2-specific word segmentation cues (e.g., the *kiwi/key we* confusion presented above). Namely, would a highly dynamic language environment support (through increased practice with language control) or hinder (through increased interference) the production of native-like L2 word segmentation cues? Thus, in this paper, we further investigated the relationship between the linguistic environment and language dominance through a reanalysis of previously published data (Gilbert et al., 2019), in which relative language dominance (based on relative verbal fluency scores) was found to successfully predict the production of prosodic cues by bilingual speakers. By adding variables relating to the predictability of participants' linguistic

Mixed-language environment increases L2 nativeness.

environment to the original statistical models, we can determine if the linguistic environment has an impact on language production over and above language dominance effects.

In Gilbert et al. (2019), 62 bilingual English-French participants produced English and French sentence pairs designed around oronyms that could represent either one bisyllabic word or two monosyllabic words, which allowed the direct comparison of the same sound sequence produced as one single word (no word boundary between syllables) or as two monosyllabic words (with a word boundary between syllables). In a nutshell, the results showed that L2 speakers were able to produce L2-specific word segmentation cues in a native-like manner, but only if their L2 proficiency was high enough to have an impact on their relative language dominance (having become more dominant in their L2 than in their L1). The results also suggested that it is harder to adapt one's production of F0 than syllable duration, to the point that some English-dominant participants even showed signs of over-relying on duration to signal the difference between conditions in French trials, maybe in an attempt to compensate for their inability to adapt their F0 production (see Gilbert et al., 2019, for a complementary discussion of the effects of age of first exposure to L2 and L2 proficiency).

The goal of the present study is to further characterize the factors that influence L2 prosody production by investigating the effect of the predictability of language exposureover and above the effects of proficiency or relative language dominance observed previously by Gilbert et al. (2019). The predictability of a speaker's linguistic environment, i.e., their language entropy, refers to how certain a speaker can be that they will be exposed to a specific language in a given context. A speaker with low linguistic entropy (highly predictable linguistic environment) usually knows what language to expect in any specific context (for

Mixed-language environment increases L2 nativeness.

instance, they might expect only English at home and with friends, and only French at school and at work). On the other hand, a speaker with high levels of linguistic entropy (unpredictable linguistic environment) does not know what language to expect in any specific context, with both languages being equally likely to be used. The observation of language entropy effects would contribute to our understanding of language exposure effects by taking into account not only raw L2 exposure levels but the contexts in which speakers encounter their languages. For instance, two speakers could share a balanced language exposure (50% time spent in L1, 50% in L2), but have different language entropy profiles (knowing exactly what language to expect in different contexts vs using both languages across contexts). We expect that the predictability of the linguistic environment will have a more limited impact on higher proficiency/exposure speakers (speakers with native or native-like proficiency) compared to lower proficiency/exposure speakers (i.e., English dominant speakers in the French condition, French dominant speakers in the English condition). Specifically, we expect unpredictable or more varied linguistic environment (higher entropy) to provide a great training ground for language control, thus favouring the development of native-like proficiency in L2 speakers. Therefore, we expect lower proficiency L2 speakers living in more unpredictable linguistic environments to produce more native-like L2 word segmentation cues than low proficiency L2 speakers living in more compartmentalized linguistic environments. High proficiency speakers, on the other hand, would likely not be significantly affected by the predictability of their language exposure since they have already achieved native or native-like proficiency.

Method

Participants

Mixed-language environment increases L2 nativeness.

Data from a subset of 55 participants from Gilbert et al. (2019; out of 62, 39 females, aged from 18 to 36 years old) were selected for reanalysis on the basis of having provided sufficient details in a language history questionnaire adapted from the Language History Questionnaire (LHQ) 2.0 (Li et al., 2013) to allow the computation of language entropy indices (Gullifer & Titone, 2018). Nineteen of the selected participants reported French as their L1, 17 reported English as their L1 and 19 reported having been exposed to both languages from birth (simultaneous bilinguals). Participants did not report any perceptual, speech, or learning impairments, and spoke North American varieties of English and French. Written informed consent was obtained from every participant, and the research protocol was approved by the Institutional Review Board of McGill University's Faculty of Medicine and Health Sciences.

Computation of the relative language dominance index

As reported in Gilbert et al. (2019), a relative language dominance index was computed based on verbal fluency tasks performed in English and in French. During this task, participants had one minute to name as many items as possible belonging to a specific category (semantic or sharing the same onset letter). Sequential bilingual participants performed the task in their L1 first while simultaneous bilinguals first performed the task in the language in which they felt most comfortable. The relative language dominance index was computed by dividing the total number of English words produced by the total number of French words produced during the entire task. In the present subset, 43 participants had a relative language dominance index above one, indicating that they performed better in English than in French, while 12 had a relative language dominance score below one, indicating that they performed better in French than in English (Birdsong, 2015; Treffers-Daller & Korybski,

Mixed-language environment increases L2 nativeness.

2015). Table 1 summarizes age of acquisition, as well as language proficiency, exposure and entropy variables for the 55 participants included in the present reanalysis as a function of their L1.

Table 1. *Age of L2 acquisition and language experience measures of participants.*

	French L1				Simultaneous				English L1			
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
Age of firs L2 exposure	6.94	2.76	3.5	11	0.41	0.91	0	3	6.03	2.81	4	15
Verbal fluency (total)												
French	48.76	12.64	25	70	49.00	10.47	26	69	34.88	8.16	21	51
English	48.29	13.27	28	69	60.23	14.91	27	103	59.19	9.22	39	74
Rel. language dominance*	1.02	0.30	0.56	1.82	1.26	0.38	0.70	2.65	1.77	0.41	1.08	2.33
Proportion daily conversations												
French	64.69	20.16	40	90	42.08	17.54	10	70	21.13	15.49	2	50
English	34.85	20.29	10	60	57.92	17.54	30	90	78.73	15.47	50	98
Entropy scores												
Internal	0.11	0.90	-1.53	1.21	0.46	0.67	-1.64	1.26	-0.57	1.03	-2.04	0.94
Work related	0.00	0.75	-0.99	1.51	0.37	0.89	-0.96	2.08	-0.49	0.66	-1.09	1.15
Media related	0.37	0.93	-1.01	1.61	-0.06	0.92	-1.23	1.30	-0.36	0.94	-1.31	1.31

Note: * Relative language dominance index = total number of English words produced during verbal fluency task divided by total number of French words produced.

Preliminary analyses revealed a significant correlation between the relative language dominance index used in Gilbert et al. (2019) and raw language exposure measures (English exposure: Spearman's rho = 0.6702, $p < 0.00001$; French exposure: Spearman's rho = -0.6703, $p < 0.00001$). Namely, increased exposure to English was associated with increased proficiency in English vs French (leading to relative language dominance indices towards the "English-dominant" end of the spectrum), while increased relative exposure to French was associated with increased proficiency in French vs English (leading to relative language dominance indices towards the "French-dominant" end of the spectrum). Therefore, raw measures of language exposure (English or French) will not be included in the present models to avoid potential collinearity issues. On the other hand, entropy measures were only

Mixed-language environment increases L2 nativeness.

marginally correlated with the relative language dominance index (Spearman's rho between -0.2378 and -0.3659), opening the door to different contributions from proficiency-based and entropy-based variables.

Computation of the language entropy score

Language entropy represents a way to quantify the uncertainty or variability of one's language environment. Scores range from predictable or compartmentalized (one language per context; low entropy score) to unpredictable or integrated linguistic environments (any language in any context; high entropy score). To compute language entropy scores, participants' language exposure was first estimated based on their answers to 16 language usage questions included in the above-mentioned language history questionnaire. The questions probed both passive exposure (reading, listening to the radio/watching TV) and active use of each language (conversations, doing arithmetic in one's head; See Supplemental Materials for a complete list of topics probed). Participants' responses were reformatted to percentages per language and used to compute language entropy scores for each variable and each participant using the Entropy package (Gullifer & Titone, 2018; see Gullifer & Titone, 2020, for a detailed description of data preparation procedures that are required for language entropy calculation and the specifics of the entropy computation itself.)

The entropy scores for each variable were then summarized using a Principal Component Analysis (PCA) to limit the number of variables to evaluate statistically while maintaining the maximum amount of information from the language usage questionnaires. Of note, this process did not *remove* variables from the analysis but can be conceived as creating groups of variables that pattern together to predict the variation between participants. Three components or variable groupings accounted for 43% of the variation in the language usage

Mixed-language environment increases L2 nativeness.

data. The first component had loadings from internal language usage factors like "talking to oneself", "remembering numbers", or "dreaming" and accounted for 24% of the variation in the language usage data. We therefore labelled this variable the "internal" language entropy score. The second component had loadings from factors related to language experiences where the language is imposed on the participant like "writing a paper" and "reading for work" and accounted for 10% of the variation in the language usage data. In these contexts, the speaker is likely not free to choose which language to use, hence the label of "external" language entropy. Finally, the third component had loadings from factors exclusively related to media usage ("listening to the radio / watching tv") and accounted for 9% of the variation in the language usage data and. We therefore labelled this variable as the "media" entropy score.

Summary of stimuli characteristics, response recording and acoustic analysis

As described in Gilbert et al. (2019), a total of 80 sentence pairs (40 per language) were created around oronyms, i.e., sequences of syllables which vary in meaning depending on their segmentation (e.g., in English, [kiwi] can be interpreted as "kiwi" or "key we"; in French, [ɔʁlɔʒ] can be interpreted as "horloge" - Eng. clock - or "or loge" - Eng. gold is housed at). Such sentences allowed the production of the same sound sequences produced as a single word (no word boundary between syllables), or as two words (with a word boundary between syllables). Within each pair, sentences were identical until the oronym, and sentence continuation provided semantic context to fit the intended interpretation (Eng.: "If you would like a kiwi I will buy one tomorrow." vs "If you would like a key we can duplicate one" 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 housed at the hotel; see Gilbert et al.

2019, for further details.) Given the constraints on the possible combinations of monosyllabic and bisyllabic words that can be included in the same sentence context, psycholinguistic factors like word frequency and the size of phonological neighbourhoods could not be controlled.

Full details regarding the recording procedures can be found in Gilbert et al. (2019). Briefly, the sentences were presented visually in separate language blocks and participants first performed the task in their L1 and then in their L2. Participants were instructed to read the sentences aloud as they appeared onscreen, as if talking to a friend, and to maintain a natural prosody. Recorded trials were included in the analyses only if they were produced without disfluencies occurring before the oronyms and without speech errors directly affecting the production of the oronyms. Trials with pauses between the two syllables of the oronyms were also removed to focus on more subtle prosodic word segmentation cues. Moreover, 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 language production task).

Recorded sentences were manually annotated by a trained bilingual coder using Praat (version 5.4.19, Boersma, 2001). For each syllable of the oronym, mean F0 (based on an autocorrelation method) and syllable duration were extracted using a custom script developed on site. F0 and duration measurements for each syllable of the oronyms were then combined into one numerical score per acoustic variable by comparing the second syllable (S2) of the oronym to the first syllable (S1; S2 divided by S1 for both F0 and duration). A duration ratio above one signifies that the second syllable of the oronym was longer than its first syllable. Conversely, a duration ratio below one means that the second syllable of the oronym was

Mixed-language environment increases L2 nativeness.

shorter than its first syllable (see Kim, 2019, for a similar use of relative prosodic measures in a production task.)

Statistical analyses

As mentioned above, the present reanalysis focuses on the effect of the predictability of language exposure (as indexed by language entropy) without including raw language data exposure (see ***Participants*** section for justification). Results were analyzed using linear mixed effects (LME) models evaluating F0 and duration ratios in English and French trials separately. The models were designed to determine the effects of condition (two-level deviation or sum-coded categorical variable: one bisyllabic-word – *kiwi* vs two monosyllabic words – *key we*) and how individual differences in relative language dominance (log-transformed) and predictability of language exposure (language entropy scores) modulated any condition effects. The impact of each of the three language entropy variables (internal, external and media entropy) was tested separately, with one variable being included as an interaction factor with relative language dominance and condition while the other two were included as control variables (not interacting with other variables). To account for the multiple comparisons being applied to the same dataset, alpha levels were adjusted using Bonferroni corrections, where p-values were considered significant only below 0.017 (0.05 divided by the number of comparisons per dataset, in this case, 3; Abdi, 2007; Bonferroni, 1936).

Importantly, given that the goal of present analyses is to determine if individual differences have an impact on how speakers use F0 and duration ratios to differentiate the two conditions (with or without a word boundary), only significant interactions between individual differences and condition will be reported in the results section, as well as the main effects of conditions. Complete model outputs can be found in Supplemental Materials. As in Gilbert et al. (2019), the random structure of the models took into account participants (intercept only) and sentence pairs produced (intercept and slope adjustments for the effect of condition; i.e., [F0 ratio ~ Condition * log(Relative language dominance index) * Internal entropy + External entropy + Media entropy + (1 | Participant) + (1 + Condition | Sentence Pair)]).

When necessary, follow-up models were fitted to support the interpretation of significant higher-order interactions. In these models, participants were split in two groups as a function of a specific variable (median split) to see if both subgroups showed the same effects. These models are presented in detail in the **Results section**. Simpler models investigating the impact of relative language dominance by itself were also fitted to ensure that the original pattern of results observed in Gilbert et al. (2019) was maintained within the present subsample (not reported here, see Supplemental Materials for statistical output tables). All models yielded comparable patterns of effects as those previously published (Gilbert et al., 2019), which justified the comparison of the present reanalysis to the originally published results.

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, alpha level set at $p < 0.05$). Estimates of model fit were obtained using the MuMIn package version 1.43.17 (Barton, 2020, providing R squared scores) and the Stats package version 4.0.3 (Sakamoto et al., 1986, providing AIC and BIC scores). Plots were generated using ggplot2 (version 2.1.0, Wickham, 2009).

Results

Mixed-language environment increases L2 nativeness.

Production of French trials

Figure 1 summarizes F0 and duration ratios produced during French trials as a function of condition, relative language dominance and language entropy scores. Visual inspection of the Figure suggests that the production of F0 and duration ratios during French trials was modulated by internal language entropy over and above the previously published effect of relative language dominance, but not by the other entropy scores. In other words, relative language dominance predicted different condition effects for participants at opposite ends of the internal entropy spectrum, while the impact of relative language dominance seems to remain constant across the other entropy scores. Furthermore, relative language dominance seems to have had opposite effects on the production of F0 and duration ratios, particularly within participants with the lowest internal entropy scores. Within these participants, a higher relative language dominance score (suggesting English dominance) seems to have led to greater duration ratio differences across condition but smaller F0 ratio differences across conditions.

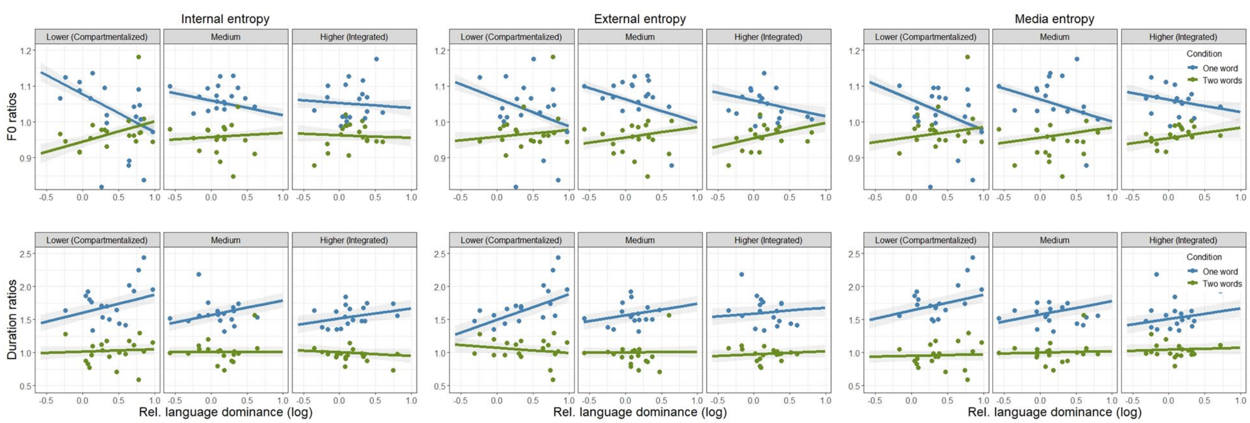

Figure 1. *Mean F0 (upper row) and duration (lower row) ratios produced during the one word (blue) and two words (green) condition of French trials as a function of relative*

Mixed-language environment increases L2 nativeness.

language dominance and three entropy scores (data from each entropy score split into three bins – panels). Points = participant averages, regression lines = effects extracted from LME models, with shaded areas representing the standard error of the mean. Note that the y-scales vary across variables and across Figures 1 and 2.

Use of fundamental frequency (F0) in French trials

Internal Entropy. As expected from Figure 1, the statistical model investigating the role of condition, relative language dominance and internal entropy score yielded a significant higher order interaction [$b = -0.0741$, $SE = 0.0140$, $t = -5.303$, $p < 0.0001$], the interpretation of which required additional models (see hereafter). The lower-level interaction between internal entropy and condition (without relative language dominance) also reached significance [$b = 0.0203$, $SE = 0.0060$, $t = 3.412$, $p = 0.0007$]. Of note, the significant main effect of condition [$b = -0.1086$, $SE = 0.0100$, $t = -10.818$, $p < 0.0001$] and the significant interaction between condition and relative language dominance (as observed in Gilbert et al., 2019) were maintained [$b = 0.0763$, $SE = 0.0136$, $t = 5.617$, $p < 0.0001$]. See Supplemental Materials for detailed model outputs.

To guide the interpretation of the higher-order interaction, participants were split into two groups based on their internal entropy scores (median split). Follow-up models were then fitted to investigate the effect of relative language dominance on condition within each participant subgroup separately. The follow-up models revealed a significant interaction between condition and relative language dominance within participants with lower internal language entropy (more predictable linguistic environment; [$b = 0.1370$, $SE = 0.0151$, $t = 9.075$, $p < 0.0001$]), where participants at the French dominant end of the spectrum produced more French-like F0 patterns (larger difference between condition) than participants at the

English dominant end of the spectrum (little to no difference between conditions, as previously observed in Gilbert et al. 2019). Interestingly, the follow-up model fitted on higher internal language entropy participants (less predictable language usage) yielded no significant interaction between condition and relative language dominance [$b = -0.0251$, $SE = 0.0247$, $t = -1.019$, $p = 0.308$]. That is, within this group, all participants produced different (French-like) F0 ratios across conditions regardless of their language dominance.

External and Media Entropy. As expected from visual inspection of Figure 1, the statistical models investigating the role of condition, relative language dominance and either external or media entropy scores failed to reveal significant higher order interactions (external entropy: [$b = -0.00057$, $SE = 0.0181$ $t = -0.314$, $p = 0.7530$], media entropy: [$b = -0.0222$, $SE = 0.0142$, $t = -1.562$, $p = 0.1180$]). In both cases, only the significant main effect of condition (external entropy: [$b = -0.1062$, $SE = 0.0101$, $t = -10.547$, $p < 0.0001$], media entropy: [$b = -0.1055$, $SE = 0.0101$, $t = -10.414$, $p < 0.0001$]) and the significant interaction between condition and relative language dominance were maintained (external entropy: [$b = 0.0928$, $SE = 0.0135$, $t = 6.865$, $p < 0.0001$], media entropy: [$b = 0.0897$, $SE = 0.0133$, $t = 6.771$, $p < 0.0001$]). See Supplemental Materials for detailed model outputs.

Use of duration in French trials

Internal Entropy. The statistical model investigating the role of condition, relative language dominance and the internal entropy scores on the production of syllable duration ratios revealed a higher-order interaction involving internal entropy [$b = 0.1995$, $SE = 0.0588$, $t = 3.390$, $p = 0.0007$], the interpretation of which required additional models (see hereafter). As observed in the F0 analyses, the model also revealed a lower order interaction between condition and internal entropy (without relative language dominance; [$b = -0.0915$, $SE =$

Mixed-language environment increases L2 nativeness.

0.0251, $t = -3.647$, $p = 0.0003$]), as well as the previously observed significant main effect of condition [$b = -0.5310$, $SE = 0.0643$, $t = -8.265$, $p < 0.0001$] and the significant interaction between condition and relative language dominance [$b = -0.2310$, $SE = 0.0573$, $t = -4.033$, $p < 0.0001$]. See Supplemental Materials for detailed model outputs.

To guide the interpretation of the higher-order interaction, follow-up models were fitted after splitting participants as a function of their internal language entropy (median split). As observed with F0 ratios, these models revealed a significant interaction between condition and relative language dominance within participants with lower internal language entropy [$b = -0.3362$, $SE = 0.0644$, $t = -5.225$, $p < 0.0001$], but not within participants with higher internal language entropy [$b = -0.0774$, $SE = 0.1019$, $t = -0.760$, $p = 0.4480$]. That is, while relative language dominance failed to predict the production of duration ratios across conditions within higher entropy participants (integrated linguistic environment; steady difference between conditions regardless of speakers' relative language dominance), in lower entropy participants (living in more predictable or compartmentalized linguistic environments), it predicted *smaller* differences between conditions for speakers at the French dominant end of the spectrum and *larger* differences between conditions for speakers at the English dominant end of the spectrum (participants with lower French proficiency; see Supplemental Materials for detailed model outputs).

External Entropy. The statistical model investigating the role of condition, relative language dominance and the external entropy scores on the production of syllable duration ratios failed to reveal any interactions involving external entropy and condition [$|b| < 0.0454$, $|SE| > 0.0266$, $|t| < 1.707$, $|p| > 0.0880$]. Only the previously observed main effect of condition [$b = -0.5556$, $SE = 0.0644$, $t = -8.626$, $p < 0.0001$] and the interaction between condition and

Mixed-language environment increases L2 nativeness.

relative language dominance reached significance [$b = -0.2213$, $SE = 0.0567$, $t = -3.901$, $p < 0.0001$].

Media Entropy. Interestingly, while the statistical model investigating the role of condition, relative language dominance and the media entropy scores on the production of syllable duration ratios failed to revealed a higher order interaction between variables [$b = 0.0419$, $SE = 0.0591$, $t = 0.709$, $p = 0.4785$], it revealed a lower order interaction between condition and media entropy (without relative language dominance; [$b = 0.1020$, $SE = 0.0203$, $t = 5.019$, $p < 0.0001$]). This result suggests that participants who are more compartmentalized in their media consumption (lower media entropy) produced larger duration ratio differences across conditions than participants who are more integrated in their media consumption, regardless of their relative language dominance. The model also revealed a significant main effect of condition [$b = 0.5816$, $SE = 0.0634$, $t = -9.180$, $p < 0.0001$] and a significant interaction between condition and relative language dominance [$b = -0.1797$, $SE = 0.0552$, $t = -3.256$, $p = 0.0012$].

Summary of French results

In sum, with regards to internal language entropy, the production of both F0 and duration ratios in French trials could be significantly predicted by participants' relative language dominance only among participants reporting a compartmentalized internal language use (lower entropy). Participants reporting an integrated internal use of their two languages (higher entropy) behaved in a native-like fashion regardless of their relative language dominance score. As previously observed by Gilbert et al. (2019), F0 and duration analyses yielded opposite effects of relative language dominance on condition, with higher relative

Mixed-language environment increases L2 nativeness.

language dominance scores participants (more English dominant) presenting smaller F0 ratio but larger duration ratio differences between conditions.

On the other hand, media consumption entropy also interacted with condition (without relative language dominance) in predicting larger duration ratio differences across conditions for participants with more compartmentalized media consumption, regardless of their language dominance.

Production of English trials

Figure 2 summarizes F0 and duration ratios production during English trials as a function of condition, relative language dominance and language entropy. Unlike in the French trials of Figure 1, visual inspection of Figure 2 suggests that language entropy has limited effects on the observed patterns of results in English trials.

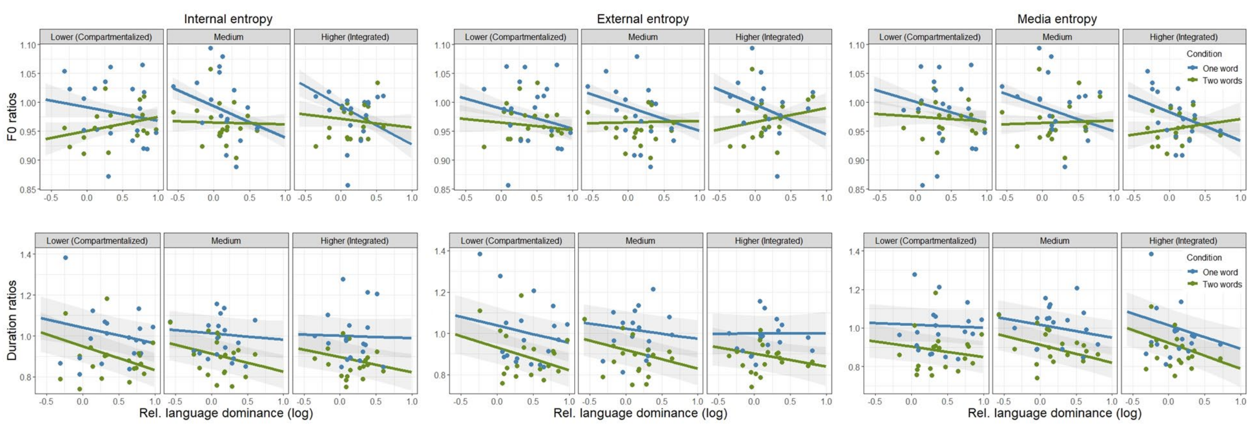

Figure 2. *Mean F0 (upper row) and duration (lower row) ratios produced during the one word (blue) and two words (green) condition of English trials as a function of relative language dominance and three entropy scores (data from each entropy score split into three bins – panels). Points = participant averages, regression lines = effects extracted from LME*

Mixed-language environment increases L2 nativeness.

models, with shaded areas representing the standard error of the mean. Note that the y-scales vary across variables and across Figures 1 and 2.

Use of fundamental frequency (F0) in English trials

As expected from visual inspection of Figure 2, the statistical models investigating the production of F0 ratios in English yielded no significant higher order interactions between condition, relative language dominance and any entropy score (internal entropy: [$b = 0.0026$, $SE = 0.0128$, $t = 0.207$, $p = 0.836$], external entropy: [$b = 0.0305$, $SE = 0.0149$, $t = 2.043$, $p = 0.0411$; above Bonferroni adjusted alpha level of 0.017], media entropy: [$b = 0.0180$, $SE = 0.0127$, $t = 1.421$, $p = 0.1553$]). Across all three models, only the previously observed main effect of condition (internal entropy: [$b = -0.0308$, $SE = 0.0063$, $t = -4.881$, $p < 0.0001$], external entropy: [$b = -0.0272$, $SE = 0.0062$, $t = -4.410$, $p < 0.0001$], media entropy: [$b = -0.0277$, $SE = 0.0061$, $t = -4.578$, $p < 0.0001$]) and interaction between condition and relative language dominance reached significance (internal entropy: [$b = 0.0500$, $SE = 0.0119$, $t = 4.202$, $p < 0.0001$], external entropy: [$b = 0.0478$, $SE = 0.0118$, $t = 4.061$, $p < 0.0001$], media entropy: [$b = 0.0465$, $SE = 0.0120$, $t = 3.871$, $p = 0.0001$]). See Supplemental Materials for detailed model outputs.

Use of duration in English trials

The statistical models investigating the role of condition, relative language dominance and the entropy scores on the production of syllable duration ratios in English trials failed to revealed any significant effects of condition either as a main effect (internal entropy model: [$b = -0.0987$, $SE = 0.0451$, $t = -2.191$, $p = 0.0340$]; external entropy model: [$b = -0.1024$, $SE = 0.0448$, $t = -2.286$, $p = 0.0275$]; media entropy model: [$b = -0.1038$, $SE = 0.0447$, $t = -2.321$, $p = 0.0254$]; all above Bonferroni adjusted alpha level of 0.017) or as part of any interaction

Mixed-language environment increases L2 nativeness.

(internal entropy model: [$|b| < 0.0525$, $|SE| > 0.0185$, $|t| < 1.466$, $|p| > 0.1428$], external entropy model: [$|b| < 0.0441$, $|SE| > 0.0183$, $|t| < 1.245$, $|p| > 0.2133$], media entropy model: [$|b| < 0.0265$, $|SE| > 0.0143$, $|t| < 0.746$, $|p| > 0.4559$]; see Supplemental Materials for detailed model outputs).

Summary of English results

In sum, entropy did not significantly interact with relative language dominance in predicting F0 or duration ratio production in English trials. Furthermore, relative language dominance successfully predicted the production of F0 ratios across conditions in English trials, but not the production of duration ratios. Such results suggest that all speakers produced similar duration ratios across condition in English, regardless of their L1, language dominance or language entropy.

Discussion

The goal of the present reanalysis was to determine if the predictability of language exposure (language entropy) affected the production of prosodic word segmentation cues among bilinguals over and above previously observed language proficiency effects (Gilbert et al., 2019), and if such effects facilitated or hindered prosodic production. The results demonstrated that increased language entropy can significantly facilitate the production of language-appropriate prosodic word segmentation cues, although these effects were not observed in all conditions. Namely, language entropy effects were observed only during French trials and different sub-components of language entropy (internal, external, media) were associated with different patterns of effects. Below, we discuss the observed patterns of results and explore their significance in terms of language dominance / proficiency effects and the different sources of variation in the linguistic environment.

Mixed-language environment increases L2 nativeness.

Language-specific effects of entropy

As mentioned above, entropy scores did not significantly interact with relative language dominance in English trials, while the previously published effects of relative language dominance were maintained in models investigating F0 ratio production. This pattern of results suggests that observed relative language dominance effects are stable enough within this set of participants to remain unchanged whether the participant lives in a compartmentalized or integrated linguistic environment.

On the other hand, significant interactions between language entropy and relative language dominance were observed during French trials. Namely, the production of F0 and duration ratios in French trials were predicted by significant interactions among relative language dominance, condition and internal entropy. Follow-up models used to guide the interpretation of the significant higher-order interactions involving internal entropy revealed that the interactions between language dominance and condition were maintained within participants with lower entropy scores (compartmentalized linguistic environments), but not within participants with higher entropy scores (integrated linguistic environments). In other words, language dominance could predict the production of F0 and duration ratios during French trials only in participants who can reliably predict what language they will need to use in different contexts (compartmentalized linguistic environments). Participants living in environments where they can encounter either language across contexts (integrated linguistic environments) produced native-like prosodic cues in French regardless of their language dominance. Thus, even English dominant (low French proficiency) participants produced native-like prosodic cues in French if they were living in unpredictable linguistic environments.

Mixed-language environment increases L2 nativeness.

Interestingly, the models also revealed that the production of duration ratios in French could be predicted by media entropy and condition regardless of relative language dominance, with more compartmentalized media consumption (lower media entropy) being associated with larger duration ratio differences (more French-like duration ratios) across conditions than a more integrated media consumption regardless of participants' relative language dominance.

On the interplay between predictability of language exposure and language dominance

When interpreting the present results in terms of relationships between language exposure and language dominance, it is important to remember that most of our participants were effectively dominant in English regardless of their reported L1. That is, most participants performed better on the English version of the verbal fluency task than they did on the French version of the same task. This dominance or proficiency imbalance in favour of English might explain why we observed significant interactions between the predictability of language exposure and language dominance during French trials, but not during English trials. That is, the present sample of participants might have been too proficient in English to reveal language entropy effects on English productions. However, their overall lower French proficiency might have allowed the observation of such language entropy effects, where the lowest proficiency French speakers benefitted from an integrated language environment, allowing them to produce more native-like word segmentation prosodic cues than proficiency-matched speakers living in linguistically compartmentalized environments. Taken together, these results confirm our hypothesis that the predictability of the linguistic environment should have a more limited impact on higher proficiency speakers.

On the different sources of variability in the linguistic environment

Mixed-language environment increases L2 nativeness.

When comparing the impact of the three entropy components included in the present analyses, it is interesting to note that they yield different patterns of effects. For instance, internal language entropy significantly interacted with relative language dominance and condition in predicting the production of word segmentation prosodic cues (F0 and duration ratios) in French. This entropy component also accounted for the greatest amount of variation in the language usage data of the three components included in the analyses (24% of the variation), loading mainly on intrinsically motivated language usage contexts like “talking to oneself”, “expressing emotions”, or “doing mental calculation”. Of note, most of these contexts do not involve overt speech production, but they nonetheless pattern together in predicting the ability to produce native-like prosodic cues. The reason *why* the predictability of language exposure in these specific contexts would predict prosody production is out of the scope of the present paper. Nonetheless, one might speculate that highly integrated language usage in these contexts may be a sign of greater linguistic flexibility, which in turn might favour the fine-tuning of prosodic production even in lower proficiency speakers.

The two other entropy components included in the present analysis accounted for a similar amount of variation in the language usage data (external and media entropy, respectively accounting for 10% and 9% of the variation) but yielded very different patterns of effects. That is, only media entropy was found to predict the production of prosodic cues within a subset of conditions in French, while external entropy did not significantly predict the production of F0 or duration ratios in any language. Of note, while both entropy scores rely on contexts that do not involve overt speech production per se, only media entropy reflects a context involving the perception of speech signals. Moreover, external entropy reflects language usage contexts where the language is imposed on participants (“writing

Mixed-language environment increases L2 nativeness.

papers" and "reading for work", both involving reading instead of speech perception) while media entropy is based mainly on "listening to the radio or watching television", which reflects linguistic choices explicitly made by the participants. Thus, we cannot determine which aspect of these two entropy components is more directly related to the observed pattern of effect (speech processing vs reading; intrinsic vs extrinsic motivation).

Taken together, the effects of all three entropy components included in the present study suggest that intrinsically motivated variability in language exposure (internal and media entropy) seems to better predict the production of language appropriate prosodic segmentation cues than imposed variability in language exposure. If considered within the context of the *Systems Framework of Bilingualism* (Titone & Tiv, 2023), these results could contribute to the description of ego-driven language dynamics by demonstrating that participants' prosodic productions are not only affected by the predictability of language exposure per se, but also by the role the participant played in creating such predictable/unpredictable linguistic environments. Moreover, on a more applied note, the present findings also suggest that L2 learners might benefit from voluntarily putting themselves in situations where language predictability is lower.

Conclusion

The objective of the present reanalysis was to further study the effect of language exposure on L2 prosody production by looking at it through the lens of language predictability as estimated by language entropy (Gullifer et al., 2018; Gullifer & Titone, 2020). The results showed that lower proficiency speakers benefitted from living in highly dynamic language environments, specifically if said dynamic environment was the result of their own linguistic choices. Higher proficiency speakers, on the other hand, were not significantly affected by the

Mixed-language environment increases L2 nativeness.

predictability of their linguistic environment, likely because they have reached a proficiency threshold making their productions immune to such modulations. These findings highlight the complexities associated with quantifying language usage in bilingual environments as well as the need to take into account the contribution of different language experience variables altogether. Nonetheless, further research would be required to determine the underlying causes of the observed effects and exactly which combination of language exposure variables (time spent and predictability of exposure) might provide the optimal context for foreign language learning.

References

Barton, K. (2020). MuMIn: Multi-Model Inference (Version R package version 1.43.17): https://CRAN.R-project.org/package=MuMIn.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. (Version R package version 1.1-6). Retrieved from http://CRAN.R-project.org/package=lme4

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

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

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345.

Mixed-language environment increases L2 nativeness.

Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics, 31*, 585-598.

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of Prosodic Boundary Correlates by Newborn Infants. *Infancy, 2*(3), 385-394. doi:10.1207/s15327078in0203_6

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and speech, 51*, 61-75.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language, 51*, 523-547.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the Comprehension of Spoken Language: A Literature Review. *Language and speech, 40*(2), 141-201. doi:10.1177/002383099704000203

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*(4), 385-400. doi:10.1016/0749-596x(86)90033-1

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24*(3), 381-410. doi:http://dx.doi.org/10.1016/0010-0285(92)90012-Q

Cutler, A., & Otake, T. (1994). Mora or Phoneme? Further Evidence for Language-Specific Listening. *Journal of Memory and Language, 33*(6), 824-844. doi:http://dx.doi.org/10.1006/jmla.1994.1039

Cutler, A., & Otake, T. (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language, 46*, 296-322.

Mixed-language environment increases L2 nativeness.

Cutler, A., & Pasveer, D. (2006). Explaining cross-linguistic difference in effects of lexical stress on spoken-word recognition. In *Proceedings of the 3rd International Conference on Speech Prosody*. Dresden, Germany.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*(2), 269-284. doi:http://dx.doi.org/10.1016/j.wocn.2005.06.002

Fry, D. B. (1955). Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal of the Acoustical Society of America, 27*(4), 765-768.

Gilbert, A. C., Lee, J. G., Coulter, K., Wolpert, M. A., Kousaie, S., Gracco, V. L., . . . Baum, S. R. (2021). Spoken Word Segmentation in First and Second Language: When ERP and Behavioral Measures Diverge. *Frontiers in Psychology, 12*, 3974. doi:10.3389/fpsyg.2021.705668

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. *The Journal of the Acoustical Society of America, 146*(6), 4255-4272. doi:10.1121/1.5134781

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*(5), 515-530. doi:10.1080/20445911.2013.796377

Grimson, A. (1980). *An introduction to the pronunciation of English*. London: Edward Arnold.

Guion, S. G., Harada, T., & Clark, J. J. (2004). Early and late Spanish–English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition, 7*(3), 207-226. doi:10.1017/S1366728904001592

Mixed-language environment increases L2 nativeness.

Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D. (2018). Bilingual experience and resting-state brain connectivity: Impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia, 117*, 123-134. doi:https://doi.org/10.1016/j.neuropsychologia.2018.04.037

Gullifer, J. W., & Titone, D. (2018). Compute language entropy with {languageEntropy} (Publication no. doi:10.5281/zenodo.1406201). https://github.com/jasongullifer/languageEntropy.

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition, 23*(2), 283-294. doi:10.1017/S1366728919000026

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in cognitive sciences, 3*(9), 323-328.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The Beginnings of Word Segmentation in English-Learning Infants. *Cognitive Psychology, 39*(3–4), 159-207. doi:http://dx.doi.org/10.1006/cogp.1999.0716

Kainada, E., & Lengeris, A. (2015). Native language influences on the production of second-language prosody. *Journal of the International Phonetic Association, 45*(3), 269-287. doi:10.1017/S0025100315000158

Kim, J. Y. (2019). Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism: Language and Cognition*, 1-18. doi:10.1017/S1366728918001220

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

Mixed-language environment increases L2 nativeness.

Lehiste, I. (1976). Influence of fundamental frequency pattern on the perception of duration. *Journal of Phonetics, 4*(2), 113-117. doi:https://doi.org/10.1016/S0095-4470(19)31231-8

Li, P., Zhang, F., Tsai, E., & Puls, B. (2013). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition, 17*(3), 673-680. doi:10.1017/S1366728913000606

Lieberman, P. (1960). Some Acoustic Correlates of Word Stress in American English. *The Journal of the Acoustical Society of America, 32*(4), 451-454. doi:10.1121/1.1908095

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38*, 465-494.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*(3), 298-305. doi:https://doi.org/10.1016/S0022-5371(81)90450-3

Mousikou, P., Strycharczuk, P., & Rastle, K. (2024). Acoustic correlates of stress in speech perception. *Journal of Memory and Language, 136*, 104509. doi:https://doi.org/10.1016/j.jml.2024.104509

O'Brien, M. G., Jackson, C. N., & Gardner, C. E. (2014). Cross-linguistic differences in prosodic cues to syntactic disambiguation in German and English. *Applied Psycholinguistics, 35*(01), 27-70. doi:doi:10.1017/S0142716412000252

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or Syllable? Speech Segmentation in Japanese. *Journal of Memory and Language, 32*(2), 258-278. doi:http://dx.doi.org/10.1006/jmla.1993.1014

Mixed-language environment increases L2 nativeness.

R Development Core Team. (2010). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Rasier, L., & Hiligsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française, 28*, 41-66.

Rietveld, A. M. C. (1980). Word boundaries in the French language. *Language and speech, 23*(3), 289-296.

Rothermich, K., Baker, R., Falkins Sharon, M., Kum, J., Capps Madison, R., & Bobb Susan, C. (2023). Reported Use of Second-Language Speech Accommodation in Everyday Interactions: The Role of Individual Differences. *Journal of Speech, Language, and Hearing Research, 66*(9), 3307-3327. doi:10.1044/2023_JSLHR-23-00167

Sakamoto, Y., Ishiguro, M., & G., K. (1986). *Akaike Information Criterion Statistics*. Dordrecht, Netherlands: D. Reidel Publishing Company.

Shen, X. S. (1990). Language by speakers of a tonal language: Chinese speakers learning French prosody. *International Review of Applied Lingusitics, 28*, 119–134.

Titone, D. A., & Tiv, M. (2023). Rethinking multilingual experience through a Systems Framework of Bilingualism. *Bilingualism: Language and Cognition, 26*(1), 1-16. doi:10.1017/S1366728921001127

Treffers-Daller, J., & Korybski, T. (2015). Using lexical diversity measures to operationalize language dominance in bilinguals. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization* (pp. 106-133). Cambridge: Cambridge University Press.

Mixed-language environment increases L2 nativeness.

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. *Bilingualism: Language and Cognition, 21*(3), 640-652. doi:10.1017/S136672891700030X

Tremblay, A., Broersma, M., Coughlin, C. E., & Choi, J. (2016). Effects of the Native Language on the Learning of Fundamental Frequency in Second-Language Speech Segmentation. *Frontiers in Psychology, 7*(985). doi:10.3389/fpsyg.2016.00985

Tremblay, A., Coughlin, C. E., Bahler, C., & Gaillard, S. (2012). Differential contribution of prosodic cues in the native and non-native segmentation of French speech. *Laboratory Phonology, 3*(2), 385-423. doi:10.1515/lp-2012-0018

Tremblay, A., Namjoshi, J., Spinelli, E., Broersma, M., Cho, T., Kim, S., . . . Connell, K. (2017). Experience with a second language affects the use of fundamental frequency in speech segmentation. *PLoS ONE, 12*(7), e0181709. doi:10.1371/journal.pone.0181709

Trofimovich, P., & Baker, W. (2006). LEARNING SECOND LANGUAGE SUPRASEGMENTALS: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech. *Studies in Second Language Acquisition, 28*(01), 1-30. doi:doi:10.1017/S0272263106060013

Trofimovich, P., & Baker, W. (2007). Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics, 28*(2), 251.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119*(1), 597-607. doi:10.1121/1.2141003

Mixed-language environment increases L2 nativeness.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.