WHAT CUE WHERE? ON THE ROLE OF PITCH AND LENGTH IN THE ACQUISITION OF L2 PROSODIC REPRESENTATIONS*

Nathan Owens and Annie Tremblay
University of Illinois at Urbana-Champaign

1. Introduction

In recent years, various studies have examined second language (L2) learners' perception of nonnative segments as a function of particular acoustic-phonetic cues manipulated in the speech signal (e.g., Gottfried and Beddor 1988; Fox, Flege, and Munro 1995). These studies have provided evidence that L2 learners rely on acoustic-phonetic cues which are either perceptually salient and/or play an important role in the native language (L1) for distinguishing between segments in the target language. Research that examines the role of acoustic-phonetic cues in the acquisition of higher-level L2 prosodic representations (e.g., syllables, feet, etc.) is, however, much less common. If acoustic-phonetic cues differ across prosodic structures, then L2 learners' extraction of these cues from the input may interact with their formation of alignment relations between the various prosodic constituents (e.g., alignment of edges, heads, etc.) represented in their interlanguage grammar (for discussion, see Carroll 2001).

We examine this hypothesis by investigating the relationship between the acoustic correlates for realizing English (primary) word stress and the prosodic structure underlying English stress placement in native speakers of Canadian French at different English proficiencies. For the sake of this study, it is assumed that the acoustic correlates of word stress that L2 learners use reflect the acoustic cues they perceive in the input, an assumption which is a reasonable

* This research was supported in part by the Department of French at the University of Illinois, the Social Sciences and Humanities Research Council of Canada (doctoral fellowship #752-2006-0178), the *Fonds québécois de la recherche sur la société et culture* (doctoral fellowship #89243), and the Elizabeth Holmes-Carr fund (Department of Second Language Studies, University of Hawai‘i). Many thanks to: Robert Bley-Vroman, Heather Goad, William O'Grady, Amy Schafer, and Bonnie D. Schwartz (dissertation committee) for their valuable feedback; Soondo Baek, Susanne Carroll, Jennifer Cole, Kamil Deen, Patricia Donegan, Zsuzsanna Fagyal, Scott Fraundorf, Angela Isaac, Scott Jackson, Tuan Lam, Eun-Kyung Lee, Ryan Shosted, Duane Watson, and the audience of the 2008 Canadian Linguistic Association Meeting for discussion; Jeffrey Steele, John Norris, Martyn Clark, and James Dean Brown for advice on proficiency testing; the Department of Linguistics at McGill University; David Faulhaber, Hunter Hatfield, Jason Price, and Stéphane Trouillet for help with the experimental set up and data analysis; Erik Hill, Mark Messer, Dustin Price, Jason Price, and Dana Wong for rating the learners' accents; and Shirley Black, Lynne Champoux-Williams, Marise Duquette, Peter Enns, Mary Gauthier, Ginette Gendron, Pascal Labbé-Thibault, Aimée Lévesque, Donna Lisney, Jennifer Mah, Kristina Maiorino, Richard Martel, Robert Myles, Irene Nasim, Heather Newell, Carolyn Samuel, Marion Steff, Georges St-Laurent, Silvia Ursini, and Joannie Vienneau for help with the recruitment of participants.

Actes du congrès annuel de l'Association canadienne de linguistique 2008.
Proceedings of the 2008 annual conference of the Canadian Linguistic Association.

basis for an initial inquiry, but which has yet to be validated by perception data. Canadian French and English represent an excellent L1-L2 pair for investigating the hypothesis we set forth, because they differ in both realization of word-level stress and prosodic structure. In this paper, we will demonstrate that there is indeed a direct relationship between the acoustic correlates of English stress that L2 learners use and the prosodic grammar governing stress placement that they acquire. This suggests that L2 learners' extraction of these cues from the input plays a determining role in their acquisition of L2 prosodic representations.

The paper is organized as follows: Section 2 provides the theoretical framework and analysis of stress that will be adopted for Canadian French and English; Section 3 describes the methodology used in the study; Section 4 presents the results; and Section 5 discusses the implications of the study and provides some concluding remarks.

2. Stress in Canadian French and English

2.1 Theoretical Assumptions

This study uses the Prosodic Hierarchy (e.g., McCarthy and Prince 1986; Nespor and Vogel 1986; Selkirk 1984), partially represented in (1), as a theoretical framework for analyzing the prosodic organization of languages. Each constituent in the hierarchy is a well-defined prosodic domain in which specific phonological processes apply. In this framework, stress is understood as the instantiation of the head of the foot, and stress placement is determined by the shape (trochaic, i.e., stressed-unstressed, or iambic, i.e., unstressed-stressed) and alignment of the foot with lower- and higher-level prosodic units.

(1)	PP	(Phonological Phrase)	*in Canada*
	|		
	PW	(Prosodic Word)	*Canada*
	|		
	Ft	(Foot)	*Cana*
	|		
	σ	(Syllable)	*Ca*
	|		
	μ	(Mora)	*Ca*

2.2 Canadian French

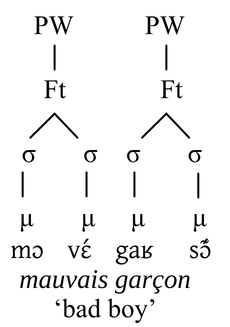
In Canadian French, the most prominent syllable is generally the last syllable not containing a schwa in the (Prosodic) Word (e.g., Walker 1984; see Paradis and Deshaies 1990 for evidence from perception data). This is shown in (2), where primary stress is represented with an acute (ˊ) accent. It has been suggested that stress may shift to the penultimate syllable if it contains a nasal vowel (i.e., /ã, ẽ, ɔ̃, œ̃/), a vowel which is inherently long (i.e., /ɛ, ø, o, ɑ/), or a syllable-final voiced fricative (i.e., /v, z, ʒ, ʁ/) (e.g., Ouellet and Thibault 1996; Paradis and Deshaies 1990). It is not clear, however, whether this shift is a purely segmental phenomenon or the instantiation of stress, because stress is realized predominantly with duration in Canadian French and rhymes containing the above segments tend to be longer than other rhymes in the language (e.g., Walker 1984).

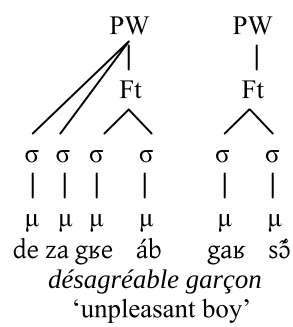
(2) merveilleux chapeau [mɛʁvɛjǿ]$_{PW}$ [ʃapó]$_{PW}$
‘wonderful hat’

Stress in Canadian French is not iterative (i.e., stressed and unstressed syllables do not alternate in long words). Nevertheless, the first or second syllable of the first Prosodic Word in a (Phonological) Phrase is aligned with a high tone (Hi) delimiting the left edge of the phrase (Jun and Fougeron 2000), as illustrated in (3). We understand this tone to be part of the intonational contour of Canadian French rather than evidence for secondary stress, because its realization is optional and its location not entirely predictable (cf. Goad and Buckley 2006).

(3) merveilleux chapeau [[mɛʁvɛjǿ]$_{PW}$ [ʃapó]$_{PW}$]$_{PP}$
‘wonderful hat’ Hi

On the basis of the patterns of word-level prominence attested in Canadian French, it has been proposed that the language has a single iambic foot, whose right edge is aligned with the right edge of the Prosodic Word (e.g., Charette 1991; Goad and Buckley 2006; cf. Montreuil 2002 and Selkirk 1978 for European French). This analysis is plausible given that prominence is word-final and predictably non-iterative in Canadian French. Additionally, it is cross-linguistically preferable as it assumes that Canadian French also has a foot. Finally, this analysis is consistent with research showing that languages with an iambic foot tend to realize stress with duration (e.g., Hayes 1995). It is therefore the analysis that will be adopted in this study. The examples in (4) illustrate the prosodic organization of words in Canadian French.

(4) a.


b.


2.3 English

In contrast to Canadian French, stress placement in English varies according to word class and syllable weight, among other factors (for discussion, see Halle and Vergnaud 1987 and Hammond 1999). In monomorphemic verbs and adjectives, the most prominent syllable is word-final if it is heavy (i.e., if it contains a long vowel or a complex coda), and otherwise penultimate. This is illustrated in (5a)–(5d) (from Halle and Vergnaud 1987:230–231). In monomorphemic nouns, on the other hand, the most prominent syllable is the penult if it is heavy (i.e., if it contains a long vowel or a simple coda), and otherwise the antepenult, as shown in (6a)–(6b) (from Halle and Vergnaud 1987:227).

(5)	a.	absurd	[əbsɜ́ɹd]$_{\text{PW}}$
	b.	direct	[dɪɹɛ́kt]$_{\text{PW}}$
	c.	solid	[sɑ́lɪd]$_{\text{PW}}$
	d.	develop	[dɪvɛ́ləp]$_{\text{PW}}$
(6)	a.	agenda	[ədʒɛ́ndə]$_{\text{PW}}$
	b.	aluminum	[əlúmɪnm̩]$_{\text{PW}}$

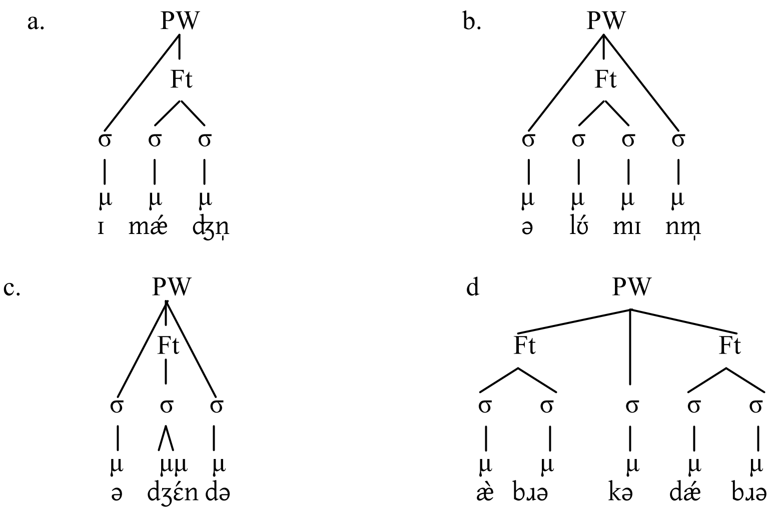
Unlike Canadian French, stress in English is iterative: in monomorphemic, even parity words with light syllables, stressed and unstressed syllables alternate, whereas in monomorphemic, odd parity words with light syllables, the initial and the penultimate syllables tend to be stressed. This is shown in, respectively, (7a) and (7b) (from Halle and Vergnaud 1987:243–245). In monomorphemic words, primary stress generally falls on the rightmost stressed syllable, and secondary stress falls on the remaining stressed syllables.

(7)	a.	onomatopoeia	[ɑ̀nomæ̀ɾopíːə]$_{\text{PW}}$
	b.	abracadabra	[æ̀bɹəkədǽbɹə]$_{\text{PW}}$

English also differs from Canadian French in its acoustic realization of word stress. For example, Lieberman (1960) reported that in the presence of a pitch accent, higher pitch is the predominant acoustic correlate of word stress in English. However, in this early study, acoustic measurements were collapsed across stress patterns, rendering difficult any comparison of the various acoustic cues of stress. In the present work, we show that the acoustic correlates of stress are, in fact, not uniform across stress patterns, with pitch being the predominant correlate of stress only in word-initial position.

In light of the above stress patterns, English has been analyzed as having a trochaic foot right-aligned in the Prosodic Word, as shown in (8a) (e.g., Hammond 1999). The distinct behavior of adjectives and verbs, on the one hand, and of nouns, on the other, has been attributed to the extraprosodicity of the final syllable in nouns (Hayes 1981): the syllable is ignored for foot construction, as illustrated in (8b). Additionally, it has been suggested that the head of the foot aligns with heavy syllables (Hammond 1999; Hayes 1995), thus explaining the sensitivity of stress placement to syllable weight, as seen in (8c). For secondary stress, it has been proposed that one foot is built at the left edge of the word, with all remaining feet being aligned with the right edge of the Prosodic Word, as shown in (8d). This is the analysis of English stress that will be adopted in this study.

(8)



2.4 The L2 Learning Problem

For speakers of Canadian French, the task of learning the generalizations underlying stress placement in English requires that they determine how lower- and higher-level prosodic constituents are aligned in English representations. This entails (among other things) (re)setting foot headedness to the left and aligning the head of the foot with heavy syllables—all the while ignoring the last syllable of nouns. Such a process is rather opaque, because words with seemingly identical stress patterns (e.g., (8a) and (8c)) can have different representations, and syllables of different shapes attract stress in different word classes (e.g., (5a)–(5b) and (6a)–(6b)).

As will be shown in the results, further complications stem from the non-uniformity of acoustic cues across stress patterns, with pitch being a reliable acoustic correlate of stress only in word-initial position. Native speakers of Canadian French must establish the correct mapping between the acoustic cues of English stress and the head of the foot before they can determine the shape of the foot and its alignment with lower- and higher-level prosodic constituents. They may have additional difficulty establishing this mapping, because the acoustic correlates of stress in Canadian French and in English differ. We will show that the L2 learners' failure to weigh the acoustic cues to English stress in a target-like fashion may, in fact, be responsible for the non-target-like generalizations they make regarding the prosodic structure of the language.

3. Method

3.1 Participants

Seventy-six French Canadian L2 learners of English (age 18–51) and 31 native speakers of Canadian or American English (age 18–55) took part in the study. L2 learners were allowed to participate only if their knowledge of languages other than English was no higher than "intermediate" (as self-reported). Similarly, native speakers of English were not allowed to participate in the study if they qualified their knowledge of French as greater than "intermediate". A language background questionnaire revealed that the L2 learners' mean age of first exposure to English was 9 years 1 month (standard deviation (SD) = 1;10) and that most L2 learners (65/72) were first exposed to English in school. The mean amount of English instruction was 9 years 7 months (SD = 3;1), and there was wide variability in the amount of time these learners had spent between in an English environment, ranging from 0 to 15 years.

On the basis of a cloze test and a read-aloud task, the L2 learners were placed into three proficiency levels. The cloze test (Brown 1980) was used to assess the L2 learners' morphosyntactic, lexical, and discourse competence, while the read-aloud task provided a measure of L2 phonological competence (for a similar procedure, see Colantoni and Steele 2007). In the latter task, the L2 learners were audiorecorded while reading an excerpt from a newspaper article. Their foreign accents were subsequently rated by five native speakers of English (3 Canadian, 2 American) on a scale from 1 to 5 (1 = very strong accent; 2 = strong accent; 3 = noticeable accent; 4 = mild accent; 5 = no accent). Composite proficiency scores were computed by summing the L2 learners' accuracy rates on the cloze test (with their total scores out of 50 being divided by 10) and the mean accent ratings they received (out of 5) on the read-aloud task. On the basis of the composite score, three groups were formed. The most advanced group included fewer participants due to the second author's intuition that few L2 learners were highly proficient in English. The individual and composite proficiency scores are presented in Table 1.

Table 1. Mean Scores (SDs) on Proficiency Measures

Proficiency Level	Cloze (/5)	Read-aloud (/5)	Global (/10)
Intermediate (n = 29)	3.1 (0.8)	2.5 (0.5)	5.6 (1.0)
Low-advanced (n = 29)	4.1 (0.3)	3.2 (0.4)	7.3 (0.3)
High-advanced (n = 18)	4.3 (0.4)	3.9 (0.4)	8.2 (0.4)

3.2 Materials

The participants were asked to produce phonologically possible English nonsense words contextualized as nouns. The test items included 12 disyllabic words with light syllables (i.e., LL, e.g., [dʌ.ʃm̩]), 12 trisyllabic words with a light, a heavy, and a light syllable (i.e., LHL, e.g., [ʃʌ.mol.fn̩]), and 24 fillers. The disyllabic words were used to determine if foot headedness had been reset to the left, and the trisyllabic words were used to determine if the head of the foot was aligned with the heavy syllable.[1] The nonsense words were produced

[1] Due to its prosodic structure, English does not have multisyllabic words that both end with a syllabic sonorant and are stressed word-finally. There is thus a possibility that the participants will not stress word-final syllables because they contain a syllabic sonorant. Unfortunately, this problem could not be avoided. The only way to determine whether L2 learners have acquired a trochaic foot is by using light syllables, because heavy syllables can form a monosyllabic foot on their own. Word-final light syllables in English (i.e., word-final syllables not considered heavy in any word class) contain either a

in the sentence "The yellow _____ was sold yesterday." This sentence contextualized the nonsense words as nouns, and the prenominal adjective decreased the likelihood that the L2 learners would use the high tone found at the beginning of the (Phonological) Phrase in (Canadian) French to produce English stress on the nonsense nouns (e.g., Jun and Fougeron 2000).

The L2 learners were expected to manifest one of the four prosodic grammars presented in Table 2. It was predicted that if the L2 learners had not reset foot headedness to the left and did not align the head of the foot with the heavy syllable (i.e., L1 transfer), they would stress the last syllable of both word types; if they had reset foot headedness to the left but did not align the head of the foot with the heavy syllable (i.e., Type A grammar), they would stress the first syllable of both word types; if they have not reset foot headedness to the left but did align the head of the foot with the heavy syllable (i.e., Type B grammar), they would stress the second syllable of both word types; finally, if the L2 learners had reset foot headedness to the left and aligned the head of the foot with the heavy syllable (i.e., target-like grammar), they would stress the penult of both word types. The remaining two logical possibilities (i.e., LL + LHL, LL + LHL) were not expected under the analyses presented earlier.

Table 2. Typology of Interlanguage Grammars

Production	Grammar	Foot Headedness	Alignment with H
(LL) + L(HL)	L1 Transfer	–	–
(LL) + (LH)L	Type A	+	–
(LL) + (LH)L	Type B	–	+
(LL) + L(H)L	Target-like	+	+

Note. The underlined syllable is the head of the foot, and the parentheses represent foot edges.

The nonsense words were audiorecorded by a phonetically trained male native speaker of American English for the creation of aural stimuli. They were recorded one syllable at a time, with each syllable being stressed. The recordings were done using a Marantz PMD 660 Solid State recorder and a Rode NT 1-A Condenser microphone at a sampling rate of 44,100 Hz and 32-bit. They were then digitally transferred to a notebook computer and manipulated with the audio editor software Audacity (<www.audacity.sourceforge.net>). The syllables were normalized for fundamental frequency, duration, and amplitude. In each stimulus, the syllables were separated by an interval of 500 milliseconds (for a similar procedure, see Guion, Harada, and Clark 2004). In order to mask subtle pitch differences that may have remained between the syllables, a 100-millisecond tone of either 440

syllabic sonorant (e.g., *button* [bʌ́.ʔn̩]) or a schwa (e.g., *comma* [kɑ́.mə]). Since schwas are never stressed in French, using a word-final schwa would not have been appropriate to identify the foot shape that French Canadian L2 learners of English use to produce English words. Word-final syllables containing a lax vowel (e.g., [ʊ], [ɪ], [ʌ], etc.) but no coda could not be used either, because they are not attested in English. Since English syllabic sonorants *can* be stressed when they are not in word-final position (e.g., *bully* [bl̩.li]), it was judged that this syllable type would be the best for the last syllable of the nonsense words in this experiment.

Hz or 500 Hz was inserted halfway through each interval, as well as 200 milliseconds before the first syllable and 200 milliseconds after the last syllable. The stimuli were pseudo-randomized and presented twice in each trial.

3.3 Procedures

All testing took place in a quiet room. Instructions and prompts were presented on a CRT monitor at a resolution of 640 × 480 pixels, and JVC HA-G101 headphones were used for audio playback. The participants were instructed to combine the syllables they would hear and produce the resulting word in the sentence they would see on the computer screen. They were told to produce the syllables exactly as heard, in the order they were heard, and without pausing between them. After producing the sentence in each trial, the participants pressed a key to move on to the next trial. Their productions were recorded on an Olympus WS-300M digital voice recorder at a sampling rate of 44,100 Hz with a Uni-Tex UM300 condenser microphone. There was a practice session of 10 nonsense words at the beginning of the experiment to ensure that the participants would understand how to carry out the task. The total duration of the experiment was approximately 12 minutes.

3.4 Data Analysis

Three sets of analyses were performed on the data. In the first analysis, the syllable with primary stress was identified in the participants' productions. This was done by the second author, a native speaker of Canadian French. Certain sentences had to be excluded from these analyses, because the participants paused between the syllables in the nonsense word, stressed all syllables equally, or changed the segments in a way that affected syllable weight (e.g., tensing of lax vowels, insertion of codas, etc.). The participants who did not have at least 4 remaining disyllabic words and 4 remaining trisyllabic words were excluded from the analyses. This led to the exclusion of 6 L2 learners in the intermediate group, 3 L2 learners in the low-advanced group, and 2 native speakers. For the remaining participants, 20% of the native speakers' data and 28% of the L2 learners' data were excluded for the above reasons. As a reliability measure, 10% of the remaining data were re-analyzed for stress placement by a native speaker of American English. The results yielded an inter-coder reliability rate of 80% for the L2 learners' data and 91% for the native speakers' data.

In the second analysis, we identified the number of participants who produced a given stress pattern at a consistency rate of ≥75%. Although this cut-off point may seem somewhat arbitrary, it was deemed high enough to establish that one of the prosodic grammars (in Table 2) had been adopted, and low enough to allow for some of the intra-speaker variability commonly encountered in interlanguage grammars. This led to the further exclusion of 9 L2 learners in the intermediate group, 8 L2 learners in the low-advanced group, 8 L2 learners in the high-advanced group, and 6 native speakers. The remaining data were reanalyzed for stress placement by a third native speaker of American English, the first author of this paper. There were very few disagreements between the original coding and the second coding, and where there was disagreement, it was either resolved by the authors or presented to a fourth native speaker of American English for final judgment. In this second analysis, we also returned to the nonsense words that had been excluded from the initial data set because of changes in syllable weight, as they provide important information about the generalizations that L2 learners had made with respect to the sensitivity of stress to syllable weight (see Section 4).

The third analysis consisted of acoustic measurements of the nonsense words produced by the consistent participants using Praat (Boersma and Weenink 2007). This analysis included both the words that matched the participants' prosodic grammar (≥75%) and those that did not (<25%). Rhymes were segmented from the words by looking at formant transitions (for a similar procedure, see Turk, Nakai, and Sugahara 2006). The mean pitch, length, and amplitude of the rhymes were calculated. Excluded from the acoustic analyses were rhymes which exhibited creaky voice or in which the participants hesitated in their pronunciation. This resulted in the further exclusion of 0.8% of the data. The nonsense words ending with a high phrase accent or boundary tone (H– or H%), both of which are independent from stress (for details, see Jun and Fougeron 2000), were also excluded from the acoustic analyses. This led to the exclusion of 6 participants (2 L2 learners in the intermediate group, 1 L2 learner in the low-advanced group, 1 L2 learner in the high-advanced group, and 3 native speakers) who consistently ended the nonsense words with a high phrase accent or boundary tone. Among the remaining participants, 5% of the data were excluded because the nonsense word ended with a high phrase accent or boundary tone. Between-group comparisons were made using the ratio of the rhyme in stressed syllable over the rhyme in the unstressed syllable for each participant. This method was chosen, because stress is relative rather than absolute. Furthermore, ratios reduce the non-relevant inter-participant variance in absolute pitch, length, and amplitude. The ratios were computed first across stress patterns, and then separately for each stress pattern. In the trisyllabic words, the final (unstressed) syllable served as the comparison point for computing the ratios.

4. Results

The initial results are presented in Figure 1 (see Table 2 for a description of each prosodic grammar). Recall that the participants had to be at least 75% consistent in their productions in order to be included in these results.

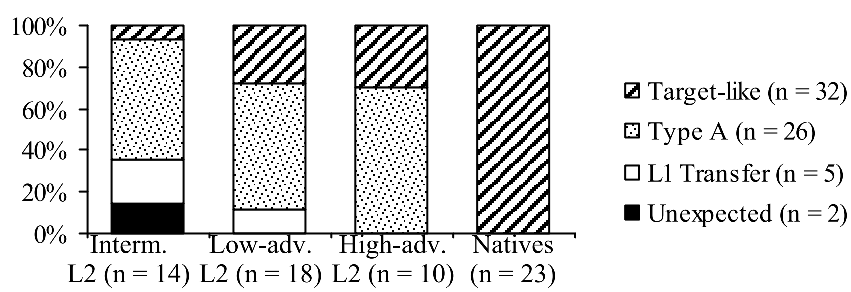


Figure 1. Within-level Distribution of Participants by Grammar Type

As can be seen in Figure 1, most L2 learners manifested a Type A grammar (i.e., $(\underline{L}L) + (\underline{L}H)L$): they had reset foot headedness to the left, but did not align the head of the foot with the heavy syllable. As proficiency increased, the percentage of L2 learners showing evidence of an L1 transfer grammar (i.e., $(L\underline{L}) + L(H\underline{L})$) decreased and eventually disappeared, but the percentage of L2 learners acquiring a target-like grammar (i.e., $(\underline{L}L) + L(\underline{H})L$) did not increase beyond the intermediate level. This contrasts with the native speakers' productions, which were target-like (i.e., $(\underline{L}L) + L(\underline{H})L$). Notably, not a single L2 learner showed evidence of a Type B grammar (i.e., $(L\underline{L}) + (L\underline{H})L$), and two L2 learners unexpectedly stressed the first syllable of disyllabic words and the last syllable of trisyllabic words. We will return to these findings in the discussion section.

In order to explain the L2 learners' typical failure to align the head of the foot with the heavy syllable, three hypotheses were considered: (i) these L2 learners do not know that CVC syllables are heavy; (ii) they do not know that heavy syllables are only licensed as the head of the foot; or (iii) they have some knowledge of both (i) and (ii), but align the trochaic foot with the left edge of the prosodic word at the expense of segmental faithfulness. In order to assess these possibilities, we examined the tokens that had been excluded from the initial analyses due to syllable weight changes. It was reasoned that if the L2 learners had some knowledge that CVC syllables are heavy and only licensed as the head of the foot, they should change the penultimate syllable from heavy to light and stress the antepenultimate syllable. Table 3 shows the results of this assessment for the L2 learners with a Type A grammar, as well as for the L2 learners with a target-like grammar and the native speakers (control).

Table 3. Percent Distribution of Excluded Trisyllabic Words

Weight Changes	Type A L2 Grammar (n = 26)[a]	Target-like L2 Grammar (n = 9)[b]	Natives (n = 23)[c]
LHL → LLL	75	86	74
Other weight changes	25	14	26

Note. [a] Total: 110 words; [b] Total: 14 words; [c] Total: 42 words.

As may be seen in Table 3, 75% or more of the participants' alterations in syllable weight consisted of heavy penultimate syllables being produced as light, with nearly all of those tokens receiving stress on the first syllable. In the case of L2 learners with a Type A grammar, the number of such words was not trivial (i.e., 110), representing about one third of the all the nonsense words they produced. It appears, then, that the L2 learners with a Type A grammar did have at least some knowledge that CVC syllables are heavy and that heavy syllables are only licensed as the head of the foot. Hence, something else must have led them to produce stress word-initially at the expense of segmental faithfulness.

We now turn to the acoustic analysis of the participants' nonsense words. Recall that this analysis included both the words that matched the participants' prosodic grammar (≥75%) and those that did not (<25%). The results of the L1 transfer group will not be discussed, because too few L2 learners showed evidence of this prosodic grammar for the acoustic analysis to be reliable. Table 4 presents the mean pitch, length, and amplitude ratios of stressed over unstressed rhymes (collapsed across stress patterns). Again, the final syllable served as the comparison point for computing the ratios in the trisyllabic words.

Table 4. Mean Pitch, Length, and Amplitude Ratios (SDs)

	Mean Pitch		Length		Amplitude	
	Disyll.	Trisyll.	Disyll.	Trisyll.	Disyll.	Trisyll.
Type A L2	1.09	1.16	0.52	0.37	1.02	1.03
(n = 24)[a]	(0.17)	(0.17)	(0.13)	(0.09)	(0.02)	(0.02)
Target-like L2	0.97	0.97	0.59	0.77	1.01	1.01
(n = 9)[b]	(0.19)	(0.21)	(0.13)	(0.15)	(0.02)	(0.01)
Native Speakers	1.07	1.06	0.62	0.82	1.03	1.03
(n = 20)[c]	(0.08)	(0.06)	(0.14)	(0.13)	(0.02)	(0.03)

[a] Total: 219 disyllabic words, 133 trisyllabic words; [b] Total: 73 disyllabic words, 82 trisyllabic words; [c] Total: 188 disyllabic words, 184 trisyllabic words

Overall, it appears that the Type A L2 learners approximated the native speakers in their use of pitch, and the target-like L2 learners were closer to the native speakers in their use of length. For pitch, two one-way ANOVAs conducted on the disyllabic and trisyllabic tokens revealed a main effect of group only in the trisyllabic words ($f(2, 50) = 6.345$, $p < .004$; disyllabic: $f(2, 50) = 2.220$, $p < .119$), with the Type A L2 learners showing a greater pitch ratio than the target-like L2 learners (Tukey $p < .004$). For length, two one-way ANOVAs conducted on the disyllabic and trisyllabic tokens revealed a main effect of group in both disyllabic ($f(2, 50) = 3.303$, $p < .045$) and trisyllabic ($f(2, 50) = 94.215$, $p < .001$) words, with the Type A L2 learners showing a smaller length ratio than both the target-like L2 learners (Tukey $p < .001$) and the native speakers (Tukey $p < .001$). The effect of group in the amplitude results did not reach significance for either disyllabic ($f(2, 50) = 2.724$, $p < .075$) or trisyllabic ($f(2, 50) = 2.064$, $p < .138$) words. For this reason, amplitude will not be discussed further.

Since the results in Table 4 were collapsed across stress patterns, there is a possibility that some of the significant differences, at least those found in the use of length, were due to segmental effects caused by the different number of tokens with a given stress pattern that each of the groups produced. We therefore examined the mean proportional differences between the stressed and unstressed rhymes by stress pattern. Note, however, that due to the large number of empty cells (e.g., whenever the participants adopted a stress pattern 100% of the time), it was not possible to conduct statistical analyses on these results.

Figures 2 and 3 show the mean proportional difference in pitch between the rhymes in the disyllabic words and the trisyllabic words, respectively. As can be seen in the figures, the Type A L2 learners' use of pitch was somewhat similar to the native speakers' when stress was produced word-initially but different from the native speakers' when stress was produced on the second syllable. The Type A L2 learners used higher pitch to realize stress, whereas the native speakers' highest pitch mean was always on the initial syllable, even when it was not stressed. By contrast, with the exception of the trisyllabic words with word-initial stress, the target-like L2 learners did not use higher pitch to realize English stress. In fact, they tended to flatten pitch differences across syllables.

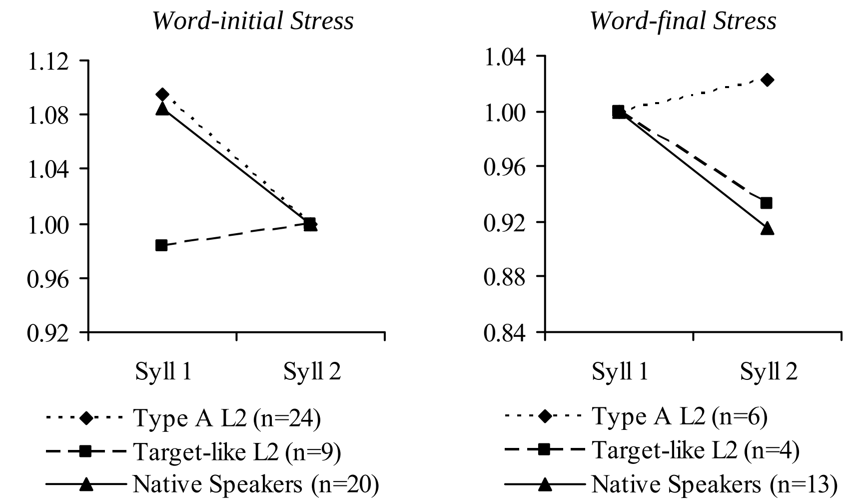


Figure 2. Mean Proportional Pitch Differences in Disyllabic Words

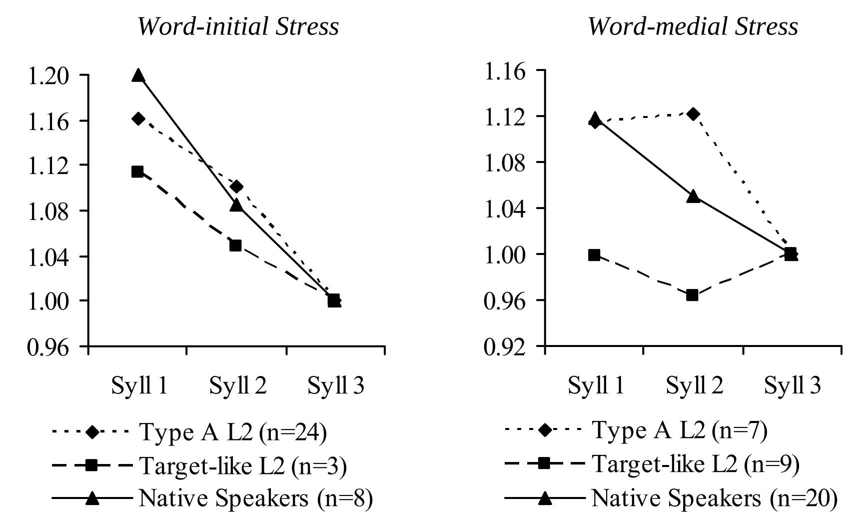


Figure 3. Mean Proportional Pitch Differences in Trisyllabic Words

Figures 4 and 5 respectively show the mean proportional differences in length between the rhymes in the disyllabic words and the trisyllabic words. Although the groups did not differ as much in their use of length as they did in their use of pitch, there was a tendency for the target-like L2 group to pattern like the native speakers and make greater use of length than the Type A L2 learners did when stress was not word-initial. When stress was produced on the first syllable of the trisyllabic words, the proportional difference between the initial and the medial syllables was also smaller for the target-like L2 learners, indicating perhaps that they attempted to lengthen the stressed syllable. Let us now turn to a discussion of the present findings and some concluding remarks.

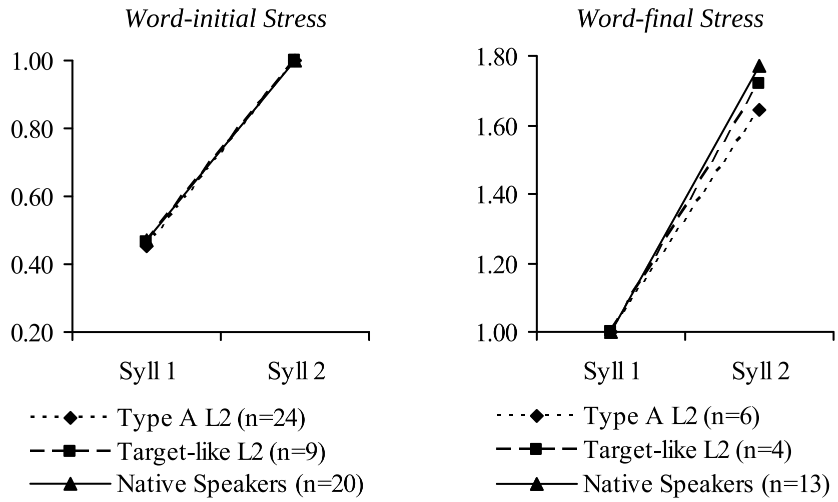


Figure 4. Mean Proportional Length Differences in Disyllabic Words

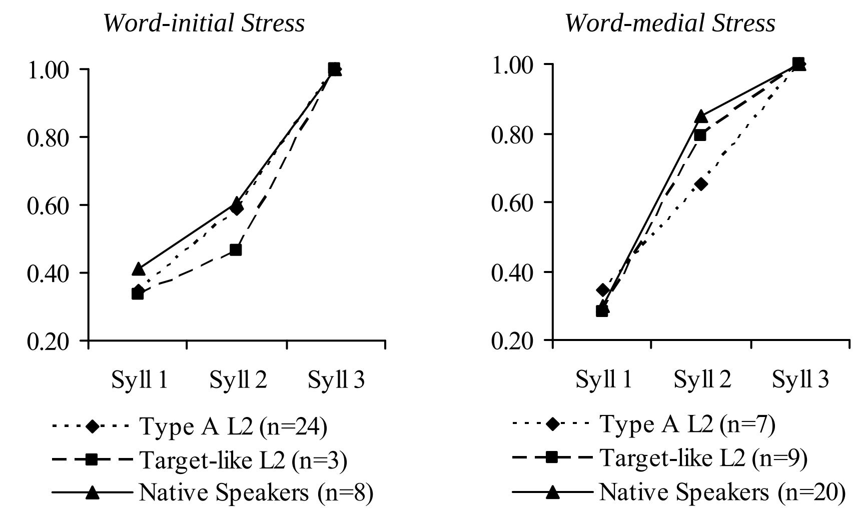


Figure 5. Mean Proportional Length Differences in Trisyllabic Words

5. Discussion and Concluding Remarks

The stress pattern results showed that the L2 learners had little difficulty acquiring the trochaic foot, but most of them failed to align the head of the foot with the heavy syllable. An examination of the tokens that contained changes in syllable weight indicated that the L2 learners with a Type A grammar had at least some knowledge that CVC syllables are heavy and that heavy syllables are only licensed as the head of the foot. Their failure to stress the heavy syllable may therefore not be a problem at the level of phonological analysis.

The acoustic analyses yielded at least three observations: (i) the L2 learners with a Type A grammar tended to use pitch to realize English word stress; (ii) the L2 learners with a target-like grammar tended use length to realize English stress; and (iii) native speakers used both pitch *and* length to realize stress in initial and non-initial positions, respectively. Importantly, the native speakers produced higher pitch on word-initial syllables irrespective of stress. If this behavior is typical, and if L2 learners produce the acoustic correlates of stress that they perceive in the input, then the L2 learners with a Type A grammar may have analyzed the presence of higher pitch on word-initial syllables as the instantiation of stress. In other words, these L2 learners' failure to align the head of the foot with heavy syllables may be due to a non-target-like mapping between the acoustic correlates of English word stress and the head of foot. This suggests a direct relationship between L2 learners' extraction of acoustic cues from the input and their acquisition of L2 prosodic representations. It is unclear, however, how L2 learners can recover from this non-target-like mapping.

It is surprising that most L2 learners used higher pitch rather than length to realize English stress, given that length is the predominant acoustic correlate of stress in Canadian French. On the other hand, studies have shown that pitch is generally more perceptible than length and amplitude cross-linguistically (for discussion, see Lehiste 1970). Moreover, higher pitch is used as an intonational marker of the left edge of the (Phonological) Phrase in (Canadian) French (Jun and Fougeron 2000). This might explain why this acoustic correlate was interpreted by many L2 learners as the instantiation of stress in English. We are confident that the Type A L2 learners' use of higher pitch on word-initial syllables was indeed the instantiation of stress, because an intonational high tone would have fallen on the adjective preceding the nonsense noun rather than on the noun itself.

Bearing the above generalizations in mind, we are left with three findings to explain: (i) the complete absence of the Type B grammar (in which the head of the iambic foot is aligned with the heavy syllable) in the L2 data; (ii) the production of unexpected stress patterns by 2 intermediate L2 learners; and (iii) the successful acquisition of a target-like grammar by some (albeit few) L2 learners. As concerns the first finding, we suggest that the iambic foot was abandoned very early (i.e., before any syllable weight generalizations had been made), because it was incompatible with too much of the input: phonologically, nouns with initial stress are extremely common in English (e.g., Clopper 2002), and acoustically, higher pitch is generally found on word-initial syllables irrespective of stress (as shown in our results). With respect to the second finding, we hypothesize that the L2 learners used a trochaic foot to produce the disyllabic words, but resorted to the L1 iambic foot with the trisyllabic words due to the higher processing load these words imposed. This is a realistic possibility, given the low proficiency level of these L2 learners. Finally, we tentatively conclude that the L2 learners who arrived at a target-like grammar (in which the head of the trochaic foot is aligned with the heavy syllable) relied more on length than on pitch to identify the location of stress, and from there established target-like generalizations about the alignment of the prosodic constituents. However, future research should establish why some L2 learners are more successful than others at weighing the correlates of English word stress across stress patterns.

References

Boersma, Paul, and David Weenink. 2007. Praat: Doing phonetics by computer (Version 4.5.18) [Computer program]. <http://www.praat.org>.

Brown, James D. 1980. Relative merits of four methods for scoring cloze tests. *Modern Language Journal* 64:311–317.

Carroll, Susanne E. 2001. *Input and evidence.* Amsterdam: John Benjamins.

Charette, Monik. 1991. *Conditions on phonological government.* Cambridge: Cambridge University Press.

Clopper, Cynthia G. 2002. Frequency of stress patterns in English: A computational analysis. *Indiana University Linguistics Club Working Papers Online* 2. Retrieved November 15, 2005 from <http://www.indiana.edu/iulcwp>.

Colantoni, Laura, and Jeffrey Steele. 2007. Acquiring /R/ in context. *Studies in Second Language Acquisition* 27:381–406.

Fox, Roberts A., James E. Flege, and Murray J. Munro. 1995. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America* 97:2540–2551.

Goad, Heather, and Meaghen Buckley. 2006. Prosodic structure in child French: Evidence for the Foot. *Catalan Journal of Linguistics* 5:109–142.

Gottfried, Terry L., and Patrice S. Beddor. 1988. Perception of temporal and spectral information in French vowels. *Language and Speech* 31:57–75.

Guion, Susan G., Tetsuo Harada, and J. J. Clark. 2004. Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition* 7:207–226.

Halle, Morris, and Jean-Roger Vergnaud. 1987. *An essay on stress.* Cambridge, MA: MIT Press.

Hammond, Michael. 1999. *The phonology of English.* Oxford: Oxford University Press.

Hayes, Bruce. 1981. *A metrical theory of stress rules*. Unpublished PhD diss., MIT.

Hayes, Bruce. 1995. *Metrical stress theory*. Chicago: University of Chicago Press.

Jun, Sun-Ah, and Cécile Fougeron. 2000. A phonological model of French intonation. In *Intonation: Analysis, modelling, and technology,* ed. Antonis Botinis, 209–242. Dordrecht: Kluwer.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America* 32:451–454.

McCarthy, John, and Alan Prince. 1986. Prosodic morphology. Report no. RuCCS-TR-32. New Brunswick, NJ: Rutgers University Center for Cognitive Science.

Montreuil, Jean-Pierre. 2002. Vestigial feet in French. In *Proceedings of the 2002 Texas Linguistics Society Conference on Stress in Optimality Theory,* eds. Augustine Agwuele, Willis Warren, and Sang-Hoon Park. Retrieved February 14, 2006 from <http://uts.cc.utexas.edu/~tls/2002tls/Jean-Pierre_Montreuil.pdf>.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology.* Dordrecht: Foris.

Ouellet, Maryse, and Linda Thibault. 1996. L'allongement prétonique: un phénomène opportuniste? In *Recherches en phonétique et en phonologie au Québec,* eds. Maryse Ouellet and Jean Dolbec, 47–61. Québec: CIRAL.

Paradis, Claude, and Denise Deshaies. 1990. Rules of stress assignment in Québec French: Evidence from perceptual data. *Language Variation and Change* 2:135–154.

Selkirk, Elisabeth O. 1978. The French foot: On the status of the mute *e*. *Studies in French Linguistics* 1:141–150.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure.* Cambridge, MA: MIT Press.

Turk, Alice, Satsuki Nakai, and Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In *Methods in empirical prosody research,* eds. Stefan Sudhoff et al., 1–28. Berlin: De Gruyter.

Walker, Douglas C. 1984. *The pronunciation of Canadian French.* Ottawa: University of Ottawa Press.