ARTICULATION OF ENGLISH "PROMINENCE" BY L1 (ENGLISH) AND L2 (FRENCH) SPEAKERS

Ting Huang & Donna Erickson

Haskins Laboratories, USA
funting.huang@gmail.com, ericksondonna2000@gmail.com

ABSTRACT

This study examines jaw and tongue blade (TB) articulation of prominence in two English sentences (one with target words containing mid front vowels, the other with all low vowels) produced by L1 and L2 (French) speakers of English. The results show that even though the phonological target vowels are kept constant in the sentence, the amount of jaw lowering and the corresponding tongue position in the sentences vary for each word. This is true for both L1 and L2 speakers. However, the patterns of the L2 speakers can be different from those of the L1 speakers; for instance, the L1 speakers show consistent patterns of Low-High jaw position for each word in a phrase in these sentences, with a step-wise lowering at TB position and one word at the lowest jaw position in the utterance, whereas L2 speakers generally do not have a consistent word with the lowest jaw position.

Keywords: articulatory prominence, jaw, tongue, L1 & L2, English & French

1. INTRODUCTION

Our work is inspired by the observation that in English, the jaw is lowest on the prominent syllables in an utterance (e.g., [1, 2, 3, 4, 5, 6]). Moreover, for emphasized syllables, regardless of vowel height, the jaw lowers more, and the distance between F1 and F2 either decreases or increases depending on whether the vowel is low or high, respectively [2]. Recent work has pursued this finding to report that the amount of jaw displacement from the occlusal plane in each syllable in an English utterance (with all low vowels) shows significant correlation with first formant frequency values as well as metrical grid levels of that sentence, (e.g., [4]). That is, a larger jaw displacement/larger F1 value indicates a more prominent syllable, and the largest jaw displacement/largest F1 indicates the word in the sentence with nuclear stress/prominence.

Based on findings from these studies on English that jaw height reflects syllable prominence, this study examines articulation of prominence in two English sentences, as spoken by L1 and L2 (French) speakers of English. English and French employ different rules for prominence implementation. For English, one syllable in each phrasal unit is more prominent than the other, with flexibility about the position of the prominence within the phrase, depending on the speaker's semantic intention. French has phrasal prominence occurring at the end of the phrase, as a function of marking phrase boundaries. Given the different rules for location of the phrasal stress, we would expect to see French speakers showing phrasal stress at the end of each phrase, while English speakers would show more flexibility about the position of the prominence in the phrase (see also [13]). Besides, since the two languages use "stress" differently, i.e. one to signal the end of the phrase and the other to signal the semantically important word, we might expect that articulatory implementation of "stress" might be different for the two languages, and that learners of English might carry over their L1 patterns to L2. A comparison of the L1 and L2 speakers of English would be useful in seeking salient articulatory features for prosody learning.

In this paper, we investigate not only jaw articulation but also tongue movement of both English L1 speakers and French L2 speakers of English. Recent work (e.g. [13, 15, 16]) has indicated that increased displacement of jaw and tongue body, and accompanying lengthened duration and increased F1 values are highly correlated with prominence in both of the languages. We wonder if French learners of English have different articulatory patterns from native speakers of English. Our questions are two-fold: what are some consistent jaw and tongue articulatory patterns for the L1 speakers? Are the L2 patterns similar to that seen for the L1 speakers?

2. METHOD

2.1. Participants

Articulatory and acoustic data were recorded with an NDI Wave articulograph at GIPSA-lab, Grenoble, for 5 French L1 speakers (3 M, 2 F) and 4 English L1 speakers (2 M, 2 F). Note that the two female L1 speakers were excluded from analysis due to some recording problems. Data from two additional English speakers (1 M, 1 F) recorded at JAIST, Japan, using 3-D EMA (Carstens AG500) were analysed as well. The data obtained respectively from NDI and 3-D EMA were pooled together for analysis (see [10]). The subjects' ages ranged from 22-70 at the time of recording. The L2 French speakers had an intermediate to advanced level in English based on informal interviews. The number of subjects was not sufficient for exploring a relation of L2 proficiency with patterns of prosodic articulation, but for finding out possible variations in prosody acquisition.

2.2. Materials

The material consisted of both French and English sentences; here we report only on the English utterances (for French utterances, see [13]). The target words in one of the sentences contain all front mid vowel [e]; the other, all low vowel [a]. The two sentences examined here are the following:

(1) *I gave dates to Kay today.*
(2) *I saw five bright highlights in the sky tonight.*

The sentences were chosen to control for vowel quality, that is, to keep the vowels as similar as possible, since vowel height affects jaw opening (for low vowels in English, the jaw opens about 2 mm lower than for mid vowels, and about 4 mm lower than for high vowels [11, 20]). The target vowels in the first sentence were the [e] part of the /ei/ diphthongs, those in the second sentence, the [a] part of the /ai/ diphthongs.

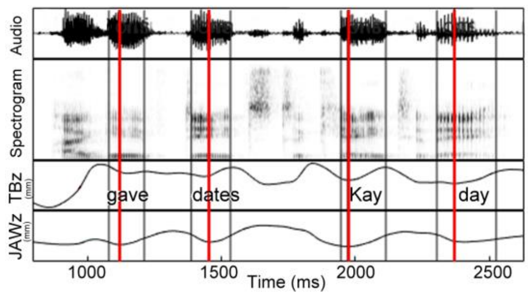
2.3. Procedure

For all recordings, one sensor was placed on the lower medial incisors to track jaw motion, and four additional sensors (upper incisors, bridge of the nose, left and right mastoid processes behind the ears) were used as references to correct for head movement. For the NDI English recordings, each speaker repeated the sentences six to nine times; for the EMA recordings, five times. All of the tokens were pooled together for a comprehensive comparison across speakers. The speakers read a randomized list from a PowerPoint display of the sentences about five to nine times; each sentence was presented separately to avoid any list effects. Before recording they practiced until they felt comfortable reading the sentences. The articulatory and acoustic data were digitized at sampling rates of 200 Hz and 16 kHz, respectively. The occlusal plane was estimated using a biteplate with three additional sensors. In post processing, the articulatory data were rotated to the occlusal plane and corrected for head movement using the reference sensors after low-pass filtering at 20 Hz.

2.4. Data Analysis

MView [18] was used for marking boundaries of vowel gestures (see Figure 1). The boundaries for target vowels were demarcated based on the steady state of vowel formants. The articulatory parameters (jaw and Tongue Blade (TB) positions relative to occlusal plane) were measured at the lowest jaw position (marked with vertical thick lines in the figure) within the vowel for each of the target words. By doing so, the articulatory movement and corresponding acoustic measures can be correctly estimated in the vowels [19].

Figure 1: A typical articulatory pattern for a sample utterance of *(I) gave dates to Kay today* by L1 speaker, I1. Upper panels show the acoustic wave form and spectrogram and bottom panels show TB and jaw tracings (mm), respectively. The vertical thick lines demarcate the point of maximum jaw opening in the vowel part of each word, the point at which the jaw and tongue measurements were made.



TB is a correlate of the tongue movement based on the following reasons. First, it was easier to displace sensors of tongue blade onto a consistent position than of tongue dorsum in the separate experiments. Second, tongue body is physically connected to jaw via digastric muscle to hyoid bone, but their coupling is non-rigid for speech [9, 14, 21]. Thus, we reported results of the tongue blade, which is a main articulator for the mid-front vowel /e/.

3. RESULTS

For all speakers, the amount of jaw lowering as well as the corresponding TB position for each syllable in the sentences varies, even though the phonological target vowels are kept constant. However, the patterns vary depending on whether the speakers are L1 or L2. In describing the results, we refer to a low jaw position as indicative of a prominent syllable (S) and a high jaw position, of a relatively less prominent syllable (W). For both sentence types (regardless of a mid or a low vowel in the target words), the jaw position varies such that there tends to be one lowest jaw position in the utterance, which is referred to here as the nuclearly prominent syllable (cf. [4]). Tongue position necessarily varies depending on vowel height. That is, for the mid vowel utterances, the jaw is low and TB is high enough for a prominent syllable. For the low vowel utterances, TB is expected to be low and jaw is relatively open. We also examine TB-Jaw (i.e., the value of the tongue blade position relative to the jaw position) to understand how different articulators collaborate in prosodically varying contexts.

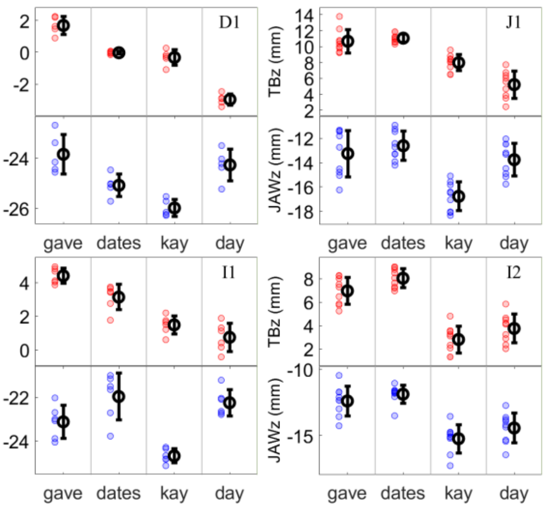
3.1. Jaw, TB and TB-Jaw patterns for sentence with mid vowels. *(I) gave dates to Kay today*.

3.1.1. L1 Speakers

Figure 2 shows Jaw and TB patterns for the four L1 speakers. The jaw patterns are as follows: for the two words in the first phrase (*gave dates*), three speakers (J1, I1, I2) show SW patterns with *gave* having a lower jaw position than *dates*, while the fourth speaker (D1), has a WS pattern, with *date*s having a lower jaw position. For the final phrase, all speakers show *Kay* having the lowest jaw position in the utterance as the nuclearly prominent word in the utterance. As for the TB, in most cases it appears unaffected by jaw lowering on prominent syllables and shows a general word by word lowering of the tongue. The slight exception is I2, who exhibits a dependency of TB movement on the jaw (R=0.86, p=0.00): when the jaw is low on *gave* and *Kay*, the corresponding TB position is low, too.

Regarding the results of TB-Jaw, i.e., the vertical distance from TB to jaw, all L1 speakers show the lowest TB-Jaw value on the final word; the jaw is high and the TB is low, so that we see both less jaw and TB movement for phrase final position, which might be interpreted as indicative of articulatory declination, i.e., a reduction of articulatory effort over the utterance. Two of the speakers (D1 and J1) show the highest TB-Jaw value (25.66mm for D1, and 24.75mm for J1) for the nuclear prominent item (*Kay*), and in this case, the jaw is lower, but the TB is raised. This is consistent with findings by [2], that prominent (emphasized) mid-vowel syllables have lower jaw and higher tongue positions. However, the other speakers show the largest TB-Jaw on either the first word of the utterance (I1), or the final word of the first phrase (I2). Note that their TB position is highest on W(ord) 1 and 2, respectively. It seems that the L1 speakers optionally raise their TB to enhance the nuclearly prominent word.

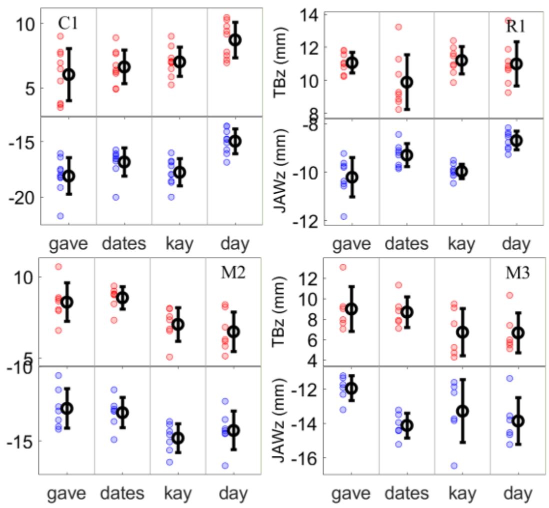
Figure 2: Average Jaw and TB positions for L1 speakers for each of the target words in *(I) gave dates to Kay today*. Since the patterns vary across speakers, y-axis values are such to allow easy visual comparison of articulatory patterns. (bottom: jaw position; top: TB position)



3.1.2. L2 Speakers

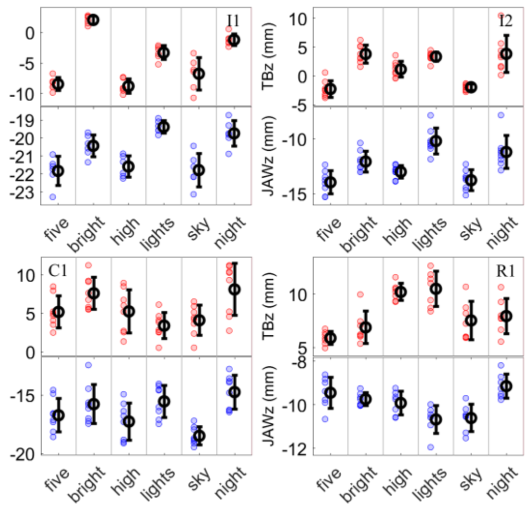
Figure 3 presents the articulatory patterns for the four L2 speakers. M2 and M3 suggest a WS SW or WS WS pattern, respectively, for jaw movement, with increased prominence on *Kay* or *dates/day*, respectively. Both M2 and M3 show a less marked lowering of the TB, as compared to that for L1 speakers. C1, R1 and M1 exhibit a SW SW pattern of jaw, which resembles the rhythmic pattern by L1.

Figure 3: Average Jaw and TB positions for L2 speakers (M1 is omitted here for space limitation) for each of the target words in *(I) gave dates to Kay today*. See caption for Figure 2 for details.



In fact, the L2 speakers exhibit many noticeable characteristics in their production of English. The L2 speakers locate the lowest jaw as indicative of a nuclearly prominent word to various positions, which is consistently on *Kay* for L1. As in Figure 3, the lowest jaw position is on *gave* for three speakers (C1, R1, M1), *dates* for M3, and *Kay* for M2. Placing the nuclear stress on *gave* is not acceptable in English rules, unless this word is emphasized. Furthermore, the TB fluctuation by L2 is less pronounced compared to that by L1. The difference in TB position (=TB_{max} - TB_{min}) throughout the utterance ranges from 1.31mm to 2.68mm across the L2 speakers, while for L1, 3.65mm to 5.84mm. It is also found that the TB movement by the L2 speakers is strongly correlated with jaw at least in three cases (C1: R=0.77, p=0.00; M2: R=0.53, p=0.00; M1: R=-0.50, p=0.01), implying that TB and jaw are interdependent in L2 production.

Figure 4: Average Jaw and TB positions for representative L1 speakers (I1 and I2; top panel) and L2 speakers (C1 and R1; bottom panel) for each of the target words in *(I) saw five bright highlights in the sky tonight.* See caption for Figure 2 for details.



3.2. Jaw and TB patterns for Sentence 2. *(I) saw five bright highlights in the sky tonight.*

Figure 4 presents the articulatory pattern of the sentences for L1 (top) and L2 (bottom) speakers, respectively. For the L1 speakers, the patterns of Jaw and TB for this sentence are such that all speakers show an SW pattern for each of the phrases; however, which word receives the nuclear prominence varies: for D1, it is *high*; for J1, *sky* and for I1 and I2, *five*. The TB pattern matches that of the jaw for the three male speakers, but less so for the one female speaker, which may be related to male-female differences in articulation strategies, given vocal tract size differences [12]. In contrast to the mid front vowel, the consecutive low-vowel production involves a marked synergy for jaw and TB, suggesting gestural coordination is vowel-dependent.

For the L2 speakers, they show tendencies for alternating patterns of jaw position for this sentence, but generally not as marked as the L1 speakers. Only three of the speakers (C1, M1, and M3) showed the SW repetitive pattern that L1 speakers showed, and they also showed possible nuclear prominence (M3 on *five*, M1 on *high*, and C1 on *sky*). As for the TB of the three L2 speakers who had SW jaw patterns, their TB tracings did not match the SW jaw patterns, as did the TB for the L1 speakers. The other two speakers (M2 and R1) showed a WS pattern in the first two phrases, placing prominence on *bright* and *lights*, which is an unacceptable production based on English stress rules (unless these particular words were emphasized). Interestingly enough, when the L2 speakers produced the undesirable WSWSSW pattern, there is a tendency that their TB and jaw are not correlated, in contrast to the L1 speakers.

4. DISCUSSION

This study has focused on the articulation (jaw and tongue blade) of English prominence, comparing L1 and L2 speakers. The data reported for the L1 speakers here support findings from previous studies: English prominence is implemented articulatorily by increased jaw lowering, regardless whether the vowel is mid or low. Moreover, English speakers are free to choose which word in the phrase and which word in the utterance has the most prominence. This could be the case when the speakers were guided to read with no pragmatic contexts. A new finding in this study concerns TB movement in English prominence. In general, the vowel height affects whether TB and jaw can form a synergy in articulation of prominence. In the case of mid front vowel, the TB and jaw are independent: the jaw shows a high-low pattern, reflecting SW word prominence, while the TB shows a step by step lowering. By contrast, the repetitive SW jaw pattern found in the low vowels perfectly matches the corresponding TB positions. We see more constraints in TB movement for the front mid vowel, which needs to be distinctive in a crowded vowel space (see also [7, 8]). New findings are also reported about L2 articulation of prominence in English sentences. Some L2 speakers show alternating patterns of high-low jaw position, but the others produce prosodic patterns that violate English stress rules, suggesting varying degrees of prosodic transfer from their native language. The errors found in the L2 production may be associated to articulatory strengthening at pre-boundary positions reported in French [15, 17].

5. ACKNOWLEDGEMENTS

Our thanks to Christophe Savariaux and Pascal Perrier at GIPSA Labs, Grenoble, France, and Jianwu Dang (JAIST, Japan) and Atsuo Suemitsu (Sapporo Health and Medical University, Japan), Japan, for use of and technical help with articulographs, the French and English subjects, and Caroline Smith for help on earlier versions of the paper. This work was partially supported by the Japan Society for the Promotion of Science (JSPS), Grants-in-Aid for Scientific Research (C) #25370444 to the second author, and partially by Haskins Laboratories.

6. REFERENCES

[1] Erickson, D. 1998. Effects of contrastive emphasis on jaw opening. *Phonetica* 55, 147-169.

[2] Erickson, D. 2002. Articulation of extreme formant patterns for emphasized vowels. *Phonetica* 59, 134-149.

[3] Erickson, D. Year. On phrasal organization and jaw opening. *Proc. Proceedings of From Sound to Sense* MIT. 24.

[4] Erickson, D., Suemitsu, A., Shibuya, Y., Tiede, M. 2012. Metrical structure and production of English rhythm. *Phonetica* 69, 180-190.

[5] Erickson, D., Kawahara, S., Williams, J., Moore, J., Suemitsu, A., Shibuya, Y. 2014. Metrical structure and jaw displacement: An exploration. Speech Prosody, 300-303.

[6] Erickson, D., Kim, J., Kawahara, S., Wilson, I., Menezes, C., Suemitsu, A., Moore, J. 2015. Bridging articulation and perception: The C/D model and contrastive emphasis. International Congress Phonetic Sciences.

[7] Flemming, E. 2004. Contrast and perceptual distinctiveness. In: Bronson (eds), *Phonetically based phonology.* Cambridge. 232-276.

[8] Flemming, E. 2006. The role of distinctiveness constraints in phonology, Technology, ed.

[9] Kaburagi, T., Honda, M. 1994. Determination of sagittal tongue shape from the positions of points on the tongue surface. *The Journal of the Acoustical Society of America* 96, 1356-1366.

[10] Kolb, A. 2015. *Software tools and analysis methods for the use of electromagnetic articulography data in speech research*. Master's thesis, Biomedical Engineering, Marquette University, Milwaukee.

[11] Menezes, C., Erickson, D. Year. Intrinsic variations in jaw deviations in English vowels. *Proc. POMA.* 060253.

[12] Simpson, A. P. 2002. Gender-specific articulatory–acoustic relations in vowel sequences. *Journal of Phonetics* 30, 417-435.

[13] Smith, C., Erickson, D., Savariaux, C. forthcoming. Articulatory and acoustic correlates of prominence in French: Comparing L1 and L2 speakers. *Journal of Phonetics Special Issue*.

[14] Stone, M. 1990. A three-dimensional model of tongue movement based on ultrasound and X-ray microbeam data. *The Journal of the Acoustical Society of America* 87, 2207-2217.

[15] Tabain, M. 2003. Effects of prosodic boundary on/aC/sequences: acoustic results. *The Journal of the Acoustical Society of America* 113, 516-531.

[16] Tabain, M. 2003. Effects of prosodic boundary on/aC/sequences: articulatory results. *The Journal of the Acoustical Society of America* 113, 2834-2849.

[17] Tabain, M., Perrier, P. 2005. Articulation and acoustics of/i/in preboundary position in French. *Journal of Phonetics* 33, 77-100.

[18] Tiede, M. 2010. MVIEW: Multi-channel visualization application for displaying dynamic sensor movements.

[19] Whalen, D. H., Chen, W.-R., Tiede, M. K., Nam, H. 2018. Variability of articulator positions and formants across nine English vowels. *Journal of phonetics* 68, 1-14.

[20] Williams, J., Erickson, D., Ozaki, Y., Suemitsu, A., Minematsu, N., Fujimura, O. Year. Neutralizing differences in jaw displacement for English vowels. *Proc. POMA.* 060268.

[21] Yehia, H., Rubin, P., Vatikiotis-Bateson, E. 1998. Quantitative association of vocal-tract and facial behavior. *Speech Communication* 26, 23-43.