Tonal Aspects of Languages 2016
24-27 May 2016, Buffalo, New York

Jaw displacement and phrasal stress in Mandarin Chinese

Donna Erickson [1], *Ray Iwata* [2], *Atsuo Suemitsu* [3]

[1] Kanazawa Medical University, Japan
[2] Kanazawa University, Japan
[2] Sapporo University of Health Sciences, Japan
EricksonDonna2000@gmail.com,rayiwata@staff.kanazawa-u.ac.jp,
sue@sapporo-hokeniryou-u.ac.jp

Abstract

Spoken languages have rhythmic structures, often phonologically described in terms of the metrical phrasal stress patterns of that language [1, 2, 3, 4]. Recent studies with English, Japanese, and Spanish suggest that rhythmic phrasal stress patterns are articulated by syllable to syllable jaw displacement variations, with concomitant syllable to syllable changes in vocal tract resonance frequencies, especially F1[5, 6, 7, 8, 9, 10, 11]. Moreover, the jaw and F1 patterns match the hierarchical (metrical) structure of the particular language. Inspired by these cross-linguistic findings, we investigate Mandarin Chinese to see if a tonal language also shows patterns of jaw displacement/F1 reflecting phrasal stress/hierarchical groupings of syllables. Acoustic and electromagnetic articulatory recordings of jaw displacement of 6 Mandarin Chinese speakers were analyzed. The results indicate that jaw displacement increases for phrase-final syllables, and optionally for phrase-initial syllables, regardless of lexical tone, with acoustic consequences of increased F1/duration, but not increased intensity or F0.

Index Terms: Mandarin Chinese, phrasal stress, jaw displacement, F1, duration, hierarchical structure

1. Introduction

Spoken languages have rhythmic structure, often phonologically described in terms of metrical/hierarchical groupings of phrasal stress (syllable weight) patterns. These patterns, according to recent cross-linguistic studies [5, 6, 7, 8, 9, 10, 11], are articulated by the speaker's changing how much s/he opens the jaw for each syllable in the utterance, resulting acoustically in patterns of increased F1/duration. A question we investigate in this paper is how does the articulation of post-lexical phrasal prosodic patterns interact with lexical prosody (e.g., tones)?

Mandarin Chinese has 4 lexical tones, Tone 1 (T1), Tone 2 (T2), Tone 3 (T3) and Tone 4 (T4), as well as a neutral tone that is unmarked, and these are articulated by changes in vocal fold tension, resulting in systematic variations in fundamental frequency (F0) [12]. There may be voice quality changes, also implemented by changes in patterns of vocal fold vibration, especially for T3, resulting in a vocal fry-type of voice [12, 13, 14]. As for Mandarin phrasal patterns, duration has been reported to be the most important acoustic cue, with duration longest in the final position, second longest in the initial position, and shortest in the medial positions [15]. In terms of phonological descriptions, structural linguistics suggests an iambic pattern, i.e. the last syllable is the strongest, the first next, and the intermediate is least stressed, which notice matches the above duration description [16]. Thus the stress pattern, in the case of three syllable phrases is said to be MWS or for four syllables, 'MWWS', where 'S', 'M' and 'W' represent Strong, Mid and Weak, respectively. Against this view, recent linguists in theoretical field argued for the existence of a trochaic left-headed foot, whether in a phrase or in a compound word [17].

This paper, an extension of an earlier pilot study [18], examines Mandarin Chinese jaw displacement patterns, along with the corresponding acoustic outputs, in order to see how articulatory patterns, e.g., jaw displacement patterns, might help better understand the language's phrasal stress/metrical patterns structure; also, we investigate to what extent jaw displacement phrasal patterns interact with lexical tones.

2. Methods

Acoustic and articulatory recordings were done using 3-D EMA (Carstens AG500 Electromagnetic Articulograph) at the Japan Advanced Institute of Science and Technology (JAIST). This paper reports on the one sensor placed on the lower medial incisors to track jaw motion. Correction for head movement was done from four sensors attached to the upper incisors, bridge of the nose, and left and right mastoid processes behind the ears, respectively. 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. The lowest vertical position (maximum displacement) of the jaw with respect to the bite plane was located for each target syllable of the utterance using the MATLAB-based custom software mview (Haskins Laboratories). Acoustic duration measurements of the vowels were made using Praat software [19]. Approximately 6 repetitions of each utterance were successfully recorded and analyzed.

Two experiments were done with a total of six Mandarin Chinese speakers. The first experiment examined short phrases from four speakers who were mostly from the Beijing Region (one middle age male (C02), one middle age female (C03), and two female graduate students (C04, C06)). C06 was actually a native speaker of Swatow (a variety of South Min), but spoke Mandarin Chinese fluently. The utterances were 4 Mandarin Chinese phrases: *jia1 na2 da4* ('*Canada*'), *ba1 na2 ma3* ('*Panama*'), *ma1 ma ma4 ma3* ('*Mother curses the horse*'), and *da4 ba3 da4 ba3* ('*a lot of*'', literally 'big grasp big grasp'). Since jaw displacement varies according to vowel quality (i.e., low vowels have larger jaw displacement than high vowels, e.g., [20]), the vowels were all the same in these utterances. The two utterances, *jia1 na2 da4* and *ba1 na2 ma3*, have 3 syllables with a tone contrast only on the final syllable: T4 vs T3, respectively. The two utterances *ma1 ma ma4 ma3* and *da4 ba3 da4 ba3* have 4 syllables with a tone contrast on the first two syllables. However, these two

 http://dx.doi.org/10.21437/TAL.2016-14

utterances differ syntactically: the first is a complete sentence with the noun phrase “mama” followed by a verb and object; the second is a phrase.

A second experiment with two additional Mandarin Chinese speakers (two male graduate students, C08 & C09) was done in order to look at phrasal patterns in longer utterances. The vowels within the sentences were also kept the same, but in addition to /a/ vowels, we used /ai/ and /ao/ vowels. The sentences were the following:

1. 他在哪儿啦？ -他在加拿大啦。他在哪儿啦？ -他在巴拿马啦。*Ta1 zai4 nar3 la? Ta1 zai4 Jia1na2da4 la. Ta1 zai4 nar3 la? Ta1 zai4 Ba1na2ma3 la. (Where is he? He is in Canada. Where is he? He is in Panama.)*
2. 赖奶奶来白海采海带。 *Lai4nai3nai0 lai2 bai2hai3 cai3 hai3dai4. (Grandma Lai comes to Baihai to pick kelp.*)
3. 曹姥姥叫姚嫂嫂找小毛。 *Cao2 lao3lao0 jiao4 Yao2sao3sao0 zhao3 xiao3mao2. (Grandma Cao asks Yao Saosao (sister-in-law) to find Xiaomao.)*
4. 小豹要咬老猫，老猫只好跑。*Xiao3bao4 yao4 yao3 lao3mao1, lao3mao1 zhi3hao3 pao3. (The small leopard wanted to bite the old cat, the old cat had to escape.)*
5. 佳佳//把/爸爸/妈妈/吓/怕/啦。*Ja1ja1 ba3 ba4ba ma1ma xia4 pa4 la. (Jiajia frightened baba and mama.)*

3. Results

3.1. Experiment 1. Short phrases

Figure 1 shows a sample jaw displacement pattern for the utterance *Mother curses the horse* (*ma1 ma ma4 ma3).* Notice that for each syllable, the jaw opens and closes, and moreover, the amount of jaw opening varies even though the vowels are all [a].

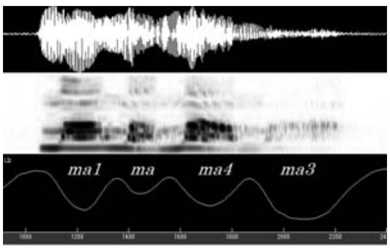

Figure 1: *Sample jaw opening pattern for ‘Mother curses the horse’ (ma1 ma ma4 ma3). The top panel is the acoustic wave form, the middle, the spectrogram and the bottom, the jaw displacement pattern. The lower the jaw curve, the greater the jaw opening. The x-axis is time in ms.*

Figure 2 shows average jaw displacement patterns for the 3-syllable utterances that differ in final tone (*jia1 na2 da4* (left side) vs. *ba1 na2 ma3* (right side)) for the four speakers. The amount of jaw displacement for each syllable is represented by the height of the vertical bar. Notice that for all four speakers, even though all the vowels are [a], the amount of jaw displacement varies per syllable. Moreover, across speakers, it varies in a systematic way: the final syllable of each utterance has the greatest amount of jaw displacement, and optionally, the initial syllable has the next greatest amount of jaw displacement. This pattern is regardless of whether the final tone is T4 (left panels) or T3 (right panels). As for the overall jaw displacement patterns of both utterances, some variations occur among the speakers: for *jia1 na2 da4*, 3 of the 4 speakers show a WMS (weak mid strong) jaw displacement pattern, while for *ba1 na2 ma3*, we see MWS for C02 and C03, WMS for C04, and MMS for C06.

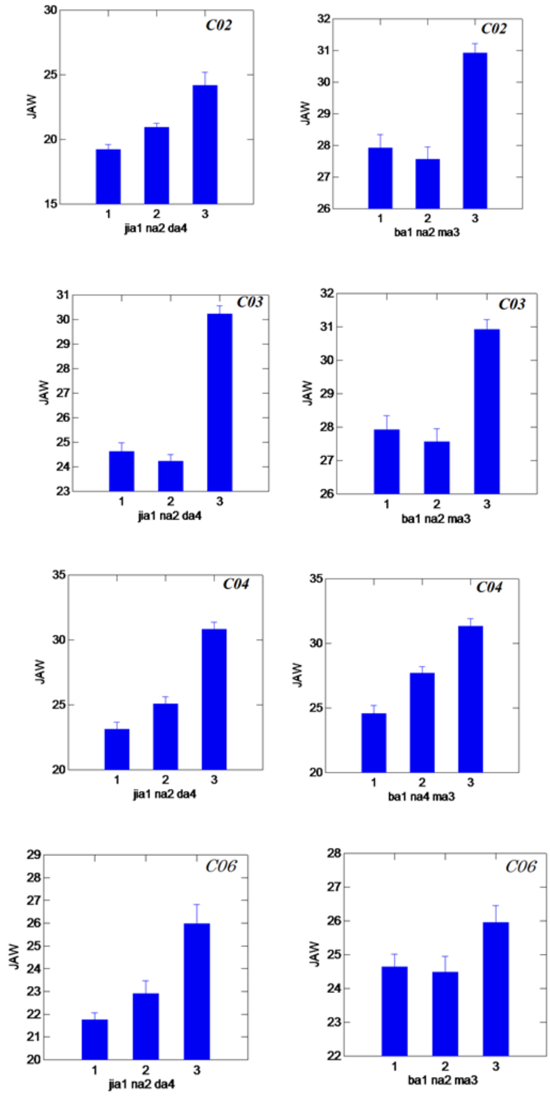

Figure 2: *Bar graphs of jaw displacement for the utterances that differ in final syllable tone (jia1 na2 da4 (left side) vs. ba1 na2 ma3 (right side)) for Speakers C02, C03, C04, and C06. The x-axis indicates the position of the syllable in the sentence, and the y-axis, the amount of jaw displacement in mm. The jaw displacement values are shown as positive in order to indicate patterns of syllable magnitude (see e.g., [21]); the y-values are set for each speaker to maximally show the jaw displacement patterns. Error bars (in this and all subsequent figures) represent standard error.*

Figure 3 shows jaw displacement patterns for utterances with four syllables, *ma1 ma ma4 ma3* (left side) vs. *da4 ba3 da4 ba3* (right side). Again, for all speakers the amount of jaw displacement varies, even though the vowels are the same; also the final syllable shows the largest jaw displacement, except for speaker C04. In addition, the first syllable tends to have larger jaw displacement than the second one, both for ma1 ma and da4 ba3 (even though the tones are different). The overall pattern for *ma1 ma ma4 ma3* seems to be MWMS for three of the speakers, and MMMS for C03. For *da4 ba3 da4 ba3*, the pattern may be MMMS for C02, MWMS for C03, SWMM for C04, and SWMS for C06.

For these short phrases, regression analyses, together with Pearson Product correlation analyses, were done with syllable jaw displacement and vowel duration measurements. The correlation analyses all showed a Bartlett Chi-square statistic of $p=0.000$. The r2 values indicate that there is a strong correspondence between jaw displacement and vowel duration: $r2=0.55$ (C02), 0.50 (C03), 0.51 (C04) and 0.43 (C06).

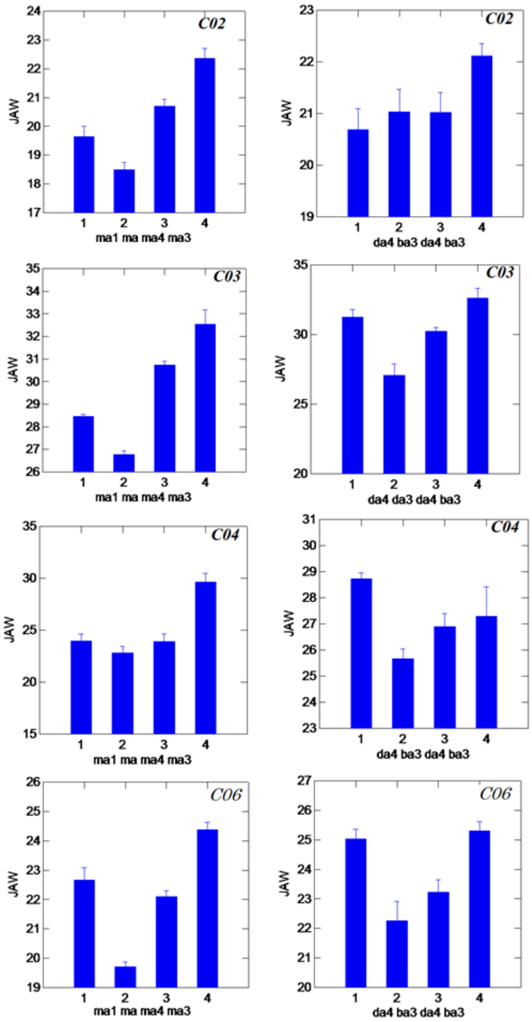

Figure 3: *Bar graphs of jaw displacement for the 4-syllable utterances: ma1 ma ma4 ma3 (left side) vs. da4 ba3 da4 ba3 (right side) for Speakers C02, C03, C04, and C06. x-axis indicates syllable position, y-axis, jaw displacement in mm.*

3.2. Experiment 1. Longer utterances

Figure 4 shows the average jaw displacement values for utterance 1, 他在哪儿啦？ -他在加拿大啦。他在哪儿啦？ -他在巴拿马啦。*Ta1 zai4 nar3 la? Ta1 zai4 Jia1na2da4 la. Ta1 zai4 nar3 la? Ta1 zai4 Ba1na2ma3 la.* (*Where is he? He is in Canada. Where is he? He is in Panama.*) This sentence can be syntactically parsed as four short phrases, with the perfective aspect *la* as extra-metrical: *{(Ta1)(zai4)(nar3)}la?/{(Ta1)(zai4)(Jia1na2da4)}la/{(Ta1)(zai4)(nar3)}la?/{(Ta1)(zai4)(Ba1na2ma3)}la./* Notice that for the circled words *Jia1na2da4* and *Ba1na2ma3*, the final syllable has the largest jaw displacement, as was also seen when spoken as a short phrase in isolation (Figure 2). Notice also the phrase-final syllable for each of the phrases tends to have the largest jaw displacement, with the initial syllable of each phrase having the next largest amount of jaw displacement. An interesting observation is that for the perfective aspect particle *la*, even though the F0 is the same for both the interrogative and declarative sentence, jaw displacement differs. For *(nar) la* (interrogative), *la* is produced with increased jaw displacement, while *la* in the declarative sentence shows no increased jaw displacement, even though occurring phrase finally. Also, a similar pattern of reduced jaw displacement was found for the final *la* in the sentence: *Ja1ja1 ba3 ba4ba ma1ma xia4 pa4 la* (not shown). Thus, jaw displacement may be affected not only by phrasal organization but also by parts of speech, such as particles.

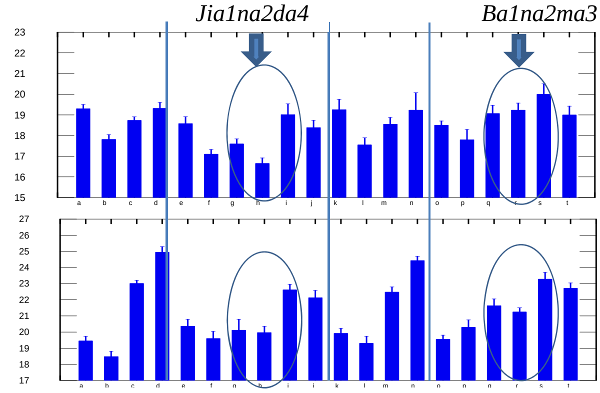

Figure 4: *Bar graphs of jaw displacement for the utterance {(Ta1)(zai4)(nar3)}la?/{(Ta1)(zai4)(Jia1na2da4)}la/{(Ta1)(zai4)(nar3)}la?/{(Ta1)(zai4)(Ba1na2ma3)}la./ for Speakers C08 (top) and C09 (bottom). x-axis indicates syllable position, y-axis, jaw displacement in mm. Phrases are marked with vertical lines; the words Jia1na2da4 and Ba1na2ma3 are circled.*

Figure 5 shows the average jaw displacement values for utterance 4. 小豹要咬老猫，老猫只好跑。*Xiao3bao4yao4 yao3lao3mao1,lao3mao1 zhi3hao3 pao3.* (*The small leopard wanted to bite the old cat, the old cat had to escape.*) This sentence can be syntactically parsed as four short phrases, *{Xiao3bao4}/{(yao4yao3)(lao3mao1)}/{lao3mao1}/{zhi3hao3 pao3}/*.

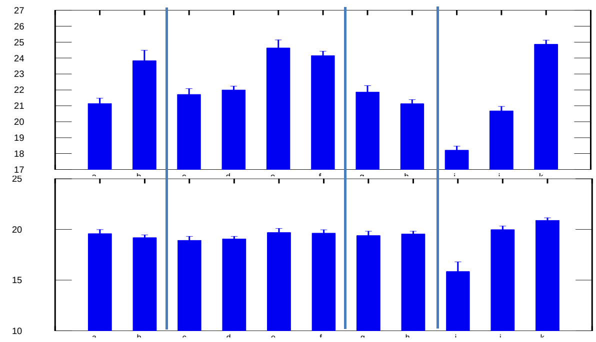

Figure 5: *Bar graphs of jaw displacement for the utterance {Xiao3bao4}/{(yao4yao3)(lao3mao1)}/{lao3mao1}/{zhi3hao3 pao3}/ for Speakers C08 (top) and C09 (bottom). x-axis indicates syllable position, y-axis, jaw displacement in mm. Phrases are marked with vertical lines.*

Notice whether the vowels in the sentence are /ao/, /ai/ or /a/, patterns of jaw displacement occur, with the final syllable having the largest jaw displacement. However, the Adj+N phrases (*Xiao3bao4, lao3mao1)* have a variable effect on jaw displacement, perhaps depending on the intended meaning of the utterance, i.e., whether the adjective or the noun receives the phrasal focus. In addition, C08 tends to show greater jaw displacement, which may be indicative of greater expressivity.

For these long phrases, regression analyses, together with Pearson Product correlation analyses, were done with the jaw displacement, F1, and syllable duration measurements. The correlation analyses all showed a Bartlett Chi-square statistic of p=0.000. The r2 values indicate that there is a correspondence between jaw displacement, F1 and syllable duration, as shown in Figure 6. However, no strong correlation was found among jaw displacement, average F0 and average intensity. Similar findings have been reported for English and Japanese, except that for these languages, interestingly no strong correlation was found with jaw displacement and duration.

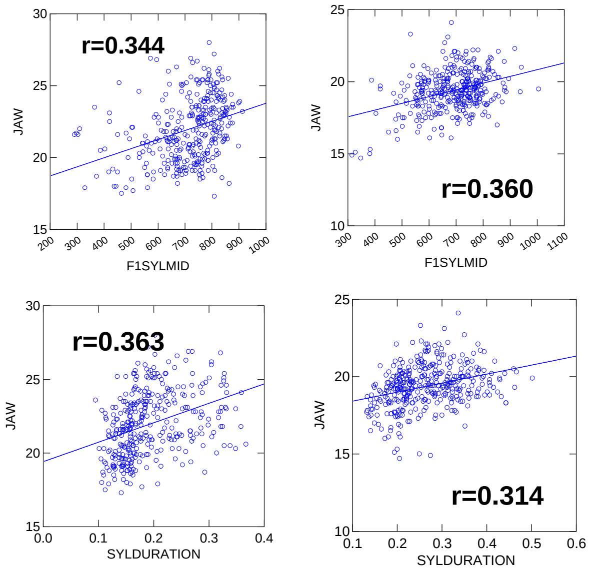

Figure 6: *Scatter plots of maximum jaw displacement vs. F1 (as measured at the acoustic midpoint of the syllable) (top row) and jaw displacement vs. syllable duration (bottom row) for C08 (left column) and C09 (right column).*

4. Discussion and Conclusions

Mandarin Chinese sentences consisting of syllables with the same vowel show varying amounts of jaw displacement, as has been found also for a number of other languages, such as English, Japanese, and Spanish. Theoretically, the amount of jaw displacement per syllable should be determined by the vowel quality, i.e., vowel height, but our results clearly show that the amount of jaw displacement is affected by phrasal organization: phrase-final syllables have the largest amount of jaw displacement, and optionally, phrase-initial syllables have increased jaw displacement. In addition, intended meaning (i.e., which word in the phrase is focused) affects jaw displacement. These results suggest that in terms of articulation, the jaw may be the primary articulator of phrasal (post-lexical) stress. The acoustic manifestations of increased jaw displacement are increased F1 and duration.

Thus, in addition to tones in Mandarin Chinese, syllables have different "weights", articulated by changes in jaw displacement and manifested acoustically by changes in vowel quality (F1) and duration. However, "syllable weights" and syllable tone are independent, e.g., for the 3-syllable utterances, the final lexical tone is T4 for *jia1 na2 da4* and T3 for *ba1 na2 ma3*, yet both words have increased jaw/F1/ duration for the final syllable.

A currently hot topic in phonology is how to describe the organization of these varying weighted syllables within utterances, within phrases. One possibility is to talk about a finite set of degrees of syllable weight, i.e., Weak, Mid and Strong) articulated by weak, mid and strong amounts of jaw displacement. An analysis of 3 and 4-syllable phrases, based on data from experiment 1, can be summarized as follows: (1) For 3-syllable utterances, 75% of the speakers use a WMS pattern for *jia1 na2 da4* while 50% use a MWS pattern for *ba1 na2 ma3* (see [22, p. 63] which reports three syllable phrases exhibiting an MWS patterns in terms of syllable duration). (2) For the 4-syllable utterances, 75% of the speakers produce the sentence *ma1 ma ma4 ma3* with a MWMS pattern. But, we do not see this pattern as clearly for *da4 ba3 da4 ba3*. A possible reason may be due to syntactic differences: The *da4 ba3 da4 ba3* utterance is a reduplicated measure, literally meaning 'big grasp big grasp,' so that the first element *da4* 'big' could be focused, whereas *ma1 ma ma4 ma3* is a sentence, where *ma1 ma* (mother) is a subject, so that it is not necessarily focused but the following verbal phrase is focused.

Yet another approach to describing syllable organization might be Yuenren Chao's Wave Theory, where a "Big wave" represents the "Major sentence or utterance", a "Small wave" a "Minor sentence or phrase" and a "Ripple word" is a single word phrase [23]. Jun [24, 25] suggests in her prosodic typologies that Mandarin Chinese has phrasal stress on intermediate phrases (ip) and intonational phrases (IP). Along these lines we might suggest a hierarchical description as shown in Figure 7, similar to that also proposed by Xu and Wang [15, figure 4],.

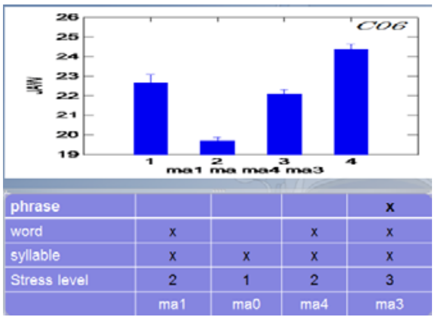

Figure 7: *Possible metrical hierarchical structure of 4-syllable phrase.*

Metrical phrase patterns can vary depending on the speaker, the communication intent, and also, of course, such things as emotion and personality can affect phrase stress patterns (e.g., [15]).

Future work will try to better understand the articulatory and concomitant acoustic characteristics of Mandarin Chinese patterns of phrasal stress, and how they interact with lexical stress. We also plan perception tests to ask listeners to indicate where they hear stress.

Finally, we would like to encourage researchers to examine the interactions between metrical structure and lexical tones; also, to comprehensively view "prosody" as embracing both laryngeal (including voice quality) and supralaryngeal articulatory activities.

5. Acknowledgements

This work was supported by the Japan Society for the Promotion of Science, Grants-in-Aid for Scientific Research (C) #25370444 and (A) #25240026. A special acknowledgement is made to the subjects, and especially to Chunyue Zhu for his help in data collection and utterance suggestions.

6. References

[1] Nakatani, L.H., O'Connor, K. D., and Aston, C. H. "Prosodic aspects of American English speech rhythm," Phonetica, 38: 84-106, 1981.

[2] Liberman, M. and Prince, A., "On stress and Linguistic rhythm," Linguistic Inquiry 8: 249-336, 1977.

[3] Selkirk, E.O., "The Syllable," Foris, Dordrecht , 1982.

[4] Hayes, B., "Metrical Stress Theory: Principles and Case Studies," The University of Chicago Press, Chicago, 1995.

[5] Erickson, D., Suemitsu, A., Shibuya, Y., and Tiede, M., "Metrical structure and production of English rhythm," Phonetica 69: 180–190, 2012.

[6] Erickson, D. and Kawahara, S., "Articulatory correlates of metrical structure" submitted to Linguistic Vanguard.

[7] Erickson, D., Kawahara, S., Shibuya, Y., Suemitsu, A., and Tiede, M., "Comparison of jaw displacement patterns of Japanese and American speakers of English: A preliminary report", Journal of Phonetic Society of Japan, 18: 88-94, 2014.

[8] Kawahara, S., Erickson, D., Moore, J., Suemitsu, A. and Shibuya, Y., "Jaw displacement and metrical structure in Japanese: The effect of pitch accent, foot structure, and phrasal stress", Journal of Phonetic Society of Japan, 18: 77-87, 2014.

[9] Kawahara, S., Erickson, D. and Suemitsu, A. "Edge prominence and declination in Japanese jaw displacement patterns: A view from the C/D model", Journal of Phonetic Society of Japan 19: 2015.

[10] Erickson, D., Villegas, J., Wilson, I., and Iguro, Y., "Spanish articulatory rhythm", Acoustical Society of Japan Fall Meeting. 319-322, 2015.

[11] Erickson, D., Villegas, J., Wilson, I., Iguro, Y., Moore, J., Erker, D., "Some acoustic and articulatory correlates of phrasal stress in Spanish", Speech Prosody 2016.

[12] Howie, J. M., "Acoustical Studies of Mandarin Vowels and Tones," Princeton/Cambridge Studies in Chinese Linguistics, Cambridge: Cambridge University Press (originally published in 1976), 2010.

[13] Erickson, D., Iwata, R., Endo, M., and Fujino, A., "Effect of tone height on jaw and tongue articulation in Mandarin Chinese," Proceedings of International Symposium on Tonal Aspects of Languages with Emphasis on Tone Languages (TAL 2004), 53-56, 2004.

[14] Endo, M., Iwata, R., Erickson, D., and Fujino, A., "Effect of tone height on mandibular and lingual articulation," Presentation at International Association of Chinese Linguistics, Aug 20-22, Aichi Prefectural University, Aichi, Japan, 2002.

[15] Xu, Y., and Wang M., "Organizing syllables into groups—Evidence from F0 and duration patterns in Mandarin," J. Phonetics 37:502-520, 2009

[16] Chao, Y-R., "A Grammar of Spoken Chinese," University of California Press, 1968.

[17] Duanmu, S., "The Phonology of Standard Chinese," Oxford University Press, 2000.

[18] Iwata, R., Erickson, D., Shibuya, Y., and Suemitsu, A., "Articulation of phrasal stress in Mandarin Chinese", Acoustical Society of Japan Fall Meeting, 207-210, 2015.

[19] Boersma, P. and Weenink, D., "Praat: Doing phonetics by computer. [Computer Software]" Department of Language and Literature, University of Amsterdam. Online: http://www.praat.org/, accessed on 15 Dec 2013.

[20] Menezes, C. and Erickson, D., "Intrinsic variations in jaw deviation in English vowels", Proceedings of International Congress of Acoustics, POMA 19: #060253, 2013.

[21] Fujimura, O., "The C/D model and prosodic control of articulatory behavior," Phonetica 57: 128-138, 2000.

[22] Lin, Maocan. "Hanyu yudiao shiyan yanjiu" [Experimental studies on Chinese tone and intonation]. Beijing: Zhongguo Shehui Kexue Press, 2012.

[23] Chao, Y.R. "Aspects of Chinese Sociolinguistics: Essays", Stanford University Press, 1976.

[24] Jun, S.-A. "Prosodic typology." In S.-A. Jun (ed.) Prosodic typology: The phonology of intonation and phrasing, 430–458 Oxford: Oxford University Press, 2005.

[25] Jun, S.-A. "Prosodic typology: By prominence type, word prosody, and macro-rhythm." In S.-A. Jun (ed.) Prosodic typology II: The phonology of intonation and phrasing, 520–539 Oxford: Oxford University Press, 2014.