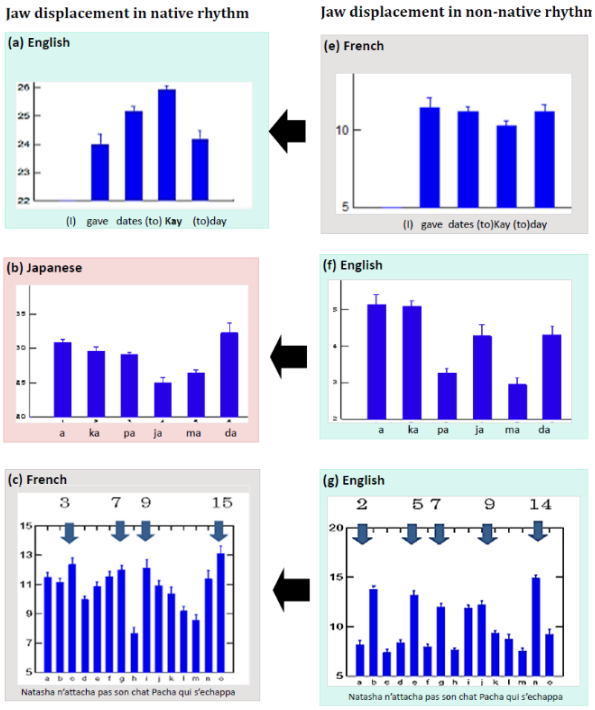
The MARRYS cap: A new method for analyzing and teaching the importance of jaw movements in speech production

Donna Erickson[1] (Haskins Labs), Oliver Niebuhr[2] (Southern Denmark U.), Wentao Gu[3] (Nanjing Normal U.), Ting Huang[4] (Haskins Labs), Puyang Geng[5] (Nanjing Normal U.)
[1]ericksondonna2000@gmail.com, [2]olni@sdu.dk, [3]wtgu@njnu.edu.cn, [4]funting.huang@gmail.com, [5]gengpuyang6@gmail.com

Research has shown that articulatory measures, more specifically, the amount of jaw lowering (i.e. mouth opening) during speaking, are closely correlated with the language-specific rhythmical structure of a spoken utterance. The difference in jaw displacement not only directly reflects varying syllable-stress levels in production [1, 2, 3, 4, 5], but also it has a strong correlation with perceptual ratings of syllable stress [6, 7]. English shows metrically strong stress by increasing jaw displacement on syllables, whereas for languages like French, Chinese and Japanese places maximal jaw displacement on the final prosodic unit (e.g., syllable) of a phrase.

These findings have significance for language learning applications. As reported in [8, 9], second language learners transfer their first language rhythmical structures in terms of jaw displacement patterns when speaking their second language, as shown in Figure 1 (reproduced from [8]). The left column shows the jaw patterns of the L1 speakers of an English, a Japanese, and a French sentence, respectively, and the right columns, the corresponding patterns of L2 speakers of these sentences. The L1 and L2 patterns are different, with the L2 patterns reflecting the rhythmic structure of their first language. The resulting foreign accent not only bears the risk of stigmatizing speakers in social respects. It often also hampers speech communication, e.g., in terms of identifying words or conveying information-structure or turn-taking signals. Furthermore, in public speaking, pronounced and dynamic jaw movements (esp. lowerings for open vowels) make a speaker sound more passionate and captivating and are, thus, important for his/her charismatic impact on listeners [11]. In addition, recent work has reported jaw movement changes associated with cross-cultural differences in expressions of attitudes, e.g., [12, 13].

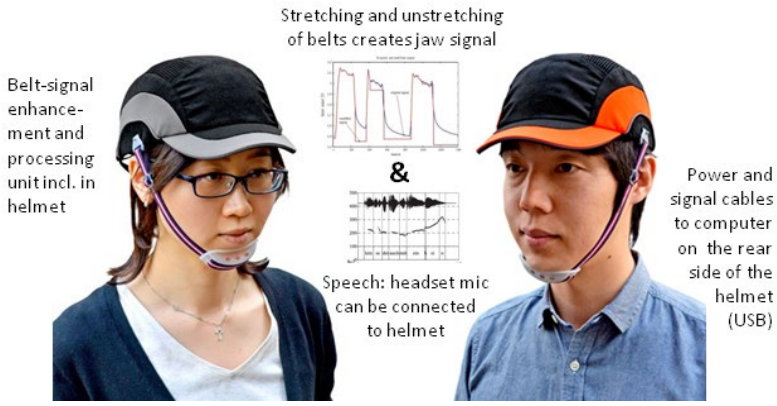
Figure 1: Jaw displacement in native and non-native rhythm.



A challenge that is addressed in this paper is how to obtain these jaw-displacement data from speakers, particularly at a larger scale. Currently, the amount of obtainable data is limited by the expense, complexity, and usability of equipment, i.e., X-ray microbeam and EMA (AG500 and NDI WAVE). EMA involves gluing sensors on various articulators, and on the mandible incisor to assess jaw displacement. Using EMA is time-laborious and skill-demanding, both of which limits the number of speakers who can be examined.

This paper reports on a new articulatory device that was developed by Niebuhr in collaboration with electrical engineers at Southern Denmark University: The Mandible Action Related Rhythm Signals (MARRYS) cap (Figure 2). Based on stereo signals of two stretch-sensor transducer belts that lead from the wearer's cheeks to a connector at the chin (cf. the RIP [14]), the MARRYS cap is more handy, mobile, and affordable than EMA and some other apparatus of similar functions. MARRYS is suitable for all kinds of phonetic analyses, be it classroom teaching/training, scientific lab studies on jaw-related vowel targets or rhythm patterns (see [10]). Specifically, it can be used to help teach more native-like expressions of attitudes for different cultures, such as dominance or friendliness. It is effective too in rhetorical trainings to assess, visualize, and improve a speaker's public-speaking performance. This study will examine the validity and precision of jaw-displacement data collected by MARRYS and present results from interaction-design research usability tests based on the established System Usability Scale (SUS) [15].

Figure 2: A configuration of Mandible Action Related Rhythm Signals (MARRYS).



References

[1] Erickson, D. and Kawahara, S. (2016) Articulatory correlates of metrical structure: Studying jaw displacement patterns. *Linguistic Vanguard* 2, pp.102-110. De Gruyter Mouton. DOI 10.1515/lingvan-2015-0025.

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

[3] Erickson, D., Iwata, R., Suemitsu, A. (2016) Jaw displacement and phrasal stress in Mandarin Chinese, *TAL 2016.*

[4] Kawahara, S., Erickson, D., Moore, J., Suemitsu, A., Shibuya, Y. (2014) Jaw displacement and metrical structure in Japanese: The effect of pitch accent, foot structure, and phrasal stress. *Journal of Phonetic Society of Japan,* 18.2, pp.77-87.

[5] Huang, Ting, Erickson D. 2019. Articulation of English 'prominence' by L1 (English) and L2 (French) speakers. *International Congress of Phonetic Sciences, Melbourne, Australia.*

[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 of Phonetic Sciences 2015.

[7] Erickson, D., Huang, T., Menezes, C. (subm) Temporal organization of spoken utterances from an articulatory point of view

[8] Erickson, D., Huang, T., Menezes, C., Kawahara, S. (2019) Using jaw movement patterns to visualize prosody, *Prosody Visualisation Challenge 2 (PVC2), International Congress of Phonetic Sciences, Melbourne, Australia,* 2019.

[9] Wilson, I., Erickson, D., Kawahara, S., Monou, T. (2019) Acquiring jaw movement patterns in a second language: Some lexical factors. *Phonetic Society of Japan, Fall Meeting.*

[10] Wilson, I., Erickson, D., Vance, T., Moore, J. (submitted) Jaw Dancing American style: A way to teach English rhythm.

[11] Niebuhr, O. (2020) "Space fighters" on stage - How the F1 and F2 vowel-space dimensions contribute to perceived speaker charisma. *Proc. 31st Conference on Electronic Processing of Speech Signals, Magdeburg, Germany*, 1-12.

[12] Geng, P., Johnson, K., Gu, W., and Erickson, D. (submitted) Fundamental frequency and vocal tract characteristics of American English speech conveying (non-)aggressive attitudes.

[13] Geng, P., Gu, W., Johnson, K., and Erickson, D. (submitted) Prosodic and articulatory characteristics of the Mandarin speech conveying dominance or submissiveness.

[14] Heldner, M. , Włodarczak, M., Branderud, P., & Stark, J. (2019). The RespTrack system. *Proc. 1st International Seminar on the Foundations of Speech, Sonderborg, Denmark, 16-19.*

[15] Lewis, J. R. (2018). The system usability scale: past, present, and future. International. *Journal of Human–Computer Interaction* **34**, 577-590.