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

Donna Erickson[1] , Oliver Niebuhr[2] , Wentao Gu [3] , Ting Huang[1] , Puyang Geng [3]

1 Haskins Laboratories, U.S.A., 2 University of Southern Denmark, Sonderborg/DK, 3 Nanjing Normal University, China

Introduction

➢ 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].

➢ Also, it strongly correlates with **perceptual ratings of syllable stress** [6, 7].

➢ English places larger jaw displacements on metrically stronger stress.

➢ French, Chinese and Japanese place maximal jaw displacement on the final prosodic unit (e.g., syllable) of a phrase.

Jaw transference from L1 to L2

➢ 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]).

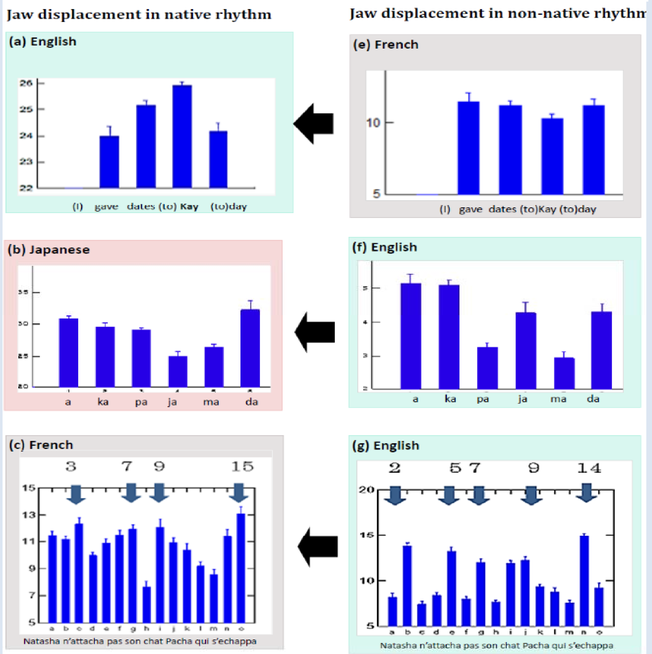


Figure 1: Jaw displacement in native and non-native rhythm. Left column shows jaw patterns of the L1 speakers of an English, a Japanese, and a French sentence, respectively; right columns, the corresponding patterns of L2 speakers of these sentences.

➢ **L1 and L2 patterns are different; L2 patterns reflect the rhythmic structure of their first language.**

Importance of good jaw prosody

➢ The resulting foreign accent of L2 speakers may **socially stigmatize speakers** and/or may **hamper speech communication**, e.g., in terms of identifying words or conveying information-structure or turn-taking signals.

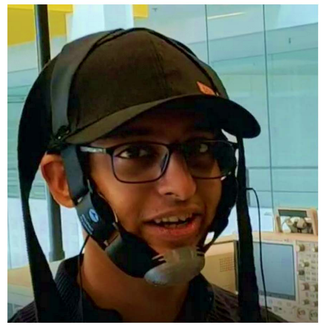
➢ For **public speaking**, pronounced and dynamic jaw movements (esp. lowerings for open vowels) make a speaker sound more passionate and captivating and are, thus, important for his/her charismatic impact on listeners [10].

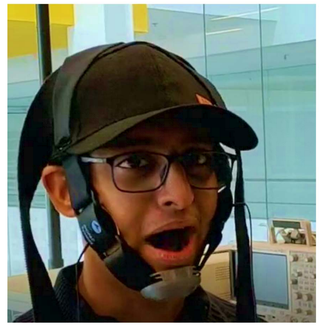
➢ In addition, jaw movement patterns are associated with cross-cultural differences in **expressions of attitudes**, e.g., [11, 12].

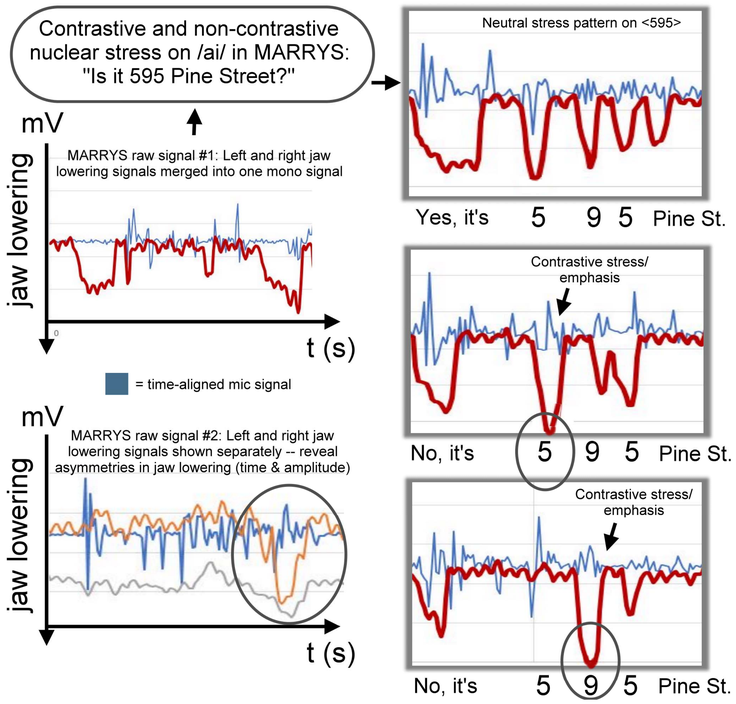
Challenge: How to collect lots of jaw data in a handy and precise way?

➢ 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 expensive, time-laborious and skill-demanding**, both of which limits the number of speakers who can be examined.
➢ Is there another, better, easier, faster way?

New solution: The MARRYS cap (Mandible Action Related Rhythm Signals)







➢ 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 [13]), the MARRYS cap is more handy, mobile, and affordable than EMA and some other apparatus of similar functions.

➢ Recordings are made at a 10 kHz sampling rate and with a 16-bit quantization on an internal SD card; thus, the MARRYS cap needs no external cables and runs with an internal battery for up to 8 hours.

➢ A microphone integrated in the cap's peak allows a **time-aligned recording of jaw and speech signals**; moreover, the speech signal is recorded independently of head movements, i.e. at a constant mouth-to-microphone distance.

➢ 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 [14]).

➢ An interesting further dimension of phonetic analysis: **left-right asymmetries** in the timing and amplitude of jaw lowering → pilot data suggest that this articulatory dimension is highly speaker-specific – interplay with other articulatory patterns? Acoustic consequences?

➢ 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.

➢ We 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].

Pilot data support the validity of the MARRYS concept

➢ So far, gender-balanced pilot speech data have been collected with **8 speakers, 4 native and 4 non-native speakers of English**.
➢ The speech-elicitation tasks were adopted from the works of Erickson and colleagues (e.g., [5, 8]).
➢ Results of the post-processed jaw data (i.e. merged left/right signals, smoothed/filtered, mV translated into millimeters, mm) show a **pleasingly high level of similarity to the EMA-based results** of Erickson and colleagues, see Figure 2.

Comparison of EMA and MARRYS data

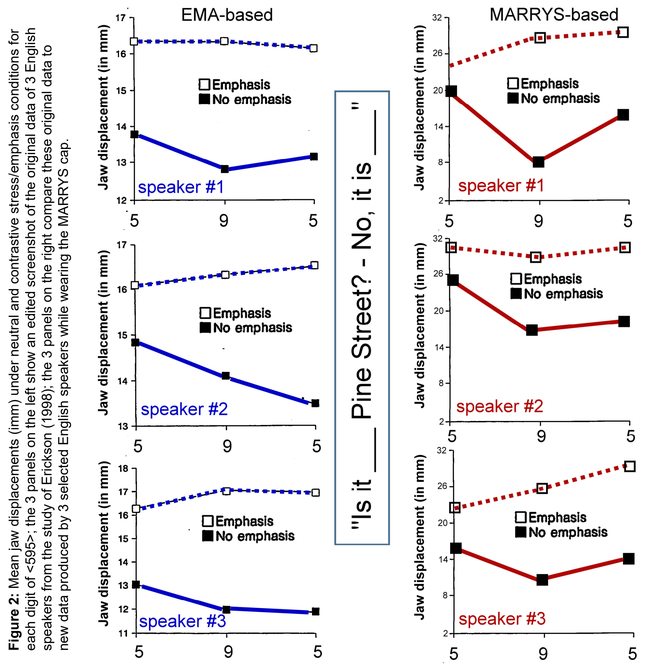


Figure 2: Mean jaw displacements (mm) under neutral and contrastive stress/emphasis conditions for each digit of <595>; the 3 panels on the left show an edited screenshot of the original data of 3 English speakers from the study of Erickson (1998); the 3 panels on the right compare these original data to new data produced by 3 selected English speakers while wearing the MARRYS cap.

Outlook

➢ Collecting additional pilot data and conducting **MARRYS vs EMA comparisons** based on a within-subjects design in **USA, Denmark, and China**;

➢ Refining the **mV-to-mm conversion**;

➢ Using belts with different degrees of elasticity for user comfort and as a means of amplitude normalization, cp. the y-axes of EMA and MARRYS in Figure 2;

➢ Getting an idea about how **effective speaker/L2 training is with the MARRYS cap**. In this context: Developing a **pedagogically informed user interface** (visualization and assessment concept).

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] 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.
- [11] Geng, P., Johnson, K., Gu, W., and Erickson, D. (submitted) Fundamental frequency and vocal tract characteristics of American English speech conveying (non-)aggressive attitudes.
- [12] Geng, P., Gu, W., Johnson, K., and Erickson, D. (2020). Prosodic and articulatory characteristics of the Mandarin speech conveying dominance or submissiveness. Proc. 10th Speech Prosody, Tokyo, Japan, 424-428.
- [13] 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.
- [14] Wilson, I., Erickson, D., Vance, T., Moore, J. (submitted) Jaw Dancing American style: A way to teach English rhythm.
- [15] Lewis, J. R. (2018). The system usability scale: past, present, and future. International. Journal of Human–Computer Interaction 34, 577-590.

Acknowledgements: The development and pilot testing of the MARRYS cap is kindly funded by the Danish Research Council (grant # 9096-000062B) and AllGoodSpeakers ApS. We are also greatly indebted to Mihaela Gruia and Shashwath Suresh who mainly built the cap.