Jaw complex: openness, prominence, and dynamics

Malin Svensson Lundmark[1], *Donna Erickson*[2]
[1] *Centre for Languages and Literature, Lund University, Sweden*
[2] *Haskins Laboratories, New Haven, CT., USA*
malin.svensson_lundmark@ling.lu.se, ericksondonna2000@gmail.com

Abstract

This paper introduces the term *jaw complex* as a way to describe the various functions of jaw openness in speech production. We here present three ongoing studies, in various stages, with preliminary results. The studies cover prominence, phonemic contexts, and dynamics, while utilizing different methodologies (EMA, MARRYS and ultrasound) appropriate for achieving each of the research goals.

Introduction

The lower jaw is constantly active during speech: in opening the oral cavity for vowel productions, and in closing for constrictions of various consonants. If anything, it is absolutely indispensable in the creation of syllables. One can go so far as to say that speech is a string of syllables, produced with the help of the jaw. This approach follows Frame/Content theory where the jaw is the frame which the rest of the articulation fill with content (MacNeilage & Davis 1990).

But, of course, the lower jaw is biomechanically integrated with the rest of the articulation, not least with the lips, and it is hardly possible to see the organs as isolated single articulators. Although we often like to describe bilabials as consisting of lip movements, vowels of tongue body movements, alveolars of the tip of the tongue, etc., one cannot overlook the fact that the jaw participates in all these aforementioned movements. For example, both manner and place of articulation correlate with jaw openness; bilabials usually display a higher jaw than coronals, except for nasals that despite place of articulation usually display a lower jaw (Lindblom 1983; Mooshammer et al. 2007; Kawahara et al., 2014). Vowel production is naturally correlated with the jaw: less mandible lowering for the tongue positions of /i e/ and more mandible lowering (= more displacement) in e.g. /a o/ (Lindblom 1971; Wood 1979; Menezes & Erickson 2013). This also greatly affects the jaw openness degree of the neighbouring constrictions (Lindblom 1983; Mooshammer et al. 2007).

Mandibular openness is also highly correlated with syllable strength, as an effect of localised hyperarticulation (de Jong 1995). Thus, a prominent syllable means faster movements, a more open cavity (Beckman & Edwards 1994), and more acceleration (Svensson Lundmark 2024) compared to a non-prominent syllable, which affects the coarticulatory degree of consonants and vowels (Lindblom et al. 2007).

That higher levels of prominence correlate with jaw opening is most evident when we compare the prosodic metrical structure of different languages (Erickson & Niebuhr 2023; Erickson et al. *accepted*). This research is inspired by Fujimura's C/D model (Fujimura 2000) which purports that prosody is the basis of spoken language, and more specifically, the patterns of mandible lowering give prosodic structure to spoken utterances. This work on jaw (mandible) lowering and language rhythm (prominence) patterns has been done on a number of languages, including English (Erickson et al. 2012), Japanese (Kawahara et al. 2014), French (Erickson and Kawahara 2016; Smith et al. 2019), Mandarin (Erickson et al. 2016), and most recently, Brazilian Portuguese (Erickson et al. *submitted*).

Thus, differences in mouth cavity opening, no matter the reason, affect the movement patterns of other articulators which are all contributing to the acoustical outcome (Lindblom 1971). But, not only spatial positions affect the inter-articulatory coordination: also structured timing relations of speech postures, and the transitions between them, as has been shown using acceleration peaks to separate postures from fast movements of lower jaw, tongue tip and lips (Svensson Lundmark & Erickson 2024).

The lower jaw is evidently not acting in isolation, with both extrinsic and intrinsic effects on speech articulation. For this reason, we introduce the phrase "jaw complex" to symbolize its complex function in speech production. Studies on the jaw complex is an area of research where we are still only seeing the tip of the iceberg. In this paper, we present preliminary results from three ongoing research studies with different approaches to jaw lowering/openness:

1. Jaw openness as an indicator of prominence (rhythm) structure
2. Jaw complex in stressed vs unstressed syllables
3. Jaw complex and tongue: introducing a new method

Material and methods

Study 1: Prominence structure

Currently, the technique that is being used to record jaw movement is EMA, an electromagnetic articulographic set up that, despite its spatio-temporal resolution, has some limitations due to being expensive, immobile, invasive, and time-consuming. A more user-friendly device for mechanical recordings of jaw lowering, the MARRYS, has been developed at University of Southern Denmark (Erickson et al. 2020; Gudmundsson et al. *submitted*). Elastic bands attached to a rigid bike helmet are moving while you speak, and these stretching's are picked up by bending sensors, transformed into signals of more or less relative mandible openness (Fig 1).

Figure 1. MARRYS helmet for mechanical measure of lower jaw displacement.

MARRYS may be used in classroom settings, fieldwork etc. To test its methodological applicability we compared it to EMA (Carstens AG501) in a pilot recording at Haskins Laboratories, New Haven, with 6 American English (AmE) speakers (Svensson Lundmark et al. 2023) (Fig 2).

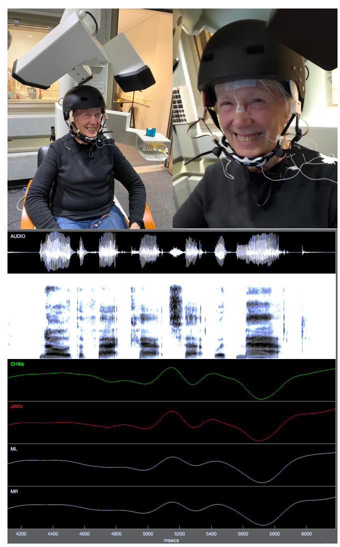

Figure 2. The EMA/MARRYS setup at Haskins Laboratories. Mview: EMA sensors placed on the MARRYS (top) and the jaw, and the MARRYS left and right sensors. The sentence: *No, the fat cat sat with PAT.*

Initial co-collection tests found no interference of signals. The AmE speakers read the following sentence, repeated six times: *The fat cat sat with Mat*. Leading questions read by the experiment leader directed the attention of the focus to either of the monosyllables, such as *Did the THIN cat sit with Mat*? (here: focus on *fat*). So far, preliminary analyses on one speaker in Mview (Tiede 2010) indicate that EMA and MARRYS signals follow each other (Fig 2). See Results section for correlation between jaw openness and prominence structure.

Study 2: Stressed vs unstressed

Here we look at dynamics of jaw openness in stressed and unstressed syllables, using an EMA corpus with 21 South Swedish speakers, recorded with a Carstens AG501 (250 Hz) at the Lund University Humanities Laboratory (detailed information: Svensson Lundmark 2020). The aggregated data set consists of ten speakers and the disyllabic words: /nana/ /mama/ /papa/, with stress on the first syllable (242 tokens). Vertical and horizontal positions of EMA sensors on the lower lip (LL), the tongue tip (TT) and the lower incisors (JW) were used to calculate *displacement* and *peak velocity* during the release from C to V in each syllable. Calculations and statistical analysis (Welch Two Sample t-test) were done in R (R core team 2021).

Study 3: Jaw complex and tongue

The third study focuses on tongue movements and jaw openness. Ultrasound has the advantage that movement of the posterior tongue surface and hyoid can be monitored, which the EMA does not allow. However, even though ultrasound is cost-effective and non-invasive, the analysis of edge detecting is sometimes fooled by imaging artefacts. Combining the high temporal accuracy and point-tracking of EMA with ultrasound has shown promising results (Kirkham et al. 2023). Here we explore combining ultrasound instead with MARRYS, applicable for research questions directly related to jaw openness. Construction and testing of the device were done at the Division of Speech and Hearing Sciences, Queen Margaret University (see more in Results section).

Results

Study 1: Prominence structure

The preliminary analyses show that in the broad focus condition, the speaker produces alternating weak and strong syllables (Fig 3). When producing a contrastive focus, the jaw is lowered for that word. However, in order to maintain the weak/strong pattern across the phrase, jaw openness of the other syllables are also affected, and an already strong syllable (as in *cat* and *Mat*) seems to attract more jaw displacement in contrast to the other words.

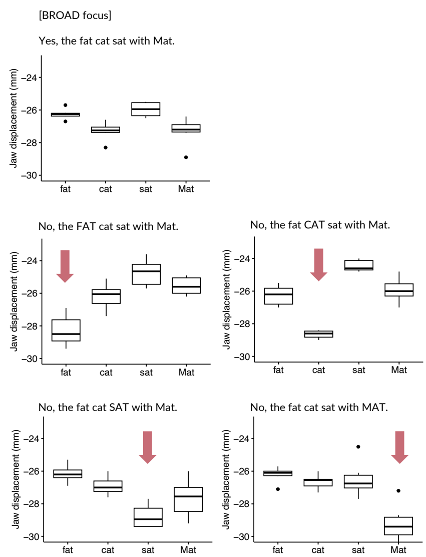

Figure 3. Results from Study 1 on one speaker. Strong/weak syllables in broad focus; below: the jaw lowers more for focus.

Study 2: Stressed vs unstressed

Openness (displacement)

Here we compare movement of the three sensors, JW, TT and LL, specifically releases in /n/, /m/ and /p/ between stressed/unstressed syllables, starting with openness. First, JW lowers slightly more for the stressed syllable than the unstressed ($t = 13.343$, $df = 406.27$, $p < .001$). No difference is found due to type of constriction, be it done with TT (as in /n/) or with LL (as in /m/ and /p/) (Fig 4).

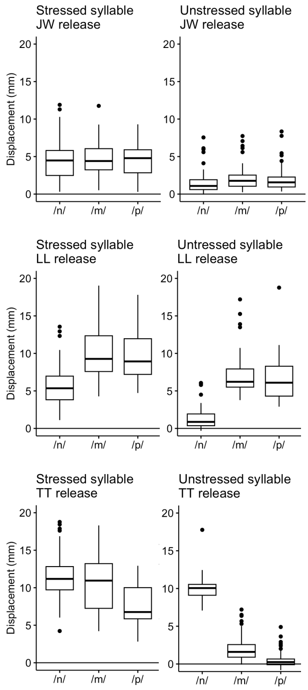

Figure 4. JW, LL, TT displacement (mm) in /n m p/ in stressed/unstressed syllables

LL follows the pattern of JW (Fig 4) and lowers more in the stressed syllable ($t = 9.8$, $df = 344.04$, $p < .001$). However, in /m/ and /p/ there is more LL displacement in both syllables, as an indication of its active role in these constrictions, independent from movement of the mandible (Fig 4). Only in /n/ does LL seem to be in a strict control unit with JW.

/n/ display a large TT displacement (Fig 4), slightly more in the stressed syllable ($t = 2.53$, $df = 63.529$, $p < .05$). However, for the bilabials, where TT is not active, there is a large difference in displacement ($t = 20.32$, $df = 113.84$, $p < .001$). This might be an indicator of TT being part of the tongue organ, and as such follows its movements, which lowers more in a stressed syllable, due to a more lowered jaw (Fig 4).

Speed (peak velocity)

JW speeds up during the stressed syllable (Fig 5). There is a significant difference between the stressed/unstressed condition ($t = -10.87$, $df = 447.21$, $p < .001$), but not between the constrictions.

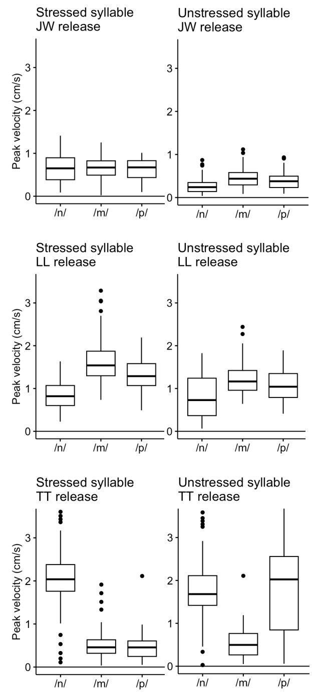

Figure 5. JW, LL, TT peak velocity (cm/s) in /n m p/ in stressed/unstressed syllables

LL is faster than JW, most evidently when it is active (the bilabials) (Fig 5). Some difference is seen between constrictions (/n/ vs /m p/), still, peak velocity of LL shows a significant difference between a stressed and an unstressed syllable ($t = -4.86$, $df = 448.05$, $p < .001$).

TT moves faster in the stressed /n/ ($t = -2.97$, $df = 150$, $p < .01$). But during stressed /m/ and /p/, TT moves almost slower than JW (Fig 5). The exception is /p/ in the unstressed syllable, where TT is moving very fast. This result is unexpected and hard to explain. The landmark seems to be in the right place for the speakers; hence the anomaly is not because of methodological reasons.

Study 3: Jaw complex and tongue

The complex movements of the tongue body are best captured with imaging techniques such as ultrasound. However, the role of the ultrasound probe needs to be evaluated. As it is the reference point in ultrasound imaging, it needs to be stable, i.e. the use of headsets. However, the placement of the probe below the chin may potentially restrict jaw movement. The interference seems to depend on vowel context (Pucher et al. 2020). Here we suggest a novel approach for addressing this relationship, by combining ultrasound with MARRYS (Fig 6). We hope this may be a low-cost approach to also understanding the jaw complex in connection with tongue movements during speech. Stay tuned for future developments and results.

Figure 6. Spare parts of an Ultrasound Stabilisation Headset attached to the MARRYS.

Discussion

Results from Study 1 are only preliminary but show promising robust patterns of jaw openness corelation with metrical tree structure. The weak/strong relations seem to be directly implemented in jaw patterns: when focused strong becomes even stronger, also affecting the weak syllables (Fig 7). It remains to be seen if all speakers show this pattern.

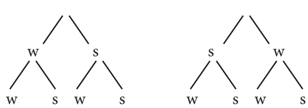

Figure 7. Metrical tree structure: a weak syllable is stronger in a strong foot.

Study 2 shows that jaw is lowered more also for stress. Lips and jaw are commonly referred to as a "control unit" because of the biomechanical connection, but they also act independently (Svensson Lundmark & Erickson 2023). Our results show that the independence, in speed and openness, occurs only when the lips are crucial for a constriction.

The tongue tip is affected by stress in any word, which we assume is because of the tongue in /a/, which can be examined further by comparing to other vowels. Both TT and LL moves faster than JW, but not much more when unstressed, hence a smaller mass cannot fully explain these dynamical patterns. TT's fast release in the word-final plosive remains to be investigated.

Conclusions

While Study 1 shows that mandible lowering interacts with metrical structure, Study 2 shows that other articulators, LL and TT, behave differently as an effect of the mandible lowering in prominent syllables. For future research, we hope Study 3 will shed light on the effect of tongue shape in the interaction between the jaw complex and other articulators.

With this paper we wish to highlight some possibilities and limitations of studying the jaw complex, and how interactive and crucial jaw openness is for speech production end results.

Acknowledgements

This work was supported by an International Postdoc grant from the Swedish Research Council (Grant No. 2021-00334). First author wishes to thank Bryan Gick and Bob Ladd for discussions, and James Scobbie for inspiration and aid with the MARRYS/UIT.

References

Beckman, M., Edwards, J. (1994) Articulatory evidence for differentiating stress categories, in Keating, P. (ed) Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III, Cambridge University Press, pp. 7–33.

De Jong, K., Beckman, M. E., Edwards, J. (1993). The interplay between prosodic structure and coarticulation, Lang Speech 36, 197–212.

Erickson, D., Kawahara, S. (2016) Articulatory correlates of metrical structure: Studying jaw displacement patterns. Linguistic Vanguard 2, pp.102-110. De Gruyter Mouton.

Erickson, D., Niebuhr, O. (2023) Articulation of prosody and rhythm: Some possible applications to language teaching, Studies in Laboratory Phonology. Language Science Presss, pp.1-45.

Erickson, D., Niebuhr, O., Gu, W., Huang, T., Geng, P. (2020). The MARRYS cap: A new method for analyzing and teaching the importance of jaw movements in speech production. Proc. of ISSP, 48-51.

Erickson, D., Rilliard, A., Svensson Lundmark, M., Rebollo Couto, L., Silva, A., de Moraes, J., Niebuhr, O. (submitted) Collecting Mandible Movement in Brazilian Portuguese.

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

Erickson, D., Svensson Lundmark, M., Huang, T. (accepted) Jaw Opening Patterns and their Correspondence with Syllable Stress Patterns

Fujimura, O. (2000). The C/D model and prosodic control of articulatory behavior. Phonetica, 57(2–4), 128–138.

Gudmundsson, V., Gönczi, K., Svensson Lundmark, M., Erickson, D., Niebuhr, O. (submitted). The MARRYS helmet: A new device for researching and training "jaw dancing". Interspeech.

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.

Kirkham, S., Strycharczuk, P., Gorman, E., Nagamine, T., Wrench, A. (2023). Coregistration of simultaneous high speed ultrasound and electromagnetic articulography for speech production research. In Proc. ICPhS, Prague.

Lindblom, B. (1983). Economy of speech gestures. In P. MacNeilage, The production of speech. Springer-Verlag.

Lindblom, B., Sundberg, J. (1971) Acoustical Consequences of Lip, Tongue, Jaw, and Larynx Movement. J. Acoust. Soc. Am. 50(4):1166-79.

Lindblom, B., Agwuele, A., Sussman, H.M., Cortes, E.E. (2007). The effect of emphatic stress on consonant vowel coarticulation. J. Acoust. Soc. Am., 121 6, 3802-13.

MacNeilage, P., Davis, B. (1990). Acquisition of speech production: The achievement of segmental independence. In: Hardcastle, W. J., Marchal, A. (eds), Speech Production and Modeling. Dordrecht: Kluwer Academic, 55-68.

Menezes, C., Erickson, D. (2013) Intrinsic variations in jaw deviation in English vowels. Proc. of International Congress of Acoustics. POMA, Vol. 19, 060253

Mooshammer, C., Hoole, P., Geumann, A. (2007). Jaw and order. Lang Speech, 50(2), 145–176.

Pucher, M., Klingler, N., Luttenberger, J., Spreafico, L. (2020). Accuracy, recording interference, and articulatory quality of headsets for ultrasound recordings. Speech Communication. 123.

R Core Team (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Smith, C., Erickson, D., Savariaux, C. (2019) Articulatory and acoustic correlates of prominence in French: Comparing L1 and L2 speakers. J Phon Special Issue. 77,

Svensson Lundmark, M. (2020). Articulation in time: Some word-initial segments in Swedish, Lunds Universitet,

Svensson Lundmark, M. (2024) Acceleration peaks as representation of activation strength. ISSP2024, Autrans.

Svensson Lundmark, M., Erickson, D. (2023) Comparing apples to oranges - asynchrony in jaw & lip articulation of syllables. In Proc. ICPhS, Prague.

Svensson Lundmark, M., Erickson, D. (2024) Segmental and syllabic articulation. A descriptive approach. J Speech Lang Hear Res.

Svensson Lundmark, M., Ericksson, D., Niebuhr, O., Tiede, M., Wei-Rong, C. (2023). A new articulatory tool: Comparison of EMA and MARRYS. PaPE, Nijmegen, The Netherlands.

Tiede, M. (2010). MVIEW: Multi-channel visualization application for displaying dynamic sensor movements.

Wood, S. (1979). A radiographic analysis of constriction locations for vowels. J Phon, 7(1), 25–43.