Interspeech 2024
1-5 September 2024, Kos, Greece

Magnitude and timing of acceleration peaks in stressed and unstressed syllables

Malin Svensson Lundmark

Lund University, Sweden
malin.svensson_lundmark@ling.lu.se

Abstract

Segment transitions can be accounted for by *acceleration peaks*, which appear at the edges of articulatory speech postures. The present study builds on previous research on timing of acceleration, and expands it by investigating magnitude of acceleration peaks at segment offset, comparing stressed and unstressed syllables. Acceleration peaks of lower lip and lower jaw are measured on ten Swedish speakers using EMA methodology.

Results show strong correlation between acceleration and velocity peak magnitude, with overall more magnitude in stressed syllables, and highest acceleration peaks on lower lip. However, timing of the acceleration peaks, as measured from minimal velocity, is not affected by how fast the articulators are moving.

The results reveal possible different functions of acceleration peak magnitude vs timing, also between the two articulators. The study stresses the significance of mapping acceleration peaks in speech for use in prosodic research.

Index Terms: acceleration peaks, stress, speech production, speech motor control, lip/jaw coordination, articulatory prosody

1. Introduction

Recent research shows that segment transitions can be accounted for by *acceleration peaks*, which appear at the edges of speech postures of primary articulators [1]. Acceleration peaks occur when a mass changes its velocity the most, which it does in connection with changing direction [2]. Specifically, for a place of articulation, an articulator moves at a given speed into a speech posture by a rapid slowing down, a deceleration, where the force added is in the opposite direction of velocity. After the speech posture, for a release of the articulation (release of constriction of the consonant), there is a positive acceleration of articulation, an acceleration peak, meaning the articulator is gaining speed because of the force added is in the same direction as velocity. The deceleration peak and the acceleration peak are important to investigate because they are moments in time signalling maximum activation of an articulator, either of it slowing down or of it gaining speed, respectively.

Investigating acceleration peaks in articulation of syllables is also motivated by previous research on articulatory prosody. Previous findings have shown that articulatory movements (e.g., in the jaw) are faster, larger and longer in duration when syllables are prominent or stressed ([3], [4], [5], [6] and [7]). Furthermore, the present production study is inspired by the concepts of *localized hyperarticulation* which addresses higher articulatory effort in a defined localized area [3], and by *embodied phonetics* where focus is on groupings of muscles and nerves to move structures rather than discrete articulators [8]. In other words, this study takes an overall approach to syllable articulation by assuming that articulation is different between stressed and unstressed syllables because of a difference in overall activation strength: a stressed syllable has more activation, more force added, than an unstressed syllable. More activation strength in stressed syllables would motivate the results on e.g. faster articulatory movements and a lower jaw as indicated by previous research ([3], [4], [5] and [7]). As such, here it is assumed that more activation strength also entails higher acceleration peaks.

In this study speech sounds are restricted to /m/ and /p/. Thus, the primary articulator that is investigated is lower lip. Mouth cavity openness is investigated by including both lower lip and mandible lowering, as motivated by previous research on jaw-lip coordination showing tendencies of both synchronous and asynchronous kinematic behavior ([9] and [10]). Including mandible lowering is also motivated by research showing correlation between jaw displacement and stress levels, e.g. [7]. Acceleration peak magnitude at the release of an onset consonant is used as a main representation of activation strength in the present study. Levels of activation strengths is achieved by comparing acceleration peak magnitude between a stressed syllable and an unstressed syllable. Furthermore, the present study builds on previous research on the timing of acceleration peaks of various articulators and the alignment to segment transitions ([1] and [11]). Hence, results include both the magnitude and the timing of the acceleration peak at the end of the consonantal speech posture, specifically at the release of the onset consonant.

Expected outcome is more activation strength (= higher acceleration peaks) in the stressed syllables than in the unstressed syllables. Higher acceleration peaks are further expected to yield faster articulatory movements. There is no prediction for how the timing of the acceleration peak is related to activation strength, as this has not been previously investigated to the best of our knowledge.

2. Method

2.1. Data collection

The method used is Electromagnetic articulography (EMA). The subset of data is from an EMA corpus with 21 South Swedish speakers producing disyllabic target words in sentence context [12], recorded at the Lund University Humanities Laboratory with the EMA system Carstens AG501 (250 Hz). Speakers read leading questions and target sentences from a prompter, each set repeated eight times. The leading questions was intended to ensure controlled prominence, i.e. that target words were given information. The words used for this study are /mama/ and /papa/, where the first syllable is stressed and the

10.21437/Interspeech.2024-746

second syllable is unstressed. EMA position data was collected from a number of articulators (see detailed information in [12]); the present study uses data from sensors on the lower lip (LL) and the lower incisors (mandible, or lower jaw, JW). For this study data from ten speakers are presented (5 female, 5 male). With some emissions and repetitions due to the experimental procedure, the subset of data used amounts to 165 tokens.

Post-processing of signals was done in Carstens software, after which the data were transferred to R [13] for calculation and analysis. Both vertical and horizontal positions were used to calculate velocity and acceleration. The acceleration was derived by computing the second-order differences of the position data using a time window of 0.02 seconds. The acceleration signal has been filtered and smoothed using a low-pass filter, the R function *loess*. Kinematic landmarks were collected semi-automatically in R. The author visually inspected each speaker's acceleration and velocity landmarks, and adjusted the time frames of the automatic script when justified.

2.2. Measurements

The present study measures the magnitude and timing of the acceleration peak at the release of the onset consonants /m/ and /p/. Levels of activation strengths is achieved by comparing acceleration peak magnitude of the onset release between a stressed syllable and an unstressed syllable. Velocity peak magnitude during the release is included as a control feature, as this has previously been proven to vary dependent on syllable strength ([3] and [4]). As a measure of timing of the acceleration peak the distance from minimal velocity is used, as this is the point where the movement presumably has changed direction and the articulator is moving in the opposite direction.

Figure 1 shows the kinematic landmarks on LL and JW used in this study. For visual purposes, the figure displays only vertical position, velocity and acceleration signals in the word /mama/. The same measures on the release of the onset consonants were collected from both syllables (the first is stressed, and the second unstressed). The landmarks are:

- Acceleration peak magnitude in cm/s^2
- Velocity peak magnitude in cm/s
- Time of acceleration peak in ms
- Time of minimal velocity in ms

The two latter landmarks were used to calculate distance – *Acceleration peak timing* – as a way to compare the acceleration peak timing between stressed and unstressed syllables, and between the two articulators: LL and JW.

2.3. Statistical analysis

All statistical tests are performed in R. Pearson correlation tests are used to assess the relationship between velocity peak magnitude and either acceleration peak magnitude or acceleration peak timing. The correlation of acceleration peak magnitude with velocity peak magnitude is motivated by assuming that more force added would yield a faster movement. Welch Two Sample t-test is used to determine if the difference between stressed and unstressed syllables is statistically significant. As post-hoc tests, to compare differences between /m/ and /p/, pairwise comparisons with t tests are performed (pairwise.t.test in R, which includes the BH adjustment method).

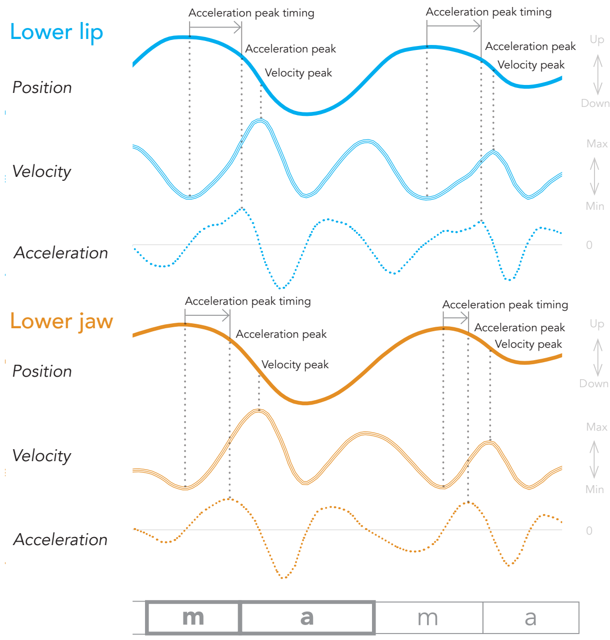

Figure 1: *The kinematic landmarks on timing and magnitude collected from the release of the onset consonant in stressed and unstressed syllables in the example target word /mama/ with signals from sensors on lower lip (LL) and lower jaw (JW)*

3. Results

Results are reported in the following order: first, on acceleration peak magnitude, on velocity peak magnitude, and on the correlation between them. Then follows results on the timing of the acceleration peak, and how that correlates with velocity peak magnitude, i.e. how fast the articulators are moving.

3.1. Acceleration peak magnitude

Figure 2 display the difference in acceleration peak magnitude between a stressed syllable and an unstressed syllable. The acceleration peak is higher in the stressed syllable both for LL (t(297.32) = -5.3, p<.001), and for JW (t(283.74) = -2.8, p<.01). Notably, as can be seen in Table 1, there is overall more acceleration peak magnitude for LL than for JW. This difference between the articulators is statistically significant in both the stressed syllable (t(244.5) = 18.9, p<.001) and the unstressed syllable (t(255.52) = 19.3, p<.001). Acceleration peak magnitude also seem to vary depending on the manner of the constriction; /m/ display higher acceleration peaks than /p/ when measured on the lower lip, in both the stressed (p<.01) and the unstressed syllables (p<.05). The same pattern is only present in JW for the unstressed syllable (p<.01), whereas for the stressed syllable it's the opposite - a higher acceleration peak in /p/ - however, this difference is not statistically significant with the present analysis method (p=.36).

3.1.1. *Correlation of acceleration peak and velocity peak magnitude*

A strong correlation between acceleration peak magnitude and velocity peak magnitude would indicate that more force added yields a faster articulatory movement. This pattern is indeed present for both LL and JW, in both stressed and unstressed

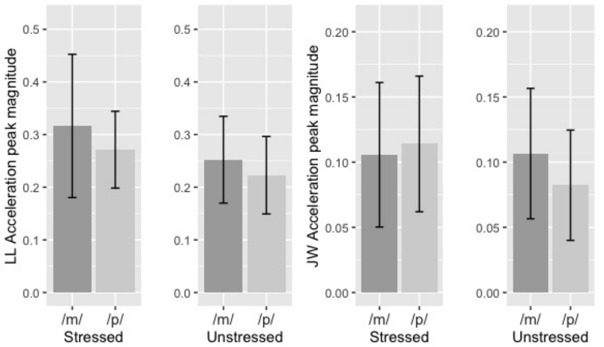

Figure 2: *Barplots showing Acceleration peak magnitude in stressed vs unstressed syllables for lower lip (LL) to the left, and lower jaw (JW) to the right.*

Table 1: *Acceleration peak magnitude in cm/s^2 (mean, standard deviation in italics).*

		LL	**JW**
Stressed	m	0.316 *(0.14)*	0.106 *(0.06)*
	p	0.271 *(0.07)*	0.114 *(0.05)*
Unstressed	m	0.252 *(0.08)*	0.107 *(0.05)*
	p	0.223 *(0.07)*	0.082 *(0.04)*

syllables (Figure 3). However, relationship strength seems to be dependent on syllable strength. The lips display a stronger correlation in a stressed syllable (/m/: r=.91; /p/: r=.79) than an unstressed syllable (/m/: r=.67; /p/: r=.7). A similar pattern can be seen for the JW landmarks between a stressed syllable (/m/: r=.92; /p/: r=.93) and an unstressed syllable (/m/: r=.85; /p/: r=.87).

A closer look at the velocity peak reveals that it follows the pattern of the acceleration peaks (Table 2): higher velocity peaks (=faster movements) are measured in the stressed syllables than in the unstressed syllables, both for LL (t(304.91) = -6.0, $p<.001$), and for JW (t(310.76) = -7.6, $p<.001$). In addition, LL is moving faster than JW as well, in both the stressed syllable (t(244.5) = 18.9, $p<.001$) and the unstressed syllable (t(253.83) = -18.8, $p<.001$). Velocity peak magnitude also seems to differ between the type of constriction, however as with the case of the acceleration peak the difference is more salient in lip movements: higher velocity peaks for LL in /m/ than in /p/ in stressed ($p<.001$) and unstressed syllables ($p<.001$) (Table 2). For JW the difference is not statistically significant (stressed: p=.81; unstressed: p=.06).

3.2. Acceleration peak timing

The *Acceleration peak timing* measures when in time the acceleration peak occur during the consonant release. Here *Acceleration peak timing* is correlated with velocity peak magnitude in order to evaluate both magnitude and timing of the acceleration peak, as these two are essentially the same landmark. A strong relationship between them would indicate that when in time maximal force is added has to do with how fast the articulator is moving. However, this does not seem to be the case. Figure 4 shows that in stressed syllables there is no correlation, or a weak negative relationship, between how fast the articulator is moving and the timing of the acceleration peak, for a stressed LL (/m/: r=.04; /p/: r=-.23) and for a stressed JW (/m/: r=-.31; /p/: r=-.15). The unstressed syllable in Figure 4, on the other hand, displays positive relationships, although weak or moderate, for an unstressed LL (/m/: r=.21; /p/: r=.24) and for an unstressed JW (/m/: r=.56; /p/: r=.13). When comparing *Acceleration peak timing* between the two articulators, a pattern emerges (Table 3). There is a statistically significant difference between a stressed and an unstressed syllable on both articulators; when measured on LL (t(299.45) = 3.3, $p<.01$) and on JW (t(183.45) = 9.6, $p<.001$). Note that the low jaw exhibits approximately twice the distance to the acceleration peak in the stressed condition (Table 3). However, only in the unstressed syllable do we find a difference between LL and JW (t(240.48) = -9.3, $p<.001$), whereas not in the stressed syllable (t(304.91) = 0.3, p=.76). Moreover, when comparing /m/ to /p/, *Acceleration peak timing* only differs on LL, both in the stressed ($p<.01$) and the unstressed syllable ($p<.001$). There is no difference in timing of the acceleration peak of the lower jaw between /m/ and /p/, either in stressed (p=.43) or unstressed syllables (p=.66).

Table 2: *Velocity peak magnitude in cm/s (mean, standard deviation in italics).*

		LL	**JW**
Stressed	m	1.599 *(0.53)*	0.636 *(0.29)*
	p	1.309 *(0.40)*	0.626 *(0.25)*
Unstressed	m	1.247 *(0.39)*	0.455 *(0.23)*
	p	1.079 *(0.34)*	0.387 *(0.20)*

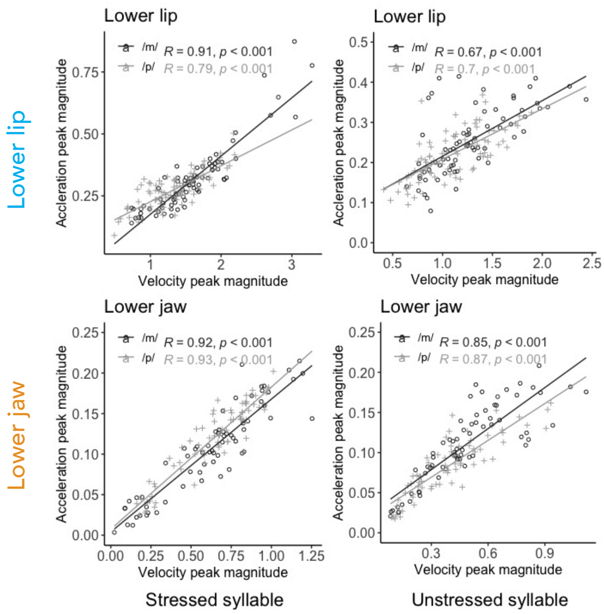

Figure 3: *Correlation of Acceleration peak magnitude and Velocity peak magnitude of lower lip (LL) and lower jaw (JW) in stressed and unstressed syllables.*

4. Discussion

Accelerations peaks seem to be able to represent activation strength as they are higher in stressed syllables, both for lower

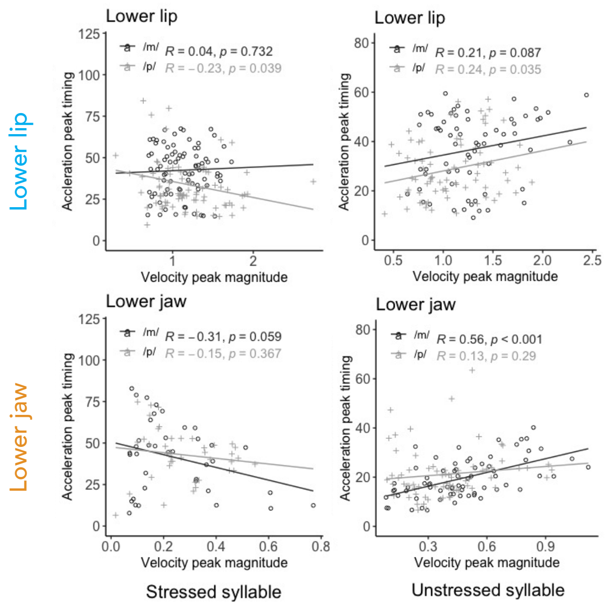

Figure 4: *Correlation of Acceleration peak timing and Velocity peak magnitude of lower lip (LL) and lower jaw (JW) in stressed and unstressed syllables.*

Table 3: *Acceleration peak timing in ms (mean, standard deviation in italics)*

		LL	JW
Stressed	m	42.40 *(14.7)*	40.92 *(25.1)*
	p	34.96 *(20.3)*	37.92 *(19.9)*
Unstressed	m	37.03 *(14.4)*	19.54 *(8.0)*
	p	28.75 *(11.8)*	20.20 *(9.9)*

lip and lower jaw. Both articulators are also faster in the stressed than in the unstressed syllables, as has been shown in previous research. The strong correlation between velocity peak magnitude and acceleration peak magnitude indicate that the amount of force added strongly affects how fast the articulators move. Notably though, the lower lip displays higher acceleration peaks overall than the lower jaw. Previous research has reported faster movements in lip than in jaw movements, which could be explained by the connection between size of mass and velocity (a smaller mass = faster movement; for a discussion see e.g.[14]). However, an alternative interpretation of these results is that the lips are faster because there is simply more force added: they are the primary articulator for constriction release of /m/ and /p/, while the mandible is a secondary articulator. Hence, the level of activation strength might be related to the purpose of the movement. This is a very specific trait of speech; that articulators are assigned different functions related to e.g. features or syllable strength. Thus, how articulator size and acceleration peak height are related (of course also in relation to duration, which is not included here) is not straightforward in speech motor control, and needs more investigation.

The main prevailing hypothesis on articulatory prominence is that the jaw as the syllable articulator governs syllable strength ([7] and [15]). In this regard, it is somewhat surprising that the acceleration peak is higher for lip than for jaw in both syllable strength conditions, suggesting that a stressed syllable is a result of a hyper-articulated lip constriction. Although the present study do acknowledge an activation strength difference on the lower jaw as well, the difference between the stressed and the unstressed syllable is simply more salient when measured on the release of the lower lip. Furthermore, although the timing of the jaw acceleration peak is not affected by the manner of articulation (see also [10]), jaw acceleration peak timing is clearly affected by syllable strength as it is approximately twice the distance when stressed. This discussion prompts further investigation into speech postures and lip/jaw coordination affected by prosody.

Timing of the acceleration peaks are not related to how fast the articulators are moving. As the study shows a very strong correlation between acceleration and velocity peaks, this means that the timing and the magnitude of the acceleration peaks are neither related in a straightforward sense. As previous research has shown, timing of the acceleration peak is aligned with segment transitions, meaning it basically makes up the consonantal segment offset [1]. Thus, for a stressed syllable, in this study, segment offset timing is not varied as a result of more activation strength. However, it is well-known that segments are prolonged in prominent position. So, what is the reason for a longer segment? Could it be that the timing of the acceleration peak, as a representation of the timing of the activation, is functional in a different sense than the level of activation strength? The present results on acceleration peak timing suggest a connection with the timing of different frequencies of phonological levels in the rhythmic and repetitive electrical activity in the brain, or perhaps with fixed time windows. Either way, the magnitude and the timing of the acceleration peak, whatever they represent, are clearly not governed by the same principles as perhaps would be expected.

Future studies should include other articulators, as well as other methodologies, to better capture the overall activation strength changes in orofacial coordination. Comparing different manner of constrictions would further yield insight into the effects of localized hyper-articulation. Limitation of the present study is that displacement of the lips and the jaw are not included which, however, is included in an upcoming paper.

5. Conclusion

This study builds on previous research on the timing of acceleration peaks, and expands this line of research by showing how both acceleration peak magnitude and timing is affected by activation strength. Specifically, lower lip constriction release in a stressed syllable displays high acceleration peaks.

Acceleration peaks are because of rapid transitions of the articulators [1]. These are represented as fast acoustical changes, i.e. segment transitions, which listeners use to decode the message ([16], [17], [18] and [19]). In other words, acceleration peaks seem to possess a unique role in both speech production and perception, which motivates the need for investigating and mapping acceleration peaks in speech, and also reveals why acceleration peaks are especially applicable for prosodic research in speech science.

6. Acknowledgements

This work was supported by an International Postdoc grant from the Swedish Research Council (Grant No. 2021-00334). The author gratefully acknowledges the Lund University Humanities Laboratory.

7. References

[1] M. Svensson Lundmark, "Rapid movements at segment boundaries," *The Journal of the Acoustical Society of America*, vol. 153, no. 3, pp. 1452–1467, 2023.

[2] D. Eager, A.-M. Pendrill, and N. Reistad, "Beyond velocity and acceleration: jerk, snap and higher derivatives," *European Journal of Physics*, vol. 37, no. 6, 2016.

[3] K. J. de Jong, "The supraglottal articulation of prominence in english: Linguistic stress as localized hyperarticulation," *The Journal of the Acoustical Society of America*, vol. 97, no. 1, pp. 491–504, 1995.

[4] D. Mücke and M. Grice, "The effect of focus marking on supralaryngeal articulation – is it mediated by accentuation?" *Journal of Phonetics*, vol. 44, pp. 47–61, 2014.

[5] W. Van Summers, "Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses," *The Journal of the Acoustical Society of America*, vol. 82, no. 3, pp. 847–863, 1987.

[6] T. Cho and P. Keating, "Effects of initial position versus prominence in english," *Journal of Phonetics*, vol. 37, no. 4, pp. 20–24, 2009.

[7] D. Erickson, A. Suemitsu, Y. Shibuya, and M. Tiede, "Metrical structure and production of english rhythm," *Phonetica*, vol. 69, no. 3, pp. 180–190, 2012.

[8] B. Gick, "How bodies talk," in *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia , Aug 2019, pp. 20–24.

[9] V. L. Gracco, "Some organizational characteristics of speech movement control," *Journal of Speech, Language, and Hearing Research*, vol. 37, no. 1, pp. 4–27, 1994.

[10] M. Svensson Lundmark and D. Erickson, "Comparing apples to oranges - asynchrony in jaw & lip artiuclation of syllables," in *Proceedings of the 20th International Congress of Phonetic Sciences*, Prague, Czech Republic , Aug 2023, pp. 1097–1101.

[11] ——, "Segmental and syllabic articulations: a descriptive approach," *Journal of Speech, Language, and Hearing Research*, 2024.

[12] M. Svensson Lundmark, *Articulation in time: Some word-initial segments in Swedish*. Lunds universitet, 2020.

[13] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/

[14] T. Cho and P. Ladefoged, "Variation and universals in VOT: evidence from 18 languages," *Journal of Phonetics*, vol. 27, no. 2, pp. 207–229, 1999.

[15] D. Erickson, M. Svensson Lundmark, and T. Huang, "Jaw opening patterns and their correspondence with syllable stress patterns," *Rhythms of Speech and Language*, accepted.

[16] K. N. Stevens and S. E. Blumstein, "Invariant cues for place of articulation in stop consonants," vol. 64, no. 5, pp. 1358–1368, 1978.

[17] J. Ohala, *The segment: Primitive or derived?* Cambridge University Press, 1992.

[18] W. Strange, "Information for vowels in formant transitions," vol. 26, no. 5, pp. 550–557, 1987.

[19] J. J. Jenkins, W. Strange, and S. A. Trent, "Context-independent dynamic information for the perception of coarticulated vowels," vol. 106, no. 1, pp. 438–448, 1999.