Rapid movements at segment boundaries – preliminary reports on manner

Malin Svensson Lundmark
Centre for Languages and Literature, Lund University, Sweden
Centre for Industrial Electronics, University of Southern Denmark, Denmark
malin.svensson_lundmark@ling.lu.se

Abstract

This paper reports on a one-to-one relation between articulation and acoustics. It explains how segment boundaries are a result of rapid movements of the articulators. In the acceleration profile, this is identified as peak acceleration, which can be measured. A previous study found that rapid movements of an active articulator – peak acceleration – correlate with the acoustic segment boundary in bilabial and alveolar nasals ([m] and [n]). The purpose of the present paper is to extend this line of research and report on some preliminary data on other manner as well ([p], [b], [l]). The results of both studies show that the one-to-one relationship between acoustics and articulation exists both in different places of articulation and in different manner of articulation.

Introduction

This study addresses a specific aspect of the articulation-acoustic relationship. It examines the mystery of two seemingly incompatible and contradictive states: a phonetic system, with physical and dynamic attributes, and a cognitive and static phonological system, a challenge that is central to the development of articulatory phonetics. The present paper concerns a specific aspect of this relationship: a one-to-one connection between movement characteristics of an articulator and the resulting acoustic segments.

Segmental articulation

Phonemes, or speech sound units, are considered to be represented in speech by segments, which are delimited by segment boundaries. The segment boundaries in turn consists of rapid articulatory movements, which result in large acoustic changes (Fant & Lindblom, 1961; Jakobson et al., 1961; Zsiga, 1994). The rapid movement changes of the articulators are local and instant. The consequence of the movements are the segment boundaries, and this causal relationship can be described:

Rapid articulatory movements → acoustic changes → phonological segment boundaries

Peak acceleration

The rapid movements occur when an articulator changes position. For example, in a nasal bilabial stop, the lips accelerate when opening the lips, and decelerate just before closing. Likewise, for a velar consonant produced after an open vowel, tongue tip speed changes can be seen just before and after contact with the palate (Svensson Lundmark, 2020).

The rapid movements are because of large speed changes which occur after a constriction (when the movement is very small), and before the moment velocity is at its highest, *peak velocity* (Figure 1).

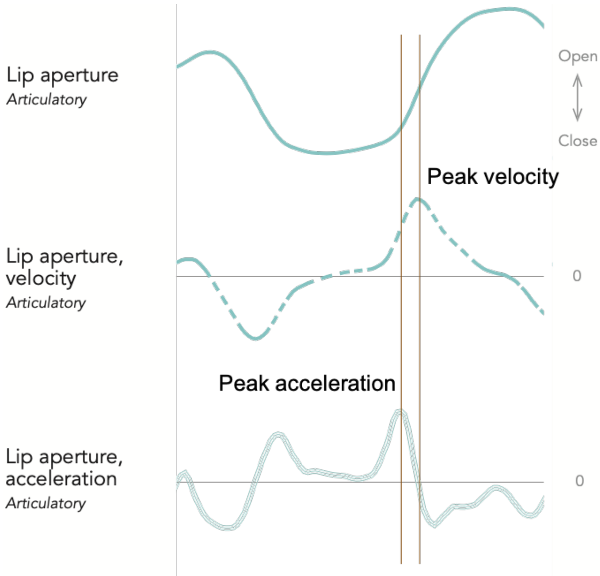

Figure 1: Lip aperture (=the distance between the upper and lower lip, top curvature), velocity profile (middle) and acceleration profile (bottom). The right vertical mark is *peak velocity* (=when velocity is at its highest) of the lip opening. The left vertical mark is *peak acceleration*, which occur between the full closure of the lips and peak velocity.

The speed changes are high velocity changes. These occur because of added energy, resulting in *peak acceleration* or *peak deceleration*, depending on type of force (Eager et al., 2016). Hence, when an articulator slows done before a constriction, there is a peak deceleration, and when the articulator moves away from the constriction, we see a peak acceleration.

As peak acceleration results in rapid movement changes of the articulator, there is a direct connection between peak acceleration and segment boundary. Hence, peak acceleration can be added to the causal relationship of an articulator and its acoustic outcome:

Peak acceleration → rapid articulatory movements → acoustic changes → phonological segment boundaries

Posture intervals

Between the deceleration peak and the acceleration peak, during the constriction, the active articulator stays at its target and has no active movement in any direction. This articulatory interval is hereafter referred to as a *posture interval* (Svensson Lundmark, 2020). A posture interval is somewhat comparable to a *steady-state* where the movement does not change over time.[1] A posture interval is thus delimited by peak deceleration and peak acceleration (Figure 2).

[1] Posture intervals are also comparable to the *"dead" intervals* between phoneme transitions, i.e. segments (Ohala, 1992).

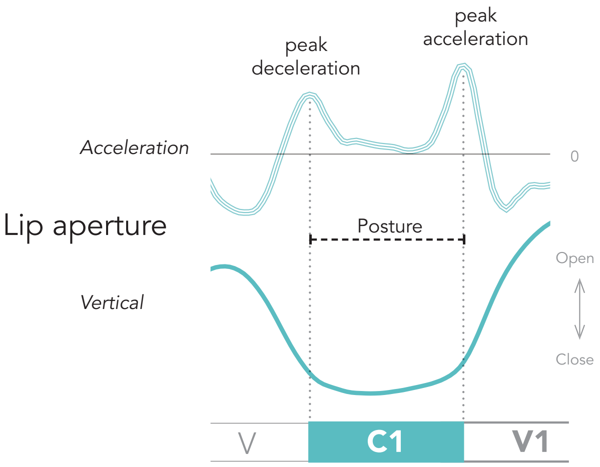

Figure 2: The acceleration profile (top curvature) of lip aperture (=the distance between vertical positions of upper and lower lip) (bottom curvature). Peak acceleration and peak deceleration frame the lip posture interval, which correlates with the bilabial C1 segment.

As the acceleration peaks occur at the same time as the acoustic segment begins and ends, the posture interval correlates with the acoustic segment. In fact, we see yet a causal relationship, this time between a posture interval of any given active articulator and the segment: a lip posture interval would result in a bilabial stop, while a tongue tip posture interval would result in an alveolar stop.

Previous findings

Svensson Lundmark (2022) examined bilabial and alveolar nasals, [m] and [n], in word-initial and word-medial positions in di-syllabic words, comparing acoustic segment duration to posture intervals. Posture intervals were measured on the active tongue tip and lip articulators, whether they were crucial articulators or not. *Active* articulators are voluntary articulatory movements, and they include e.g. the lower lip, the tongue body, the tongue tip. *Crucial* articulator refers to an active articulator where the onset and coda of a syllable is articulated (Fujimura, 2000). For example, the crucial articulator for [n] is the tongue tip, while for [m] is the lower lip.

In Svensson Lundmark (2022) correlation tests were performed on the strength of relationship between the posture interval and the segment duration of both word-initial and word-medial [m] and [n]. Results showed a very strong positive relation between the posture interval of the crucial active articulators, lip and tongue tip, respectively, and [m] and [n], suggesting that the strength of relationship is not dependent on place of articulation. There were some differences between word-initial and word-medial segments, where word-initial velar [n] did not show as strong a relationship as bilabial [m]. Nevertheless, the results indicate that segments boundaries can be determined by timing of articulation, i.e. peak acceleration and peak deceleration (Svensson Lundmark, 2022).

Research questions

As the relationship between peak acceleration and segment boundary has so far only been found in nasal stops, the goal of the present paper is to extend this line of research and present a preliminary report on other manner as well. For example, in voiceless stops, the releasing of air flow might have an effect on the moment in the movement where acceleration is at its highest. Phonation, on other hand, would for example not affect movement characteristics of the crucial articulator, as it is done with a different articulator. The research questions guiding this preliminary study on different manner are as follows:

- Is there a strong relation between lip posture intervals and all bilabials, irrespective of manner?
- Is there a strong relation between tongue tip posture intervals and all alveolars, irrespective of manner?

Method

The present paper includes results on acoustic and kinematic data. This includes acoustic segment duration on the first and the second consonant in a word onset consonant-vowel-consonant (CVC) sequence, and kinematic measurements on lip aperture (=distance between upper and lower lip) and tongue tip.

Speech material

The material was initially recorded for a dissertation project and contains approximately 3000 tokens (Svensson Lundmark, 2020). The corpus consists of target words placed in a low-prominence context, with leading questions placing a narrow (contrastive) focus on the last element of the sentence, instead of on the target word. The target words of the corpus are disyllabic word accent pairs where the CVC onsets are identical apart from the tone (e.g. /mánen/ and /mànar/).

The aggregated data set used for this study consists of six target words, all Accent 2 words, which amounts to 778 tokens. The target word onsets (CVC) of the aggregated data set consist of either of the vowels [iː], [ɑː] and [a], as well as bilabial and alveolar consonants: [m], [b], [p], [l] or [n] (Table 1). The bilabial consonants are always mora-sharing geminates in word-medial position (C2): [mː] and [pː]. In case of a long vowel, the word-medial consonant (C2) is always post-syllabic (marked in parentheses in Table 1).

Table 1: Number of tokens of each target word onset in aggregated data set.

Target word onset	No. of tokens
/mam/	137
/pap/	144
/ma:(n)/	140
/ma:(l)/	147
/bi:(l)/	72
/ba:(l)/	138

Procedure

18 Southern Swedish speakers (12 female) were recorded with electromagnetic articulography (EMA); a Carstens AG501 at the Lund University Humanities Lab. Articulatory data was recorded at 250 Hz, and audio was recorded simultaneously using an external condenser microphone (a t.bone EM 9600) at a sampling rate of 48 kHz.

The articulatory data was collected from sensors placed on the upper lip, the lower lip and the tongue tip (approximately one cm from the tongue tip). To correct for head movements three additional sensors were used: one behind each ear and one on the nose ridge.

The articulatory data was transported to R (R core team, 2015) for further analysis, where it was smoothed using a locally weighted regression with the *loess* R function. So that the smoothing should not cause any distortion, we use a low span (0.1), a value that was determined by visually inspecting the result.

The author segmented the acoustic data manually in Praat (Boersma & Weenink, 2018) using ProsodyPro (Xu, 2013). Segment boundaries were established by examining formant transitions of F1 and F2. In case of doubt, segmentation was determined by perceptual judgment. The oral stop segments include both the complete occlusion and the release burst.

Articulatory measurements

Lip aperture was calculated in R using the three-dimensional Euclidian distance between the sensors on the upper and the lower lip. Articulatory landmarks were then automatically collected from the acceleration profile of the lip aperture calculation, and on the vertical tongue tip position. The landmarks were collected from all target words of the aggregated data set, according to Table 2.

Table 2: The collected articulatory landmarks on the word-initial (C1) and the word-medial (C2) consonants.

		Lips	Tongue tip
C1	Start	Peak deceleration	Peak deceleration
	End	Peak acceleration	Peak acceleration
C2	Start	Peak deceleration	Peak deceleration
	End	Peak acceleration	Peak acceleration

Segmental information collected from textgrid files were used as reference for the time windows (e.g. the script was to search for *peak deceleration in the vicinity of the C1 offset*). Peak acceleration landmarks were without distortion as the acceleration signal was smoothed, and therefore they were also visually prominent in R (see Figure 2 for reference). As a result, the landmarks could be retrieved by the script without much manual adjustment.

Statistical analysis and predicted results

Pearson correlation tests have been used to assess the relationship between segment duration and the posture intervals. Predicted outcome of the correlation test are seen in Table 3. A strong positive relation is expected between the bilabial consonant and a posture done with the lips. The same results are expected for alveolar consonants and postures on the tongue tip. The statistical tests were performed in R (R core team, 2015).

Table 3: Predicted outcome of the Pearson correlation tests.

	Lip posture	Tongue tip posture
/m/ /b/ /p/	Strong positive relation	Weak relation
/n/ /l/	Weak relation	Strong positive relation

Results

Word-initial segment

All target word onsets consisted of a word-initial bilabial consonant. Hence, for all target words the strength of the relationship between the segment and the lip posture interval (the crucial articulator) should be high, while the relationship with the tongue tip posture is expected to be weak. The results from Pearson correlation tests confirm the predictions. The word-initial segment (C1) has a strong positive relationship with the lip posture, in all six target words, irrespective of manner (Figure 3). The relationship is the strongest for /mam/ with a near to perfect relationship (r =.94) between the segment and the lip posture interval, while the weakest relationship is found in /pap/ (r =.69) (Table 4).

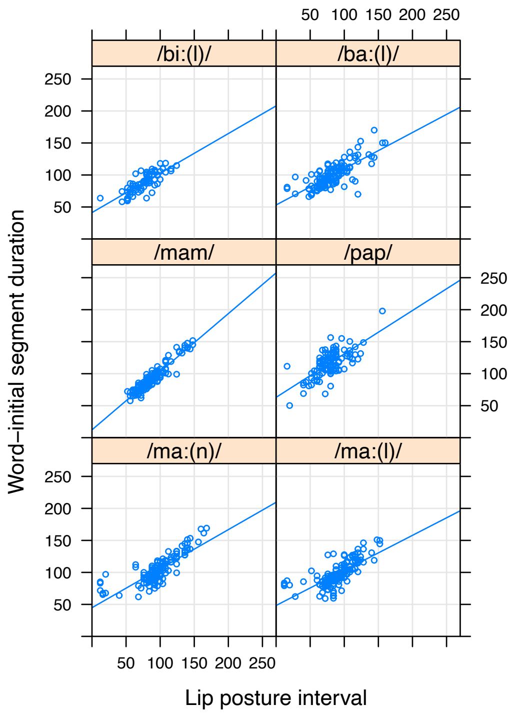

Figure 3: Correlation between acoustic word-initial segment duration (y axis, in ms) and lip posture interval (x axis, in ms).

Table 4: Correlation results between acoustic word-initial (C1) segment and posture interval (=distance between peak deceleration and peak acceleration).

C1		Lip posture			Tongue tip posture		
		df	r	p	df	r	p
[m]	/mam/	132	.94	<.001	77	.47	<.001
	/ma:(l)/	143	.71	<.001	86	.45	<.001
	/ma:(n)/	132	.81	<.001	102	.53	<.001
[p]	/pap/	116	.69	<.001	77	.47	<.001
[b]	/bi:(l)/	66	.79	<.001	55	.35	<.01
	/ba:(l)/	130	.76	<.001	96	.40	<.001

As none of the target words include a word-initial alveolar segment, the segment is not as correlated with the tongue tip posture, although there is a positive correlation (Figure 4). The results show that the C1 segment and the tongue tip posture are weakly, or at best moderately, correlated (Table 4).

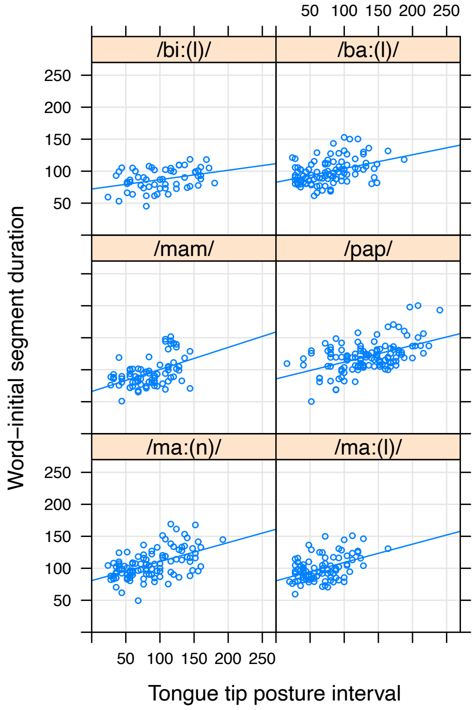

Figure 4: Correlation between acoustic word-initial segment duration (y axis, in ms) and tongue tip posture interval (x axis, in ms).

Word-medial segment

The aggregated dataset consists of both bilabial and alveolar word-medial segments, which enables a comparison of both place and manner of articulation.

Figure 5 displays the correlation between the crucial articulator lip posture interval and the acoustic segments [m] and [p], which both display an equally strong positive relation: r =.88 and r =.87, respectively (Table 5).

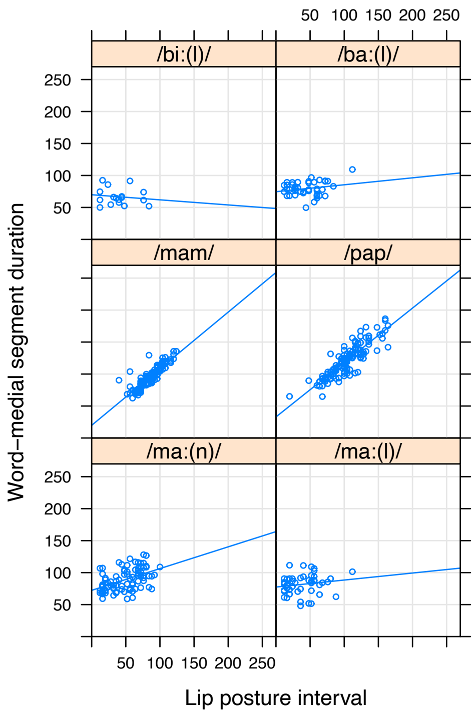

Figure 5: Correlation between acoustic word-medial segment duration (y axis, in ms) and lip posture interval (x axis, in ms).

Table 5: Correlation results between acoustic word-medial (C2) segment and posture interval (=distance between peak deceleration and peak acceleration).

C2		Lip posture			Tongue tip posture		
		df	r	p	df	r	p
[m]	/mam/	124	.88	<.001	116	.22	<.05
[p]	/pap/	141	.87	<.001	118	.61	<.001
[n]	/ma:(n)/	87	.48	<.001	116	.82	<.001
[l]	/ma:(l)/	46	.16	.257	93	.66	<.001
	/bi:(l)/	15	-.14	.597	28	.77	<.001
	/ba:(l)/	40	.22	.157	77	.70	<.001

In the remaining target word onsets, [l] and [n] display a weak positive relation with the lip posture interval, as expected (Figure 5). There's even a negative relation in word onset /bi:(l)/ (Table 5). However, a low number of data points may affect this result. Moreover, in the word onset /ma:(n)/ does the lip posture and the segment [n] display a moderately positive relation (r =.48).

The posture intervals of the crucial articulator tongue tip show a positive relation with the segments [n] and [l] (Figure 6). The correlation appears to be strong or moderate depending on the word onset (Table 5); the strongest relation is found in /ma:(n)/ (r =.82).

The relationship between the segment [m] and the tongue tip is weak, as expected, however the relationship between the tongue tip and the segment [p] is moderately positive (r =.61) (Table 5).

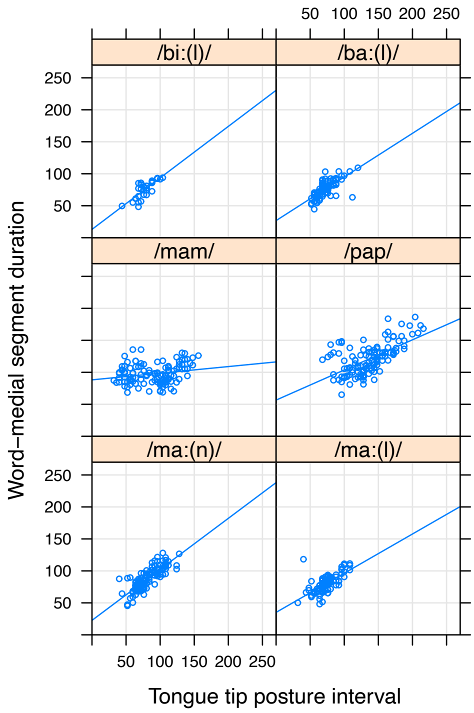

Figure 6: Correlation between acoustic word-medial segment duration (y axis, in ms) and tongue tip posture interval (x axis, in ms).

Discussion

The results confirm previous findings (Svensson Lundmark, 2022) in that consonantal segment boundaries are highly correlated with the acceleration profile of the crucial articulator, irrespective of the articulator (=here, place of articulation). Hence, the distance between peak deceleration and peak acceleration, *the* *posture interval*, has a strong positive relation with the segment – if the posture interval is measured on the crucial articulator. In addition, the results of the present study suggest that the correlation also occurs across different manner of articulation. This means that the proposed one-to-one relationship between peak acceleration and acoustic segment boundary appears to be present in both bilabials and alveolars, in voiced and voiceless, as well as in plosives, nasals and approximants.

The expectation of a weak relationship between a non-crucial articulator and a segment was not entirely met. A weak positive relation was found between the non-crucial articulator tongue tip and word-medial [m], as expected. However, between the non-crucial articulator lip and the word-medial [n] the strength of the relationship was unexpectedly moderately positive. This result is in line with previous findings (Svensson Lundmark, 2022). A strong relationship between lip movement characteristics and the [n] segment duration may be related to the fact that the lips and jaw work together as a control unit (Gracco, 1988; Bose & van Lieshout, 2012). In the jaw opening of an open vowel the lower lip moves as well, irrespective of any status as a crucial or non-crucial articulator. Hence, during construction of the preceding or following consonants, the position of the lip is affected by the position of the jaw. The jaw shows different degrees of opening depending on manner in velar constrictions (Mooshammer et al., 2007). Therefore, it can be assumed that the lower lip, and as a consequence the degree of lip opening, is also affected by different manner of tongue tip articulations. This may explain why lip posture has a stronger relation to [n] than to [l] in this study, as the jaw has been shown to have a lower degree of opening (=less jaw opening) in [n] than in [l] (Mooshammer et al., 2007).

Furthermore, prediction of a weak relation did not occur for any of the word-initial segments [b] or [m] in correlation with the tongue tip. This result was unexpected, but still in line with previous findings (Svensson Lundmark, 2022). Although the active tongue tip is not a crucial articulator at word onset, it seems to be more active than at word-medial position. There can be several explanations for this finding. On the one hand, since each target word follows a verb that contains the final VCV sequences /ade/, the tip of the tongue is a crucial active articulator in the sequences before the target word onsets. Hence, a non-local effect of the tongue tip may have affected the word-initial segments in this study.

On the other hand, it is unclear what word onset entails for active articulators that are non-crucial. Is the tongue tip preparing to move already at word onset, if it's a crucial articulator later in the word? What does that say about motoric planning? And what happens at the syllabic level? That the word-initial bilabial segment correlates better with the acceleration of the lips in a closed syllable (/mam/) than in an open syllable, regardless of manner (/ma:(n)/, /ma:(l)/, and /bi:(l)/, /ba:(l)/, respectively) indicates a difference between syllable structures and/or between vowels. Future studies on the activity of articulators may be able to tell us a lot about speech planning and memory storage in coordination with motor sequences.

Regarding the non-voiced plosive in /pap/, no relations proved to be weak. The large acoustic changes during bursts occur due to the release of complete closure. Since the burst is the result of a rapid articulatory movement (the release) it is reasonable to assume that a peak acceleration occurs before the burst. A possible reanalysis where the segment duration only includes the complete closure, and not the release burst, may possibly provide more justice to these results. Furthermore, since listeners hear acoustic differences between different bursts (Stevens & Blumstein, 1978), it can be assumed that the speed of movement at the release in stop consonants would be different for different manner of articulation. It remains to be seen if the timing of peak acceleration shows this difference, and further how it is related to voice onset time, VOT.

Conclusion

The present paper reports on a one-to-one relation between articulation and acoustics, where phonological segment boundaries are a result of rapid articulatory movements. In the acceleration profile, this is identified as peak acceleration, which is measured and correlated with acoustic segment boundaries. The purpose is to build on previous research and report on a wider range of place and manner of articulation. The results suggest that: a) acoustic segments can be determined by timing of an articulator; b) peak acceleration results in the acoustic changes constituting the segment boundary; c) the one-to-one relationship is not only present in different places of articulation, but also in different manner of articulation.

The next step is to further test the proposed causal one-to-one relationship between peak acceleration and the large acoustic changes at segment boundaries, and to continue to include different manner and places of articulation. The present approach makes it possible to examine and to predict acoustic segment duration regardless of the reason of segmental variation. The approach also enables an alternative method of phoneme analysis that is not necessarily based on segmental phonology or orthographic segmentation, which is unfortunately a possible source for circular arguments.

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 Lab.

References

Bose, A., & van Lieshout, P. (2012). Speech-like and non-speech lip kinematics and coordination in aphasia: Movement kinematics and coordination in aphasia. *International Journal of Language & Communication Disorders,* 47(6), 654–672. https://doi.org/10.1111/j.1460-6984.2012.00171.x

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program] Version 6.0.37. Retrieved from http://www.praat.org/

Eager, D., Pendrill, A.-M., & Reistad, N. (2016). Beyond velocity and acceleration: Jerk, snap and higher derivatives. *European Journal of Physics,* 37(6), 065008. https://doi.org/10.1088/0143-0807/37/6/065008

Fant, G., & Lindblom, B. (1961). Studies of minimal speech sound units. *Speech Transmission Laboratory: Quarterly Progress Report*, vol. 2, 1–11.

Fujimura, O. (2000). The C/D Model and Prosodic Control of Articulatory Behavior. Phonetica, 57(2–4), 128–138. https://doi.org/10.1159/000028467

Gracco, V. (1988). Timing factors in the coordination of speech movements. The Journal of Neuroscience, 8(12), 4628–4639. https://doi.org/10.1523/JNEUROSCI.08-12-04628.1988

Jakobson, R., Fant, G., & Halle, M. (1969). *Preliminaries to speech analysis,* (8th printing). MIT Press.

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

Mooshammer, C., Hoole, P., & Geumann, A. (2007). Jaw and Order. Language and Speech, 50(2), 145–176. https://doi.org/10.1177/00238309070500020101

Ohala, J. J. (1992). The segment: Primitive or derived? In Gerard J. Docherty & D.Robert Ladd, Papers in Laboratory Phonoloy II: Gesture, segment, prosody. Cambridge University Press.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/.

Svensson Lundmark, M. (2022). Rapid movements at segment boundaries [Manuscript in preparation]. Centre for languages and literature, Lund University.

Svensson Lundmark, M. (2020). *Articulation in time: Some word-initial segments in Swedish.* Lund University. https://portal.research.lu.se/ws/files/84061604/Malin_Svensson_Lundmark_Articulation_in_time.pdf

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. The Journal of the Acoustical Society of America, 64(5), 1358–1368. https://doi.org/10.1121/1.382102

Xu, Y. (2013). ProsodyPro—A Tool for Large-scale Systematic Prosody Analysis. in Proc. Tools and Resources for the Analysis of Speech Prosody 2013, Aix-en-Provence, pp. 7–10, 2013.

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, 22(2), 121–140. https://doi.org/10.1016/S0095-4470(19)30189-5