COMPARING APPLES TO ORANGES - ASYNCHRONY IN JAW & LIP ARTICULATION OF SYLLABLES

Malin Svensson Lundmark[1] & Donna Erickson[2]

[1]Lund University, [1]University of Southern Denmark, [2]Haskins Laboratories, Yale University
[1]malin.svensson_lundmark@ling.lu.se, [2]ericksondonna2000@gmail.com

ABSTRACT

This study examines jaw and lip aperture characteristics of syllables. High correlations were found between velocity and displacement of opening/ closing of lips and jaw, similar to previous studies. However, lip aperture opening velocity changed as a function of manner and voicing of syllable onset while jaw opening velocity, except for a plosive onset, did not change. In addition, differences were found in timing between jaw opening and lip aperture opening, and the amount of difference varied as a function of syllable type (open/closed) and voicing/ manner of the consonant. Our interpretation of the asynchronous patterns of lip and jaw articulation is that the jaw is the syllable articulator and the lips are the syllable onset/coda articulators. The findings of this study have application to language acquisition as well as to stuttering research.

Keywords: syllable, peak velocity, displacement, jaw articulation, lip aperture

1. INTRODUCTION

This paper is about the separate linguistic functions of jaw and lip articulations. Jaw-lip coordination is a skill that infants need to acquire in order to produce CV(C) syllables. Studies by Green et al. [1] and [2] suggest that infants first acquire jaw articulation before lip articulation, findings supported by fetal studies showing that mandibular movement occurs before lip muscles are developed (e.g. [3], [4]). These findings are in coherence with MacNeilage's Frame-Content theory ([5]) that babies start producing syllable-type jaw movements at the age of about 6 months, about the time they start chewing. Green et al. [2] report that 1- and 2-year-old children have jaw movements similar to those of adults, but their lip movements show considerably more variability than adults. A similar finding was reported on variability in 4-year-olds' lip movements compared with those of 7- year-olds [6]. A suggestion is that children "rely" on jaw (syllable) production, later adding articulation for syllable onsets, and then codas [2].

Research has shown that correlations between velocity and displacement for both the lips and jaw are strongly significant ([7], [8], [9]). But, less work has been done with regard to opening-closing functions of the two articulators, apart from in Gracco ([10]), where examination of voiced and unvoiced bilabials suggested that the lip closing were more related to consonant production, and the jaw opening with vowel production. [11] reported that the lower lip tends to move with the jaw, but upper lip shows more variability. However, lip opening/closing articulation, i.e., "lip aperture", which includes both upper and lower lips, is rarely the focus of studies. Few studies have researched differences in lip or jaw velocity as a function of manner and voicing of syllable onset consonant, or of syllable type (open vs closed syllable); also, very few findings have been reported about differences between opening and closing movements of syllable articulations. This is surprising considering the suggested different purposes of the jaw and the lips in speech, which would yield more of a rhythmic pattern in the jaw cycle as compared to the recurrent nature of the lips revisiting a place of articulation.

In this study we examine (1) the relationship between velocity and displacement of opening/ closing of lips and jaw for syllable productions, (2) the difference between opening and closing velocity/displacement patterns, and (3) the relative timing of peak velocity of lip articulators compared with that of the jaw articulation.

2. METHOD

The material consists of 566 stressed word initial syllables in the target words: /mama/, /papa/, /maːlar/ and /baːlar/, as spoken by 18 Swedish speakers. The aggregated dataset is part of a larger corpus of EMA data on 21 Swedish speakers and about 6000 sentences [12].

The word initial syllables have vocalic nuclei of either [a] or [ɑː]. The syllable onsets are bilabial [m], [p] and [b]. The coda consonants [m] and [p] in /mam/ and /pap/ are bilabial mora-sharing geminate, while the open syllables /baː/ and /maː/ are followed by a lateral [l] in the next syllable. All words are produced with the Swedish word accent 2 (a tonal rise in the stressed vowel). The target words are placed in statements preceded by leading questions, to ensure a non-focused elicitation.

2.1. Procedure

Speakers were recorded at 250 Hz with electromagnetic articulography (EMA); a Carstens AG501 at the Lund University Humanities Lab. Audio was recorded simultaneously using an external condenser microphone (a t.bone EM 9600); sampling rate 48 kHz. The speaker read a leading question and a target sentence from a computer screen in a random order, each set of target words appearing eight times.

The first author segmented the acoustic data manually in Praat [13] using ProsodyPro [14]. The textgrid files were used in R [15] as reference time windows for collection of the articulatory data.

2.1.1. Articulatory data and measurements

Articulatory data were collected from six sensors: two placed on the upper and lower lips at the vermilion border, one on the lower incisor (jaw), and three sensors placed on the midline of the tongue. To correct for head movements three additional sensors were used: one behind each ear and one on the nose ridge. All data post-processing was done in the Carstens software or in R, where additional analyses were done (see below). Only the sensor on the lips and the jaw were further analyzed in this study.

Peak velocity and displacement on the lips were measured on lip aperture, which was calculated in R using the three-dimensional Euclidian distance between the upper and the lower lip sensors. Peak velocity was calculated during opening at syllable onset (C1), and closing at syllable coda (C2) (Fig. 1). Displacement was measured at C1 and C2 from the maximum distance between the lips during the vowel.

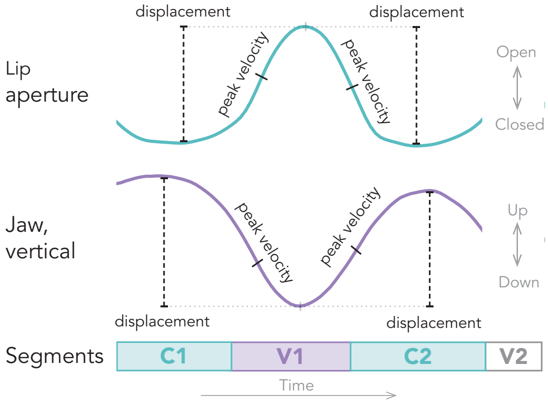

Figure 1: Articulatory measurements and calculations on lip aperture and jaw (example word /mama/ by one speaker).

For the jaw, peak velocity was calculated on the two-dimensional movement (vertical and horizontal) during jaw lowering (opening) and closing (Fig. 1). Displacement was measured as the distance between the highest and lowest vertical position during jaw opening and jaw closing. Notice this is different from studies which measure jaw displacement from the occlusal bite plane (e.g.,[16]). In addition, the relative timing of peak velocity was calculated by subtracting the time of lip peak velocity from jaw peak velocity.

2.3. Statistical analysis

All statistical tests are performed in R [15]. Pearson correlation tests are used to assess the relationship between the two dependent variables peak velocity and displacement. A one-way ANOVA is used to determine whether there is a statistically significant difference between the mean values for the four word initial syllables. As post-hoc tests we use pairwise comparisons with t tests (*pairwise.t.test* in R, which includes the BH adjustment method).

3. RESULTS

3.1. Syllable onset - jaw opening

The main finding from the correlation tests on the jaw is that velocity and displacement have a linear relationship (Fig. 2). The further the jaw moves, the faster it moves. This pattern is seen for all four of the word initial syllables. Thus, the relationship between displacement and velocity is constant, regardless of differences in syllable onset, i.e., /m/, /p/ and /b/.

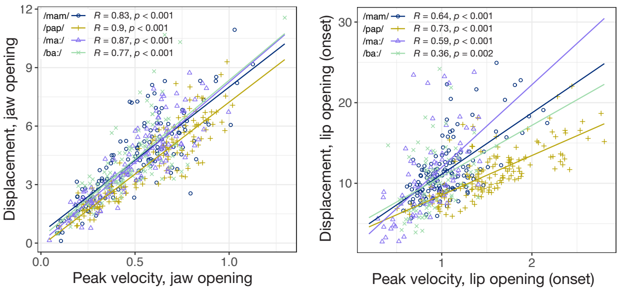

Figure 2: Jaw opening (left) and lip opening (right): displacement (mm) and peak velocity (cm/s).

However, the word initial syllable onsets show some variation in terms of their mean values of peak velocity and displacement, where the jaw appears to be the fastest in /pap/, but moves the longest distance in /mam/ (Table 1). A one-way ANOVA revealed statistically significant differences in jaw peak velocity among the word initial syllable onsets ($F(3,526) = 10.44$, $p < .001$). Pairwise comparisons using t tests with pooled SD found that the mean value of peak velocity was significantly faster for the plosive onset of /pap/ compared to the other word initial syllable onsets; /mam/ ($p < .05$), /ma:/ ($p < .001$) and /ba:/ ($p < .001$) (Table 2). However, for jaw displacement, no statistically significant differences were found between word initial syllable onsets, as determined by the one-way ANOVA ($F(3,526) = 0.511$, $p = .675$).

Dependent variables	Word onsets	Count	Syllable onset				Syllable coda			
			Jaw opening		Lip opening		Jaw closing		Lip closing	
			mean	*sd*	mean	*sd*	mean	*sd*	mean	*sd*
Peak velocity (cm/s)	/mam/	137	0.498	*0.228*	1.06	*0.310*	0.642	*0.246*	1.99	*0.629*
	/pap/	144	0.567	*0.244*	1.56	*0.473*	0.597	*0.247*	1.59	*0.946*
	/maː/	147	0.468	*0.208*	0.806	*0.287*	0.514	*0.216*	1.37	*0.531*
	/baː/	138	0.427	*0.186*	0.856	*0.262*	0.438	*0.204*	1.06	*0.351*
Displacement (mm)	/mam/	137	4.24	*2.06*	11.6	*3.68*	4.04	*1.93*	12.2	*3.58*
	/pap/	144	4.06	*1.98*	11.3	*3.17*	3.84	*2.06*	12.5	*3.44*
	/maː/	147	4.12	*1.89*	10.6	*4.96*	2.01	*1.57*	1.93	*2.03*
	/baː/	138	3.94	*1.94*	10.3	*4.03*	1.82	*1.42*	1.72	*1.88*

Table 1: Mean and standard deviation of the dependent variables: peak velocity (cm/s) and displacement (mm).

	Peak velocity		Displacement	
	Jaw open	Lip open	Jaw open	Lip open
/pap/ vs /mam/	.012	.000	.760	.570
/pap/ vs /maː/	.000	.000	.800	.350
/pap/ vs /baː/	.000	.000	.760	.180
/mam/ vs /maː/	.241	.000	.760	.180
/mam/ vs /baː/	.120	.000	.760	.120
/maː/ vs /baː/	.141	.220	.760	.570

Table 2: Syllable onset opening patterns of jaw and lips: p-values of pairwise comparisons (ANOVA post-hoc test).

3.2. Syllable onset - lip opening

The relationship between peak velocity and displacement for lip movements, compared to jaw opening, appears to be more affected by factors such as the manner and voicing of the onset consonant and whether the syllable has a coda or not (Fig. 2). Although the correlations are all significant, the r values are more robust for /pap/ (r = .73) and /mam/ (r = .64), moderate for /maː/ (r = .59), and weak for /baː/ (r = .36). In addition, velocity differ depending on manner, as well as voicing: during the lip opening, the lips move the fastest in /pap/ (1.56 cm/s, Table 1).

A one-way ANOVA revealed statistically significant differences in peak velocity in the word initial syllable onsets (F(3,562) = 141.9, p < .001). The post-hoc test found that the mean value of lip aperture peak velocity was indeed significantly different between all word initial syllable onsets (p < .001), except for when there was no coda consonant, i.e., between /baː/ and /maː/ (p = .22) (Table 2). Only marginally statistically significant differences were found on lip aperture displacement (F(3,429) = 2.298, p = .077).

3.3. Syllable coda - jaw closing

The results on the jaw closing movement are similar to those for jaw opening: as the jaw displacement is larger, so is the movement faster (Fig. 3). However, the strongest relation is for the closed syllables, /mam/ (r = .87) and /pap/ (r = .85). The disyllabic word initial open syllables, i.e. /maː/ and /baː/ preceding a syllable with a lateral consonant onset, show a less strong relationship (r = .63 and r = .68, respectively). For these words, the jaw perhaps closes only midway, and thus a shorter distance, whereas for the closed syllables, the jaw closes more completely, a longer distance (Fig. 3).

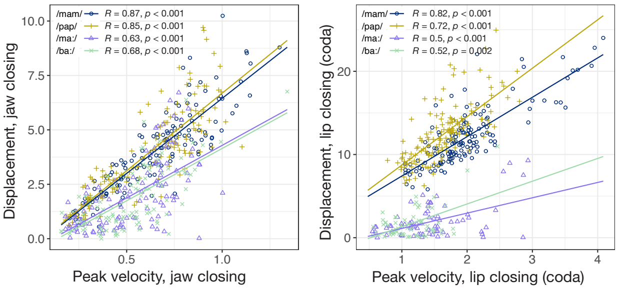

Figure 3: Jaw closing (left) and lip closing (right). Displacement (mm) and peak velocity (cm/s).

The ANOVA confirms that there are statistically significant differences in jaw peak velocity (F(3,562) = 14.35, p < .001), as well as in jaw displacement (F(3,495) = 52.38, p < .001). Interestingly, there is a tendency for the jaw to have the fastest closing for the coda of /mam/ (0.642 cm/s) which is opposite to the pattern of jaw opening (Table 1), where /pap/ shows the fastest jaw opening. Pairwise comparisons using t tests, however, found no statistically significant difference between /pap/ and /mam/ for either peak velocity (p = .123) or displacement of jaw closing (p = .44). Significant differences in jaw closing between the closed syllables (/mam/ and /pap/) and the open syllables (/maː/ and /baː/) were found for the mean values of peak velocity and displacement (Table 3).

	Peak velocity		Displacement	
	Jaw close	Lip close	Jaw close	Lip close
/pap/ vs /mam/	.123	.000	.440	.470
/pap/ vs /maː/	.003	.000	.000	.000
/pap/ vs /baː/	.000	.000	.000	.000
/mam/ vs /maː/	.000	.000	.000	.000
/mam/ vs /baː/	.000	.000	.000	.000
/maː/ vs /baː/	.248	.000	.450	.680

Table 3: Syllable coda closing patterns of jaw and lips: p-values of pairwise comparisons (ANOVA post-hoc test).

3.4. Syllable coda - lip closing

Lip closing displays a strong or very strong displacement-peak velocity correlation in /pap/ (r = .72) and in /mam/ (r = .82) (Fig. 3). Moreover, coda /m/ seems to be faster than /p/, while /p/ has a slightly larger lip opening than /m/ (Table 1). As /maː/ and /baː/ are open syllables with no syllabic coda, lip displacement/peak velocity is expected to be a by-product of the syllabic or vocalic articulations.

A one-way ANOVA revealed that there are statistically significant differences between the words in both peak velocity ($F(3,562) = 84.94$, $p < .001$), and in lip displacement ($F(3,417) = 360.7$, $p < .001$). A post-hoc test found that the mean value of lip closing peak velocity was significantly different between all word onsets ($p < .001$), while lip closing displacement was only significantly different between the closed and the open syllables (Table 3).

3.5. Relative timing of peak velocity

Regarding the relative timing of peak velocity of the jaw and the lips, a one-way ANOVA reveals statistically significant differences between the words during opening ($F(3,533) = 40.64$, $p < .001$), but not during closing ($F(3,416) = 0.179$, $p = .911$). Post-hoc tests show differences in timing during opening both between the type of syllables (open or closed), as well as between types of syllable onsets/codas, i.e., /mam/ and /pap/ ($p < .001$): in /mam/ the jaw peak velocity is timed with the lips, while in /pap/ the jaw peak velocity is instead about 15 ms after the lip peak velocity. When comparing syllable types, the jaw peak velocity occurs before the lip peak velocity in the open syllables, while after in the closed syllables. This difference is about 30 ms, but we also need to look at inter-speaker differences.

4. DISCUSSION AND CONCLUSION

Similarities and differences are observed for jaw and lip articulation of syllables. For both lip aperture and jaw, we see that the faster the movement, the larger the displacement, a relationship that follows the law of nature. However, for lip aperture opening, velocity varies significantly depending on both manner and voicing, and also whether the syllable is open or closed; but for jaw opening, generally there are no such effects. For coda articulation, lip aperture and jaw are both sensitive to syllable-type combinations, i.e., open-closed or closed-open syllables.

The finding for syllable onsets of essentially no difference in jaw opening velocity/displacement patterns due to initial consonant, yet differences in lip aperture due to manner and voicing of syllable onsets, suggests the two articulators have different functions for speech. These results are in accordance with the findings of Gracco ([10]). Although in [10] /p/ was faster during both lip closing and opening, while we found faster /m/ in lip closing. We suspect this difference between our results to be related to the prosodic characteristics of the syllable (in [10] this was not controlled for) which in our case is always a stressed syllable. Thus, the different functions of the lip and the jaw seem to be dependent on syllable prominence; the specific velocities of different manner of articulation may even rely on the level of jaw displacement. Hence our proposal that the jaw is the syllable articulator while the lips are the segmental articulators. See e.g. work by Svensson Lundmark on segmental articulations [17], and work by Erickson and colleagues ([18] and [19]) which proposes that the jaw is the syllable articulator, where the amount of jaw displacement manifests the degree of syllable prominence, once the intrinsic vowel effects have been factored out ([16] and [20]). This interpretation of the jaw as the syllable articulator fits in with pre- and post-natal studies showing that jaw articulation develops before lip articulation.

About timing differences of peak velocity between jaw-lip aperture, specifically, that timing during opening is affected by both syllable type, i.e, open or closed, and manner and voicing of the initial consonant—as far as we know, this specific combination has not been reported in previous studies. Future work is needed to explore timing of lip-jaw opening/closing, especially, with regard to children's acquisition of syllables; also, how this may relate to stuttering in terms of synchronization of jaw-lip articulation for syllable production.

A final note on our data set, which is limited in that the segmental structure does not entail us to look at the lip opening of both onset and coda (e.g., comparing a release burst in /baba/ to /papa/), or vice versa: lip closing of both coda and onset. We are left with comparing apples to oranges, i.e., two halves of recurrent lip constrictions to the more rhythmic jaw cycle. In order to compare the syllable articulator to the segmental articulators, the correct approach would be to compare the jaw cycle (opening + closing) to two constriction cycles: closing and opening of both the onset and of the coda consonant. Moreover, as the lip opening of the jaw cycle is correlated with the type of vowel (narrow, open, rounded), the lip constrictions movement patterns are dependent on not only the jaw displacement/velocity, but also on the type of vowel. Hence, for vowel production, we suggest a combined approach of the jaw, lip and the tongue to be necessary.

5. ACKNOWLEDGMENTS

This work was supported by an International Postdoc grant from the Swedish Research Council (Grant No. 2021-00334), and by an infrastructure grant from the Swedish Research Council (SWE-CLARIN, 2018–2024; Grant No. 2017-00626). The authors gratefully acknowledge the Lund University Humanities Lab.

6. REFERENCES

[1] Green, J. R, Moore, C. A, Higashikawa, M, Steeve, R.W. 2000. The physiologic development of speech motor control: Lip and jaw coordination. *J. Speech Lang. Hear. Res.* 43, 239–255.

[2] Green, J. R., Moore, C. A., Reilly, K. J. 2002. The Sequential Development of Jaw and Lip Control for Speech, *J. Speech Lang. Hear. Res.* 45.1., 66–79. doi:10.1044/1092-4388(2002/005).

[3] Gasser, R.F. 1967. The development of the facial muscles in man. *American Journal of Anatomy* 120, 357–376.

[4] Humphrey, T. 1964. Some correlations between the appearance of human fetal reflexes and the development of the nervous system. *Progress in Brain Research* 4, 93–135.

[5] MacNeilage, P. F., Davis, B.L. 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.

[6] Sharkey, S. G, Folkins, J. W. 1985. Variability of lip and jaw movements in children and adults: Implications for-the development of speech motor control. *J. Speech Lang. Hear. Res.* 28, 8–15.[PubMed: 3982001].

[7] Kelso, J. A. S., Vatikiotis-Bateson, E., Saltzman, E. L., and Kay, B. 1985. A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling, *J. Acoust. Soc. Am.* 77, 266-280.

[8] Munhall, K. G., Ostry, D. J., Parush, A. 1985. Characteristics of velocity profiles of speech movements. *J. Experimental Psychology* 11.4, 457–474.

[9] Kollia, H. B., Gracco, V. L., Harris, K. S. 1995. Articulatory organization of mandibular, labial, and velar movements during speech. *J. Acoust. Soc. Am.* 98, 1313 (1995); doi: 10.1121/1.413468, https://doi.org/10.1121/1.413468

[10] Gracco, V. 1994. Some organizational characteristics of speech movement control. *J. Speech and Hearing Research* 37, 4–17

[11] Van Lieshout, P. H. H. M. 2015. Jaw and lips. In: Redford, M. A. (ed.) *The Handbook of Speech Production*, John Wiley & Sons, Inc. 79–108.

[12] Svensson Lundmark, M. 2020. Articulation in time. Some word-initial segments in Swedish. Lund University

[13] Boersma, P., & Weenink, D. 2018. Praat: Doing phonetics by computer [Computer software]. Version 6.0.37. Retrieved 3 February 2018 from http://www.praat.org/.

[14] Xu, Y. 2013. ProsodyPro—A tool for large-scale systematic prosody analysis. 4. In: Bigi, B. & Hirst, D. (eds.), *Proceedings of tools and resources for the analysis of speech prosody* (TRASP 2013), Aix-en-Provence, France 2013, 7–10. Laboratoire Parole et Langage. http://www2.lplaix.fr/~trasp/Proceedings/TRASP2013_proceedings.pdf.

[15] R Core Team. 2015. R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

[16] Erickson, D. 2002. Articulation of extreme formant patterns for emphasized vowels. *Phonetica* 59, 134–149.

[17] Svensson Lundmark, M. 2023. Rapid movements at segment boundaries. *J. Acoust. Soc. Am.* 153 (3).

[18] Erickson, D. and Kawahara, S. 2016. Articulatory correlates of metrical structure: Studying jaw displacement patterns. *Linguistic Vanguard* 2, 102–110.

[19] Erickson, D. and Niebuhr, O. 2022. Articulation of Prosody and rhythm. *Nordic Prosody 13.*

[20] Williams, J. C., Erickson, D., Ozaki, Y., Suemitsu, A. Minematsu, N., Fujimura, O. 2013. Neutralizing differences in jaw displacement for English vowels, *Proc. of International Congress of Acoustics*. *POMA 19, 060268.*