

# The influence of accentuation and onset complexity on gestural timing within syllables

*Sandra Peters, Felicitas Kleber*

Institute of Phonetics & Speech Processing, Ludwig-Maximilians-University, Munich, Germany

{sandra|kleber}@phonetik.uni-muenchen.de

## Abstract

This paper presents results from a production experiment using electromagnetic articulography. The main aim of the study was to investigate how phrasal accent and the number of onset consonants influence the gestural timing of syllable constituents in German. Five speakers of German with sensors attached to the tongue tip, tongue body and lower lip were recorded reading sentences with either accented or unaccented target words that contained simplex (one consonant) and complex (two consonants) onsets. The nucleus was always /a/ and the coda consonant was always /p/. We analyzed acoustic segment duration and gestural overlap (in terms of lag measurements). Onset complexity influenced both CV and VC overlap and accentuation affected gestural overlap to a greater extent than acoustic vowel duration. However, the extent of overlap differed between segment sequences and accentuation patterns: while for CC and VC sequences trends for greater overlap in deaccented than in accented condition were found, CV overlap decreased with deaccentuation. Shorter plateau durations in this context explain the diminished CV overlap in a prosodically weak context. The findings are discussed with respect to the predictions made by articulatory phonology regarding gestural timing and with respect to timing stability in weak versus strong prosodic contexts.

**Index Terms:** Phrasal accentuation, incremental shortening, Articulatory Phonology, EMA

## 1. Introduction

This study addresses the challenge of investigating the timing of articulatory gestures between all syllable constituents (i.e. onset, nucleus, and coda) in two prosodic conditions and drawing the connection between gestural coordination and variation in acoustic duration.

Segments tend to be shorter in prosodically weaker contexts (i.e. unstressed syllables, deaccented words, at low prosodic boundaries) than in prosodically stronger contexts (i.e. lexically stressed syllables, accented words, at high prosodic boundaries). In particular, vowels were shown to differ in duration with respect to lexical stress and phrasal accent [1]. For consonants, lengthening has been found in the vicinity of high prosodic boundaries [2, 3]. Most studies showing prosody-dependent duration differences are based on acoustic measurements. Acoustic shortening, however, may not always be due to a reduction of the articulatory gestures, but instead to a greater degree of overlap between two gestures that occur in prosodically weak contexts [e.g. 3]. In addition, prosodic reduction effects may also be seen in spatial instead of temporal reduction. In particular, [4] showed that lax as opposed to tense vowels were spatially but not temporally reduced and thus not acoustically shortened.

Segmental duration is not only influenced by prosody but also depends on a couple of other factors. For example, vowels are shorter before obstruents than before sonorants. The duration of the vowel, which usually constitutes the syllable's nucleus, is also affected by the number of the adjacent consonants: vowels flanked by consonant clusters (henceforth complex) tend to be shorter than vowels flanked by single consonants (henceforth simplex). This phenomenon – commonly known as incremental compensatory shortening [5] – seems to be both language specific and dependent on syllable position. According to studies conducted in the framework of Articulatory Phonology (e.g. [6, 7]) vowels are shortened only by preceding but not by following clusters. This prediction arises from the assumption that onset clusters are globally timed, i.e. consonants are organized around a stable midpoint of the cluster – the c-center – causing vowel remote consonants to move away from the nucleus and vowel adjacent consonants to move towards the vocalic nucleus resulting in an increased overlap between the consonant and the following vowel (henceforth CV overlap) [8]. Coda clusters, on the other hand, are assumed to be locally timed, i.e. consonants are in sequential order, therefore not causing any increase in overlap between the vowel and the following consonant (henceforth VC overlap) and vowel shortening. But [9] found a trend towards more vowel shortening before coda clusters in sonorant, accented tokens as opposed to other contexts as well as an influence of accentuation on acoustic vowel duration. In addition, [10] found that onset manner had an influence on VC(C) timing, i.e. the vowel adjacent coda consonant shifted towards the vowel if the onset consonant was a lateral but not if the onset consonant was a nasal.

Previous studies using electropalatography (EPG) and electromagnetic (midsagittal) articulography (EM(M)A) found that prosodic strength affects duration and overlap of consonantal onset clusters. For example, [2] and [3] both found that overall cluster gestures in prosodically weaker contexts were shortened. With respect to overlap, [3] found that consonantal overlap decreased in stressed position and at higher prosodic boundaries. [11], however, reported more overlap when the target words were uttered in focus position than when produced in unaccented position. [11] explained their findings with a shortening of the gesture's plateau in prosodically weaker contexts (instead of indirect shortening due to more overlap).

The aim of this study was to further investigate the interplay between accentuation and articulatory timing patterns with respect to all syllable constituents, i.e. coda and onset consonants (C) as well as the vocalic nucleus (V). We extend our study to prosodic effects on CC, CV and VC overlap and plateau shortening. In particular we are interested in the effects of accentuation and cluster coordination on vowel shortening and VC organization. Three hypotheses were tested:

**H1** There is more CV overlap in words with complex onsets than in words with simplex onsets.

**H2** There is more overlap in the deaccented condition compared to the accented condition concerning all syllable constituents, i.e. CC, CV and VC sequences.

**H3** If there is, however, less overlap in deaccented tokens, than plateau durations should be shortened.

## 2. Method

### 2.1. Speech Material and experimental set-up

The test items were non-existent monosyllabic words (representing monomorphemic words) that contained the lax vowel /a/. The vowel was preceded by either simplex (/n/) or complex (/kn/) onsets. The coda was kept simplex containing the labial stop /p/. All test words were produced in sentence medial position in the carrier phrase “*Melanie’s Omi* [target word] *imitiert ein Lied.*” (“*Melanie’s grandma* [target word] *imitates a song.*”) with only one intermediate (and thus one intonation) phrase. This test sentence was chosen as it allows for optimal tracking of the articulators. Each test word indicated the name of Melanie’s grandma. The test sentence contained one pitch accent that was either on *Melanie’s* (i.e. the target word was deaccented) or on the target word. In order to elicit the test words in accented and deaccented position, we first familiarized the participants with two questions that are each associated with one prosodic pattern. “*Wessen Omi* [target word] *imitiert ein Lied?*” (“*Whose grandma* [target word] *imitates a song?*”) was used to evoke the pitch accent on *Melanie’s* in the deaccented condition. The question “*Welche Omi von Melanie imitiert ein Lied?*” (“*Which grandma of Melanie imitates a song?*”) was used to trigger a pitch accent on the target word in the accented condition. In addition, during the recordings, blue coloured letters drew attention to the words that were to be produced with a pitch accent. The experimenter ensured that subjects corrected any incorrect prosodic patterns.

Articulatory recordings were made in a sound attenuated booth at the Institute of Phonetics in Munich. The movement of speech articulators was tracked using 3D Electromagnetic Articulography (EMA, AG 501) [12]. In total, twelve EMA receivers were attached to various parts of each speaker’s head. The sensors included in the current analysis were those attached to the tongue tip (TT), tongue back (TB) and to the upper and lower lips (LA). Seven repetitions of each sentence were presented isolated in randomized order on a computer screen.

### 2.2. Participants

Five speakers (3 females, 2 males) of Southern Standard German aged between 19 and 25 were recorded. None of them reported any hearing or speaking disorders. Three out of the five participants were undergraduate students of phonetics, but they were naïve as to the purpose of the experiment. One participant was the first author of this paper.

### 2.3. Data analysis

The acoustic data was automatically segmented and labeled using MAuS [13]. Segment boundaries were hand-corrected using praat [14] where necessary and the target word’s accentuation pattern (i.e. accented vs. deaccented) was manually labeled. For the present study, we only analyzed temporal articulatory measures, i.e. specific moments in time of vertical and horizontal movements of the relevant sensors.

The physiological data was labeled using Emu [15]. The target words’ components /k/ and /a/ labels were set based on the vertical movements of the tongue back. For /p/, the lip aperture was calculated as the Euclidean distance between upper and lower lips. For /n/ the tangential velocity of the tongue tip was used. The plateau’s onsets and offsets were defined on the basis of changes in the articulators’ velocity, which are interpolated values and represent the 20% threshold of the difference between two adjacent maxima in the velocity signal. Prior analyses have shown that vowels vary acoustically depending on the presence or absence of a pitch accent on the target word and also on the number of consonants within a cluster (see e.g. [9, 16]). Because of this (potential) variability, we did not use normalization methods usually applied in c-center studies as they normalize on anchor points using either the following coda consonant [7] or the acoustic vowel midpoint (e.g. [17]). Instead we used the normalization method described in [18]. Following this method, we first determined the lag between two neighboring segments and then normalized the lag on the entire duration of the two gestures. For measuring  $CC_{lag}$ , the first consonant’s ( $C1_{on}$ ) plateau offset ( $P_{off}$ ) was subtracted by the second consonant’s ( $C2_{on}$ ) plateau onset ( $P_{on}$ ) and divided by the entire gesture ( $G$ ) as described in the following equation. Using the start and end of the gesture’s plateau was found to be the most stable timing measure, i.e. the one with the lowest variation coefficient [19].

$$CC_{lag} = \frac{P_{on}[C2_{on}] - P_{off}[C1_{on}]}{G_{off}[C2_{on}] - G_{on}[C1_{on}]} \quad (1)$$

The same procedure was applied to the  $CV_{lag}$  and  $VC_{lag}$  with the exception of subtracting the consonant’s ( $C2_{on}$ ) plateau offset by the vowel’s ( $V$ ) plateau onset and the vowel’s plateau offset by the coda consonant’s ( $C1_{off}$ ) plateau onset (cf. equations (2) and (3)), respectively. Thus, lower lag values indicate more overlap.

$$CV_{lag} = \frac{P_{on}[V] - P_{off}[C2_{on}]}{G_{off}[V] - G_{on}[C2_{on}]} \quad (2)$$

$$VC_{lag} = \frac{P_{on}[C1_{off}] - P_{off}[V]}{G_{off}[C1_{off}] - G_{on}[V]} \quad (3)$$

Plateau durations were calculated by subtracting the plateau onset by the plateau offset of each gesture. Figure 1 schematically displays the specific landmarks used for analysis.

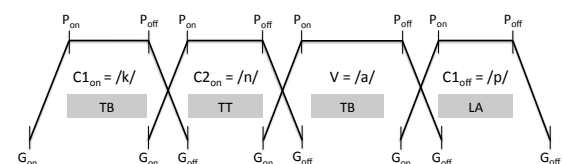


Figure 1: Schematic illustration of gestures and landmarks used in the present analysis.

For the statistical analyses we conducted repeated measures ANOVAs.  $CC_{lag}$ ,  $CV_{lag}$ ,  $VC_{lag}$  and plateau durations of all syllable constituents each served as the dependent variable. ONSET COMPLEXITY (simplex vs. complex) and ACCENTUATION (accented vs. deaccented) were the independent variables. Speaker was entered as a random factor.

### 3. Results

#### 3.1. Acoustic vowel durations

Prior to the articulatory analyses, we tested whether onset clusters and different prosodic patterns acoustically influenced lax vowels. The results show that, especially in the accented condition, there was vowel shortening due to complex onsets (mean = 66.9ms) compared to simplex onsets (mean = 76.3ms). There was no vowel shortening due to deaccentuation. In spite of this result, there may be articulatory shortening (see [4] for discussion).

#### 3.2. $CC_{lag}$ and plateau durations

In order to account for the influence of ACCENTUATION on the temporal organization of onset clusters and plateau durations, we analyzed the  $CC_{lag}$  between /k/ and /n/ in accented and deaccented words (i.e. only words containing complex onsets were used).

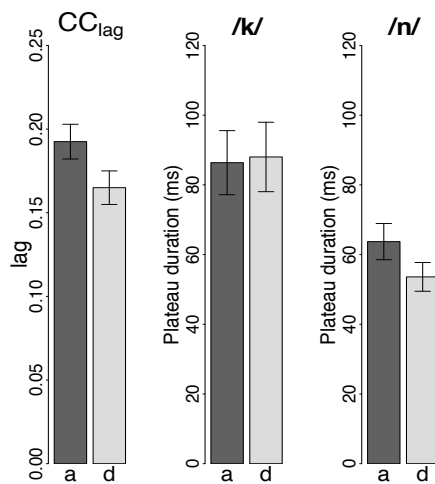


Figure 2:  $CC_{lag}$  (left) between /k/ and /n/, plateau durations for /k/ (middle) and /n/ (right) in accented (a; dark grey) vs. deaccented (d; light grey) tokens.

ACCENTUATION did not significantly affect the timing of onset clusters. However, a trend towards more overlap in terms of lower lag values can be observed in the deaccented condition in the left plot of Figure 2. This means that there was a tendency towards more overlap in deaccented than in accented words.

Concerning plateau duration of /k/ and /n/, there was no significant effect of ACCENTUATION. While the middle plot of Figure 2 shows that plateau durations of /k/ were about the same in accented and deaccented condition, the right plot of Figure 2 indicates tendencies for shorter plateau durations in deaccented words although, again, this was not statistically significant.

#### 3.3. $CV_{lag}$ and plateau durations

In order to investigate the influence of ONSET COMPLEXITY and ACCENTUATION both on the timing between the vowel adjacent onset consonant ( $C2_{on}$ ) and the vowel and on plateau durations, we analyzed words containing both simplex (/n/) and complex (/kn/) onsets.

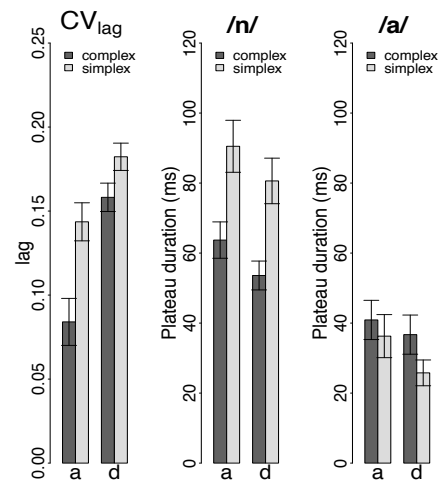


Figure 3:  $CV_{lag}$  (left) between /n/ and /a/, plateau durations for /n/ (middle) and /a/ (right) in accented (a; dark grey) vs. deaccented (d; light grey) tokens.

ACCENTUATION had a significant ( $F[1,4]=18.6$ ,  $p<0.05$ ) influence on the timing of  $C2_{on}$  and the following vowel. There were higher lag values in the deaccented condition (cf. Figure 3). This means, that, in general, there was less overlap in deaccented words compared to accented words.

The independent variable ONSET COMPLEXITY had no significant effect, although, again, a trend towards more overlap in terms of lower lag values can be observed in complex onsets compared to their simplex counterparts (cf. Figure 3), indicating a shift towards the vowel. This trend was much more pronounced in the accented than in the deaccented condition. That is, although there was no significant interaction, the effect of ACCENTUATION, tended to be greater in words containing complex onsets as opposed to simplex onsets (cf. left plot of Figure 3).

Concerning plateau duration of /n/, there was a significant influence of ONSET COMPLEXITY ( $F[1,4]=12.3$ ,  $p<0.05$ ). The plateau duration was shortened in complex onsets compared to simplex onsets in both prosodic conditions. There was no effect of ACCENTUATION.

For /a/, however, only the independent variable ACCENTUATION significantly influenced the plateau duration ( $F[1,4]=67$ ,  $p<0.01$ ). That is, the vowel's plateau was shortened in deaccented words although there was no shortening effect due to deaccentuation in the acoustics. ONSET COMPLEXITY did not effect the nucleus' plateau duration.

#### 3.4. $VC_{lag}$ and plateau durations

In order to investigate the influence of ONSET COMPLEXITY and ACCENTUATION both on the timing between the vowel and the following consonant ( $C1_{off}$ ) and on plateau durations, again, we differentiated between simplex (/n/) and complex (/kn/) onsets.

Concerning VC sequences, the results have shown that neither ONSET COMPLEXITY nor ACCENTUATION had a significant influence on the timing between the vowel and  $C1_{off}$ . However, again, a trend towards more overlap in terms of lower lag values in the deaccented condition can be observed (cf. left plot of Figure 4). That is, there was a

tendency towards more overlap in words produced without a pitch accent compared to words produced with a pitch accent.

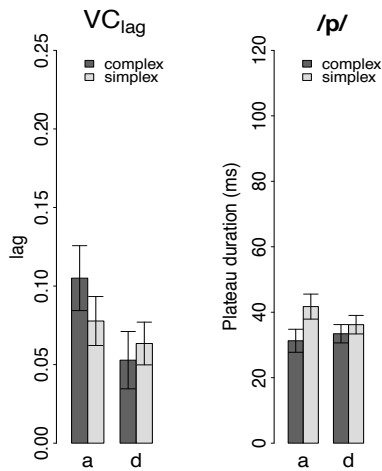


Figure 4:  $VC_{lag}$  (left) between /a/ and /p/, plateau durations for /p/ (right) in accented (a; dark grey) vs. deaccented (d; light grey) tokens.

The vowel plateau duration was significantly influenced by ACCENTUATION (cf. right plot of Figure 3), as described in the section above. Concerning /p/, there was a significant influence of ONSET COMPLEXITY ( $F[1,4]=60.8$ ,  $p<0.01$ ). This means that the coda's plateau was shortened as the number of onset consonants increased.

#### 4. Discussion and Conclusion

The results of the present study are somewhat tentative as only some of the main and none of the interaction effects reached significance. This may be due to the general weak temporal compression effects in lax vowels (see for example [4]). In addition, there were only five speakers producing seven repetitions of four tokens (accented vs. deaccented; simplex vs. complex), i.e. the statistical power may have been too low. Nevertheless, clear trends showed that accentuation differently affects the gestural timing within a syllable depending on onset complexity. In particular, accentuation significantly influenced CV overlap and the vowel's plateau duration whereas onset complexity had a significant effect on the plateau duration of the vowel-adjacent consonants.

Concerning CC sequences, there was a trend towards more overlap in deaccented words. This finding partially supports hypothesis H1 and is in line with [2] and [3]. Despite this trend we found tendencies towards shorter plateau durations of  $C2_{on}$  in the deaccented condition, although [2] found gestural lengthening of this consonant in lexically stressed position. Thus, contrary to the predictions made in hypothesis H3, plateau shortening may also co-occur with more overlap.

There were two findings regarding CV sequences: first, consonants preceding vowels are shifted into the vowel in complex onsets confirming hypothesis H1. This result supports the predictions made by Articulatory Phonology [8]. The second one is that the CV lag is greater in deaccented than in accented tokens, which can be explained by shorter plateau durations of both, /n/ and /a/. This result contradicts hypothesis

H2 and confirms hypothesis H3 and thus is in line with findings described in [11].

In addition, we found that the cluster dependent shift tends to be greater in accented than in deaccented tokens. This may be explained by additional lengthening of the plateau duration in accented as opposed to deaccented tokens, which may then result in a greater shift into the vowel. This is not to say that articulatory lengthening in principal results in more overlap. Articulatory strengthening should cause greater coarticulatory resistance [20] resulting in less overlap (e.g. there was a tendency towards less overlap in accented CC sequences). In this particular case of incremental onset shortening, however, the shift towards the vowel seems to be stronger in the accented condition, because temporal overlap between longer plateau gestures (due to accentuation) is reached faster. This means that the predictions made by Articulatory Phonology depend not only on syllable position (i.e. onset vs. coda clusters) or the cluster's composition (cf. [16]), but also on prosodic structures.

Concerning VC sequences, there was a trend towards more overlap in deaccented words, which is again in line with [2] and [3] and partially supports hypothesis H1. Despite this trend, we found shorter plateau durations of the vowel in the deaccented condition. This again indicates that shorter plateau durations can also co-occur with increased overlap. In addition, the onset complexity influenced the VC timing. There was more CV overlap but less VC overlap in the accented condition. Moreover, the onset composition also affected the plateau duration of the coda consonant, i.e. there were shorter plateau durations of  $C1_{off}$  as the number of onset consonants increases. This indicates that onsets and codas do not behave independently of each other within a syllable (see also [21]) and supports the findings presented in [10].

This study allows no direct comparison of plateau shortening between segments of different articulators, since the tongue tip is flexible resulting in short gestures and long plateau durations for nasals, whereas the tongue back moves slowly resulting in long vowel gestures and short plateau durations for vowels (cf. middle and right plot of Figure 3). For this reason, the results concerning plateau durations have to be extended by also taking into account gesture durations.

In summary, two major conclusions arise from this study. The first one is that lax vowels may be articulatorily compressed due to deaccentuation without showing acoustic shortening. The second one is that syllable constituents are affected differently by prosodic conditions and that onsets and codas do not behave independently of each other within a syllable. These findings need to be included in theories modeling syllable structure such as Articulatory Phonology.

#### 5. Acknowledgements

This research was supported by European Research Council grant number 295573 'Sound change and the acquisition of speech' to Jonathan Harrington.

## 6. References

- [1] Lehiste, I., "Suprasegmentals." MIT Press, Cambridge, 1970.
- [2] Byrd, D. & Choi, S., "At the juncture of prosody, phonology, and phonetics — The interaction of phrasal and syllable structure in shaping the timing of consonant gestures." *Papers in Laboratory Phonology 10*. Mouton de Gruyter, 2010.
- [3] Bombien, L., Mooshammer, C., Hoole, P., Rathcke, T. & Kühnert, B., "Articulatory strengthening in initial German /k/ Clusters under prosodic variation." In *Proceedings of the 16th International Conference of Phonetic Sciences*. Saarbrücken, pp.457-460, 2007.
- [4] Mooshammer, C. and Geng, C., "Acoustic and articulatory manifestations of vowel reduction in German." *Journal of the International Phonetic Association*, vol.28(2), pp.117-136., 2008.
- [5] Katz, J., "Compression effects in English." *Journal of Phonetics*, vol.40(3), pp.390-402, 2012.
- [6] Byrd, D., "C-centers revisited." *Phonetica*, vol.52, pp.263-282, 1995.
- [7] Marin, S. and Pouplier, M., "Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model." *Motor Control*, vol.14, pp.380-407, 2012
- [8] Browman, C. and Goldstein, L., "Some notes on syllable structure in articulatory phonology." *Phonetica*, vol.45, pp.140-155, 1988.
- [9] Peters, S. and Kleber, F., "Compensatory vowel shortening before complex coda clusters in the production and perception of German monosyllables." *Journal of the Acoustical Society of America*, vol.134(5), p.4202, 2013.
- [10] Peters, S. and Kleber, F., "Articulatory mechanisms underlying incremental compensatory vowel shortening in German." Abstract submitted to the 10th International Seminar on Speech Production, Cologne, Germany, accepted.
- [11] Mücke, D., Grice, M. and Kirst, R., "Prosodic and lexical effects on German place assimilation." Poster Presentation at the 8th International Seminar on Speech Production, 8 - 12 December 2008, Strasbourg, France.
- [12] Hoole, P., Zierdt, A. and Geng, C., "Beyond 2D in articulatory data acquisition and analysis." *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, pp.265-268, 2003.
- [13] Schiel, F., "Maus goes iterative.", *Proceedings of the 14th International Conference on Language Resources and Evaluation*, pp.1015-1018, 2004.
- [14] Boersma, P. and Weenink, D., "Praat: Doing phonetics by computer. Software program.", 1992.
- [15] Harrington, J., "Phonetic Analyses of Speech Corpora." Chichester: Wiley-Blackwell, 2010.
- [16] Marin, S., "The temporal organization of complex onsets and codas in Romanian: A gestural approach." *Journal of Phonetics*, vol.41, pp.211-227, 2013.
- [17] Pouplier, M., "The gestural approach to syllable structure: Universal, language- and cluster-specific aspects." In: Fuchs S., Weirich M., Pape D. and Perrier P., (Eds.), *Speech planning and dynamics*. New York, Oxford, Wien: Peter Lang, pp.63-96, 2012.
- [18] Bombien, L., "Segmental and prosodic aspects in the production of consonant clusters – On the goodness of clusters." Ph.D. Dissertation. Ludwig-Maximilians-University, Munich, 2011.
- [19] Oliviera, L., Yanagawa, M., Goldstein, L. and Chitoran, I., "Towards standard measures of articulatory timing." *The Journal of the Acoustical Society of America*, vol.116(4), pp.2643-2644, 2004.
- [20] Cho, T., "Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English." *Journal of Phonetics*, vol.32, pp.141-176, 2004.
- [21] Hawkins, S. and Nguyen, N., "Predicting syllable-coda voicing from the acoustic properties of syllable onsets." In *SWAP-2000*, pp.167-170, 2000.