



## Phonetic indices of syllabic organization in German stop-lateral clusters

**Stavroula Sotiropoulou**, Department of Linguistics, University of Potsdam, Potsdam, Germany, [stsotiro@uni-potsdam.de](mailto:stsotiro@uni-potsdam.de)

**Adamantios Gafos**, Department of Linguistics, University of Potsdam, Potsdam, Germany; Haskins Laboratories, New Haven, CT, USA, [gafos@uni-potsdam.de](mailto:gafos@uni-potsdam.de)

---

Using articulatory data from five German speakers, we study how segmental sequences under different syllabic organizations respond to perturbations of phonetic parameters in the segments that compose them. Target words contained stop-lateral sequences /bl, gl, kl, pl/ in word-initial and cross-word contexts and were embedded in carrier phrases with different prosodic boundaries, i.e., no phrase boundary versus an utterance phrase boundary preceded the target word in the case of word-initial clusters, or separated the consonants in the case of cross-word sequences. For word-initial cluster (CCV) onsets, we find that increasing C1 stop duration or the lag between two consonants leads to earlier vowel initiation and reduced local timing stability across CV, CCV. Furthermore, as the inter-consonantal lag increases, C2 duration decreases. In contrast, for cross-word C#CV sequences, increasing inter-consonantal lag does not lead to earlier vowel initiation and robust local timing stability is maintained across CV, C#CV. In other words, in CCV sequences within words, local perturbations to segments have effects that ripple through the rest of the sequence. Instead, in cross-word C#CV sequences, local perturbations stay local. Overall, the findings indicate that the effects of phonetic perturbations on coordination patterns depend on the syllabic organization superimposed on these clusters.

---



## 1. Introduction

This paper investigates the relation between syllabic organization and inter-segmental coordination in German stop-lateral clusters. Specifically, we examine the effect of syllabic organization on the spatiotemporal coordination patterns of stop-lateral-vowel sequences under different prosodic boundary conditions. Articulatory data were used to register the spatiotemporal form of stop-lateral clusters composed of the same segments parsed in two contrasting syllabic organizations, e.g., word-initial onsets such as /kl/ in /kla:gə/ versus cross-word /k#l/ as in /pək#lɑ:gə/ where only the lateral is in the onset of the syllable /lɑ:/; we will refer to the former prosodification of the /kl/ sequence as global organization and to the latter as local organization (because in the former all three segments /kla:/ are part of one syllable whereas in the latter only the lateral is prosodified as part of that syllable). Some previous studies on German have contributed valuable investigations of spatiotemporal coordination patterns in various consonant clusters (Pouplier, 2012; Brunner, Geng, Sotiropoulou, & Gafos, 2014) but not under different prosodic boundary conditions as we do here. Other studies have investigated intra- and inter-segmental timing patterns under prosodic variation in a subset of clusters in German (Bombien, Mooshammer, Hoole, 2013; Bombien, Mooshammer, Hoole, & Kühnert, 2010) but with a different focus. In terms of materials, these latter studies do not include voiced stop-lateral clusters (/bl/, /gl/), which are prototypical onsets in German, and focus on the effects of prosodic variation and cluster composition on the duration of the consonants and on intra-cluster timing (overlap) but not on how prosodic variation affects coordination patterns in clusters with different syllabic organizations (global versus local). The present study, for the first time, combines the scope of investigation of the previous studies on German. Namely, we examine intra- and inter-segmental properties, such as consonant lengthening and overlap between the consonants of the cluster respectively, of voiced and voiceless stop-lateral clusters under prosodic variation. We also investigate for the first time how the effects of such prosodic variation on these properties depend on the syllabic organization of the consonants (global versus local).

As will be argued in detail in the ensuing analysis of the effects of prosodic variation on intra- and inter-segmental properties turns out to be particularly revealing about the mode of organization in these sequences. This is so because, as will be amply demonstrated in what follows, the mode of global versus local organization presiding over some sequence of segments, is made manifest in a multiplicity of ways (pleiotropy; Sotiropoulou, Gibson, & Gafos, 2020) when properties of these segments or their relations with one another are locally varied. More specifically, the consequences of such local variation on the rest of the sequence are particularly revealing of the mode of organization presiding over such sequences. When local perturbations to segments or relations between adjacent segments have effects that ripple through the rest of the sequence, this is evidence that organization is global. If instead local perturbations stay local, with no consequences for the rest of the whole, this indicates that organization is local.

## 1.1. Background

The main concern of the present study is the relation between syllabic organization and spatiotemporal coordination patterns. Since 1988, when Browman and Goldstein reported facts which indicated that there may be a lawful relation between syllabic organization and the temporal unfolding of the consonants and vowels that partake in that organization (Browman & Goldstein, 1988), a number of follow-up studies have investigated inter-segmental coordination and its relation to syllables in several languages (Honorof & Browman, 1995, Byrd, 1995, and Marin & Pouplier, 2010 on English; Pouplier, 2012 and Brunner et al., 2014 on German; Marin, 2013 and Marin & Pouplier, 2014 on Romanian; Shaw, Gafos, Hoole, & Zeroual, 2009, 2011 on Moroccan Arabic; Hermes, Mücke, & Grice, 2013 on Italian; Pastätter & Pouplier, 2015 and Hermes, Mücke, & Auris, 2017 on Polish). The working hypothesis in these studies has been that, for languages like English and German which admit complex onsets, as the number of consonants in the onset increases from CV to CCV to CCCV, the consonant cluster acts as a unit (hence, it is globally organized) in the way it is coordinated with the subsequent syllable nucleus vowel. The metric proposed to reflect this global organization refers to an interval that spans from the temporal midpoint (the so-called c-center) of the consonant(s) (hence, again, global because this landmark is computed by taking into account all prevocalic consonants) until some landmark usually found at the end of the nucleus vowel, the so-called c-center-to-anchor interval (Browman & Goldstein, 1988; Honorof & Browman, 1995; Byrd, 1995; Marin & Pouplier, 2010; Shaw & Gafos, 2015). In the present study, we refer to this interval as the *global timing interval* because it is an interval that spans the part of the phonetic string that involves the vowel and a landmark computed on the basis of the entire (hence global) preceding consonant cluster. The specific prediction that has been pursued about this interval is that when consonants are added in the onset from CV to CCV to CCCV, the global timing interval remains relatively stable.

We can contrast this prediction to a different one that is best illustrated by considering languages that do not admit complex onsets like Moroccan Arabic and Berber. In these languages, word-initial clusters are syllabified as C.CV where the dot indicates a syllabic boundary (Dell & Elmedlaoui, 2002; Gafos, Roeser, Sotiropoulou, Hoole, & Zeroual, 2020). Thus, the syllable in Moroccan Arabic is hypothesized to exhibit a local span of organization. This local organization has been indexed by an interval spanning from the immediately prevocalic consonant to the end of the vowel, the so-called right-edge-to-anchor interval (Shaw et al., 2009 on Moroccan Arabic; Goldstein, Chitoran, & Selkirk, 2007 and Hermes, Auris, & Mücke, 2015 on Berber). In the present study, we refer to this as the *local timing interval* because it is an interval that spans the part of the phonetic string that involves only the vowel and its immediately prevocalic consonant (to the exclusion of any other consonants that may come before that consonant). The prediction in Moroccan Arabic or Berber that contrasts with that expected in languages like English or German is that as the number of consonants increases from CV to CCV, the local timing interval

remains relatively stable. In sum, local timing interval stability is expected for simplex onsets C.CV where the ‘.’ indicates syllabic boundary and global timing interval stability is expected for complex onsets CCV.

These interval stability patterns, global timing and local timing interval stability, have been considered as the phonetic correlates of syllabic structure and they have been used in follow-up studies to verify syllabic structure in different languages and segmental contexts. We refer to these presumed phonetic correlates as the stability-based heuristics for syllabic structure. However, in several languages, hypothesized complex onsets have not provided consistent evidence for the expected stability of the global timing interval (Hermes et al., 2013 on Italian; Marin & Pouplier, 2010 on English; Pouplier, 2012 and Brunner et al., 2014 on German; Marin, 2013 on Romanian). Similarly, the expected stability of the local timing interval in Moroccan Arabic has not been consistently verified (Shaw et al., 2009). Therefore, a number of studies on the relation between syllabic organization and spatiotemporal coordination patterns have shown that the stability-based heuristics related to syllabic organization often break down. These results largely challenge the idea that syllabic organization has consistent phonetic manifestations in the articulatory record. In sum, what appeared at first as a promising hypothesis has in more recent work been questioned by several sub-results depending on the clusters and the phonetic properties of the segments in that cluster. Granting that cluster-specific phonetics affects the temporal coordination patterns between segments, what role then, if any, does syllabic structure play in articulation?

A first attempt at an elaboration of the directness of the relation between stability-based heuristics and syllabic structure was made by Shaw et al. (2009, 2011) and Gafos, Charlow, Shaw, and Hoole (2014). One main result from the Shaw et al. (2011) study is that in the set of Moroccan Arabic data examined, evidence for local timing interval stability was pervasive but with certain consistent exceptions seen in some corners of the datasets where the local timing interval was less stable than or equally stable to the global timing interval. For example, in the data examined in Shaw et al. (2011), the expected local timing interval stability for Arabic and other simplex onset systems was not met for the /sk/ cluster. Closer examination of the more restricted areas of the datasets that did not exhibit the expected stability patterns revealed that the duration of the prevocalic /k/ gesture across the relevant /kulha/, /skulha/ datasets as well as the vowel duration (of the following /u/) were notably different between the /kulha/ and /skulha/ contexts. Shaw et al. (2011) conducted simulations in order to examine the behavior of the stability-based heuristics as phonetic parameters such as consonant and vowel duration are varied (or scaled, in their terms) in different syllabic organizations. The phonetic parameters (whose effects on the stability of intervals) were investigated consisted of the degree of shortening of the prevocalic consonant and a parameter reflecting degree of vowel compression in two syllabic organization contexts, CCV (complex onset) and C.CV (simplex onset). Their simulation results showed that,

even though as observed in the experimental data the local timing interval may lose its stability advantage over the global timing interval (as for example, when the prevocalic consonant shortens from CV to CCV), the two syllabic organizations still make different predictions about the way stability-based heuristics behave as the chosen phonetic parameters were scaled. These results thus suggest a new perspective on the relation between syllabic structure and phonetic indices. According to that perspective, dubbed the dynamic invariance view in Shaw et al. (2011), the relation between syllabic structure and phonetic indices cannot take the form of statements along the lines of “complex onset organization is manifested by global timing interval stability and simplex onset organization is manifested by local timing interval stability.” This static view may seem to make strong predictions, but as we have reviewed it meets challenges in the data reported so far across different studies: There are clear cases where the expected stability-based patterns are not met in the experimental data. In contrast, the dynamic invariance view proposed in Shaw et al. (2011) distinguishes syllabic organizations by looking at the way in which stability-based heuristics respond to phonetic perturbations (for a precursor, see also Shaw et al., 2009; for a formal treatment, see also Gafos et al., 2014). The predictions are not about which interval, local or global timing, is the most stable. Rather, the predictions are about how interval stabilities change as parameters such as duration of the prevocalic consonant or the vowel are varied. More specifically, the key idea is that it is these patterns of change (in stabilities) that reveal the underlying syllabic organizations. Invariance is not to be found in terms of one interval which remains stable regardless of segmental parameter modifications. For example, in the simplex onset organization, simulations from Shaw et al. (2011) indicate that as prevocalic consonant duration varies (and specifically reduces from CV to CCV), beyond a certain degree of reduction in duration of that prevocalic consonant, the global timing interval may emerge as the most stable interval, even though the underlying organization is that of simplex onsets, which as reviewed above are expected to show local timing interval stability throughout according to the static view; the simulations, in other words, show that omnipresent local timing interval stability cannot be guaranteed under such segmental variability conditions. In contrast, for complex onsets, the same scaling of the phonetic parameter of prevocalic consonant duration resulted in a different prediction in terms of interval stabilities. It is these different patterns of change in interval stabilities that reveal the underlying syllabic organization. This dynamic invariance view can thus distinguish between different syllabic organizations even in cases where the stability-based heuristics of the static view break down.

However, the predictions of Shaw et al. (2009, 2011) and Gafos et al. (2014) have so far remained largely untested with experimental data beyond the data considered in developing the dynamical invariance view in these studies. Nevertheless, whatever work there is in other languages does suggest that interval stabilities are modulated by various factors. For example, Hermes et al. (2017) used simulated data on Polish where variability was introduced in the

anchor landmark (see also Shaw et al., 2011; Gafos et al., 2014). The computational model they used provides an evaluation of how well the empirical data fits the simulated complex onset organization as certain parameters are varied. The results showed that the complex onset organization provides a reasonable match to the empirically observed interval stability pattern but, after some threshold of variability, the stability pattern changes to one that does not correspond to what is expected of complex onsets.

Following the example of Shaw et al. (2011) using simulated data, the present study investigates the relation between syllabic structure and phonetic indices as phonetic parameters are scaled using actual experimental data on German. Scaling phonetic parameters is a crucial aspect in the Shaw et al. (2011) approach as well as in our present experimental study, because arguably it is only under such scaling that predictions of distinct syllabic organizations may be revealed. Using simulations, Shaw et al. (2011) investigated changes in the shortening of the prevocalic consonant and vowel compression and studied the consequences of such variability in these parameters for stability indices of syllabic organization. Here, we pursue a parallel approach experimentally. That is, we have pursued the consequences of varying three seemingly unrelated parts of the sequence over which syllabic organization is assessed: the lag between the constrictions of the two consonants, what we refer to as the interplateau interval or IPI, the duration of the initial stop consonant, and the duration of the prevocalic consonant duration in word-initial CCV and cross-word C#CV sequences where # indicates word boundary. Different degrees of IPI and C1 plateau duration were elicited by embedding the target words with the consonant cluster of interest in carrier phrases with varying prosodic boundary strength. Here, we followed the experimental design of Byrd and Choi (2010) on English who found that as the prosodic boundary strength increases, the lag between the two consonants increases with the effect being stronger for heterosyllabic CC clusters than tautosyllabic CC onset clusters. For C1 duration, Byrd and Choi found that both in onsets and in heterosyllabic clusters, C1 lengthens with increasing boundary strength. Byrd and Choi's (2010) prior work on English thus provided us with reasonable evidence that an experimental design along their lines, but implemented in German, may induce analogous variability in the parameters we seek to vary. If this turns out to be the case, it would imply that we have generated the appropriate experimental test to pursue the simulation predictions of Shaw et al. (2011) in German.

Therefore, a prerequisite for the main goal of our study is to examine effects of prosodic boundary strength on the duration of the initial stop, on the duration of the prevocalic lateral, and on the IPI in consonant clusters of different syllabic affiliation in German. These are the three parameters we attempted to vary in the present study as a result of our experimental manipulations using different boundary strengths. We did so because, ultimately, the main aim is to examine if and how such variability in these parameters affects the stability-based heuristics in clusters of contrasting syllabic affiliations. Furthermore, beyond pursuing so far untested

predictions of the Shaw et al. (2011) perspective, we will also extend the Shaw et al. (2011) study in the following sense. The present study does not only consider the effects of prosodic variation on stability-based heuristics, but also on other measures, such as the relation between duration of the prevocalic consonant and IPI, and vowel initiation with respect to prevocalic consonant(s), for the first time.

The paper is organized as follows. First, an outline of the methodology and the German stimuli is provided in Section 2. A detailed analysis of the data including the changing of phonetic parameters under prosodic strengthening and the spatiotemporal coordination patterns that emerge in clusters of different syllabic affiliation follows in Section 3. Section 4 summarizes the results and argues that syllabic organization makes predictions regarding the way various phonetic indices change as phonetic parameters are scaled. We conclude in Section 5 with implications of our results for the issue of the relation between syllable structure and phonetic indices related to spatiotemporal coordination patterns. We argue that a joint consideration of various measures, such as prevocalic consonant shortening, vowel initiation, and compensatory effects between IPI and prevocalic consonant duration proves to be highly informative for understanding the ways in which syllabic organization is expressed in the phonetics. Specifically, syllabic organization (word-initial versus cross-word sequences) makes different predictions about the way phonetic indices respond to perturbations of phonetic parameters such as IPI or C1 lengthening.

## 2. Methods

### 2.1. Subjects

Articulatory data using the Carstens AG501 Articulograph were collected from five native German subjects (vp01, vp02, vp03, vp04, vp05), one male and four females, between 20 and 35 years old. All subjects reported no speech or hearing problems. They provided written informed consent prior to the investigation and they were reimbursed for their participation. The experiment took place at the Speech Lab at the University of Potsdam. All experimental procedures were approved by the Ethical Committee of the University of Potsdam (application number 62/2016).

### 2.2. Speech material

The corpus consists of real disyllabic words in German starting with consonant clusters (CC) or single consonants (C) with stress on the initial syllable. CC-initial words begun with stop-lateral clusters. Their paired single consonant-initial words begun with a lateral such that in a CV~CCV pair the prevocalic consonant remained the same across CV~CCV words (e.g., /la:gə/~ /pla:gə/). The cluster occurred word-initially as in CCV sequences and across words as in C#CV where # indicates a word boundary. Thus, for the cluster /pl/, the word /plage/ corresponds to a word-initial CCV sequence and the two-word combination /knap#la:gə/

corresponds to the C#CV sequence. Both of the CCV and C#CV sequences were paired with the same CV sequence as a separate stimulus, e.g., /la:gə/ for the example above. The word-initial stop-lateral clusters consist of /bl, gl, pl, kl/ where the initial stop is either a voiced or a voiceless stop. The cross-word sequences consist of only the voiceless stop-lateral clusters /pl, kl/ because German has final devoicing and thus voiced stops do not occur in the word-final position. The vowel following the cluster or the single consonant of interest is always a tense vowel. The postvocalic consonant was maintained the same across the two words within a pair.

Each stimulus word was recorded in two prosodic conditions with varying boundary strength preceding the stimulus word. For the cross-word sequences, the prosodic boundary is between the two consonants of the cluster and thus it precedes the second word. The boundary strength is increasing from no boundary (or word boundary) to utterance phrase boundary. The word boundary condition was elicited by embedding the stimulus word in the carrier phrase *Ich sah \_\_\_ an* ('I looked at \_\_\_') with the stimulus word location indicated by the "\_\_\_." The utterance phrase boundary condition was elicited by embedding the stimulus word in the carrier phrase *Zunächst sah ich Anna. \_\_\_ sagte sie* ('First I saw Anna. \_\_\_ she said'). For the cross-word C#CVs, the carrier phrases were *Ich sah \_\_\_ an* for the word boundary condition and *Zunächst sagte Anna . . . war das nächste Wort* ('First Anna said . . . was the next word') for the utterance phrase boundary condition. In these phrases, the stimulus word is always preceded by a low vowel. During the experimental session, these phrases were presented in a randomized order; there were eleven or twelve blocks per session so that each session lasted approximately 1.5 hours. Thus, each subject produced eleven or twelve repetitions of each item ( $N = 9$ ) in two prosodic conditions yielding a total of approximately 200 tokens per subject. **Table 1** presents the list of complex onsets (CCV), simplex onsets (C#CV), and the respective singleton (CV) words.

We illustrate our materials with examples of the two contrasting organizations, local organization versus global organization. The local organization context is C#CV, where the two consonants belong to separate words, as in *Ich sah "pack Lage" an* or *Zunächst sagte Anna "pack." "Lage" war das nächste Wort*. Here, the first C of the C#CV (the /k/ in *pack*) is not part of the same syllable as the downstream vowel and hence no global organization presides over the k#la sequence. Instead, a local organization presides, where the vowel is related only to its immediately prevocalic consonant (/l/ in *Lage*). This C#CV context is to be contrasted with the global organization in the CCV context as in the utterances *Ich sah "Klage" an* or *Zunächst sah ich Anna. "Klage" sagte sie*. In the CCV case, the first C (/k/ in *Klage*) is part of the same syllable with the subsequent vowel and hence global organization presides over the /kla/ sequence. In a global organization, the vowel is related to both tautosyllabic consonants preceding the vowel, that is, the vowel /a/ is globally organized with the entire cluster in /kla/.

Cluster	CCV	C#CV	CV
pl	Plage 'plague'	knapp Lage 'scarce location'	Lage 'location, position'
bl	blasen 'to blow'	—	lasen 'to read' (pret.)
gl	gleite 'to glide' (imper.)	—	leite 'to lead' (imper.)
kl	Klage 'complaint'	Pack Lage 'box position'	Lage 'location, position'

**Table 1:** German stimuli.

### 2.3. Data acquisition

The data were acquired by means of Electromagnetic Articulography (EMA) using the Carstens AG501 Articulograph. The system tracks the three-dimensional movement of sensors attached to various structures inside and outside the vocal tract. During recording the raw positional data are stored in the computer which is connected to the articulograph. The stimuli were prompted by another computer which also triggered the articulograph to start recording. The subject sat on a chair in a sound-proof booth and was instructed to read the sentences appearing on a computer monitor at a comfortable rate. The articulatory data were recorded at a sampling rate of 250 Hz. Audio data were also recorded at 48 kHz using a t.bone EM 9600 unidirectional microphone.

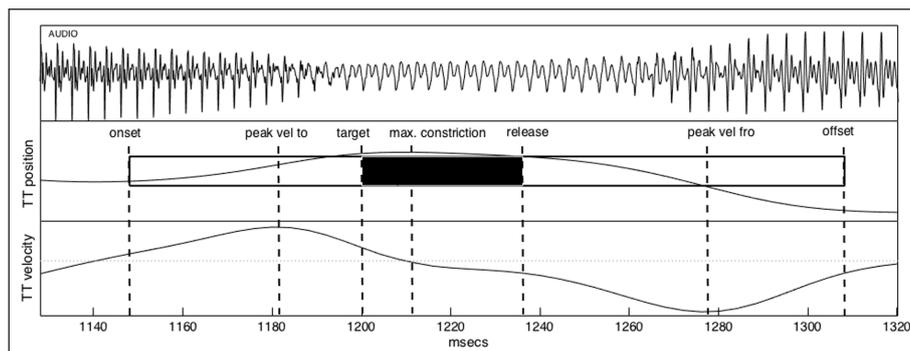
We now describe the placement of the sensors. Three sensors were placed midsagittally to the tongue: the tongue tip (TT) sensor attached 1 cm posterior to the tongue tip, the tongue mid (TM) sensor attached 2 cm posterior to the TT, and the tongue back (TB) sensor attached 2 cm posterior to the TM. Additional sensors were attached to the upper and lower lip and to the low incisors (jaw). Reference sensors were attached on the upper incisor, behind the ears (left and right mastoid), and on the bridge of the nose. In a post-processing stage, the data were corrected by subtracting the head movement captured from the reference sensors on the upper incisor and on the left and right mastoid. The data of the reference sensors were filtered using a cut-off frequency of 5 Hz, while the rest of the sensors' data were filtered using a cut-off frequency of 20 Hz. At the final stage, the data were rotated according to the occlusal plane of each subject.

### 2.4. Articulatory segmentation

Articulatory segmentation consists in identifying the points in time where characteristic events such as onset of movement, achievement of target, and movement away from the target for a consonant or a vowel take place. For each consonant in the cluster of a cluster-initial or singleton consonant-initial word, the consonant(s), the subsequent vowel, and the postvocalic consonant temporal

landmarks were measured using the primary articulator(s) involved in their respective production. Thus, velar consonants (/k, g/) were measured using the most posterior TB sensor, coronals (/t, l, z/) using the TT sensor, and labials (/p, b/) using the lip aperture (LA). LA is a derivative signal using the Euclidean distance between upper and lower lip sensors. The tense vowel, which is a low vowel, following the consonant(s) of interest, was measured using both the TM and the TB sensors. Landmark identification for both the consonantal and vowel gestures was based on the tangential velocities of the corresponding positional signals (or derived signals for the case of LA).

The articulatory segmentation of the data was conducted using the Matlab-based *Mview* software developed by Mark Tiede at Haskins Laboratories. Its segmentation algorithm first finds the peak velocities (to and from the constriction) and the minimum velocity within a user-specified zoomed in temporal range. The achievement of target (*target*) and the constriction release (*release*) landmarks were then obtained by identifying the timestamp at which velocity falls below and rises above a 20% threshold of the local tangential velocity peaks. **Figure 1** illustrates an example parse of the coronal gesture of the prevocalic lateral of the word /lɑ:gə/ using the TT sensor. The panels from top to bottom illustrate the acoustic signal of the zoomed in portion of the utterance along with the TT movement trajectory and the TT vertical velocity profile. The black filled box corresponds to the constriction phase, or plateau, of the gesture delimited by the *target* and *release* landmarks (left and right side of the black filled box). The left and right side of the white box indicate the initiation (*onset*) and end of the gesture (*offset*) calculated as the timestamps at which velocity rises above and falls below a 20% threshold of local tangential velocity peaks. The landmarks *peak vel to* and *peak vel fro* indicate peak velocity towards the *target* and from the *release* landmarks respectively. The *max. constriction* landmark corresponds to the minimum velocity.



**Figure 1:** Parsing of a gesture using *Mview*. Panels from top to bottom: acoustic signal of the zoomed in portion of the word *Lage*, tongue tip (TT) movement trajectory in the superior-inferior dimension (vertical), tongue tip (TT) vertical velocity signal. The black filled box indicates the constriction phase, or plateau, of the gesture for the /l/ delimited by the *target* and *release* landmarks (located at the timestamps of the left and right edges of the black filled box). The (timestamps of the) left and right edges of the longer white box indicate the initiation and end of the /l/ gesture. *Peak vel to*, *peak vel fro*, and *max. constriction* correspond to the peak velocity towards the target, away from the release, and minimum velocity respectively.

## 2.5. Statistical analysis

We used R Studio version 3.3.1 (RStudio Team, 2015) and the *lmer* package (Bates, Maechler, Bolker, & Walker, 2015) to perform linear mixed effects analyses for the effect of C1 voicing, C1 place, prosodic condition, interval type, and cluster size on the temporal intervals of interest such as IPI, C1 plateau duration, C2 plateau duration, and interval duration. The voicing factor is included in our analyses, even if it is only applicable in the CCV context (not in C#CV due to final devoicing), because there is evidence from prior work that IPI is shorter in voiced stop-lateral than in voiceless stop-lateral clusters in German (Bombien & Hoole, 2013; Pouplier, 2012). In the event that such differences in IPI are also observed in our data, they would present another opportunity to study how such local variability affects or does not affect spatiotemporal coordination patterns in the rest of the CCV sequence. In all analyses, we fitted separate models to the CV~CCV and CV~C#CV subsets of the data. This was done because a single model for the whole dataset CV~CCV~C#CV would not converge and the driving questions of this paper do not concern a direct comparison of effects between CV~CCV and CV~C#CV.

Models were tested for main effects and all interactions. Subject and repetition were treated as random factors. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality. *P*-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. For post-hoc comparisons, significance was determined using the *Tukey* adjusted contrast using the *multcomp* package (Hothorn, Bretz, & Westfall, 2008) and the *lsmeans* package (Lenth & Hervé, 2015). To address relations between continuous variables Pearson's correlations were used.

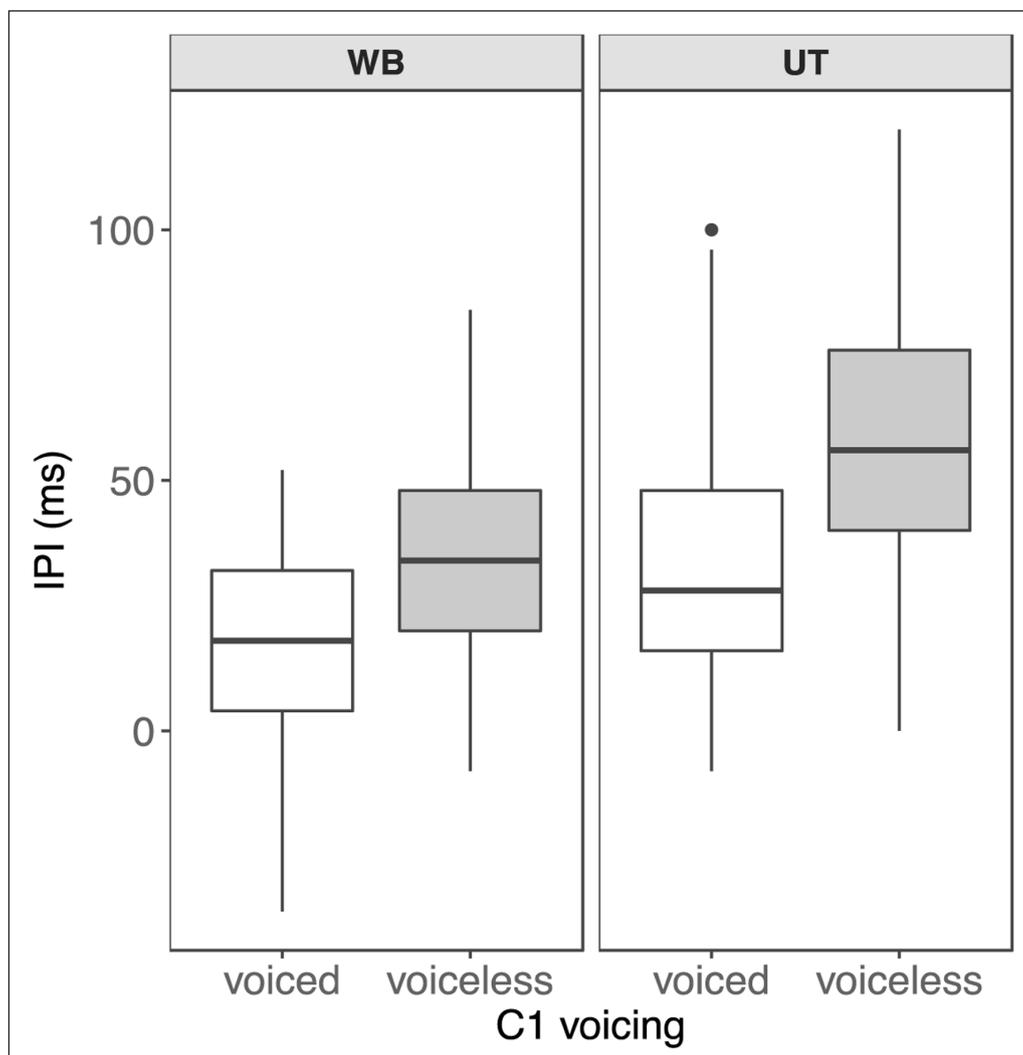
## 3. Results

### 3.1. Interplateau interval

The interplateau interval (henceforth, IPI) in a C1C2V sequence is defined as the lag between the release of the initial consonant C1 and the target of the second consonant C2, that is, C2 target – C1 release, in milliseconds. Positive IPIs indicate no temporal overlap between the plateaus of the two consonants, while negative IPIs indicate temporal overlap.

We begin by quantifying IPI as a function of prosodic condition, to examine whether our experimental design successfully elicited variability of IPI (the presence of such variability is a crucial prerequisite of our approach as developed in the ensuing sections). For word-initial CCV sequences ( $N = 393$ ), we fitted a linear mixed effects model with IPI as a dependent variable and prosodic condition (word boundary, utterance phrase boundary), voicing of the initial stop (C1 voicing) (voiced, voiceless), C1 place of articulation (labial, velar) as fixed effects. As random effects, we had random intercepts for subject and repetition, as well as by-subject and by-repetition random slopes for the effect of prosodic condition. Model comparisons showed a significant main effect of C1 voicing ( $\chi^2[1, N = 393] = 72.40, p < 0.0001$ ) and prosodic

condition ( $\chi^2[1, N = 393] = 8.38, p = 0.003$ ) on IPI. There is no effect of C1 place on IPI and no interaction between C1 voicing and prosodic condition. The post-hoc analysis showed that IPI increases significantly from the word boundary condition (henceforth, WB) to the utterance phrase boundary condition (henceforth, UT) (estimate = 21 ms,  $p = 0.0001$ ). Furthermore, IPI is 27.3 ms larger in C1 voiceless stop-lateral clusters than in C1 voiced stop-lateral clusters ( $p < 0.0001$ ) across prosodic conditions. **Figure 2** illustrates IPI in CCV as a function of prosodic condition and C1 voicing. Mean and standard deviation for IPI in C1 voiced and C1 voiceless stop-lateral clusters when C1 is a velar or a labial are shown in **Table 2**.



**Figure 2:** Interplateau interval (IPI) in CCV as a function of prosodic condition and C1 voicing.

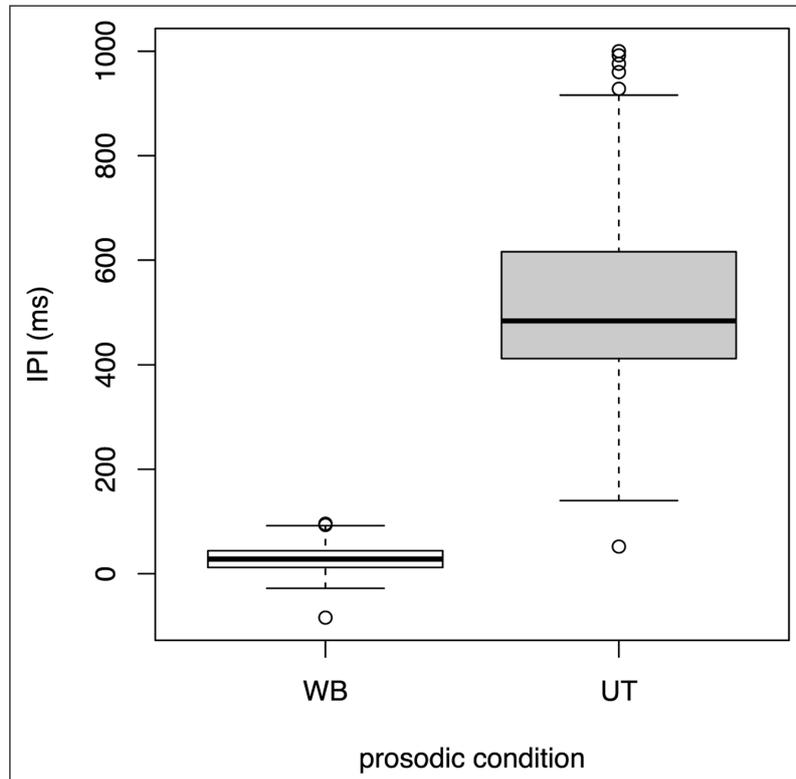
IPI (ms)	WB		UT	
	Mean	SD	Mean	SD
C1 voiceless	35.9	21	61.5	28.3
labial	44	19.5	64.7	29.1
velar	27.8	19.5	58.1	27.4
C1 voiced	17.6	17.9	35.1	26.5
labial	25.9	15.3	33	22.8
velar	9.7	16.7	36.8	29.5

**Table 2:** Mean and standard deviation of IPI in CCV as a function of prosodic condition, voice, and place.

For cross-word C#CV sequences, the IPI or lag between consonantal plateaus is obscured by the pause duration which separates the two consonants in the utterance phrase boundary conditions. The notion of IPI is thus not identical across word-initial and cross-word sequences. However, in what follows, we do report results on IPI or the lag between consonantal plateaus for cross-word sequences as well. This is because studying how the segmental sequence responds to perturbations in IPI duration (Sections 3.5 and 3.6) illustrates rather clearly the presence of two distinct organizations (global organization in CCV versus local organization in C#CV).

For cross-word sequences, a linear mixed effects model was fitted with IPI as a dependent variable and prosodic condition (word boundary, utterance phrase boundary), C1 place of articulation (labial, velar) as fixed effects. As random effects, we had random intercepts for subject and repetition, as well as by-subject and by-repetition random slopes for the effect of C1 place. Model comparisons showed a main effect of prosodic condition on IPI ( $\chi^2[1, N = 189] = 294.8, p < 0.0001$ ). There is no effect of C1 place on IPI and no interaction between C1 place and prosodic condition. The post-hoc analysis showed that IPI increases significantly from the word boundary condition (henceforth, WB) to the utterance phrase boundary condition (henceforth, UT) (estimate = 512 ms,  $p < 0.0001$ ). IPI increases as a function of prosodic condition for C#CV and this increase is to a greater extent than that seen for word-initial CCV (this is not surprising of course given that in C#CVs the two consonants of the cluster belong to different words which are separated by a pause). **Table 3** shows IPI as a function of prosodic condition in C#CV sequences. For the C#CV sequences, the effect of voicing of the initial stop cannot be examined because of final devoicing reported in German. Thus, only voiceless C#CVs are available for

analysis. **Figure 3** illustrates IPI in C#CVs as a function of prosodic condition. It can be seen that with increasing boundary strength, from WB to UT, IPI increases (which is expected, as described above).



**Figure 3:** Interplateau interval (IPI) in C#CV as a function of prosodic condition.

IPI (ms)	WB		UT	
	Mean	SD	Mean	SD
C1 voiceless	26.9	26.2	590	272
labial	29.8	21.4	580	262
velar	23.9	30.3	600	285

**Table 3:** Mean and standard deviation of IPI in voiceless C#CV sequences as a function of place.

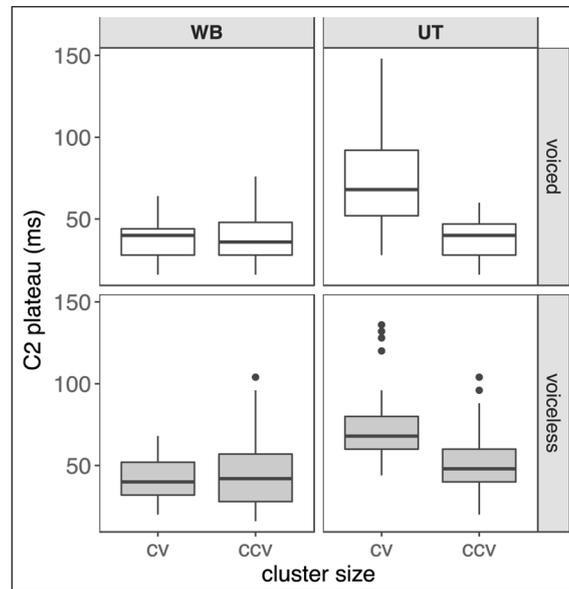
To sum up, IPI in both word-initial complex onsets and cross-word sequences increases with prosodic strengthening, corroborating the results in Byrd and Choi (2010) on English. Thus, the

crucial point for now is that our experimental design successfully elicited different degrees of IPI. The demonstration of such variability in phonetic parameters is the crucial prerequisite of our approach, as pointed out in the introduction, which aims to harness this variability (here, in terms of IPI) by assessing how segmental sequences in different syllabic organizations adapt or respond to such perturbations in the phonetic parameters of the segments that take part in those organizations. Moreover, for stop-lateral complex onsets, there is an effect of C1 voicing of IPI in that across prosodic positions IPI for voiceless stop-lateral clusters is greater than IPI for voiced stop-lateral clusters. This too is a desired effect which can also be harnessed in assessing the effects of variability in IPI on the spatiotemporal form of the sequences in which these segments are included. Recall that the effect of C1 voicing on IPI cannot be examined in cross-word sequences due to final devoicing.

### 3.2. Prevocalic lateral duration

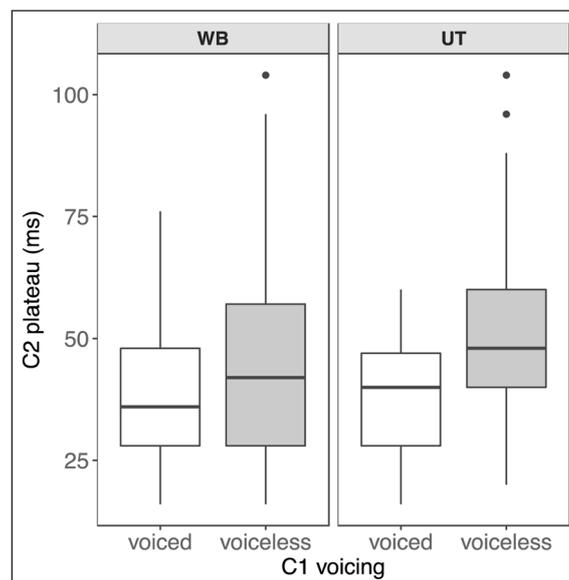
We turn next to examine differences in prevocalic lateral duration when the lateral occurs as the single consonant in simple CV sequences and when it occurs in a cluster as the prevocalic consonant of a complex onset in CCV versus in a cross-word C#CV (where CC is not a complex onset). Lateral duration is calculated as the plateau duration of the lateral, meaning the interval between C release and C target: C plateau = C release – C target.

We begin with assessing differences in prevocalic lateral (plateau) duration from CV to CCV. A linear mixed effects model was fitted with lateral duration (log transformed to better approximate a normal distribution) as a dependent variable. The variables cluster size (CV, CCV), C1 voicing (voiced, voiceless), prosodic condition (word boundary, utterance phrase boundary), C1 place of articulation (labial, velar) were used as fixed effects. As random effects, we had random intercepts for subject and repetition, as well as by-subject and by-repetition random slopes for the effects of cluster size and prosodic condition. There is no four-way interaction and no interaction between cluster size, prosodic condition, and C1 place. There is an interaction between cluster size, C1 voicing, and prosodic condition ( $\chi^2[4, N = 754] = 86.5, p < 0.0001$ ). The post-hoc multiple comparisons analysis showed that for both voiced and voiceless stop-lateral clusters, the lateral's duration does not change significantly from CV to CCV in the word boundary condition (C1 voiced: estimate =  $-0.01$ ; C1 voiceless: estimate =  $-0.01$ ), but it does decrease in the utterance phrase boundary condition (C1 voiced: estimate =  $0.67, p = 0.0002$ ; C1 voiceless: estimate =  $0.43, p < 0.002$ ). Additionally, across CCV, the lateral's duration is shorter after a voiced than a voiceless stop in the word boundary (estimate =  $0.12, p = 0.02$ ) and in the utterance phrase boundary condition (estimate =  $0.29, p < 0.0001$ ). **Figure 4** illustrates prevocalic lateral duration from CV to CCV per voicing and per prosodic condition. The duration of the lateral decreases from CV to CCV in the utterance (UT) phrase boundary condition for both voiced and voiceless stop-lateral clusters, but it remains the same in the word boundary (WB) condition.



**Figure 4:** Prevocalic lateral plateau duration as a function of cluster size, C1 voicing, and prosodic condition.

**Figure 5** shows prevocalic lateral plateau duration as a function of C1 voicing and prosodic condition in CCV sequences.



**Figure 5:** C2 lateral plateau duration as a function of C1 voicing and prosodic condition in CCV sequences. The C2 lateral in CCV is shorter when the initial stop is voiced than when it is voiceless in both word boundary (WB) and utterance phrase (UT) boundary conditions.

**Table 4** presents mean and standard deviation for prevocalic lateral plateau duration in CV and CCV when the initial stop in the cluster is voiced or voiceless in two prosodic conditions.

Lateral duration (ms)	WB		UT	
	Mean	SD	Mean	SD
C1 voiceless				
CV	41.5	11.9	74.1	20.7
CCV	44.2	19.3	50.6	16.2
C1 voiced				
CV	37.2	10.9	73.8	34.5
CCV	39.3	14.4	37.1	11.2

**Table 4:** Mean and standard deviation for prevocalic lateral plateau duration from CV to CCV when the initial consonant in CCV is a voiced or a voiceless stop in two prosodic conditions.

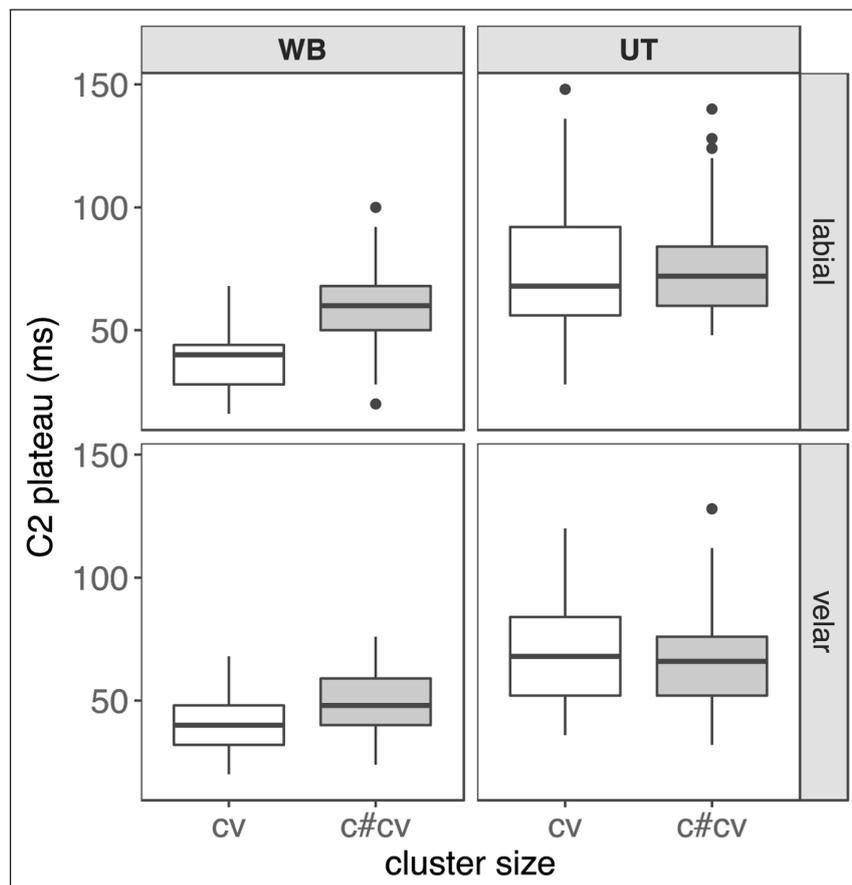
Consider now prevocalic lateral duration from CV to C#CV. **Table 5** presents mean and standard deviation for prevocalic lateral plateau duration in CV and C#CV in the word boundary (WB) and utterance phrase boundary (UT) conditions.

Lateral duration (ms)	WB		UT	
	Mean	SD	Mean	SD
CV	41.5	11.9	74.1	20.7
C#CV	53.9	16	76	31.4

**Table 5:** Mean and standard deviation for prevocalic lateral plateau duration in the CV and C#CV contexts in two prosodic conditions.

In what follows, we evaluate statistically the change in prevocalic lateral plateau duration from CV to C#CV. For the cross-word sequences and their respective singletons ( $N = 320$ ), we fitted a linear mixed effects model with lateral duration (log transformed to better approximate a normal distribution) as a dependent variable. As fixed effects, we used cluster size (CV, C#CV), prosodic condition (word boundary, utterance phrase boundary), and C1 place (labial, velar). As random factors, we used random intercepts for subject and repetition, as well as by-subject

random slopes for the effect of cluster size, prosodic condition, and C1 place. There is an interaction between cluster size, prosodic condition, and C1 place ( $\chi^2[4, N = 320] = 18.17, p = 0.001$ ), which means that the way the lateral duration changes from CV to C#CV depends on the prosodic condition and on C1 place. The post-hoc analysis showed that in the word boundary condition the lateral duration increases significantly from CV to C#CV for C1 labial (estimate =  $-0.41, p = 0.003$ ) but not for C1 velar (estimate =  $-0.22$ ). In the utterance phrase boundary condition, the lateral duration does not change significantly from CV to C#CV for either C1 labial (estimate =  $0.02$ ) or velar (estimate =  $0.03$ ). **Figure 6** illustrates prevocalic lateral duration from CV to C#CV for labial and velar stop-laterals in two prosodic conditions, word boundary and utterance phrase boundary.



**Figure 6:** Prevocalic lateral plateau duration from CV to C#CV in the word boundary (WB) and utterance phrase boundary (UT) conditions for C1 labial (/p#l/) and C1 velar (/k#l/). The lateral in C#CV is longer than in CV only in WB and only for the labial. In the UT condition, there is no significant difference with respect to C2 lateral duration from CV to C#CV.

To summarize, for word-initial stop-lateral clusters, the prevocalic lateral in CCV is shorter than the single lateral in CV across prosodic conditions. However, there are inconsistencies within prosodic conditions. In the word boundary condition, the lateral has the same duration in CV and CCV regardless of the stop's voicing. In the utterance phrase boundary condition, the lateral in the cluster CCV is shorter than the single lateral in CV and its shortening is greater when the initial stop is voiced than when it is voiceless. For laterals in the stop-lateral CCV only (no CV) context, across prosodic conditions, the lateral in CCV is shorter when the initial stop is voiced than when it is voiceless. For cross-word sequences, the lateral in C#CV is longer than the single lateral in CV for C1 labial but not C1 velar in the word boundary condition. In the utterance phrase boundary condition, for both C1 labial and velar, the duration of the lateral in C#CV does not differ from that of the lateral in CV. Lengthening or no change in lateral duration from CV to C#CV across prosodic conditions is not what we observed for CV~CCV earlier. Clearly, lateral duration adjusts differently (from its single instance in CV) when the lateral is part of a complex onset versus when it is not (CCV versus C#CV).

### 3.3. IPI – C2 lateral compensatory relation

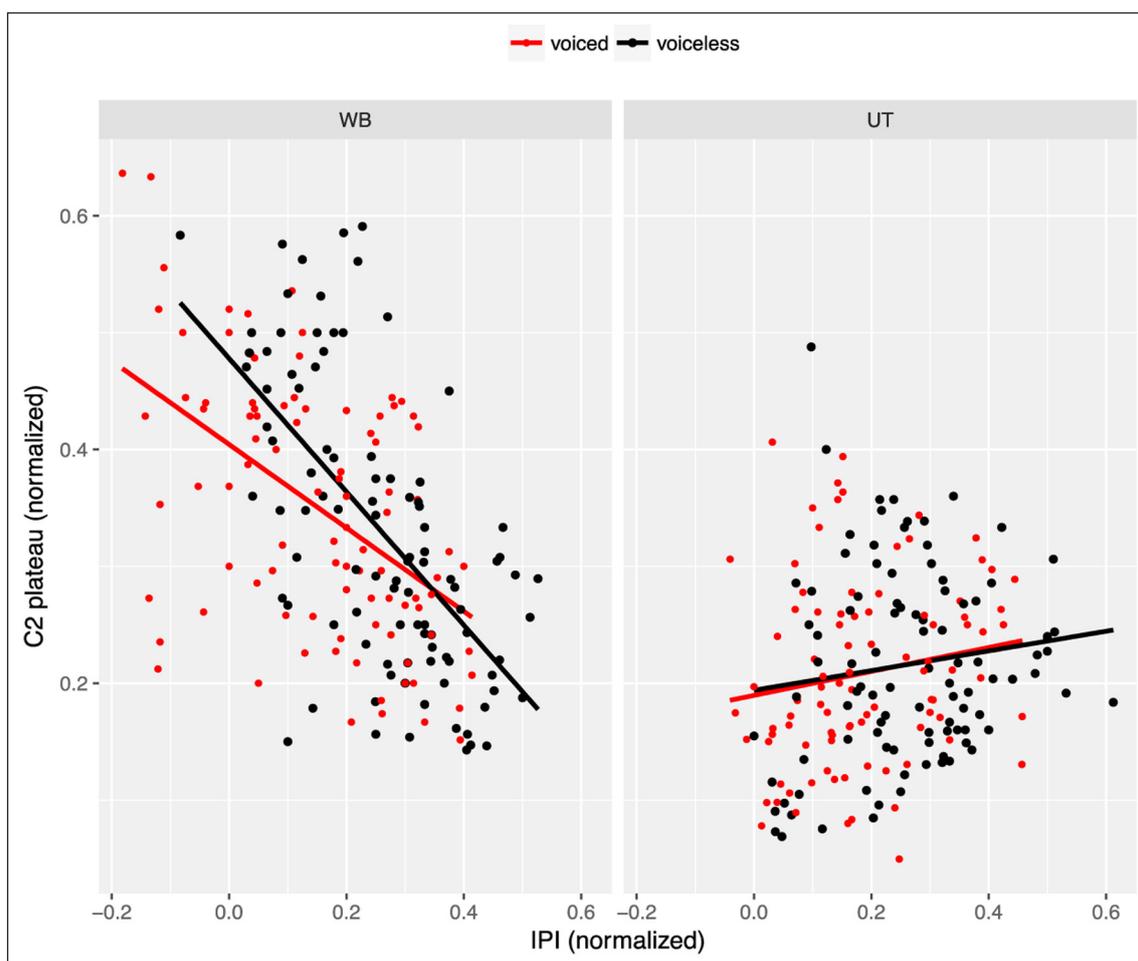
We turn here to examine the relation between IPI and C2 lateral plateau duration in word-initial stop-lateral complex onsets across and within prosodic conditions.<sup>1</sup> Our aim is to assess the presence of compensatory effects in the CCV string, that is, effects where spatiotemporal modification in one local region of the string comes systematically with a change in another part of the string. The presence of such effects would indicate that the organization of the different parts of CCV is not independently planned and produced and thus such effects, if present, offer evidence for global organization.

For the word-initial stop-lateral complex onsets CCV ( $N = 412$ ), there is no correlation between IPI and C2 lateral duration across prosodic conditions when looking at raw values ( $r(410) = 0.028$ ,  $p = 0.56$ ). This result holds true when looking also within each prosodic condition: for the word boundary condition, ( $r(207) = -0.07$ ,  $p = 0.28$ ); for the utterance phrase boundary, ( $r(201) = 0.07$ ,  $p = 0.26$ ). Normalized IPI and C2 lateral plateau duration were also calculated to compensate for effects related to inter-speaker variability (cf. Bombien, 2011). Raw IPI and C2 lateral plateau measures exhibit substantial variability as can be seen by their ranges in **Figure 2** and **Figure 5**. Some of this variability may be speaker-specific and derives from simple continuous scaling of IPI as a function of rate (e.g., the longer the CC, the longer the IPI). Hence, it seems useful to also examine a normalized measure of IPI. When looking at normalized

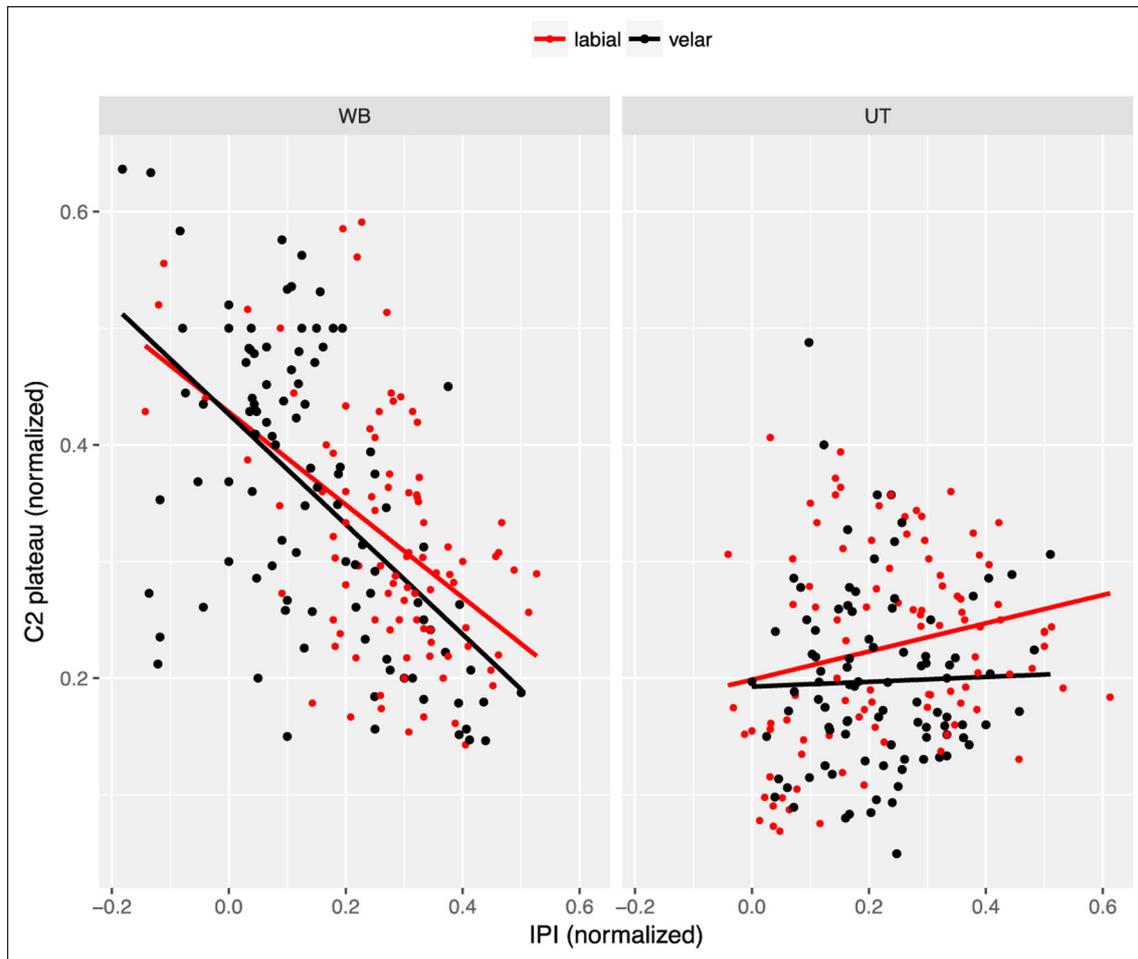
---

<sup>1</sup> The relation between IPI and C2 lateral plateau will not be investigated for cross-word sequences. The notion of IPI is not meaningful in that condition as IPI is obscured by the pause duration which separates the two consonants at the utterance boundary.

values, there is a weak negative correlation across prosodic conditions ( $r(410) = -0.34, p < 0.0001$ ), such that as IPI increases, C2 lateral duration tends to decrease. However, there are inconsistencies within prosodic condition. For the word boundary condition, there is a moderate negative correlation ( $r(207) = -0.58, p < 0.0001$ ) but for the utterance phrase boundary, there is no correlation ( $r(201) = 0.14, p = 0.04$ ). **Figure 7** illustrates the relation between normalized IPI and C2 lateral plateau for CCV in each prosodic condition (WB, UT). **Figure 8** illustrates the relation between normalized IPI and C2 lateral plateau for labial and velar stop-laterals in each prosodic condition.



**Figure 7:** Scatterplots showing the relation between normalized C2 lateral duration and normalized IPI for voiced and voiceless stop-lateral CCV in two prosodic conditions. In the word boundary condition (WB), there is a moderate negative correlation ( $r(207) = -0.58, p < 0.0001$ ) such that as IPI increases, C2 lateral plateau tends to decrease. In the utterance phrase boundary condition (UT), there is no correlation ( $r(201) = 0.14, p = 0.04$ ).



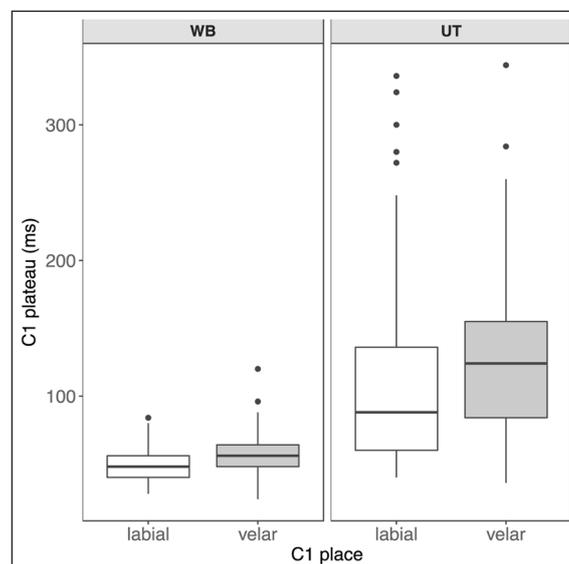
**Figure 8:** Scatterplots showing the relation between normalized C2 lateral duration and normalized IPI for labial and velar stop-lateral CCV in two prosodic conditions.

To sum up, using the raw values of the variables, there is no compensatory relation between IPI and C2 lateral plateau across and also within prosodic conditions. However, a compensatory relation does emerge when looking at the normalized values of the two variables IPI and C2 plateau. Across prosodic conditions, there is overall a weak relation between IPI and C2 lateral such that as IPI increases, C2 lateral duration tends to decrease. This compensatory relation is stronger in the word boundary condition and it weakens at the utterance phrase boundary. To identify a compensatory relation between two variables, variability along both variables individually is required. From the distribution of the data seen in **Figure 7**, substantial variability of the lateral duration can be observed (y-axis) only in the word boundary condition. The variability of the lateral duration decreases with prosodic strengthening as can be seen in the shrinking of the y-axis values from WB to UT. Thus, the lack of a compensatory relation in the

condition of prosodic strengthening is not surprising. As argued in Section 4, the compensatory IPI-C2 lateral plateau duration relation, in contexts where it is found, serves as a hallmark of global organization.

### 3.4. C1 plateau duration

In this subsection, we examine the effect of C1 voicing and prosodic condition on the duration of the C1 initial stop in stop-lateral clusters. For word-initial complex onsets CCV, we fitted a linear mixed effects model with C1 plateau duration as dependent variable. The dependent variable was log transformed to better approximate a normal distribution. C1 voicing (voiced, voiceless), prosodic condition (word boundary, utterance phrase boundary), C1 place (labial, velar) were modeled as fixed effects. As random effects, we had intercepts for subject and repetition, as well as by-subject and by-repetition random slopes for the effects of C1 place and prosodic condition. There is no main effect of C1 voicing on the duration of C1 plateau. However, there is a main effect of prosodic condition ( $\chi^2[1, N = 404] = 14.16, p = 0.0001$ ) and a main effect of C1 place on the duration of C1 plateau. Specifically, C1 plateau is shorter in the word boundary than in the utterance phrase boundary condition (estimate = 0.71,  $p < 0.0001$ ) and velars are longer than labials (estimate =  $-0.17, p = 0.02$ ). There is no interaction between C1 voicing, prosodic condition, and C1 place. **Figure 9** plots C1 plateau duration as a function of C1 place and prosodic condition.



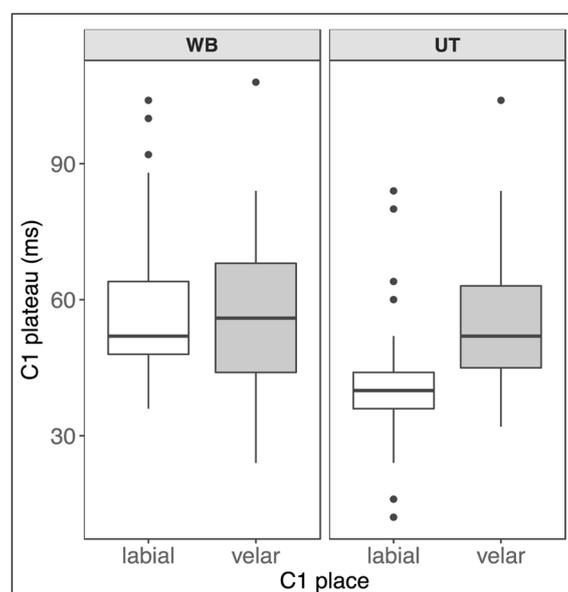
**Figure 9:** C1 plateau duration as a function of C1 place and prosodic condition in CCV. C1 plateau is longer for velars than labials and longer in UT than in WB.

Mean and standard deviation for C1 stop plateau duration in milliseconds in C1 voiced and C1 voiceless stop-lateral CCV sequences are provided in **Table 6**.

C1 stop duration (ms)	WB		UT	
	Mean	SD	Mean	SD
C1 voiceless	53.5	14.8	112.3	58.1
C1 voiced	53.4	14.4	120.4	62.2
C1 labial	49.4	11.9	109	67.4
C1 velar	57.3	16.2	126.5	55.4

**Table 6:** Mean and standard deviation for C1 plateau duration as a function of C1 voicing and C1 place in each prosodic condition.

For the cross-word stop-lateral clusters C#CV ( $N = 189$ ), a linear mixed effects model was fitted with C1 plateau duration as a dependent variable. The dependent variable was log transformed to better approximate a normal distribution. Prosodic condition (word boundary, utterance phrase boundary) and C1 place (labial, velar) were used as fixed effects. We had random intercepts for subject and repetition, as well as by-subject and by-repetition random slopes for the effect of prosodic condition. The results showed a significant main effect of prosodic condition ( $\chi^2[1, N = 189] = 15.9, p < 0.0001$ ) and a significant main effect of C1 place ( $\chi^2[1, N = 189] = 9.57, p = 0.001$ ). Specifically, C1 plateau duration is shorter in the utterance phrase boundary than in the word boundary condition (estimate =  $-0.18, p = 0.01$ ) and C1 plateau is longer for velars than labials (estimate =  $-0.14, p = 0.002$ ). **Figure 10** illustrates C1 plateau duration for C#CV as a function of prosodic condition and C1 place.



**Figure 10:** C1 plateau duration in milliseconds for C#CV sequences as a function of prosodic condition and C1 place.

Mean and standard deviation for C1 plateau duration in milliseconds in voiceless C#CV sequences when C1 is velar or labial are provided in **Table 7**.

C1 stop duration (ms)	WB		UT	
	Mean	SD	Mean	SD
C1 voiceless	57	17.2	48	16.3
p#l	56.9	16	40.1	13.2
k#l	57	18.5	56	15.4

**Table 7:** Mean and standard deviation for C1 plateau duration in C#CV simplex onsets when C1 is a labial p#l or a velar k#l in three prosodic conditions.

Summarizing, there is an effect of C1 place on the duration of the initial stop across prosodic conditions and across CCV~C#CV with velars being longer than labials (a result that has been reported before for German in Bombien & Hoole, 2013). For word-initial clusters, there is no effect of C1 voicing on the duration of the initial stop in word-initial stop-lateral complex onsets; rather, there is an effect of prosodic condition on the duration of the initial stop. This means that the duration of the initial stop increases with prosodic strengthening in word-initial stop-lateral clusters. However, the duration of the initial stop remains approximately the same across prosodic conditions in cross-word C#CV sequences. Thus, in cross-word sequences, the initial stop is somewhat shorter in the prosodic strengthening compared to the word boundary (control) condition. Such differences in C1 plateau duration as a function of prosodic strengthening in clusters of different syllabic affiliations are informative as they provide the prerequisite perturbations under which we can assess whether, as a result of these perturbations, spatiotemporal readjustments in (other parts of) the CCV string occur. As we discuss in Section 4, C1 lengthening in the utterance phrase boundary condition together with IPI lengthening create the setting under which effects related to syllabic organization emerge.

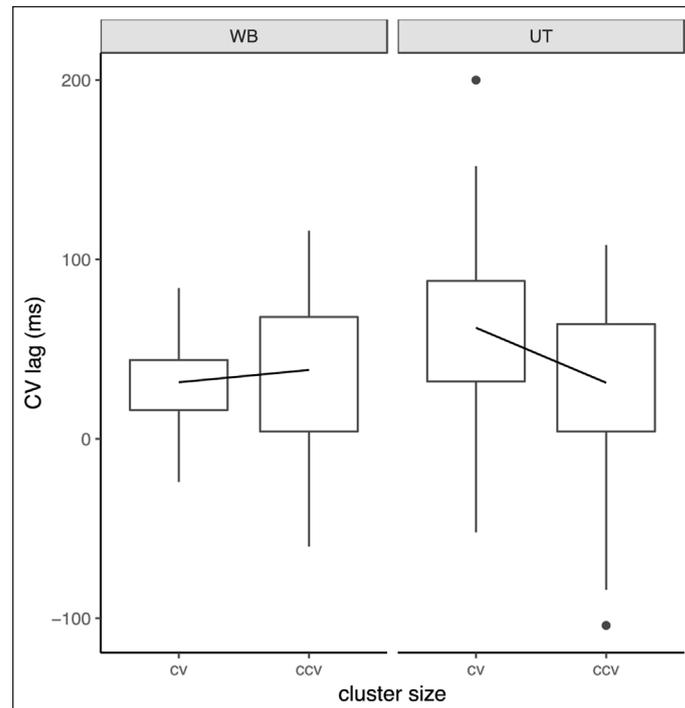
### 3.5. Vowel initiation with respect to prevocalic lateral

The c-center organization pattern prescribes that the vowel starts somewhere around the c-center of the prevocalic onset cluster (Browman & Goldstein, 1988; Browman & Goldstein, 2000; Honorof & Browman, 1995; Nam & Saltzman, 2003; Gafos, 2002). We say ‘somewhere around the c-center’ because the literature is somewhat ambiguous on the matter, depending on whether one interprets any relevant statement to be a statement about observed movement properties versus underlying phonological demands which may or may not have directly observed physical consequences (depending on various parameters). Thus, for example, Browman & Goldstein

(1988:150) write, “let us make the following assumption: the (temporal) interval from the c-center to the final consonant anchor point is a measure of the activation interval of the vocalic gesture where: (a) the c-center corresponds to a fixed point early in the vocalic activation...” but also that “We also assume that the actual movement for the vocalic gestures begins at the achievement of target of the first consonant in a possible initial cluster” (Browman & Goldstein, 1988:150). Honorof and Browman (1995: **Figure 1**, p. 552) show the beginning of the vowel activation window to be at the c-center of the prevocalic consonantal cluster (made out of three consonants). Nam and Saltzman (2003) assume a default phasing for the CV relation of 50 degrees and show the V starting somewhat after the c-center of a single consonant. Gafos (2002), in his Optimality Theoretic interpretation, using constraints referring to both spatial and temporal properties of gestures, employs an alignment constraint requiring the V to start at the c-center of the consonant or prevocalic consonant cluster. Again, as noted above, no empirical study has explicitly sought to quantify when the vowel starts in reference to its preceding cluster in any systematic way.

Here, we quantify (for the first time in the relevant literature) the vowel start or what we refer to as vowel initiation in word-initial stop-lateral CCV and cross-word C#CV sequences in two prosodic conditions: word boundary (WB) and utterance phrase boundary (UT). To do so, the vowel gesture was parsed in each token by using the tongue back sensor based on the tangential velocity of the signal. The time interval between the so-obtained gestural onset landmark of the vowel and the gestural target landmark of the preceding lateral was then used to define our dependent variable, CV lag.

For the CV~CCV pairs ( $N = 712$ ), we fitted a linear mixed effects model with CV lag as a dependent variable. Cluster size (CV, CCV), C1 voicing (voiced, voiceless), prosodic condition (word boundary, utterance phrase boundary), C1 place (labial, velar) were used as fixed effects. As random effects, we had random intercepts for subject and repetition, as well as by-subject random slopes for the effect of prosodic condition. There is a three-way interaction of cluster size, C1 voicing, and prosodic condition ( $\chi^2[4, N = 712] = 20.8, p = 0.0003$ ) and a three-way interaction of cluster size, C1 place, and prosodic condition ( $\chi^2[4, N = 712] = 18.2, p = 0.001$ ). The post-hoc analysis showed that in the word boundary condition, the CV lag does not change significantly from CV to labial CCV (estimate =  $-5.9$  ms) and from CV to velar CCV (estimate =  $-13$  ms) while in the utterance phrase boundary the CV lag changes significantly for both CV to labial CCV (estimate =  $38$  ms,  $p < 0.0001$ ) and CV to velar CCV (estimate =  $19.3$  ms,  $p = 0.003$ ). Additionally, in the word boundary condition the CV lag does not change significantly from CV to voiced CCV (estimate =  $-2.4$  ms) and from CV to voiceless CCV (estimate =  $-13.9$  ms). In the utterance phrase boundary condition, the CV lag decreases significantly from CV to voiced CCV (estimate =  $40$  ms,  $p < 0.001$ ) and from CV to voiceless CCV (estimate =  $19$  ms,  $p = 0.003$ ). **Figure 11** illustrates the CV lag across CV~CCV in two prosodic conditions (WB, UT).

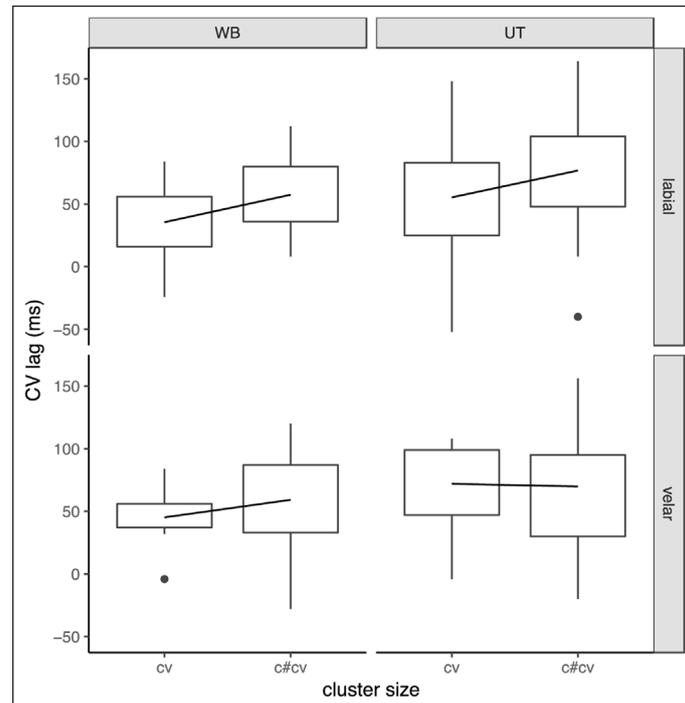


**Figure 11:** CV lag in milliseconds between the target of the lateral and vowel initiation in lateral-vowel sequences (CV) and stop-lateral sequences (CCV) in two prosodic conditions, word boundary (WB) and utterance phrase boundary (UT). The CV lag remains the same across CV~CCV in WB, but it decreases from CV to CCV in the UT condition.

With prosodic strengthening from WB to UT, the vowel is found to start earlier with respect to the target of the prevocalic lateral in the word-initial CCV sequences than in CV (across C1 place and C1 voicing). The exact opposite pattern is observed in cross-word sequences with prosodic strengthening as we will see below. Vowel initiation does not change with respect to the target of the prevocalic lateral in cross-word C#CV sequences compared to CV in the UT condition.

We now turn to vowel initiation differences between CV and C#CV in two prosodic conditions. For the CV~C#CV pairs ( $N = 321$ ), we fitted a linear mixed effects model with CV lag as a dependent variable. Cluster size (CV, C#CV), prosodic condition (word boundary, utterance phrase boundary), C1 place (labial, velar) were used as fixed effects. We had random intercepts for subject and repetition and by-subject and by-repetition random slopes for the effect of prosodic condition. There is an interaction between cluster size, prosodic condition, and C1 place ( $\chi^2[4, N = 321] = 12.94, p = 0.01$ ). The post-hoc analysis showed that in the word boundary condition the CV lag is longer in C#CV than in CV for both labial (estimate =  $-23.5$  ms,  $p = 0.0006$ ) and velar C1 in C#CV (estimate =  $-32$  ms,  $p < 0.0001$ ), while in the utterance phrase boundary condition the CV lag does not change significantly from CV to C#CV for either

labial (estimate =  $-7$  ms) or velar C1 in C#CV (estimate =  $-16$  ms). **Figure 12** plots the CV lag in lateral-vowel (CV) and voiceless labial and velar stop-lateral (C#CV) sequences in two prosodic conditions, word boundary and utterance phrase boundary.



**Figure 12:** CV lag in milliseconds between the target of the lateral and vowel initiation in lateral-vowel sequences (CV) and voiceless labial and velar stop-lateral sequences (C#CV) in two prosodic conditions, word boundary (WB) and utterance phrase boundary (UT).

To sum up, the CV lag, which describes the proximity of vowel initiation to the target of the prevocalic lateral, decreases with prosodic strengthening across CV and word-initial stop-lateral CCV sequences. The exact opposite pattern is observed across CV and cross-word C#CV sequences, where the CV lag remains the same with prosodic strengthening. Thus, with prosodic strengthening the vowel is found to start earlier in word-initial complex onsets but no vowel initiation difference is observed in cross-word sequences composed of the same consonants.

Recall the crucial methodological diagnostic in our approach: To uncover evidence for global organization presiding over a sequence of segments (here, CCV), properties of the segments or their relations with one another must somehow be locally varied. The consequences of such variation on the rest of the sequence can then be used to unveil the mode of organization. When local perturbations to segments or relations between adjacent segments have effects that propagate to the rest of the sequence, this is evidence that the organization presiding over that

sequence of segments is global. In our case, the UT condition causes the CC part of the CCV sequence to expand (by lengthening of the first C and increase in the duration of the IPI between the two consonants). As a consequence of that expansion, the rest of the sequence responds so that the vowel starts earlier in the cluster (than what would be the case if the CV lag did not change between WB and UT).

Let us see how this general idea of readjustments to local perturbations subsumes both what was thought before to be a standard index of global organization as well as other indices documented in our present work. In a well-known take (Browman & Goldstein, 1988), when a consonant is added in front of a CV to form a larger CCV which is organized as one syllable (i.e., global organization), the consonantal material in front of the vowel expands with the vowel compensating for that expansion by overlapping more with its preceding consonant in CCV compared to the CV context. This is the classic c-center diagnostic which has met issues in German (Pouplier, 2012; Brunner et al., 2014) and other languages (Shaw et al., 2011; Gafos et al., 2014) where evaluations of its validity have been undertaken. What we wish to point out here is that this diagnostic can be restated in terms of and subsumed under the revised approach pursued in the present paper: As the consonantal portion before the vowel expands (here, due to the addition of a C in front of a CV) in CCV, the lag between the prevocalic C and the vowel decreases to compensate for that expansion. In fully parallel terms, when in our German data, the duration of the first consonant in CCV and the gap between the end of the constriction of the first and the beginning of the second consonant (the so-called IPI) increase, the vowel starts earlier to compensate for this expansion in the CC part of the CCV sequence. Another compensatory relation between IPI and C2 lateral duration was found in the word boundary condition in German stop-lateral CCV sequences. Specifically, as IPI increased the C2 lateral decreased to compensate for that increasing IPI. Hence, even though the specifics which enter into these different compensatory relations may change depending on the contexts examined and even depending on the clusters (as we will show later), the hallmark of global organization is entirely general: Local perturbations result in readjustments in some other portion of the CCV in globally organized (but not in locally organized) CCV sequences. This reformulation of what it means for sequences of segments to be globally organized bypasses the issues with the original formulation while at the same time elevating the search of phonological organization at the level of compensatory relations among the elements that partake in that organization.

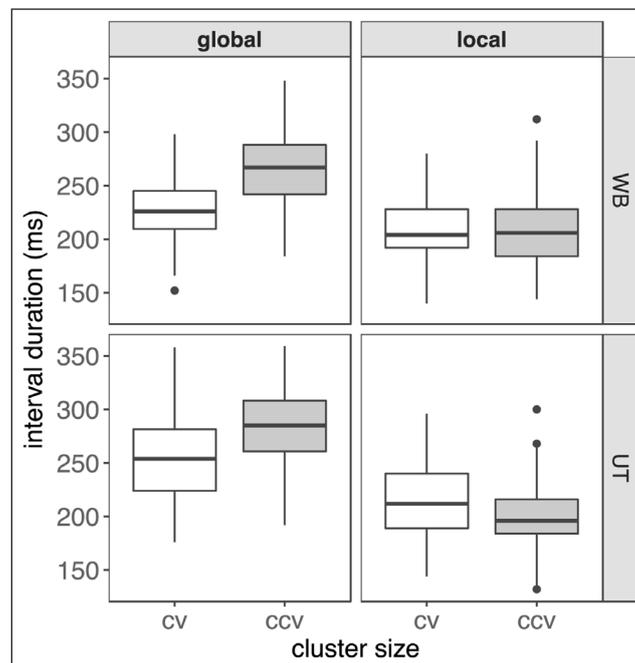
### 3.6. Stability-based heuristics

Past assessments of syllabic organization make use of the stability of certain intervals (described below) computed across CV~CCV stimuli pairs. For the stability analysis, two intervals were calculated for each stimulus observation. We first describe the right-delimiting landmarks (henceforth, anchors) used in defining these intervals. Three different anchors were used in order

to assess the robustness of results: the target of the constriction of the postvocalic consonant ( $C^{\text{tar}}$ ), the maximum constriction of the postvocalic consonant ( $C^{\text{max}}$ ), and the temporal midpoint of the vowel plateau ( $V^{\text{mid}}$ ). For each such anchor, we define two intervals, left-delimited by two different landmarks that are found on the consonantism before the vowel, the c-center and the right-edge (as used in several prior studies, e.g., Shaw et al., 2009; Shaw et al., 2011; Gafos et al., 2014; Hermes et al., 2015; Shaw & Gafos, 2015). The two intervals left-delimited by these two landmarks were the c-center to anchor interval, which stretches from the temporal midpoint of the consonant(s) to the anchor, and the right-edge to anchor interval, stretching from the constriction release of the (immediately) prevocalic consonant and the anchor. We refer to these two intervals as global timing and local timing, respectively; ‘global’ as the first interval is left-delimited by the c-center landmark whose computation implicates all consonants before the vowel as opposed to ‘local’ for the second interval which is left-delimited by the constriction release of just the immediately prevocalic consonant. Next, we evaluate statistically how the duration of the two interval types changes as the number of consonants increases from CV to a cluster with different syllabic affiliation CCV and C#CV.

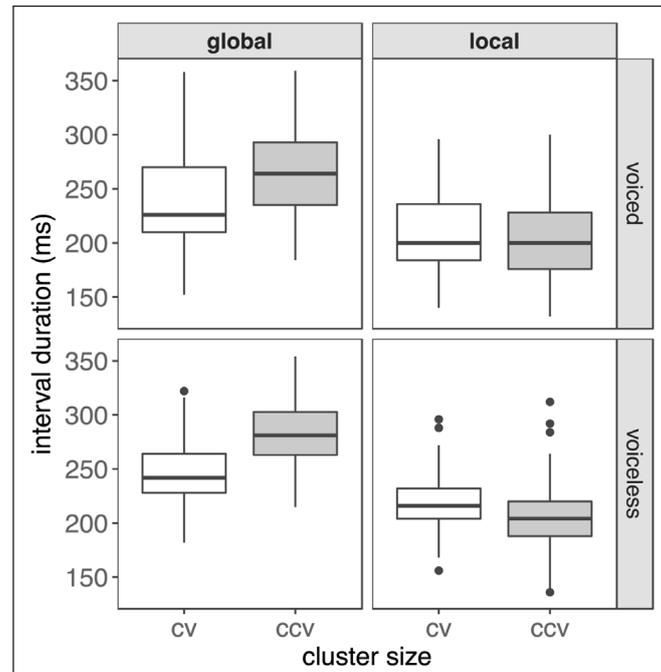
For word-initial sequences, from CV to CCV, we fitted a linear mixed effects model with interval duration as a dependent variable (log transformed to better approximate a normal distribution). Cluster size (CV, CCV), interval type (global, local), prosodic condition (word boundary, utterance phrase boundary), C1 place (labial, velar), C1 voicing (voiced, voiceless) were used as fixed effects. As random effects, we had intercepts for subjects and repetition, as well as by-subject and by-repetition random slopes for the effects of interval size and prosodic condition. The results show an interaction of interval type, cluster size, prosodic condition, and C1 voicing ( $C^{\text{tar}}$ : ( $\chi^2[8, N = 702] = 64.3, p < 0.0001$ );  $C^{\text{max}}$ : ( $\chi^2[8, N = 702] = 30.21, p = 0.0001$ );  $V^{\text{mid}}$ : ( $\chi^2[8, N = 692] = 214.8, p < 0.0001$ ) and an interaction of interval type, cluster size, prosodic condition, and C1 place ( $C^{\text{tar}}$ : ( $\chi^2[8, N = 702] = 657.8, p < 0.0001$ );  $C^{\text{max}}$ : ( $\chi^2[8, N = 702] = 776.5, p < 0.0001$ );  $V^{\text{mid}}$ : ( $\chi^2[8, N = 692] = 465.05, p < 0.0001$ ). The post-hoc analysis showed that, in the word boundary, the global timing interval increases significantly from CV to CCV across voiced and voiceless CCV (voiced:  $C^{\text{tar}}$  estimate =  $-0.11, p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.10, p < 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.32, p = 0.0001$ ; voiceless:  $C^{\text{tar}}$  estimate =  $-0.17, p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.16, p < 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.33, p < 0.0001$ ), while the local timing interval does not change significantly from CV to CCV across voiced and voiceless CCV. In the utterance phrase boundary, the global timing interval increases from CV to CCV across voiced and voiceless CCV (voiced:  $C^{\text{tar}}$  estimate =  $-0.11, p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.09, p = 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.24, p = 0.0009$ ; voiceless:  $C^{\text{tar}}$  estimate =  $-0.12, p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.11, p < 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.21, p = 0.003$ ), while the local timing interval does not change from CV to voiced CCV but it does decrease from CV to voiceless CCV ( $C^{\text{tar}}$  estimate =  $0.11, p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $0.10, p = 0.0004$ ;  $V^{\text{mid}}$  estimate =  $0.21, p = 0.002$ ).

Additionally, the post-hoc analysis for the interaction of cluster size, interval type, prosodic condition, and C1 place showed that in the word boundary, the global timing changes significantly from CV to CCV for both C1 places, labial and velar, (labial:  $C^{\text{tar}}$  estimate =  $-0.14$ ,  $p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.12$ ,  $p < 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.27$ ,  $p = 0.01$ ; velar:  $C^{\text{tar}}$  estimate =  $-0.20$ ,  $p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.19$ ,  $p < 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.54$ ,  $p = 0.0001$ ) while the local timing does not change significantly from CV to labial CCV but it does change from CV to velar CCV ( $C^{\text{tar}}$  estimate =  $-0.05$ ,  $p = 0.03$ ;  $C^{\text{max}}$  estimate =  $-0.06$ ,  $p = 0.05$ ;  $V^{\text{mid}}$  estimate =  $-0.32$ ,  $p = 0.002$ ). In the utterance phrase boundary, the global timing changes significantly from CV to CCV for both C1 places (labial:  $C^{\text{tar}}$  estimate =  $-0.11$ ,  $p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.10$ ,  $p = 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.18$ ,  $p = 0.07$ ; velar:  $C^{\text{tar}}$  estimate =  $-0.18$ ,  $p < 0.0001$ ;  $C^{\text{max}}$  estimate =  $-0.18$ ,  $p < 0.0001$ ;  $V^{\text{mid}}$  estimate =  $-0.39$ ,  $p = 0.0006$ ). In the same condition, the local timing interval does not change from CV to velar CCV but it does decrease from CV to labial CCV ( $C^{\text{tar}}$  estimate =  $0.06$ ,  $p = 0.01$ ;  $C^{\text{max}}$  estimate =  $0.06$ ,  $p = 0.04$ ) apart from the  $V^{\text{mid}}$  anchor where the local timing does not change significantly from CV to CCV for both C1 places. **Figure 13** plots the duration of the two intervals, global timing and local timing, with  $C^{\text{tar}}$  as anchor for the two prosodic conditions, word boundary (WB) and utterance phrase boundary (UT) across speakers, as a function of the number of consonants (CV, CCV).



**Figure 13:** Duration (in milliseconds) of the two intervals, global timing and local timing, for CV (white) and CCV (grey) words in two prosodic conditions. In word boundary (WB), the global timing interval increases from CV to CCV, while the local timing interval remains stable. In utterance phrase boundary (UT), both global and local timing intervals change from CV to CCV.

**Figure 14** plots the duration of the two intervals, global timing and local timing, with  $C^{\text{tar}}$  as anchor for voiced and voiceless stop-lateral clusters across speakers, as a function of the number of consonants (CV, CCV).

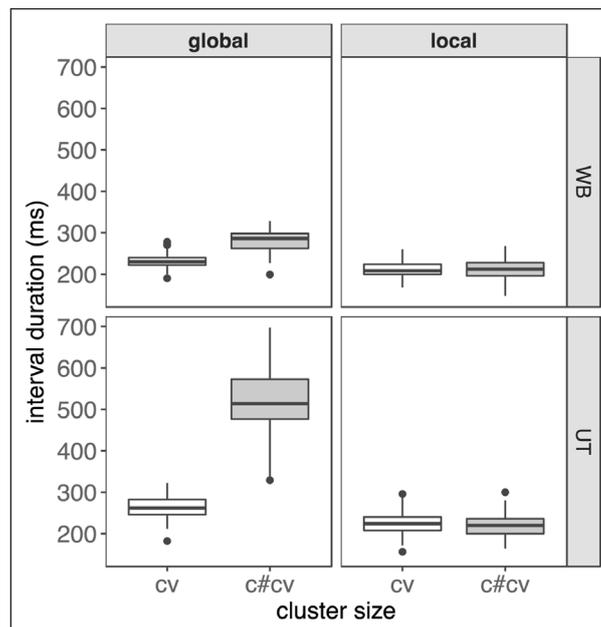


**Figure 14:** Duration (in milliseconds) of the two intervals, global timing and local timing, for CV (white) and CCV (grey) words. For voiced stop-lateral CCV words (top), the difference in duration between CV and CCV for the global timing interval is greater than the difference for the local timing interval. For voiceless stop-lateral CCV words (bottom), the duration difference between CV and CCV for the global timing interval is also greater compared to the one for the local timing interval.

Overall, for both voiced and voiceless stop-laterals, the global timing interval changes more from CV to CCV than the local timing interval across prosodic conditions. However, with prosodic strengthening, local timing ceases to be as stable especially when a voiceless stop precedes the /IV/ sequence. For both C1 places, the global timing interval changes more from CV to CCV than the local timing interval across prosodic conditions. However, with prosodic strengthening, local timing ceases to be stable especially when a labial stop is added to the /IV/ sequence. Across CV~CCV, with prosodic strengthening, the global timing interval changes less while the local timing interval changes more compared to the word boundary condition.

For cross-word sequences, we fitted a linear mixed effects model with interval duration as dependent variable. Interval duration was log transformed to better approach a normal distribution. Interval type (global, local), cluster size (CV, C#CV), prosodic condition (word

boundary, utterance phrase boundary), C1 place (labial, velar) were used as fixed effects. As random effects, we had intercepts for subjects and repetition, as well as by-subject and by-repetition random slopes for the effect of interval type. There is no four-way interaction, but there is a three-way interaction of cluster size, interval type, and prosodic condition ( $C^{\text{tar}}$ :  $\chi^2[4, N = 299] = 254, p < 0.0001$ ;  $C^{\text{max}}$ :  $\chi^2[4, N = 299] = 255.3, p < 0.0001$ ;  $V^{\text{mid}}$ :  $\chi^2[4, N = 282] = 789.3, p < 0.0001$ ). The post-hoc analysis showed that in the word boundary condition, the global timing interval increases significantly from CV to C#CV ( $C^{\text{tar}}$ : estimate =  $-0.18, p < 0.0001$ ;  $C^{\text{max}}$ : estimate =  $-0.16, p < 0.0001$ ;  $V^{\text{mid}}$ : estimate =  $-0.35, p < 0.0001$ ), while there is no effect on the local timing interval ( $C^{\text{tar}}$ : estimate =  $-0.004$ ;  $C^{\text{max}}$ : estimate =  $0.0007$ ;  $V^{\text{mid}}$ : estimate =  $-0.02$ ). In the utterance phrase boundary condition, the global timing interval increases significantly ( $C^{\text{tar}}$ : estimate =  $-0.70, p < 0.0001$ ;  $C^{\text{max}}$ : estimate =  $-0.65, p < 0.0001$ ;  $V^{\text{mid}}$ : estimate =  $-0.97, p < 0.0001$ ), while there is no effect on the local timing interval ( $C^{\text{tar}}$ : estimate =  $0.01$ ;  $C^{\text{max}}$ : estimate =  $0.005$ ;  $V^{\text{mid}}$ : estimate =  $0.04$ ). **Figure 15** plots the duration of the two intervals, global timing and local timing, with  $C^{\text{tar}}$  as anchor for the two prosodic conditions, word boundary (WB) and utterance phrase boundary (UT) across speakers, as a function of the number of consonants (CV, C#CV).



**Figure 15:** Duration (in milliseconds) of the two intervals global timing and local timing for CV (white) and voiceless C#CV (grey) words in two prosodic conditions. Across prosodic conditions, word boundary (WB) and utterance phrase boundary (UT), the difference between CV and C#CV for the global timing interval is far greater than the difference for the local timing interval.

To summarize, across word-initial CV~CCV sequences, local timing interval stability was observed across three different anchors, across prosodic conditions, and regardless of the initial stop's voicing. When a voiceless stop precedes the lateral-vowel string, however, both the global and the local timing intervals were perturbed to a larger extent than when the initial stop was voiced. Across voicing, for word-initial CV~CCV, local timing interval stability was found across three different anchors within each prosodic condition, word boundary, and utterance phrase boundary. However, with prosodic strengthening, the global and local timing intervals change to some extent as a consonant is added to the CV string. This means that with prosodic strengthening, the local timing interval stability begins to weaken and the global timing interval stability tends to improve. On the other hand, across CV~C#CV, local timing interval stability was consistently observed across three different anchors and across prosodic conditions. When looking within each prosodic condition, the local timing interval stability as well as the degree of change of the local timing interval from CV to C#CV remains minimal and remarkably invariant. Thus, with prosodic strengthening, the local timing interval remains the same, while the global timing interval changes to a large extent across CV~C#CV. How the above results of the stability-based heuristics fit together with the rest of the results from vowel initiation, C2 lateral plateau duration, and the IPI-C2 lateral compensatory relation, as seen in the previous sections, will be addressed in Section 4.

### 3.7. Interim summary of findings

We summarize here, in **Table 8**, the findings as reported in the preceding. Our study begins by documenting a set of basic results concerning the phonetic parameters of IPI, C1 stop duration, and lateral duration (the first three rows in **Table 8**). These are the three parameters that make up the CC part of the CCV syllable. Demonstrating that these parameters vary systematically as a function of our prosodic condition is prerequisite, in our approach, to the study of the consequences of their variability for the rest of the segmental sequence (that may be globally or locally organized depending on context). Therefore, the findings regarding these first three parameters lay the foundation for the rest of our results which concern relational properties. These relational properties, tabulated the last three rows of **Table 8**, concern the IPI-C2 lateral duration relation, the initiation of the vowel with respect to its preceding consonantism, and the relative stability of the local and global timing intervals. These last three areas of our findings play a major role in the view developed in the next section which underscores the non-uniqueness of phonetic indices expressing syllabic organization and the presence of inter-dependencies among parameters in the sequence of segments that are organized globally.

	<b>Word-initial CCV</b>	<b>Cross-word C#CV</b>
IPI	lengthens from WB to UT	lengthens from WB to UT
C1 stop	lengthens from WB to UT	shortens from WB to UT
C2 lateral from CV to CCV	no change in WB shortens in UT	lengthens in WB no change in UT
IPI-C2 relation	moderate in WB no relation in UT	<i>not applicable</i> <i>not applicable</i>
V initiation from CV to CCV	no change in WB earlier in UT	later in WB no change in UT
Stability patterns	local stability in WB local stability worsens in UT	local stability in WB local stability in UT

**Table 8:** Summary of the main findings on each variable of interest (first column) for the two sequence types, word-initial CCV (second column) and cross-word C#CV (third column), under the two prosodic conditions, the control or word boundary (WB) and the utterance phrase boundary (UT) condition.

#### 4. Multiplicity of phonetic indices for phonological organization

We have compared how word-initial versus cross-word sequences, as in word-initial /kl/ in /kla:gə/ versus cross-word /k#l/ in /pak#la:gə/, respond to perturbations under different prosodic conditions. The word-initial stop-lateral clusters instantiate the global organization because they are prototypical complex onsets in German and other Germanic languages. The cross-word stop-lateral clusters are not considered complex onsets as the first consonant constitutes the final consonant of the first word and the second consonant constitutes the first consonant of the second word; German is well-known for its strict avoidance of across word resyllabification processes (Wiese, 1996; McCarthy & Prince, 1993). In /pak#la:gə/, the stop /k/ is coordinated with the lateral which in turn enters into a coordination relation with the vowel but the stop and the vowel are not part of the same phonological unit and are thus not coordinated with one another (as would be the case in global organization in /kla:gə/ where the entire CCV would be one syllable). Such /k#l/ sequences instantiate the local organization. In pursuing this comparison, instead of seeking or promoting a single, privileged index of (local versus global) mode of organization (i.e., as in stability-based indices), we have studied the consequences of perturbations of phonetic parameters, such as IPI or C1 lengthening, throughout the segmental sequences over which the presumed organizations preside. What we learn from this comparison is that clusters under the two different organizations respond differently to perturbations of phonetic parameters such as IPI or C1 lengthening. In this section, we put together the effects pointing to this conclusion.

#### 4.1. IPI – C2 lateral duration relation

Word-initial stop-lateral CCV sequences in the control condition (word boundary or WB) exhibit a compensatory relation such that as IPI increases C2 lateral duration decreases. This compensatory relation is a first indication of global organization. Specifically, when the lag between the plateaus of the two consonants, what we call IPI, in a CCV increases (a perturbation which would push the first consonant farther away from its tautosyllabic vowel), shortening of the second consonant compensates by bringing the vowel to overlap more with its tautosyllabic cluster. The CCV sequence thus seems to be organized globally: If each segment in a CCV were planned independently of the other segments, then an increase or decrease in the duration of that segment or inter-segmental interval is not predicted to result in a decrease or increase in the duration of the other. If, instead, the segments are planned as a group (globally), such compensatory relations are expected.

As we have seen, the two variables that enter into the compensatory relation, IPI and C2 lateral duration, individually show sufficient variability in the WB condition. However, in the UT condition, the range of variability in IPI and C2 lateral duration shrinks. This precludes or at least reduces the chances for such a compensatory relation to be expressed. In other words, the effects of prosodic strengthening on the duration of the first C in CCV as well as on the duration of the IPI between the two consonants freeze the extent of variability in these two parameters, thus precluding the manifestation of a compensatory relation between the two. Emphatically, however, this does not mean that there are no indications for global organization in the UT condition. To the contrary, we find strong and perhaps even stronger indications of global organization in that condition, to be addressed next in subsections 4.2 and 4.3.

#### 4.2. Stability-based heuristics

Our focus here turns to the ways the global and local timing intervals change from CV to CCV. Diverging patterns of such change will be revealed as a function of syllabic structure in the condition of prosodic strengthening (UT condition), that is, the condition where variability is introduced by changing phonetic parameters in the cluster. Previous studies using stability-based heuristics on German (e.g., Brunner et al., 2014) failed to find robust evidence for global organization because they did not include such a condition in their experimental designs. Our result here corroborates the same larger theme as seen in our other findings: The mode of organization of any given segmental sequence is manifested by introducing perturbations or variability in that sequence. Such variability allows one to observe whether local perturbations to segments or relations between segments have effects that ripple through the rest of the sequence in attestation of the global model of organization presiding over the entire sequence.

Let us recall again the definitions of the two relevant intervals. The local timing interval is defined as the interval between the release of the prevocalic consonant and the anchor and the

global timing interval is defined as the interval between the c-center of the consonant(s) and the anchor. For word-initial clusters, local timing interval stability is observed across CV, CCV. This same result has been seen for some clusters also in Brunner et al. (2014) and in Pouplier (2012). However, with prosodic strengthening (a condition not included in previous studies on German) the local timing interval becomes less stable while the global timing interval becomes more stable compared to the control condition (word boundary). Let us first disentangle effects of prosodic strengthening from effects of syllabic structure that pertain to this result. As seen in 3.2, the lateral is shorter in CCV than in CV in the UT condition, presumably due to differential lengthening based on the lateral's proximity to the prosodic boundary; the lateral is shorter in CCV than in CV at least in part due to the former being farther from the prosodic boundary than the latter. This lateral shortening from CV to CCV affects the calculation of the c-center landmark, which left-delimits the global timing interval. A shorter prevocalic segment (in our case, the lateral) causes the c-center landmark to occur earlier than if the same segment would be longer, all else being equal. Therefore, any change in the lateral duration caused by something other than syllabic organization could be responsible for the global timing interval becoming gradually more stable with prosodic strengthening. This means that the improved global timing interval stability in the UT condition may not be driven exclusively by syllabic organization. As opposed to the global timing interval, the local timing interval and the worsening of its stability in the UT condition is not affected by the lateral shortening from CV to CCV. Lateral shortening from CV to CCV in itself does not affect the local timing interval, which is left-delimited by the release of the lateral, by shortening it. Shortening of the local timing interval from CV to CCV can occur either by shortening and shifting the lateral towards the vowel or by just lengthening the lateral. Just shortening the lateral without shifting it towards the vowel would result in lengthening of the local timing interval. Therefore, the crucial parameter responsible for the instability of the local timing interval is shifting of the lateral towards the vowel which occurs in UT, the condition of prosodic strengthening, i.e., as IPI and C1 lengthen.<sup>2</sup>

---

<sup>2</sup> It is worth noting here that closer inspection of the data reveals further fine-grained differences in interval stabilities. Thus, local timing interval shortening is greater for voiceless than for voiced stop-lateral clusters. This implies that the lateral is shifted somewhat more towards the vowel when the initial stop is voiceless than when it is voiced. Why may this be so? It seems plausible to attempt to relate the degree of shortening of the local timing interval depending on the stop's voicing to the duration of the lateral which is shorter after voiced than after voiceless stops. Specifically, a shorter lateral in contexts where the lateral is not shifted to the vowel to a large extent serves to bring the vowel to overlap more with its preceding cluster. Such shortening may thus be seen as a signature of global organization in the case of voiced stop-lateral clusters. Shortening of the lateral is not as necessary in voiceless stop-lateral clusters (hence the longer lateral in voiceless than in voiced stop-laterals), because the lateral is shifted to the vowel more in voiceless than in voiced stop-laterals. Hence, overlap between the vowel and its preceding cluster can be achieved in different ways depending on the segmental composition of the cluster (either by shortening the lateral in the context where it cannot be shifted closer to the vowel or by shifting the lateral closer to the vowel in the context where the lateral does not shorten). We leave a more systematic investigation of these asymmetries and the validity of our suggested interpretation to future work.

Once again, the familiar crucial methodological theme of our study emerges. It is when we introduce variability via the prosodic boundary manipulation (a manipulation not present in earlier studies on German as in Brunner et al., 2014) that the local timing interval ceases to be as stable and thus evidence for global organization in German onsets begins to emerge.

A markedly different pattern is observed for cross-word sequences. Specifically, for cross-word sequences, local timing interval stability is observed across prosodic conditions and within prosodic conditions as the number of consonants increases from CV to C#CV. The local timing interval remains remarkably stable (difference of 0 to 2 ms from CV to C#CV) in the UT condition, while the global timing interval changes to a great extent. As we have reported in 3.2, there are no lateral duration differences between CV and C#CV in the UT condition that could affect any of the local and global timing intervals. Therefore, when variability is introduced by changing phonetic parameters in the cluster, such as increasing the lag between the plateaus of the consonants, syllabic organization makes different predictions about the way the local timing interval changes. For word-initial clusters, the local timing interval progressively becomes less stable, while it remains remarkably stable for cross-word sequences.

### 4.3. Vowel initiation

We summarize here our results on vowel initiation with respect to the target of the prevocalic lateral in CV and in clusters of different syllabic affiliation, complex onsets CCV, and cross-word C#CV sequences, in two prosodic conditions.

For word-initial clusters, the CV lag, which quantifies the interval between the target of the prevocalic lateral and vowel initiation, does not change between CV and CCV in the control condition (word boundary or WB); i.e., there is no difference in vowel initiation from CV to CCV in the control condition. However, in the UT condition, the vowel starts earlier with respect to the target of the prevocalic lateral in CCV as seen by a reduction of the CV lag compared to CV. Once again, when we look at CV, CCV sequences statically, that is, just at the WB condition, German does not offer much evidence for the expected increased overlap between the vowel and its preceding consonantism in CCV. But when, unlike in previous work on German (Brunner et al., 2014), we introduce perturbations to these sequences by placing them in different prosodic conditions, that is, comparing WB to the UT condition, then evidence for this increased overlap between the vowel and the CCV emerges.

Consider now CV and cross-word C#CV sequences. With respect to vowel initiation in cross-word C#CV sequences compared to single CV sequences, the CV lag is either longer or remains the same in C#CV compared to CV across prosodic conditions. Specifically, in the UT condition, vowel initiation does not change with respect to the prevocalic lateral across CV~C#CV. This is exactly the opposite of what happens between CV and CCV (without a word boundary) sequences where the vowel occurs earlier with respect to the prevocalic lateral in the cluster CCV than in CV across prosodic conditions.

Summarizing, for word-initial clusters, vowel initiation occurs earlier in CCV than in CV in the UT condition. Specifically, vowel initiation occurs earlier with respect to the target of the lateral in the cluster CCV case as the CV substring in CCV is compressed due to prosodic strengthening. The first C and the IPI in CCV lengthen due to prosodic strengthening. When these effects take place, what we observe is that the rest of the string (the inner CV) has to shorten to compensate for the first C's and IPI's increased length. Once again, these are signatures of a global organization. If each segment in the CCV were planned independently of the other segments, such compensatory adjustments are not expected. In other words, that the lag between vowel initiation and the target of the prevocalic lateral decreases with prosodic strengthening is another species of compensatory effect, just like the IPI-C2 lateral duration relation, which serves to indicate the presence of a global organization by bringing the vowel to overlap more with the cluster than it would otherwise be the case if vowel initiation were not to occur earlier, as in the cross-word sequences. For cross-word sequences, where no organizing principle prescribes that the vowel overlaps with the cluster as a unit, there is no change in vowel initiation from CV to C#CV when expansion (IPI lengthening) occurs due to prosodic boundary strength in the UT condition. Thus, the vowel initiation patterns differ as a function of syllabic structure and these differences become apparent as phonetic parameters are varied as in contexts of increasing IPI.

## 5. Conclusion

We studied stop-lateral word-initial clusters /bl, pl, gl, kl/ and cross-word sequences /p#l, k#l/ produced by five native speakers of German in two prosodic conditions. These two conditions induced a host of perturbations on various phonetic properties in the consonants of these clusters and their relation to the following vowel (see Byrd & Choi, 2010 and Fougeron & Keating, 1997 for similar effects of prosodic boundaries in English). The main finding can be summarized by saying that the effect of these perturbations on the spatiotemporal coordination patterns depends on the syllabic organization superimposed on these clusters. Changing phonetic parameters in consonant clusters leads to different inter-segmental coordination patterns depending on whether these clusters are word-initial complex onsets or cross-word sequences which do not combine to form a syllable onset in German. In short, syllabic organization is reflected in the spatiotemporal patterns of the segments partaking in this organization. However, to reveal these patterns, we must go beyond stability-based indices and specifically beyond static statements along the lines of “complex onset is reflected in global timing stability” and “simplex onset in local timing stability.” As has been argued before (Shaw et al., 2011, 2009) and as we have seen in our results and in results from previous work on German (Brunner et al., 2014; Pouplier, 2012), such stability-based heuristics are unreliable by themselves in consistently diagnosing syllabic organization. Moreover, apart from looking at stability-based heuristics, we extended our investigation to other measures and specifically to how these other measures as well as their relation to one another are affected as phonetic parameters are scaled. As we have seen, joint consideration of several measures points to the conclusion that syllabic organization is not reflected only in stability-

based heuristics (indeed, exclusive attention to these heuristics is unreliable or points to the wrong conclusions as evidenced in our results) but also in other measures such as the duration of the prevocalic consonant depending on the voicing of the initial stop, vowel initiation in relation to the cluster, and compensatory effects between IPI (the lag between the two consonants) and duration of the prevocalic consonant. Specifically, as we have seen, increasing IPI and C1's duration in word-initial stop-lateral clusters leads to earlier vowel initiation from CV to CCV and reduced local timing interval stability, with both effects indicating the presence of global organization. Furthermore, a compensatory relation between IPI and C2 lateral plateau is also observed within CCV when the variability of these two parameters is sufficient (word boundary condition) so as to allow such an effect to emerge. The presence of such compensatory effects indicates that the organization of the different parts of CCV is not independently planned and produced and thus such effects offer evidence for global organization. In other words, we find that the global organization presiding over the segments partaking in these tautosyllabic CCVs is pleiotropic (Sotiropoulou et al., 2020), that is, simultaneously expressed over a set of different phonetic parameters rather than via a privileged metric such as c-center stability or any other such given single measure (employed in prior works).

Markedly different effects are observed when we turn to cross-word stop-lateral clusters. In the latter, increasing the lag between two consonants leads to no change with respect to vowel initiation across CV~C#CV and robust local timing interval stability. This scaling of phonetic parameters (e.g., increasing the lag between the two consonants in CC and C#C) is the crucial methodological diagnostic that has allowed us to reveal how different syllabic organizations (CC as a word-initial complex onset versus the same cluster across words in C#C) modulate the spatiotemporal coordination of the same segmental material.

To conclude, our study further informs research on the relation between syllabic organization and phonetic indices. Going beyond prior work, joint consideration of various measures, such as prevocalic consonant shortening, vowel initiation, and compensatory effects between inter-consonantal lag and prevocalic consonant duration proves to be highly informative. As we have seen, the relation between syllabic organization and phonetic indices is not expressed in static statements which remain valid all along regardless of, say, variability in the lag between the two consonants. Rather, as the lag between the two consonants varies, different syllabic organizations respond to that variability differently. In the complex onset organization, as the lag between consonantal plateaus increases, the local timing interval ceases to be as stable and the vowel starts progressively earlier to establish more overlap with the cluster. In the same clusters but now parsed not as complex onsets, as the lag between consonantal plateaus increases, the local timing interval remains remarkably stable and there is no change on the vowel initiation pattern. Overall, we have seen that revealing the systematic but at times subtle effects syllabic organization has in our data requires the joint consideration of several measures and the disclosure of compensatory relations which offer telling indicators of the mode of phonological organization.

---

## Acknowledgements

This work has been funded by the European Research Council (AdG 249440) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 317633480 – SFB 1287, Project C04.

## Competing interests

The authors have no competing interests to declare.

## Author contribution

Both authors contributed equally to this work.

---

## References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Bombien, L. (2011). *Segmental and prosodic aspects in the production of consonant clusters* (Doctoral dissertation). Ludwig-Maximilians-Universität Munich, Germany.
- Bombien, L., & Hoole, P. (2013). Articulatory overlap as a function of voicing in French and German consonant clusters. *Journal of the Acoustical Society of America*, 134(1), 539–550. DOI: <https://doi.org/10.1121/1.4807510>
- Bombien, L., Mooshammer, C., & Hoole, P. (2013). Articulatory coordination in word-initial clusters of German. *Journal of Phonetics*, 41, 546–561. DOI: <https://doi.org/10.1016/j.wocn.2013.07.006>
- Bombien, L., Mooshammer, C., Hoole, P., & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, 38(3), 388–403. DOI: <https://doi.org/10.1016/j.wocn.2010.03.003>
- Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, 45(2–4), 140–155. DOI: <https://doi.org/10.1159/000261823>
- Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée*, 5, 25–34.
- Brunner, J., Geng, C., Sotiropoulou, S., & Gafos, A. (2014). Timing of German onset and word boundary clusters. *Journal of Laboratory Phonology*, 5(4), 403–454. DOI: <https://doi.org/10.1515/lp-2014-0014>
- Byrd, D. (1995). C-Centers revisited. *Phonetica*, 52, 263–282. DOI: <https://doi.org/10.1159/000262183>
- Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics—The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In C. Fougeron, B. Kühnert, M. d’Imperio & N. Vallée (Eds.), *Papers in Laboratory Phonology 10*. Berlin: Mouton de Gruyter.

- Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht, Netherlands: Kluwer. DOI: <https://doi.org/10.1007/978-94-010-0279-0>
- Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America*, *101*, 3728–3740. DOI: <https://doi.org/10.1121/1.418332>
- Gafos, A. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory*, *20*(2), 269–337. DOI: <https://doi.org/10.1023/A:1014942312445>
- Gafos, A., Charlow, S., Shaw, J. A., & Hoole, P. (2014). Stochastic time analysis of syllable referential intervals and simplex onsets. *Journal of Phonetics*, *44*, 152–166. DOI: <https://doi.org/10.1016/j.wocn.2013.11.007>
- Gafos, A., Roeser, J., Sotiropoulou, S., Hoole, P., & Zeroual, C. (2020). Structure in mind, structure in vocal tract. *Natural Language and Linguistic Theory*, *38*, 43–75. DOI: <https://doi.org/10.1007/s11049-019-09445-y>
- Goldstein, L., Chitoran, I., & Selkirk, E. (2007). Syllable structure as coupled oscillator modes: Evidence from Georgian versus Tashlhiyt Berber. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the sixteenth International Congress of Phonetic Sciences (ICPhS)* (pp. 241–244). Saarbrücken, Germany.
- Hermes, A., Auris, B., & Mücke, D. (2015). Computational modelling for syllabification patterns in Tashlhiyt Berber and Maltese. In S. Fuchs, M. Grice, A. Hermes, L. Lancia & D. Mücke (Eds.), *Proceedings of the 10th International Seminar on Speech Production (ISSP)* (pp. 186–189). Cologne, Germany.
- Hermes, A., Mücke, D., & Auris, B. (2017). The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics*, *64*, 127–144. DOI: <https://doi.org/10.1016/j.wocn.2017.05.004>
- Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: The case of ‘impure s’. *Phonology*, *30*, 1–25. DOI: <https://doi.org/10.1017/S095267571300002X>
- Honorof, D., & Browman, C. P. (1995). The centre or edge: How are consonant clusters organised with respect to the vowel? In K. Elenius & P. Branderud (Eds.), *Proceedings of the 13th International Congress of the Phonetic Sciences (ICPhS)* (pp. 552–555). Stockholm, Sweden.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. DOI: <https://doi.org/10.1002/bimj.200810425>
- Lenth, R., & Hervé, M. (2015). *Package ‘lsmeans’*. Retrieved from [cran.r-project.org/web/packages/lsmeans/lsmeans.pdf](https://cran.r-project.org/web/packages/lsmeans/lsmeans.pdf).
- Marin, S. (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics*, *41*, 211–227. DOI: <https://doi.org/10.1016/j.wocn.2013.02.001>
- Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*(3), 380–407. DOI: <https://doi.org/10.1123/mcj.14.3.380>
- Marin, S., & Pouplier, M. (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. *Journal of Phonetics*, *42*, 24–36. DOI: <https://doi.org/10.1016/j.wocn.2013.11.001>
- McCarthy, J. J., & Prince, A. S. (1993). Generalized alignment. In *Yearbook of Morphology*, 79–153, Kluwer. DOI: [https://doi.org/10.1007/978-94-017-3712-8\\_4](https://doi.org/10.1007/978-94-017-3712-8_4)

- Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure, *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)* (pp. 2253–2256). Barcelona, Spain.
- Pastätter, M., & Pouplier, M. (2015). Onset-vowel timing as a function of coarticulation resistance: Evidence from articulatory data, *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*. Glasgow, UK.
- Pouplier, M. (2012). The gestural approach to syllable structure. In S. Fuchs, M. Weirich, D. Pape & P. Perrier (Eds.), *Universal, language and cluster-specific aspects. Speech planning and dynamics* (pp. 63–96). Frankfurt am Main: Peter Lang AG.
- RStudio Team. (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from <http://www.rstudio.com/>.
- Shaw, J., & Gafos, A. (2015). Stochastic time models of syllable structure. *PloS One*, 10(5). eCollection 2015. DOI: <https://doi.org/10.1371/journal.pone.0124714>
- Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2009). Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability. *Phonology*, 26(1), 187–215. DOI: <https://doi.org/10.1017/S0952675709001754>
- Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure. *Phonology*, 28, 455–490. DOI: <https://doi.org/10.1017/S0952675711000224>
- Sotiropoulou, S., Gibson, M., & Gafos, A. (2020). Global organization in Spanish onsets. *Journal of Phonetics*, 82, 1–22. DOI: <https://doi.org/10.1016/j.wocn.2020.100995>
- Wiese, R. (1996). *The phonology of German*. Oxford, UK: Oxford University Press.

