Pleiotropy of Phonetic Indices in the Expression of Syllabic Organization

A thesis submitted for the degree of

Doctor of Philosophy

by

Stavroula Sotiropoulou

Supervisor: Prof. Dr. Adamantios Gafos, Universität Potsdam

January 24, 2019

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-54639
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-546399

Acknowledgments

This page contains personal data and was removed from the document.

Erklärung

Ich erkläre, dass die Arbeit selbständig und ohne Hilfe Dritter verfasst wurde und bei der Abfassung alle Regelungen guter wissenschaftlicher Standards eingehalten wurden.

Stavroula Sotiropoulou

....

Datum Unterschrift

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Topic and main claim

This dissertation is concerned with the relation between qualitative phonological organization in the form of syllabic structure and continuous phonetics, that is, the spatial and temporal dimensions of vocal tract action that express syllabic structure. The main claim of the dissertation is twofold. First, we argue that syllabic organization exerts multiple effects on the spatio-temporal properties of the segments that partake in that organization. That is, there is no unique or privileged exponent of syllabic organization. Rather, syllabic organization is expressed in a pleiotropy of phonetic indices. Second, we claim that a better understanding of the relation between qualitative phonological organization and continuous phonetics is reached when one considers how the string of segments (over which the nature of the phonological organization is assessed) responds to perturbations (scaling of phonetic variables) of localized properties (such as durations) within that string. Specifically, variation in phonetic variables and more specifically prosodic variation is a crucial key to understanding the nature of the link between (phonological) syllabic organization and the phonetic spatio-temporal manifestation of that organization. The effects of prosodic variation on segmental properties and on the overlap between the segments, we argue, offer the right pathway to discover patterns related to syllabic organization. In our approach, to uncover evidence for global organization, the sequence of segments partaking in that organization as well as properties of these segments or their relations with one another must be somehow locally varied. The consequences of such variation on the rest of the sequence can then be used to unveil the span of organization. When local perturbations to segments or relations between adjacent segments have effects that ripple through the rest of the sequence, this is taken as 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.

Let us consider some specific examples illustrating the main claim. If perturbations such as lengthening the duration of the initial consonant (C1) or increasing the interval between two consonants in a consonant-consonant-vowel (CCV) sequence are compensated by shortening of other elements in the CCV string, such compensatory effects (something lengthens and in response to it something shortens) are indications of global organization that serves to bring the vowel to overlap with the cluster. For example, when the lag between the plateaus of the two consonants in a CCV, what we call interplateau interval (IPI), increases, our Spanish data indicate that the lateral in the second position (or C2) of the C1C2 cluster shortens to compensate for the IPI increase, thus bringing the vowel to overlap more with its tautosyllabic cluster (than it would otherwise be the case if no such compensatory relation existed). The CCV sequence thus seems to be organized globally. This IPI-C2 lateral duration compensatory relation is found in our Spanish data (e.g., in /klapas/) but not in the Moroccan Arabic data (e.g., in /klat/) where the same sequence of segments (same as in Spanish) is claimed not to be organized as a complex onset but rather with the first consonant in the CCV as part of a separate unit from that of the CV.

Another example can be illustrated from our German study by observing vowel initiation with respect to the prevocalic consonants, that is, the timing of the vowel-specific tongue movements in relation to the preceding consonants (this dissertation makes systematic use of this measure for the first time on the topic of syllable structure and its phonetic manifestations). When perturbations cause some parts of the hypothesized syllable to expand (C1 or IPI lengthening), we observe that as a result of this expansion the inner CV substring in CCV gets compressed. Specifically, when C1 and IPI in CCV lengthen due to prosodic strengthening (by embedding target words starting with a CCV in carrier phrases of increasing prosodic boundary), the lag between the target of the prevocalic lateral and the start of vowel-specific tongue movement across stop-lateral clusters in German decreases; i.e., the rest of the string (the inner CV) shortens to compensate for the C1's and IPI's increased length. Once again, this is a signature of global organization. Earlier vowel initiation in CCV with increasing boundary strength is another species of compensatory effect, just like the IPI-C2 lateral duration relation, which serves to indicate the presence of a global organization. If, instead, each segment in a CCV was planned independently of the other segments (that is, not global, but local organization), then an increase or decrease in the duration of a segment 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.

Let us highlight some key ingredients in the main claim of the thesis that the above examples have helped to bring out. First, even with our only two examples presented so far, we have referred to different measures drawn from a number of dependent variables to argue for the same concept of global organization. To wit, we have used the presence of a relation involving the interplateau interval (IPI) between the two consonants and the second consonant's duration but also patterns related to vowel initiation with respect to the prevocalic cluster to argue that when some parts of the CCV are altered for whatever reasons, other parts of the CCV sequence respond to such alterations in ways that reveal global organization. This multiplicity of phonetic measures is related to our pleiotropy claim in this dissertation: instead of employing a single privileged metric to diagnose syllabic structure across all contexts or all clusters, the informative indices, via which syllabic structure is expressed in the phonetic record, are non-unique and they are to be found throughout the domain of the string over which organization is assessed. We argue that the relevant issue is not finding an index that 'works' (or even that works most of the time). Rather, the relevant issue is finding how phonetic measures respond to alterations or as we have called them above perturbations which the CCV sequence may undergo for a variety of reasons (such as C duration increase due to whether the C is voiceless versus voiced or due to the presence of a strong boundary before the first consonant). As we will see in detail below, a single phonetic index fails to reliably diagnose syllabic structure even within one language. Consider a simple case: /bla/ versus /pla/ in German. Both word-initial clusters are considered complex onsets in German. Brunner *et al.* (2014) and Pouplier (2012) have presented data in German indicating that the interval whose stability is thought to be characteristic of complex onsets does not in fact remain stable (with the addition of a consonant from CV to CCV) for the word-initial /pl/ cluster. However, the expected stability of the same interval is found for the word-initial /bl/ cluster (as seen in Pouplier 2012; Brunner *et al.* 2014 had no stimuli with /bl/ clusters but had stimuli with /pl/ clusters for this labial stop-lateral context). This shows that a single phonetic index or here (interval) stability index fails to reliably diagnose syllabic organization; it works successfully for some clusters but not for others. The same logic that one phonetic index fails to diagnose syllabic organization extends to across languages. Consider again a simple case: /bla/ in German versus /bla/ in Spanish. Even though the phonological structure superimposed on these segments is uniform in that in both languages /bla/ is uncontroversially a single syllable, the segments involved across the two languages exhibit different phonetic properties which also affect their temporal organization in the cluster. In German, the two consonants are highly overlapped in this context. But in Spanish, this is not so. Spanish consonant-to-consonant timing in /bl/ resembles in terms of its interplateau interval German /pl/; the two consonants are separated by a substantial interplateau gap (more details and an illustration of a /gl/ cluster in Spanish and German follows in the next section). The phonological organization may be the same but the language-particular phonetics are different in /bl/ across German and Spanish.

A single index of global organization is bound to fail also across languages just as it fails within German. In this example involving German and Spanish, it fails as follows. Under the high overlap conditions in the German /bl/, the stability measures turn out to be well behaved: the expected interval for global organization is the most stable. But in Spanish, in terms of stability measures, the /bl/ datasets (or any voiced stop-lateral datasets for that matter) strikingly point to the opposite pattern, namely, to local or simplex onset organization even though these clusters are uncontroversially syllable onsets. Once again, a single index of global organization is not feasible.

In the new perspective of this dissertation, we claim that there is no need to seek a single phonetic index which captures syllabic organization. Rather, the mode of organization is revealed by identifying compensatory relations among phonetic variables as we have illustrated in our earlier examples of such relations. An advantage of our approach is that it enables reliable diagnosis of syllabic structure even in cases where the usual stability indices either fail to offer clear indications of the mode of organization or even point to the wrong mode of organization.

1.2 Rationale of the experimental studies design

The dissertation investigates the relation between syllabic structure and inter-segmental temporal coordination patterns by employing data from three languages: Spanish, Moroccan Arabic and German. The set of languages was chosen for specific reasons which we make explicit in what folllows.

Spanish and German were chosen because they share the same (relevant) phonology but differ in (again) the relevant phonetics. Sharing relevant phonology means for us that Spanish and German share the same syllabic structure in that they admit complex onsets. Word-initial stop-lateral consonant clusters are said to be the prototypical onset clusters for syllables both in Spanish and German. The words "globo" in Spanish and "glauben" in German consist of two syllables, /glo.bo/ and /glau.ben/ respectively where '.' indicates syllabic boundary. The cluster /gl/ in both languages belongs to the first syllable of the word and forms the onset of that syllable.

Differing in the relevant phonetics means that the voicing contrast in, for example, across /gl/ versus /kl/ clusters, which are found in both languages ("glauben-Klage" in German and "globo-clono" in Spanish), is implemented differently. Specifically, the implementation of voicing in German is based on a Voice-Onset-Time (VOT) short lag versus long lag distinction with voiced stops having a short VOT and voiceless stops having a long VOT (Jessen 2002). The voicing distinction in Spanish, instead, is based on a voicing during closure (or negative VOT) versus short lag VOT distinction with voiced stops being fully voiced and voiceless stops having a short VOT (Lisker & Abramson 1964, Bombien & Hoole 2013). In effect, for both languages, clusters such as /bl/, /pl/ are prototypical onsets but the segments that enter into these clusters have different phonetic properties. An overarching aim of much past and current work in the field of phonetics and phonology is to determine what may be language-independent hallmarks of phonological organization; in this case, organization characterizing complex onsets. Given this aim, such a choice of languages where the phonological structure across the two languages remains the same but the phonetics are different offers highly appropriate chances to reach that aim. We thus investigate how this difference in phonetic properties or language-specific phonetics affects spatio-temporal coordination of the segments that are part of these clusters. Because the phonological organization is claimed to be the same across the two languages (again, in both languages, /bl/ is a complex onset), any similarities in how segments are organized across the two languages (despite the distinct phonetic properties of the segments across the two languages) offer clues on how qualitative phonological organization relates to the continuous phonetics and specifically on how syllabic organization is manifested on groups of segments with varying phonetic identities.

Let us delve more into what this difference in (relevant) phonetics contributes to our investigation. The differing implementation of voicing contrast across the two languages affects the intra-cluster timing of word-initial clusters in these two languages differently in that overall the lag between onset consonants (IPI) in Spanish is greater than in German. Specifically, in Spanish, voiced stop-lateral clusters exhibit an IPI of around 40 ms, while voiceless stop-lateral or voiceless stop-tap clusters exhibit an IPI of around 60 ms. This rather large lag between stop-lateral and stop-tap clusters is known as an 'open transition' (Gafos 2002 for theory and simulations, Gibson *et al.* 2017 for an articulatory study on Spanish, Malmberg 1965, Colantoni & Steele 2005, Bradley 2006 for acoustic studies on Spanish) and it describes a period of no vocal tract constriction between the release of the stop and the achievement of target of the lateral or tap. In German, however, voiced stop-initial clusters exhibit a very short, sometimes negative IPI, while voiceless stop-initial clusters exhibit an IPI of around 40 ms (Bombien & Hoole 2013, our study as well). That is, robust open transitions are not attested in German, at least not for voiced stop-initial clusters. Figure 1.1 illustrates the IPI difference between a /gl/ cluster in our German and Spanish data. In German, there is temporal overlap between the plateaus of the consonants, while in Spanish there is no overlap between the plateaus.

Overall, onset consonants in Spanish are timed farther apart than onset consonants in German. This is then how intra-cluster timing is affected by the implementation of voicing in these two languages and it is this difference in language-particular phonetics that we

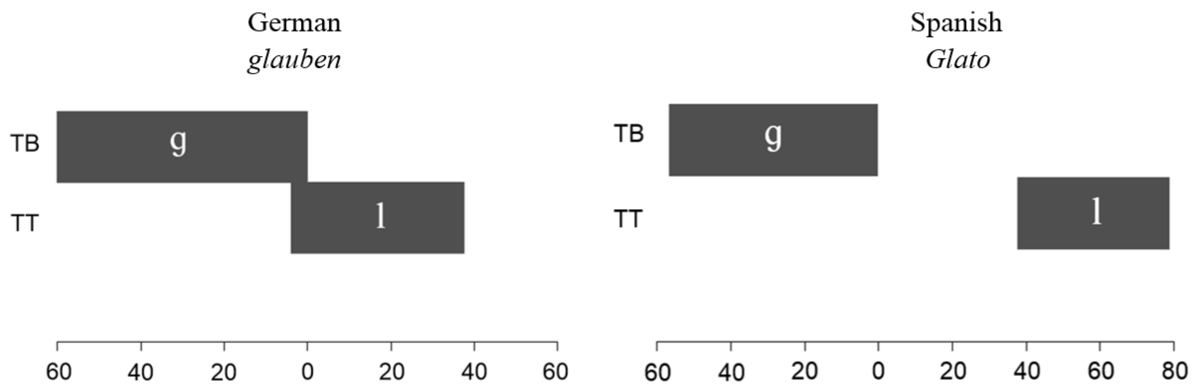

Figure 1.1: Representative schemas of gestural overlap for the cluster /gl/ in German (lefthand side) and Spanish (righthand side). Spanish shows a substantial IPI or time lag between the plateaus (constriction phases) of the two consonants, as shown by the separation between the dark rectangles in the righthand schematic, whereas in German the plateaus overlap in time.

exploit in seeking the presence of global organization constraints in the two languages. Moreover, voicing in Spanish has an effect also on the duration of the stop in that voiced stops are shorter than voiceless stops (this is not the case in our German cluster data). This relates to the fact that voiced stops in Spanish are fully voiced and maintaining voicing during closure is disfavored (Ohala 1983).

We now explain how we employ these phonetic differences across Spanish and German to reveal similarities in the global organization of their complex onsets, despite the language-particular differences in their phonetics. The key to revealing these similarities turns out to be the identification of the consequences of variability in individual phonetic dimensions for the rest of the segmental sequence. For our purposes, the large range of IPIs found between the two consonants in CC Spanish clusters and the variability in duration in the C1 of the cluster as a result of voicing (e.g., /gl/ with a short /g/ closure versus /kl/ with a longer /k/ closure) in this language offer sufficient sources of variability which allow the identification of compensatory relations across the sequence of segments over which organization is assessed. The presence of such compensatory relations is a telling indicator of global organization.

An example from Spanish is given first. In Spanish CCV sequences, we find a compensatory relation between IPI (the lag between the two consonants) and C2 lateral duration: increasing the IPI or the lag between the constrictions of the two prevocalic consonants (those of the stop and the lateral) results in shortening of the second consonant (C2, the lateral). If each of these, in principle, independent parts of the phonetic substance of a stop-lateral cluster were timed independently of one another, this compensatory relation would not be expected. It is the presence of this relation that serves as one indicator (among others that will be discussed in detail) of the global nature of organization. Crucially, identifying this compensatory relation relies on there being sufficient variability in the two relevant variables, IPI and C2 lateral duration, so that the inverse correlation between the two can be detected.

Let us now see how variability in German also turns out to be crucial in identifying compensatory relations which are indicators of global organization. For German, IPI is overall short compared to Spanish and C1 duration does not seem to change as a function of voicing. In German, these parameters do not vary sufficiently (or at least not as much as they vary in Spanish) to allow for the observation of the consequences of their variability on the rest of the segmental sequence. Therefore, for German, we must seek other ways of introducing variability in these phonetic parameters, IPI and C1 duration. We will specifically employ changing the higher level prosodic structure of the utterance in which our target stimuli are embedded so that we can induce local modulations just before or during the part of the utterance corresponding to our target sequences. Three distinct prosodic conditions corresponding to no phrase boundary (word boundary or control condition) before the target sequence, an intonational phrase boundary before the target sequence and an utterance phrase boundary before the target sequence will be employed. One way to describe the effect these conditions have on our data is to say that in the control condition, there was no pause preceding the target word; in the intonational phrase boundary, a short pause preceded the target word; and in the utterance phrase boundary condition, a long pause preceded the target word. We will show that it is possible, using this experimental design, to cause C1 and IPI of word-initial clusters in German to lengthen under prosodic strengthening. Once again, the method is the same as in Spanish. But for German we need to work harder to induce the needed variability via careful choice of experimental design. For German, it is this prosodic variation and the modulations it causes on the CCV string which induce sufficient perturbations to the segmental sequence so as to make evident the presence of global organization constraints. Specifically, with increasing boundary strength, the lag between vowel initiation and target of the prevocalic lateral in CCV decreases; i.e., the vowel starts earlier with respect to the prevocalic lateral in CCV in contexts of prosodic strengthening. When C1 and IPI in CCV lengthen due to prosodic strengthening, we find that the rest of the string (the inner CV) has to shorten to compensate for C1's and IPI's increased length. In other words, earlier vowel initiation in CCV with increasing boundary strength is another kind of compensatory effect, just like the IPI-C2 lateral duration relation observed in Spanish, 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).

We now introduce the rationale for including Moroccan Arabic in our study. German and Spanish share the same phonology (in terms of the syllabic structure they assign to sequences like /gla/ or /kla/, for example) but the segments in these two languages exhibit different phonetic properties related to the implementation of voicing as explained above. Moroccan Arabic is another language used in this dissertation specifically in comparison with Spanish. Moroccan Arabic imposes a different phonology from that of Spanish (and German) for the same clusters but has the same voicing systems as Spanish (and thus different from that of German). While the /gl/ cluster in /globo/ is part of the onset in Spanish, the consonants of the /gl/ cluster in /glih/ in Moroccan Arabic do not form part of the onset, and thus they do not belong to the same syllable (there is by now substantial theoretical and experimental evidence that the syllabic parse of words like /glih/ in Arabic is /g.lih/; see Gafos *et al.* 2018 and references therein). Whereas the phonology differs between Spanish and Moroccan Arabic in how clusters are assigned syllabic parses, the implementation of voicing in the segments of these clusters is the same across the two languages. This is the converse scenario of what we set up for German and Spanish above, where the phonology was the same but the phonetics were different. For the Spanish, Moroccan Arabic duo, the phonology differs but the phonetics are the same. Across the two languages, voiced stops show lead voicing versus short lag VOT (for voiceless) stops. Furthermore, intra-cluster timing in both languages shows robust open transitions. That is, in both Moroccan Arabic and Spanish, the consonants in a cluster are timed further apart than consonants of a word-initial cluster in German. Therefore, a comparison between Moroccan Arabic and Spanish word-initial clusters allows us to examine the effects of contrasting syllabic organization on the temporal coordination of segments which share similar phonetic properties.

For German, instead of using another language as the contrast case for a different syllabic organization, we compare clusters with the same phonetic properties but of a different syllabic affiliation by staying within the (same) language. Specifically, we compare word-initial stop-lateral clusters in German to cross-word stop-lateral clusters in German. This way we have clusters with the same phonetic properties, since we stay within a language, but these clusters have a different syllabic affiliation. Word-initial stop-lateral clusters in German are syllable onsets but cross-word stop-lateral clusters in German are not syllable onsets but coda-onset sequences (same cluster, different phonological organization). It would not have been ideal to use Spanish cross-word clusters for the different phonology condition in Spanish. This is so because of two reasons. First, the set of word-final consonants is limited in Spanish to sonorants and /s/ (at least for non-loanwords). That is, no word-final stop consonants are permitted (Harris 1983, Steriade 1999). It is such word-final stops that would be needed in the first C position of a C#CV so as to compare for example a cross-word /p#lV/ condition to a word-initial /plV/ (which is of course attested). Second, it is unclear whether in Spanish there is resyllabification of the word-final consonant in the C#CV context so as to join the onset of the following syllable (see Shepherd 2003 for a literature review on Spanish; see next paragraph on the issue of resyllabification in German). In sum, cross-word clusters in German offer feasible and fruitful, as we will see, data comparisons. But for Spanish, cross-word clusters do not offer a (safely) usable data source for the same cluster, different phonology comparison.

Conversely, it would not have been ideal to pair German with Arabic. That is, it would not be ideal for German to use Arabic as a comparison for the different phonology case as we did for Spanish. Whereas it is true that in German the /k/ in /kla/ is part of the same syllable as the /la/ and that in Arabic for the same sequence /kla/ the /k/ is not part of the same syllable, the phonetic properties of the segments in the two languages are markedly different. Hence, any differences we might observe in the phonetics of a /kla/ sequence across the two languages cannot be ascribed to syllabic organization alone. Moreover, the resyllabification concerns raised with cross-word clusters in Spanish do not apply to German. In German, resyllabification across word boundaries is considered impossible (Wiese 1996, McCarthy & Prince 1993).

Overall, then our comparison sets can be summarized as follows: the comparison between Spanish word-initial and Arabic word-initial clusters represents the case of the same phonetics, but different phonology across two languages. The comparison between German word-initial and German cross-word clusters represents the case of the same phonetics, but different phonology within the same language.

In what follows, section 1.3 begins to present the background literature on the relation between syllabic structure and phonetic indices by reviewing the first key studies. Section 1.4 reviews subsequent studies on the topic and brings out issues raised by that work which in turn anticipate some of the developments to follow in this thesis. Finally, section 1.5 spells out the organization of the thesis in terms of the following Chapters.

1.3 Background

The original insight which has inspired much work on the relation between syllables and their phonetics is expressed by the idea that syllabic organization is reflected in the way segments are temporally coordinated with each other (Browman & Goldstein 1988). Cross-linguistically, there seems to be a preference that syllables consist of consonant-vowel or CV sequences (Clements 1990, Blevins 1995). As the number of consonants increases from CV to CCV, two different syllabic organizations become apparent: in one of these, both consonants combine to form part of the onset with the vowel as the nucleus of the syllable and in the other organization the sequence of consonants (CC) cannot be part of the onset. In the first case, the CCV constitutes one syllable, while in the latter case the first C in the CCV does not belong to the same syllable as the rest of the CV. Germanic and Romance languages, like English, German, French, Spanish, and Italian have been claimed to admit consonant clusters as complex onsets. For example, stop-sonorant clusters form legitimate onsets across all these languages. But in Moroccan Arabic (Dell & Elmedlaoui 2002), Berber (Dell & Elmedlaoui 1996, Prince & Smolensky 1993), and Japanese (Kubozono 1989, Blevins 1995), these clusters cannot form complex onsets. The key insight which appeared for the first time in the work of Browman & Goldstein (1988) was that whether a CCV sequence forms one syllable or not may be reflected in the timing patterns between the segments that partake in that sequence. In the case of Germanic and Romance languages, the vowel in a CCV is the nucleus of the syllable which includes both consonants (hence, complex onset), while in the case of Moroccan Arabic, Berber, and Japanese the vowel is the nucleus of a syllable which has only one consonant as its onset (hence, simplex onset). Multiple consonants are affiliated with the vowel in the first case versus a single consonant is affiliated with the vowel in the second case. The insight from Browman & Goldstein (1988) was that the temporal patterning of the CCV string would show more overlap across all segments in CCV in the first case, due to the multiple affiliation conditions, versus less overlap in the second case; less because in this latter case only the second C is affiliated with the vowel.

The study of Browman & Goldstein (1988) was conducted using articulatory data of one native speaker of American English. Sequences such as "pea splots" and "peas plots" were recorded in order to examine whether the temporal coordination of the /s/ (word-initial in "splots" and word-final in "peas") with the following segments depended on the syllabic affiliation, namely, whether the timing of /s/ with the following segments differed between the case where /s/ constituted the coda consonant of the preceding syllable as in "peas plots" and the case where it was the initial consonant of the onset in "pea splots". The cluster /spl/ was compared with the same consonants as single CV sequences /sV, pV, lV/ or with the two consonant cluster /sp/, using stimuli as in [pi sats], [pi pats], [pi lats], [pi spats], [pi splats]. Articulatory data were collected by attaching sensors on the tongue, lips and lower teeth of the subjects, and trajectories of the respective articulators during the utterance were analyzed. The sensor on the tip of the tongue was used to capture the movement related to the production of /s/ and /l/ and the lip sensor to capture the movement related to the production of /p/. Recall that the key idea was to examine the relationship between the initial consonants and the following nucleus vowel since they belong to the same syllable. However, the vowel could not be reliably parsed in this study's datasets and thus an alternative was needed to assess the relation between the vowel and the prevocalic consonants. Specifically, the achievement of target of the postvocalic consonant was considered as a rough approximation of the acoustic vowel offset. The crucial dependent measures then were expressed by the temporal distance between the prevocalic consonants, /s, sp, spl/ in [pi sats], [pi spats], [pi splats], and the postvocalic consonant /t/. In Browman & Goldstein's (1988:142–143) words: "While we initially intended to examine the relationship between the achievement of target for the initial consonantal gestures (in words like 'splot') and the achievement of target for the vocalic gestures, a re-inspection of both the horizontal and vertical movements of all the tongue pellets convinced us that an attempt to identity the vocalic movement with a single dimension of a single pellet was too gross an approximation for the present analysis. …Therefore, instead of attempting to decompose the consonant and vowel effects on these pellets, we examined the relationship between the intervocalic consonantal gesture(s) and the trans-vocalic consonantal gestures." The approach made explicit in this quote has been followed in many subsequent studies building on Browman & Goldstein's work. In this dissertation, we will depart from this tradition by aiming to quantify vowel-specific movement (among other measures to be discussed in detail later) more directly instead of using timepoints before and after the vowel as delimiters of the vowel and its relation to the surrounding consonants. We will, however, also use the measures that such work has used so as to establish continuity with this work and better understand when and why some measures may be useful (sometimes) or not useful (sometimes).

Given the above stimuli set up and measurement choices, a main result from Browman & Goldstein's study was that the interval spanning from the consonantal center (c-center) to an anchor point remained relatively stable as the consonantal environment was changing from singleton CV to more complex clustering as in CCV, CCCV. The c-center is computed by first calculating the temporal midpoint between the left and the right-edge of the plateau (or constriction phase) for every consonantal gesture and then taking the mean of all the midpoints of the consonantal gestures in a sequence. The relative stability of the c-center-to-anchor interval was seen by comparing the duration of this interval with two other intervals, left-edge-to-anchor and right-edge-to-anchor. These were defined as follows. The left-edge (le) is defined as the plateau onset of the initial consonant in the cluster and the right-edge (re) as the plateau offset of the prevocalic consonant. A common anchor landmark is used to right-delimit intervals between the consonant(s) and the vowel starting with the left-edge, c-center and right-edge landmarks. The aim is to examine how intervals delimited by these landmarks change as the number of consonant increases from CV to CCV to CCCV. The intervals from the le, the cc and the re to an anchor point have been described with the terms le-to-anchor (le2a), cc-to-anchor (cc2a) and re-to-anchor (re2a) intervals, where anchor is a landmark of the following vowel or coda consonant. Figure 1.2 illustrates the intervals le2a, cc2a and re2a defined above.

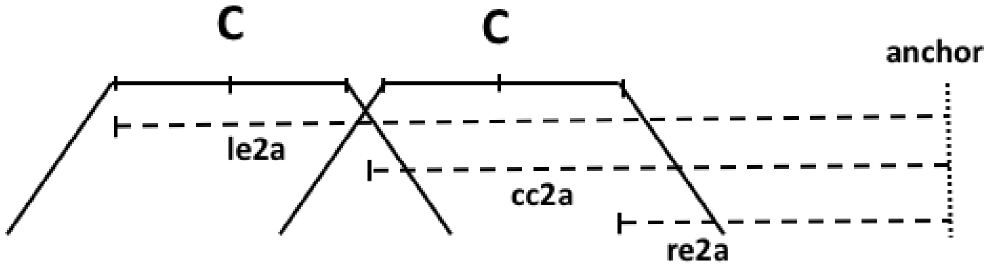

Figure 1.2: Three intervals spanning over a sequence of a consonant cluster followed by a vowel. These three intervals are delimited on the left by the left-edge (le), the c-center (cc) and the right-edge (re) landmarks and to the right by the anchor landmark which is chosen to be a time point close to the end of the vowel. The intervals are known as left-edge-to-anchor (le2a), c-center-to-anchor (cc2a) and right-edge-to-anchor (re2a).

Browman & Goldstein (1988) found that the duration of the interval stretching from the c-center to the target achievement of the postvocalic consonant, which was used as the anchor landmark, had a lower standard deviation than the other two intervals stretching from the left-edge of the initial consonant and from the right-edge of the prevocalic consonant to the anchor landmark. In other words, as the number of consonants increases from CV to CCV to CCCV, the cc2a interval showed greater stability than the le2a and re2a intervals.

The finding just reviewed concerns temporal stability of intervals within a word. To assess whether the cc2a interval describes a relation between segments which are specifically syllabically affiliated, Browman & Goldstein examined the relation between onset consonants and another consonant to which the onset consonants are not affiliated as part of the same syllable. The rationale was that if the stability of the cc2a interval is characteristic of segments which are part of the same syllable, then cc2a would not appear as the stable interval when the consonants over which c-center is calculated are not part of the same syllable as in sequences of consonants across word boundaries. In order to examine this hypothesis, the anchor was now the word-initial /p/ of the preceding word [pi] in the sequences [pi splats], [pi plats], [pi lats] and so on, while the onset consonants were of course the /l/, /pl/ and /spl/ of the words [lats, plats, splats]. The midpoint of the bilabial gesture for /p/ of the /pi/ was compared to the different landmarks left-edge, right-edge and c-center of the initial consonant(s) of the second word. Specifically, it was the standard deviation of the above three intervals le2a, cc2a, re2a that was compared in order to see which interval showed the lowest standard deviation or was the most stable. The comparisons of the three intervals le2a, cc2a, re2a did not result in the cc2a interval being the most stable; it was rather the le2a interval with left-edge being a landmark on the initial consonant of the second word that appeared as the most stable interval across sequences. This result suggests that, across words, the achievement of target of the initial consonant of a consonantal sequence seems to be affiliated with the initial consonant of the preceding word. It is thus no longer the c-center of the consonant cluster at the start of the second syllable (or word) that enters into a stable temporal organization when the anchor consonant is chosen not to be in the same syllable as that cluster. This result then offers converging evidence that the stability seen in the previous paragraph, concerning the relation between the prevocalic consonantism in [splats], [plats], [lats] and the anchor, reflects something about the whole sequence of segments being within the same syllabic unit.

We have so far reviewed results on the temporal coordination of onsets supporting the idea that the stability of the cc2a interval captures the global organization characteristic to complex onsets. What about syllable codas? How is a coda consonant coordinated with its preceding vowel, that is, the vowel that is the nucleus of the syllable to which this consonant belongs? For codas, Browman & Goldstein examined sequences such as [pis plats], [pip lats] and so on in contrast to [pi splats], [pi plats]. The initial consonant of the onset in the second word in the latter two sequences (/s/ in [splats], /p/ in [plats]) is the coda consonant of the initial word in the former two sequences (/s/ in [pi**s**], /p/ in [pip]). Recall that in the sequence [**pi s**plats], the most stable relation between the initial /s/ (shown in bold) of the second word and the initial /p/ (shown in bold) of the first word was found to be best described using the le2a interval. The question that arises now is which landmark may best capture the relation between the coda /s/ in /pi**s**/ (shown in bold) and the initial /p/ (shown in bold) of the first word in [**pis** plats]. The findings with different postvocalic consonants show that the interval from the left-edge of the postvocalic consonant /s/ in /pis/ to the anchor consonant, which is the preceding /p/, seems to remain relatively stable. Relatively stable means that the standard deviation of the interval from the c-center of the /s#pl/ cluster (where '#' indicates word boundary) in [pis plats] to the preceding anchor consonant /p/ in /pis/ across utterances showed more variability than the standard deviation of the le2a, indicating that it is only the first postvocalic consonant that is coordinated stably with the preceding vowel of its syllable. Therefore, immediate post-vocalic consonants are organized based on a sequential relation to the preceding vowel; i.e., there is left-edge organization of the /s/ with the preceding vowel regardless of whether /s/ is a coda consonant as in [pi**s** plats] (shown in bold) or the initial consonant of the onset of the following word as in [pi **s**plats] (shown in bold).

Overall, Browman & Goldstein (1988) aimed to identify temporal coordination patterns that may differ between onsets, as in CCV contexts, and codas, as in VC contexts. The results suggested that there are indeed different patterns of coordination for onset and coda consonants. We highlight the key finding of Browman & Goldstein concerning within word sequences of consonants and vowels as in CCCV, e.g., /splats/. The crucial point here is that the c-center of the CCC is a global property of the entire prevocalic cluster. This is so because it is computed by taking the mean of the midpoints of every consonant in the cluster. It is an interval defined using this landmark and extending till the end of the vowel (the anchor) that remains more stable (in comparison to other intervals). This finding suggests that all onset consonants are organized globally with respect to the vowel. The resulting hypothesis or better conjecture (given the limited data presented in that study) is thus one about the presence of a lawful relation between phonological syllabic structure and temporal organization of the segments that make up the syllable.

Observe now that although the results as well as the main hypothesis reviewed above all refer to relations between the consonants and the vowel, the actual gesture of the vowel was not measured in Browman & Goldstein (1988). The landmark used as a proxy to the vowel was, it will be recalled, actually a landmark located on the postvocalic consonant, e.g., the acoustic onset of /ts/ in /plats/ was taken as demarcating the end of the vowel gesture. That is, unlike what was done for the consonantal gestures where the landmarks measured were legitimately located on the constriction intervals of these gestures, for the vowel no such landmark was used. This was so due to difficulties in separating individual gestures of different segments, which use the same articulator; the relevant excerpt from Browman & Goldstein's (1988:142–143) on this point was given verbatim in the beginning of this section. Let us expand on this now as it introduces a point of difference between our work and all subsequent work on the topic. Vowels are mainly produced using (relatively) slow movements of the tongue. In CV or VC combinations where the consonant is using part of the tongue as well, as in the case of /s/ or /l/ where the tongue tip and blade raise to form a constriction somewhere around the alveolar ridge or /k/ where the tongue back raises to form a constriction around the velum, then the movements for the consonantal gesture and those for the vowel would be coarticulated (overlapped in time). This coarticulation makes it not straightforward to track the movements corresponding to the individual segments in the kinematic record. It was for this reason that Browman & Goldstein chose the post-vocalic consonant as an anchor point and moreover made explicit that this practice is valid under the assumption that "the (temporal) interval from the c-center to the final 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 and b) the final consonant anchor point, which is the effective achievement of the target for the final consonant, corresponds to the end of activation of the vowel, that is, articulatory vowel offset" (Browman & Goldstein 1988:150). The part of the assumption about when the vowel's activation starts (namely, at the c-center of the prevocalic cluster) deserves highlighting. This assumption has never been evaluated in any work we know of, despite the wide adoption of this quantification practice in virtually all work that builds on Browman & Goldstein's (1988). Our dissertation takes up the task of explicitly quantifying vowel-specific tongue movements (along with other measures) in an aim to provide a more direct quantitative expression of the relation between the vowel and its preceding consonants in CCV.

1.4 Subsequent studies and preview of what is to come in this dissertation

Following Browman & Goldstein's (1988) landmark study, a number of subsequent studies investigated further the relation between syllabic organization and phonetic indices in English and other languages. Here, we turn to review these studies which stand in the background of the work we undertake in this dissertation. In this review, more specifically, we highlight issues met in pursuing the original ideas of Browman & Goldstein (1988) with more extensive datasets and underscore apparently diverging results which anticipate the key developments to follow in this dissertation.

We organize the studies to be reviewed in terms of two classes: studies that contribute (mostly) theoretical developments to the original ideas of Browman & Goldstein (1988) and studies that contribute (mostly) empirical extensions of these ideas to other languages.

Let us begin by reviewing studies on further theoretical developments. The results of Browman & Goldstein (1988) can be expressed in terms of temporal shifts in the consonants preceding the vowel and these temporal shifts have been given formal expressions in some theoretical frameworks. On the one hand, the initial consonant in the CCV context has been observed in the original Browman & Goldstein (1988) results to shift leftwards away from the anchor landmark compared to its position in the CV context. On the other hand, the prevocalic consonant in CCV has been observed to shift rightwards towards the anchor landmark compared to its position in the CV context. This leftward and rightward shifts are what give rise to the so called "c-center effect", which means that it is the c-center landmark, the temporal midpoint of the consonant(s), that maintains a stable relation with the vowel. This is not the case for coda consonants where only the postvocalic consonant is timed with the vowel and thus the coda consonants occur in a sequential order after the vowel.

These coordination patterns of onsets and codas have been captured using the notion of phase (Browman & Goldstein 2000, Goldstein *et al.* 2006) in the coupled oscillators model of Nam & Saltzman (2003) and Goldstein *et al.* (2009). Based on the hypothesis that the gestures in a CV sequence start synchronously while VC gestures are sequentially organized (Löfqvist & Gracco 1999), the coupled oscillators model expresses the two different patterns by referring to two universal modes of coordination "in-phase" and "anti-phase". In-phase captures the synchronous coordination in CV, while anti-phase captures the sequential coordination in VC sequences (Browman & Goldstein 2000, Goldstein *et al.* 2006). Let us see now how the coupled oscillators model offers a formal expression of the different coordination patterns in the syllable, the global organization as described by the c-center organization in onsets and the local organization in codas. In the model, each gesture is associated with a planning oscillator and gestures are coupled pairwise to one another. The oscillators start moving at random initial phases until they settle into a stable pattern. Which stable patterns emerge turns out to be related to the way in which the oscillators are related to one another. Specifically, if in a CCV sequence both consonantal gestures are coordinated in-phase with the vowel, this means that each consonant is required to start synchronously with the vowel. However, this would result in the consonants overlapping with one another and thus in one of the consonants not being perceived (e.g., if /pl/ was the cluster, /l/ would not be perceived as it would be completely masked by the bilabial closure in front of the alveolar constriction for the /l/). To avoid this outcome, the consonants must somehow keep apart from one another. Therefore, in addition to the in-phase coordination of each consonantal gesture with the vowel, Browman & Goldstein (2000) posit that there is an anti-phase coordination between the two consonants in CCV so as to ensure perceptual recoverability of the prevocalic consonants. The two in-phase coordination relations of each consonant with the vowel are now in competition with the anti-phase coordination among the two consonants. The competition between these three coupling forces or bonds as Browman & Goldstein (2000) refer to them, two in-phase and one anti-phase, in the CCV onset results in the leftward and rightward shift of the initial and prevocalic consonant respectively compared to their positions as single consonants in a CV context. It is thus this competition between the coupling forces that accounts for the c-center effect described above in the coupled oscillators model analysis of these findings. Now, for coda consonants, in a VCC context, only the immediately post-vocalic consonant enters in a phasing relation with the vowel. The final C does not have a phasing relation to the vowel. Therefore, there is no competition in this context. This accounts for the sequential organization in codas. Further asymmetries between coordination modes have also been claimed to fall out from the model. Thus, quantifications of the relative stabilities of in-phase and anti-phase modes in this model indicate that the in-phase one is more stable than the anti-phase mode. This has been hypothesized to account for asymmetries in the cross-linguistic patterning of syllable onsets versus codas, namely, the typological fact that syllables with onsets (CV) are less marked than syllables with codas (Blevins 1995) and also for observations related to the development of syllabic forms in language acquisition where syllables with onsets tend to emerge before syllables with codas.

A subsequent theoretical and simulation study by Gafos (2002) brought these ideas in contact with constraint-based models of phonological theory, synthesizing dynamical representations (gestures) with constraint-based grammars. Specifically, Gafos (2002) has shown using an Optimality Theory account how the observed temporal relations seen in the empirical results can be derived from the interaction of violable and competing CV, CC, and VC coordination constraints, where these constraints are given formal expressions in terms of alignment constraints which synchronize some landmark in the lifetime of one gesture to another landmark in the lifetime of another gesture.

We now turn to review studies which have contributed to expanding the empirical range of consonant-vowel combinations and languages as part of the larger theme of the relation between syllabic structure and temporal organization. These studies used the previously explained (section 1.3) stability-based heuristics le2a, cc2a, re2a in order to quantify the phonetic expression of syllables and to interpret the findings. For complex onsets in English, two subsequent studies by Honorof & Browman (1995) and Byrd (1995) provided further empirical support for the results in the Browman & Goldstein (1988) study, including the result that as the number of consonants increases from CV to CCV to CCCV, it is the cc2a interval that remains the most stable in comparison to the le2a and re2a intervals. Honorof & Browman (1995) analyzed kinematic data of four American English speakers producing the onset clusters /sp/, /pl/, and /spl/ and the coda clusters /ps/, /sps/, /lp/, and /lps/. A main result was that the cc2a interval was the most stable one in onsets but not in codas. Byrd (1995) in her study with five American English speakers also examined the temporal coordination between consonants and the subsequent vowel in C#CV and CC#CV where '#' indicates word boundary. In this case, only the immediately prevocalic consonant constitutes the onset of the CV syllable, while the preceding consonants constitute the coda consonants of the preceding word, e.g., as in [bakt#kab]. In this organization, where only the prevocalic consonant is expected to be timed with the vowel, Byrd (1995) found that the re2a interval, as opposed to the le2a and cc2a, remained the most stable interval as the number of consonants (which were not part of the same syllable as the following CV) increased from C#CV to CC#CV. Both studies, Honorof & Browman (1995) and Byrd (1995), thus indicated that syllabic organization can be fruitfully expressed by patterns among interval stabilities (or what we refer to in this dissertation as stability-based heuristics). These studies thus have served as the go-to sources in subsequent works which continued to pursue actively, in both English and other languages, the hypothesis that complex onset organization is reflected in cc2a stability, while simplex onset organization (where consonant clusters do not combine to form an onset) is reflected in re2a stability.

Goldstein *et al.* (2009) investigated word-initial /pl, sp/ clusters in English. Temporal predictions concerning the leftward shift of the initial consonant away from the vowel and the rightward shift of the prevocalic consonant towards the vowel in a CCV compared to their timing with the vowel as single consonants in CV were pursued. As we have discussed, in a CCV configuration, it is expected that for complex onsets the c-center landmark of the cluster establishes a stable relation with the vowel. The rightward shift was determined by calculating the interval between the achievement of targets of the prevocalic consonant and the subsequent vowel in CCV and in CV and then taking the difference between these two intervals across the two contexts. The leftward shift was determined by calculating the interval between the achievement of targets of the initial and the subsequent vowel in CCV and in CV and then taking the difference between these two intervals across the two contexts. The results on /pl, sp/ showed that the degree of rightward and leftward shift was different depending on the cluster. Specifically, regarding the rightward shift, /p/ in /sp/ was found to shift rightwards as expected but /l/ in /pl/ showed minimal shift towards the vowel. The former result conforms with the temporal prediction of complex onsets, but the latter result does not. Goldstein *et al.* (2009) proposed that the ability of the prevocalic consonant to overlap with the vowel was a crucial parameter that affected the degree of leftward shift towards the vowel. Whereas /p/ as a bilabial can overlap with the vowel because the two involve independent articulators (lips for the /p/ versus tongue for the vowel), the lateral can be said to resist overlap with the following vowel due to the American English /l/ consisting of two gestures (Browman & Goldstein 1992): a tongue tip constriction and a tongue back lowering, with the latter being the same articulator as for vowels. Because /l/ involves a tongue back gesture, completion of that tongue gesture is required before the gesture of the vowel can start. Using the coupling oscillator model described in the beginning of section 1.3, Goldstein *et al.* (2009) proposed to account for the low overlap between the lateral and the vowel in /pl/ by specifying that both gestures associated with the lateral, the tongue tip and the tongue back gesture, are tightly coupled with the vowel. Such a tight link of the lateral with the vowel results in only a small degree of shift of the lateral towards the vowel (low overlap) compared to the /sp/ cluster where no such a double link between the /p/ and the vowel exists. Overall, the results of Goldstein *et al.* (2009) were a first indication that when looking at different clusters observed effects may differ from the expected predictions for complex onsets.

In a different study on Georgian, Goldstein *et al.* (2007) investigated the sequences

/r∼kr∼tskr/ using articulatory data of two native speakers. Georgian is considered to admit complex onsets. Only the prediction of rightward shift was assessed in Goldstein *et al.* (2007) by calculating the interval from the target of the prevocalic consonant to the target of the vowel across /r∼kr∼tskr/. The results were inconclusive, because one of the speakers showed the expected rightward shift of the prevocalic consonant as the number of consonants increased from CV to CCV to CCCV, but the other speaker showed no rightward shift of the rhotic across these sequences. In a subsequent study of Georgian, Goldstein *et al.* (2009) examined front-to-back /t∼tk/ and back-to-front clusters /b∼kb/ produced by two native speakers of Georgian. Intra-cluster timing in Georgian clusters has been found to differ as a function of the order of place of articulation with front-to-back clusters exhibiting a shorter lag than back-to-front clusters (Chitoran *et al.* 2002). Goldstein *et al.* (2009) on Georgian examined how the order of place of articulation in the clusters affects the expected shifts between the consonants and the vowels as the sequence is modified from CV to CCV. The results indicated the expected (for complex onsets) rightward shift of the prevocalic consonant to the vowel in front-to-back clusters. However, for back-to-front clusters, there was only a small degree of shift of the prevocalic consonant and not to the same direction across the two speakers. The study of Goldstein *et al.* (2009) continued to offer indications that when looking at various clusters with different phonetic properties (here, the place of articulation order of the consonants in the cluster), observed effects may differ from the expected predictions for complex onsets. However, the Georgian results must be qualified because data from only two speakers were examined.

An empirically more solid study by Marin & Pouplier (2010) on American English surpassed previous studies substantially in terms of the number of participants (seven native speakers of American English) and range of analyses. Complex onsets /sp, sk, sm, pl, kl/ and complex codas /sk, sp, lk, lp, ms, ks, ps/ were used to assess temporal characteristics of (the expected) global organization in onsets and (the expected) local or sequential organization in codas. A stability analysis was performed using the intervals le2a, cc2a, re2a introduced earlier. Maximum constriction of the postvocalic consonant was used as an anchor for the case of onsets and the maximum constriction of the prevocalic consonant was used as an anchor in the case of codas. It was found, as predicted, that the cc2a interval was more stable than the le2a and re2a for complex onsets, while for codas the le2a interval was the most stable interval for /sC/ clusters (as expected) but for lateral-stop clusters, the cc2a interval was the most stable among the rest (not expected). Furthermore, a shift analysis was conducted as in Goldstein *et al.* (2009) in order to assess the leftward and rightward shift of the consonants in the cluster compared to their presence in initial position as single consonants both in onset and in codas. The lag between the maximum constriction landmark of the relevant consonant and the postvocalic consonant (for onsets) or prevocalic consonant (for codas) was calculated in the cluster and in the singleton contexts and their difference indicated the degree of shift. For onset clusters, the expected rightward shift of the prevocalic consonant to the vowel was observed across /sC/ and stop-laterals clusters. However, the degree of rightward shift was cluster-specific. Within /sC/ clusters, there was a greater rightward shift for /sp/ than for /sk/ and a greater rightward shift for /sk/ than for /sm/. For stop-laterals, there was greater rightward shift for /kl/ than for /pl/. For codas, the /sC/ clusters showed the expected pattern of no shift of the postvocalic consonant towards the vowel, while the lateral-stop clusters exhibited a shift of the lateral towards the vowel (not expected). Finally, the acoustic vowel duration was measured in simple CV and VC sequences and in the respective cluster sequences CCV and VCC. The vowel in CCV is expected to be shorter than in CV which results from the rightward shift of the prevocalic consonant towards the vowel. Such a shift of the postvocalic consonant in VCC towards the vowel compared to VC is not expected in codas and thus the acoustic vowel duration is not expected to change between VC and VCC. The results of vowel duration were as expected for onsets but not as expected for all coda clusters, as the vowel in /l/-initial coda clusters was shorter than in singleton /l/ coda. Overall, the study of Marin & Pouplier (2010) successfully used the stability-based heuristics to diagnose syllabic organization for complex onsets. The temporal predictions of the complex onset organization were largely attested in the English complex onsets this study examined, although there was some evidence for cluster-specific effects related to the degree of rightward shift of the prevocalic consonant to the vowel. The results on coda clusters were not always as predicted and provided evidence that cluster-specific properties affect temporal coordination patterns of coda segments.

German, one of the languages we take up in this dissertation, was investigated by Pouplier (2012) for the first time with respect to the issues that are of concern in the general area of the relation between syllabic structure and temporal organization. German is a language which like English admits complex onsets. In Pouplier (2012), a similar set of analyses was used as in Marin & Pouplier (2010) on English to assess the temporal coordination patterns of clusters in onset and coda position. We only report results on onsets, because these are the relevant results for the purposes of the current dissertation. The clusters consisted of the onsets /bl/, /pl/, /gm/, /km/, and /sk/. A stability analysis was performed using the stability-based heuristics with anchor being the acoustic midpoint of the tautosyllabic vowel. The results showed that no clear cc2a stability was found. The stability of the cc2a interval was not sufficiently different from that of the re2a interval, as would be the expected prediction of the hypothesis that complex onsets show cc2a interval stability. This result then from German questions the validity of the stability-based heuristics as infallible diagnostics to syllabic organization. A shift analysis in Pouplier (2012) also provided puzzling results but in line with the results of the stability-based heuristics in the same study. Different degrees of shift were observed based on the cluster identity. The /sk/ onset cluster showed the expected rightward shift of the prevocalic consonant to the vowel, equal in magnitude to the leftward shift of the initial consonant. However, for the onset clusters /bl, gm, km/ a very small rightward shift was observed; for /pl/ there was no rightward shift at all and in fact a small shift to the opposite (from that predicted) direction was observed. Pouplier (2012) examined also intra-cluster timing, as was done for Georgian in other studies (Chitoran *et al.* 2002, Goldstein *et al.* 2009), which was quantified by the lag between the consonantal plateaus in the cluster. A relation between intra-cluster timing and cc2a stability was found such that clusters with shorter lags between consonantal plateaus or more overlap among their two consonants were more likely to exhibit cc2a stability than clusters with larger lags or less overlap. This observation, to which we will return as we develop our proposal as well as when we present our experimental results across German and Spanish, appears to resonate well with the idea that in complex onsets the vowel and its preceding consonants are more overlapped with one another (in a way that expresses their inclusion in the same syllable). If this idea is on the right track and if interval stability patterns are, at least to some extent, indicative of syllabic organization, then the observation that more overlap in the cluster goes along with better cc2a stability would seem to be an expected result. However, it still remains to be understood how other clusters which make perfectly legitimate onsets are not highly overlapped and showing cc2a stability (e.g., from our work Spanish /pl, kl/ clusters are a case in point). Putting this aside for now, overall the study of Pouplier (2012) probed further the idea of the cluster-specificity of effects that have been observed in previous studies on English and Georgian and showed that, for German, stability-based heuristics fail to diagnose syllabic structure consistently since the expected cc2a stability was not found even though for the clusters investigated (e.g., /pl, bl/) there is solid phonological evidence for their onset status (Wiese 1996).

A follow up study on German by Brunner *et al.* (2014) investigated the clusters /pl, gl, kv, sk/ word-initially, word-medially, and across a word boundary under two accentuation conditions and with different vowels following the clusters. The accentuation conditions differed with respect to whether the target word bore sentence accent or not by using the carrier phrases "Ich sah [target word] an" and "Ich sah mit Tom, nicht mit Anna [target word] an" respectively. The different accentuation was an attempt to assess whether an accented or de-accented target word modulates the temporal organization of the segments in the onset of that word (which it did not, for the measures used in that study). For the vowel manipulation, the target words consisted of consonant clusters followed by vowels with different length or tenseness properties, that is, the vowel following a C or a CC word-initial onset was lax in one word pair (/lɛtsə∼plɛtsə/) and tense in another word pair (/lɑːgən∼plɑːgən/). A stability analysis was performed using the stability-based heuristics le2a, cc2a and re2a with the anchor being the target achievement of the postvocalic consonant. The expected cc2a stability was only partially found for some clusters word-initially and word-medially (/gl, sk/). For other clusters such as /pl, kv/ the results indicated re2a stability, contrary to the expectations of complex onsets. The cc2a stability was more likely to occur if the cluster was in the tense vowel context than in the lax vowel context. Furthermore, recall that for clusters across words, the prediction is that only the prevocalic consonant is temporally organized with the subsequent vowel as expressed by the re2a stability. Brunner *et al.* (2014) also considered interval stabilities for these clusters and their finding was that not all cross-word clusters exhibited the expected re2a stability. Brunner *et al.* (2014) also considered vowel duration and its relation to the interval stability results. Vowel duration was measured acoustically and also articulatorily using the re2a interval which is considered as an articulatory approximation of the extent of the vowel gesture. In some of the clusters (/kv, sk/), the vowel was shorter in the cluster than in the corresponding CV context (that is, /v/-, /k/-initial sequences corresponding to /kv, sk/ clusters) in both word-initial and word-medial positions as expected, but in other clusters (e.g., /pl, gl/) there was no difference in vowel duration between CV and CCV. Recall that one of the predictions in terms of temporal shifts of the two consonants in a CCV would be that the second C moves towards the vowel in CCV compared to CV resulting in shorter vowel duration in CCV than in CV. This is not what was found across all clusters investigated. Turning to cross-word clusters, vowel shortening was observed for some clusters which was not as expected. Intra-cluster timing was also examined with similar conclusions as in Pouplier (2012): clusters with shorter lags between consonantal plateaus were more likely to exhibit cc2a stability than clusters with larger lags. Thus, again as in Pouplier (2012), the results indicated that stability-based heuristics fail to consistently diagnose syllabic organization. There is, instead, further evidence that cluster-specific properties affect the stability of intervals as well as other measures such as vowel duration.

Hermes *et al.* (2013) investigated the timing of word-initial voiceless stop-sonorant and /sC/ clusters in Italian. Stop-sonorant clusters are legitimate onset clusters in Italian but in word-initial /sC/ clusters, the /s/ is considered extra-syllabic in certain morphosyntactic contexts (Bertinetto 2004). Unlike in other studies, a stability analysis using the stability-based heuristics was not carried out. Instead, a shift analysis was performed to assess the leftward and rightward shift of the initial and prevocalic consonants respectively compared to their positions in the CV context. The target achievement of the vowel following the consonant(s) of interest was used as an anchor. Then the stability of the interval indicating rightward shift was evaluated across word pairs (e.g., /lina/∼/plina/, /rema/∼/krema/) using the relative standard deviation. For the voiceless stop-sonorant clusters, both a leftward shift of the initial consonant and a rightward shift of the prevocalic consonant were found. This result is in line with the predictions of the complex onset organization. As we have seen, when one considers the relative timing of the consonants in /li/, /pi/ and /pli/, then in /pli/ the /l/ is shifted to the right (in comparison to its timing with the vowel in /li/) and the /p/ is shifted to the left (in comparison to its timing with the vowel in /pi/). For /sC/ clusters, there was a substantial leftward shift of the initial /s/ in the cluster compared to its position in /sV/, while there was only a minimal or no rightward shift of the prevocalic consonant towards the vowel compared its position as a single C in CV. This result is in line with the simplex onset organization where there is a local organization of the prevocalic consonant with the subsequent vowel to the exclusion of the initial consonant in the cluster. In the Italian case, then, the evidence presented from phonetic measures seems to be in good agreement with the phonological analysis. The /s/ in /sC/ is considered extrasyllabic. The results show that there is a substantial leftward shift of the /s/ and that there is minimal or no rightward shift of the prevocalic C. In contrast, the /p/ in /pl/ is part of the onset and shows some leftward shift symmetric to the rightward shift of /l/.

Marin (2013), on Romanian, a Romance language that admits complex onsets, carried out a stability analysis where the intervals le2a and re2a were defined slightly different than in previous studies. The left-edge landmark was the constriction midpoint of the initial consonant (as opposed to the target achievement in previous studies) and the right-edge landmark was the constriction midpoint of the prevocalic consonant (as opposed to the release landmark of the same consonant in previous studies). The cc2a interval was defined as in previous studies. The constriction midpoint of the postvocalic consonant was used as an anchor. The results showed the expected cc2a stability and rightward shift of the prevocalic consonant in sibilant-initial onset clusters /sk, sp, sm/ but also patterns of unexpected re2a stability and no rightward shift of the prevocalic consonant for the stop-initial onset clusters /ps, ks, kt, kn/. Native speakers of Romanian classify both the latter and the former clusters as complex onsets. Therefore, once again, assuming all these clusters are complex onsets, the result is that stability-based heuristics fail to reliably diagnose syllabic organization.

Hermes *et al.* (2017) investigated stop-liquid word-initial complex onsets /pr, kr, kl, pl/ produced by three native speakers of Polish. A stability analysis was carried out using the intervals le2a, cc2a and re2a. The maximum constriction of the postvocalic consonant was used as an anchor. The results showed the expected cc2a stability only for the /pl/ cluster, while the rest of the clusters /pr, kr, kl/ showed le2a stability. Furthermore, the degree of overlap in the clusters was investigated by calculating the lag between the maximum constrictions of the two consonants in the clusters /pl, kl, pr, kr/. Large lags were overall observed across clusters and there was considerable variability based on the cluster identity. In addition to the actual experimental data, Hermes *et al.* (2017) used simulated data on Polish where variability was introduced in the anchor landmark (as previously done in Shaw *et al.* 2011 to be reviewed later in this section). The computational model they used provides an evaluation of how well the empirical data fits the simulated complex onset organization when the variability of certain parameter increases. The results showed that complex onset organization allowed for variability in the data to some extent but after some threshold the pattern for complex onset organization ceased to be stable. The results regarding variability in Polish clusters are in line with the findings from Pastätter & Pouplier (2015) who examined CV overlap in the clusters /mʃ, pʃ, ps, ks, ʃm, ʃp, sp, sk, pl, kl, ml, vl, pn, kn/ produced by six native speakers of Polish. CV overlap was calculated as the lag between the peak velocities of the prevocalic consonant and the postvocalic consonant in CVC and CCVC. The results indicated that the degree of CV overlap varied based on the coarticulatory resistance of the immediately prevocalic consonant. Specifically, clusters with sibilant prevocalic consonants, which are highly coarticulatory resistant, showed no change in CV overlap compared to the non clustered CV context. Clusters with /m, p, k/ as prevocalic consonants, which are of low coarticulatory resistance, exhibited large degree of CV overlap compared to the CV context. Finally, clusters with laterals or nasals as prevocalic consonants showed an intermediate degree of CV overlap compared to the former two groups of clusters. Therefore, coarticulatory resistance of the prevocalic consonant may be a cluster-specific factor which affects the timing between the consonant and the vowel. Such a factor may have contributed to the fact that stability-based heuristics failed to diagnose syllabic organization in Polish (Hermes *et al.* 2017). In Polish, the expected stability of the cc2a interval was found only for the /pl/ cluster in Hermes *et al.* (2017), but for the clusters /kl, kr, pr/ le2a stability was found. Goldstein *et al.* (2006, 2009) and Shaw *et al.* (2011) have stated that stability-based heuristics and specifically the c-center-to-anchor stability is influenced by segmental properties. This is because the pattern related to c-center-to-anchor stability is based on the assumption that the initial consonant in the cluster is shifted leftwards compared to its timing with the vowel in a simple CV sequence and the prevocalic consonant is shifted rightwards compared to its timing with the vowel in a simple CV sequence. These leftward and rightward shifts of the consonants in the cluster are supposed to be symmetrical in a c-center organization pattern. However, such a symmetrical leftward and rightward shift, is not always met since it is affected by segmental properties such as coarticulatory resistance (Pastätter & Pouplier and Hermes *et al.* 2017 on Polish) and the gestural composition of the segments involved in the clusters (Goldstein *et al.* 2009 on English /pl/ versus /sp/).

We have so far reviewed studies from a number of languages that admit complex onsets. As we have seen, the expected stability of the cc2a interval has not consistently been found for complex onsets in these languages. Similarly, the expected re2a stability has not been consistently found in cross-word clusters. Other measures related to predictions of the complex onset organization such as vowel compression in clusters, leftward and rightward shift of the initial and prevocalic consonants in the clusters also seemed to provide inconclusive evidence about the syllabic organization. Moreover, these studies present strong evidence for cluster-specific properties affecting temporal coordination patterns in the variety of clusters so far investigated. Overall, the evidence presented, if taken at face value, appears to be at best inconclusive and at times, as expressed in some of these works, outright contradictory to the idea that complex onsets can be identified by some phonetic property in the kinematic record which is shared across clusters and languages. As stated in Brunner *et al.* (2014:449): "The results of the present study suggest that the C-center hypothesis must be revised, if it is regarded to be valid for all natural languages. It is possible that syllabic parsing does not percolate through to the observables of speech production, the motion of electromagnetic coils in the present case."

The studies reviewed above were carried out on languages which all admit consonant clusters as syllable onsets (with the exception of the /sC/ clusters in Italian, which as discussed behaved differently from other clusters in that language which are legitimate onsets; see Hermes *et al.* 2013). We now turn to review studies which have focused on languages which do not admit consonant clusters as onsets. Berber (Goldstein *et al.* 2007, Hermes *et al.* 2011, Hermes *et al.* 2017) and Moroccan Arabic (Shaw *et al.* 2009, Shaw *et al.* 2011, Shaw *et al.* 2015) are languages which allow for maximally one consonant to be part of any syllable onset. Thus, for example, the words /bka/ 'he cried' and /glih/ 'to grill' in Moroccan Arabic are said to be parsed as /b.ka/ and /g.lih/ where only the prevocalic consonant forms the onset of the CV syllable (Dell & Elmedlaoui 2002). More generally, any C.CV (where again the '.' stands for a syllable boundary) in these languages instantiates the simplex onset syllabic organization because even though there is a consonant cluster, it is maximally the immediately prevocalic consonant that is allowed to be an onset of the second syllable. As reviewed earlier, this syllabic organization was initially investigated by Byrd (1995) using C#CV clusters across a word boundary in English where the '#' indicates word boundary. The two consonants in a C#CV sequence do not form part of the onset; only the prevocalic consonant is the onset of the CV syllable. Recall that for this context, Byrd (1995) found that the re2a interval remained the most stable with the addition of an extra consonant from CV to C#CV as opposed to the le2a and re2a intervals. Subsequent studies in languages that do not admit clusters as onsets have thus looked for re2a stability or other such indications that the clusters at issue are not complex onsets.

Let us begin in this branch of studies with Goldstein *et al.* (2007) who investigated the sequences /mV∼s.mV∼ts.mV/ (where again the '.' indicates syllabic boundary) produced by one speaker of Tashlhiyt Berber. The rightward shift of the prevocalic /m/ towards the subsequent vowel was quantified across these sequences. Specifically, the lag from the target achievement of /m/ to the target achievement of the subsequent vowel was measured across /mV∼s.mV∼ts.mV/. A prediction of the simplex onset organization is that there is no rightward shift of the prevocalic consonant to the vowel with the addition of consonants preceding the prevocalic consonant. This is the opposite of the complex onset organization where rightward shift is predicted as a result of the competition in the onset since in a complex onset both consonants in CCV are expected to be temporally organized with the vowel. Such a competition is not expected in a simplex onset organization where only the prevocalic consonant in a C.CV sequence is temporally organized with the vowel. Indeed, Goldstein *et al.* (2007) found no rightward shift of the prevocalic /m/ to the vowel with the addition of initial consonants in line with the prediction of a simplex onset organization.

The study of Shaw *et al.* (2009) investigated word-initial /db, kt, gl, sb∼ksb, kd∼bkd/ clusters produced by one native speaker of Moroccan Arabic, another language with a hypothesized simplex onset organization. The clusters were paired with single consonants as onsets with the same prevocalic consonant as in /bV, tV, lV/ (corresponding to /db, kt, gl/ in the list above) and so on. A stability analysis was performed using the le2a, cc2a, re2a intervals. Two different anchor landmarks, a consonantal and a vocalic one, were used to assess the robustness of the results. The consonantal anchor was the maximum constriction of the postvocalic consonant. The vocalic anchor was the gestural offset of the vowel gesture. Interval stabilities across each word pair or word triplet, e.g., /tab/∼/ktab/ or /bula/∼/sbulha/∼/ksbulha/, were computed. Using the consonantal anchor, the results indicated that the re2a interval had the lowest relative standard deviation across word sets; thus, re2a stability was observed as expected. Using a vocalic anchor, the results showed different patterns, sometimes compatible with the complex onset organization. A first clue for what may be causing the divergence in the results with the second anchor was that the relative standard deviation of the intervals le2a, cc2a, re2a was substantially greater with the vocalic anchor than with the consonantal anchor. This means that there was increased variability in the intervals when using a vocalic anchor.

This increased variability was ascribed to the fact that the gestural offset for a vowel, which was the timestamp of the vocalic anchor, tends to be more variable to demarcate with automatic algorithms than the timestamp of the consonantal anchor, which was on the postvocalic consonant. For some word sets, e.g., /bulha/∼/sbulha/∼/ksbulha/, /bula/∼/sbula/, /bal/∼/dbal/ and /tab/∼/ktab/, the cc2a interval emerged as the most stable, against the predictions of a simplex onset organization. In order to make sense of these conflicting results, Shaw *et al.* (2009) used a computational model to study how patterns of temporal stability respond to increased variability in interval durations. In the model simulations, variability was introduced in all intervals uniformly using random noise. The key findings indicated that under certain conditions of variability, the simplex onset organization, that is, the parse /C.CV/, can be expressed in cc2a stability as opposed to re2a stability. The model thus revealed how stability-based heuristics are modulated under different conditions of variability in a way that potentially allows one to make sense of apparently contradictory results. Thus, the reason why with the different quantification of intervals using the vocalic anchor showed cc2a stability is because using that anchor introduced more variability in the quantified intervals and more variability leads to cc2a stability. The phonological parse is the same all along, /C.CV/, with a simplex onset as part of the syllable headed by the vowel. But the interval stabilities change depending on how one quantifies these intervals. Overall, then, Shaw *et al.* (2009) was the first study to demonstrate by a model that stability heuristics are not infallible indicators of syllabic organization. It is an explicit prediction of the model developed in Shaw *et al.* (2009) that simplex onset organization can be manifested, under certain conditions, via cc2a stability which was previously thought to be a hallmark of complex, not simplex, onset organization.

A study by Hermes *et al.* (2011) on Berber obtained similar results to the study on the same language we reviewed earlier (Goldstein *et al.* 2007). Consonant clusters with a vocalic and a consonantal nucleus were registered from three native speakers of Tashlhiyt Berber. The corpus consisted of sequences with a vocalic nucleus, e.g., /fik/, /k.fik/, /tk.fik/ or a consonantal nucleus, e.g., /fnk/, /k.fnk/, /tk.fnk/ with /n/ (and /k/ in the case of the third form) being the nucleus across these sequences. Once again, the diagnosis of syllabic organization was pursued by quantifying intervals and their stabilities. The interval from the target of the rightmost consonant to the target of the coda consonant and the interval from the c-center of the consonant(s) preceding the syllable nucleus to the target of the coda consonant were calculated. The stability of each of these intervals was then assessed by calculating the relative standard deviation of these intervals across the sequences /fik/, /k.fik/, /tk.fik/ and /fnk/, /k.fnk/, /tk.fnk/. The results showed that the interval spanning from the rightmost consonant to the coda consonant or in other words the re2a interval did not change with the addition of consonants preceding the nucleus, meaning that there was no rightward shift of the rightmost consonant. Furthermore, the interval from the rightmost consonant to the anchor had the lowest relative standard deviation compared to the cc2a interval which means that the former was the most stable interval. These results are thus in line with the predictions of the simplex onset organization where only the rightmost consonant is temporally organized with the nucleus of the syllable (be it a vowel or a consonant, as is the case in Berber where consonants can also be syllabic nuclei). The hypothesis that it is only the rightmost (that is, immediately prenucleic) consonant that is directly related to the nucleus predicts that this local relation is not perturbed by the addition of consonants preceding that consonant (because these added consonants do not establish a relation to the same nucleus). This prediction is indeed borne out in the Berber data. A follow up study of Hermes *et al.* (2017) using an extended version of the corpus in Hermes *et al.* (2011) investigated the sequences /f∼kf/, /t∼kt/, /k∼lk/ produced by three native speakers of Tashlhiyt Berber. Intra-cluster timing was calculated as the lag between the targets of the two consonants in the clusters. Furthermore, the intervals le2a, cc2a and re2a were used to assess syllabic organization. The maximum constriction of the postvocalic consonant was used as an anchor. The relative standard deviation was used to assess the stability of the above intervals with the addition of a consonant from CV to C.CV. The re2a interval showed the lowest relative standard deviation compared to the rest of the intervals indicating re2a stability across all the sequences /f∼kf/, /t∼kt/, /k∼lk/ in Berber. As an addition to the empirical data, Hermes *et al.* (2017) used simulated data on Berber where variability was introduced in the anchor landmark (as previously done in Shaw *et al.* 2011 to be reviewed in the next paragraph). The computational model they used provides an evaluation of how well the empirical data fits the simulated simplex onset organization when the variability of certain parameter increases. The results showed than even with increased variability in the anchor, the pattern for simplex onset organization remains stable.

Another study of Shaw *et al.* (2011) investigated the word sets /lan∼flan∼kflan/, /bulha∼sbulha∼ksbulha/, /kulha∼skulha∼mskulha/ produced by four native speakers of Moroccan Arabic. The stability-based heuristics le2a, cc2a, re2a were used to perform a stability analysis. The anchor landmark was the maximum constriction of the postvocalic consonant. Overall, across word sets and across speakers, the expected re2a stability was found. However, for some speakers, cc2a stability was found in limited subsets of the data (e.g., the word pair /kulha∼skulha/). It was observed that for this pair of words, the duration of certain segments across the two words of the pair changed substantially from /kulha/ to /skulha/. For example, the duration of the prevocalic consonant /k/

was substantially shorter in the latter than the former word. In an extension of the approach taken in Shaw *et al.* (2009), a computational model was used to understand the exceptional patterns. In model simulations, instead of using an agnostic source of variability as was done in Shaw *et al.* (2009) by using anchor variance to induce more or less variability in the simulated intervals, in this study Shaw *et al.* (2011) investigated how stability-based heuristics respond to variation of phonetic parameters such as duration of the prevocalic consonant and syllable duration in a given syllabic organization. Specifically, variation or scaling of the phonetic parameters in the model was now tailored to the sources of variation observed in the experimental data, such as shortening the duration of the prevocalic consonant from CV to CCV and syllabic compression of the VC portion across CVC∼CCVC. The simulation results showed that as phonetic parameters are scaled, the two contrasting syllabic organizations (simplex and complex) make different predictions about the way stability-based heuristics respond to such perturbations. It was specifically found that the unexpected cc2a stability seen in the experimental data is in fact a prediction of the simplex onset syllabic organization under the sources of variation modeled. The underlying syllabic organization remains the same (simplex onset, as appropriate for Moroccan Arabic) but its stability-based pattern changes as phonetic parameters are varied. Once again, then, we see that unqualified stability heuristics along the lines of simplex onset organization always surfaces with re2a stability and complex onset organization always surfaces with cc2a stability cannot be used as diagnostics of syllabic organization.

The study of Shaw *et al.* (2015) is an extension of the 2009 study, reporting further simulations across Moroccan Arabic with English datasets. Specifically, Shaw *et al.* (2015) evaluated experimental data against data simulated from models of syllable parses representing the two contrasting organizations, complex and a simplex onset organization. Datasets were drawn from the Shaw *et al.* (2009) study on Moroccan Arabic and from the x-ray microbeam archive on American English (Westbury 1994). Sources of variability were introduced in the simulated data by increasing the variance of a common anchor shared by all intervals whose stability properties were studied. Model simulations replicated several key experimental results, including the fact that stability-based heuristics fail to diagnose syllabic structure, and exposed the range of conditions under which such heuristics remain valid.

The studies of Shaw *et al.* (2011) and Shaw *et al.* (2015) are particularly relevant to our concerns since it is this approach that this dissertation embraces. Specifically, the experimental designs for the Spanish and the German studies in this dissertation aim at eliciting different degrees of phonetic parameter values for the same sequence of segments (within the same language). For example, our German study (Chapter 3) successfully elicits different degrees of interplateau interval and duration of the initial consonant in CC clusters by embedding target words with these clusters in sentences of increasing prosodic boundary strength. Such parameter modifications within the same clusters of the same language are crucial prerequisites for the modeling studies above but have not been pursued in any systematic way so far in experimental data. It is such variation of phonetic parameters which enables us, as we will make explicit in the following Chapters, to find evidence for global organization over and above what stability-based heuristics have led previous researchers to infer.

In sum, stability patterns or stability-based heuristics, as we reviewed in this section, often fail to diagnose syllabic structure due to various reasons. The central aim of the current dissertation is to provide a different approach to the relation between syllabic structure and phonetic indices which proposes that syllabic structure is reflected in several phonetic indices and thus not necessarily in a single stability pattern. Our approach, as we will demonstrate, can reliably diagnose syllabic structure even in cases where the commonly used stability-based heuristics fail to do so.

1.5 Organization of the thesis

This thesis is organized as follows. Chapters 2 and 3 present the two main empirical studies on Spanish and German along with the experimental results, their analyses and interpretation of the main observed patterns.

In both of these Chapters, we take up the relation between syllabic structure and phonetic indices in ways that extend previous studies on this topic in at least two ways. First, we examine the data using a number of different analyses, as opposed to relying on one or two analyses only (such as a stability analysis and or a lag analysis, as in the studies reviewed in the previous section). Second, we explicitly take up cases which previous work has either left unresolved or are expected from previous work to be problematic with respect to the quest of relating syllabic structure and phonetic indices. Recall that German has been a language where doubts have been expressed on the viability of finding a lawful relation between syllabic structure and phonetic indices (see Brunner *et al.* 2014:449). There are no studies addressing the main topic of this dissertation with respect to Spanish, a prototypically complex onset language. However, on the basis of what we have reviewed from other languages and we will indeed verify with our data, Spanish is expected to present issues for finding the expected indices of complex onsets because CC clusters in Spanish are produced with large interplateau lags between the two consonants and such lags have been linked in previous work (Pouplier 2012, Brunner *et al.* 2014) to lack of cc2a interval stability (thought to be an index of complex onsets). With this background in mind, we will see that in both our German and Spanish studies, using different analyses based on a variety of measures and eliciting parameter changes by embedding the clusters in different contexts, enables us to uncover crucial evidence for complex onset organization over and above what previous studies based stability-based heuristics have reported. In this section, a brief summary of our two studies, Chapter 2 and Chapter 3, is provided. For each of these, we describe below the material recorded and the gross structure of our experimental designs.

The final Chapter, Chapter 4, after a recapitulation of the main lessons learned in the core experimental Chapters, offers a unifying view of the expression of syllabic organization in the phonetic record, and points to directions for future work aimed at further testing and sharpening of the proposed view.

1.5.1 Chapter 2: Global organization in Spanish onsets

In Chapter 2, the experimental work is based on articulatory data from Spanish and Moroccan Arabic. These two languages were chosen, as reviewed in section 1.2, because they share similar phonetics but have different phonology. For example, consider the segmental sequence /kla/. In saying that the two languages share similar phonetic properties we refer to the fact that a /kl/ in both Spanish and Moroccan Arabic is produced with a substantial interplateau lag between the two consonants and that the voicing implementation system of Spanish is the same as that of Moroccan Arabic in that the voiceless-voiced opposition uses the same main phonetic exponents in the two languages (voicing during closure for the voiced versus relatively short lag VOT for the voiceless). In saying that the two languages have different phonology, we refer to the fact that /kla/ constitutes a single syllable in Spanish but not so in Moroccan Arabic where the /k/ is not parsed as part of the same syllable as /la/. Therefore, the comparison between Spanish and Moroccan Arabic allows us to investigate the effect of syllabic structure on the temporal coordination patterns in the same groups of segments with the same or similar phonetic properties. Articulatory data of six native speakers of Spanish were collected using the Carstens AG501 articulograph. Target words consisted of word-initial voiced and voiceless stop-lateral /bl, pl, gl, kl/ and voiceless stop-rhotic /pɾ, kɾ, tɾ/ clusters. The target words were embedded in the carrier phrase "Di [target word] por favor" meaning "say [target word] please". For Moroccan Arabic, articulatory data were collected using the Carstens AG500 articulograph. For the purposes of the current study, we selected from a large articulatory dataset target words with word-initial stop-lateral clusters /bl, kl/ produced by four native speakers of Moroccan Arabic (this data was registered, but not analyzed, in the Shaw *et al.* 2011 study). The target words were embedded in the carrier phrase /ʒibi [target word] ħnaja/ meaning "bring [target word] here". A considerable part of the Chapter is dedicated to Spanish, because the temporal coordination patterns of Spanish word-initial clusters has not been investigated before using articulatory data. After a detailed analysis of the Spanish data, an analysis of the temporal coordination patterns of word-initial clusters in Moroccan Arabic follows with the results contrasted to those found for Spanish. Finally, the Chapter concludes by drawing implications about the relation between syllabic structure and phonetic indices on the basis of our results from Spanish and Arabic.

1.5.2 Chapter 3: Global organization in German onsets

Chapter 3 presents experimental work based on articulatory data in German. Unlike in Spanish, a language where the relation between syllabic organization and phonetic indices in consonant clusters has not been pursued before with articulatory data, in German there is already a set of studies which have provided relevant results. As reviewed in the previous section, these results largely challenge the idea that syllabic organization has consistent phonetic manifestations in the articulatory record. In our German study, we take a different approach from the previous studies on German by explicitly seeking to manipulate various phonetic parameters in clusters with different syllabic affiliation via prosodic variation. We then examine how syllabic organization responds to such perturbations in the CCV strings. Articulatory data of five native speakers of Standard German were collected using the Carstens AG501 articulograph. Target words consisted of stop-lateral clusters in the word-initial position and across words in order to address syllabic organization. Word-initial stop-lateral clusters consisted of voiced and voiceless stops as in /bl, gl, kl, pl/, while cross-word clusters consisted of only the subset with voiceless stops /kl, pl/ due to the final devoicing rule in German (the first consonant in the cluster had to be a possible word-final consonant). The target words were embedded in carrier phrases with varying prosodic boundary strength, i.e., pauses of varying duration preceded the target word in the case of word-initial clusters, or separated the consonants in the case of cross-word clusters. Using voiced and voiceless word-initial stop-lateral clusters, as we do for Spanish, we examine how modulations in the CCV string, induced by the implementation of the voicing contrast in German which is different from that in Spanish, affects the rest of the spatio-temporal organization of the CCV sequence. Furthermore, we investigate how modulations in the CCV string induced by prosodic variation affects the rest of the spatio-temporal organization of the CCV sequence in word-initial onsets versus cross-word clusters in German.

1.5.3 Chapter 4: Conclusion and outlook

Chapter 4 offers a unified view of the relation between syllabic organization across the range of our data and points to some directions for future work.

A main lesson from our studies is that global organization is manifested via a number of phonetic indices and the extent to which any given index expresses that organization depends on coincidental phonetic properties of the segments that partake in this organization. In other words, there is no unique or general index that best expresses syllabic organization even in this (limited) set of languages and conditions we have examined. A striking example of this non-uniqueness is illustrated by stability-based heuristics which have been heavily employed in previous work (see Hermes *et al.* 2013 on Italian, Shaw *et al.* 2009 on Arabic, Marin & Pouplier 2010 on English, Pouplier 2012 and Brunner *et al.* 2014 on German, Marin 2013 and Marin & Pouplier 2014 on Romanian, Hermes *et al.* 2017 on Berber and Polish) as an index of syllabic organization. When considering the behavior of interval stabilities in our data, one is either forced to the wrong conclusion or to no conclusion. German illustrates the first case. As we have seen, interval stabilities, both in our data as well as in previous work (Brunner *et al.* 2014), point to the conclusion that German does not show evidence for global organization. This is the wrong conclusion, because as we will see in further analyses of our data the evidence for global organization is rather compelling. Let us also consider Spanish. As we will see, interval stabilities in Spanish show a puzzling pattern. Voiced stop-lateral clusters show local timing interval stability, whereas voiceless stop-lateral clusters show global timing interval stability. No clear conclusion can thus be reached, if attention is focused on the one index of interval stabilities. In the approach taken in this work, the nature of the constraints imposed by a global organization seems to be expressed by a number of phonetic indices and in terms of compensatory relations among indices. The following examples offer illustrations of such compensatory relations in our data. For Spanish, when we look at the interval stabilities, we do not find the expected global timing interval stability for voiced stop-laterals, but we do find it for voiceless stop-laterals. However, the compensatory relation observed between interplateau interval (IPI) and duration of the C2 prevocalic lateral is valid for both these groups of clusters. As IPI increases, the duration of the C2 prevocalic lateral decreases in order to compensate for this increased IPI duration and in order to bring the vowel to overlap with the cluster as prescribed by a global organization. This compensatory relation between IPI and C2 lateral duration is observed across voiced and voiceless stop-lateral clusters regardless of the stability patterns they exhibit. For German, with increasing boundary strength, we find reduced local timing interval stability (an indirect evidence for global organization) which was more evident for voiceless than for voiced stop-lateral clusters. However, across both groups of clusters, we find that as IPI increased, the CV lag between target of the lateral and vowel initiation decreased. Across both voiced and voiceless stop-laterals, vowel initiation occurs earlier in the cluster as IPI increased regardless of the stability pattern that these two groups of clusters exhibit.

The second and final overarching theme of the thesis, applying across the range of German and Spanish datasets we examine, concerns the role of variation in the diagnosis of the nature of phonological organization. We argue that variation is crucial to understanding the nature of the link between (phonological) syllabic organization and the phonetic spatio-temporal manifestation of that organization. The consequences of prosodic variation on the segmental properties in a CCV string, such as lengthening of the initial consonant and lengthening of the lag between the two consonants constitute examples of local variation. When local perturbations to segments or relations between adjacent segments have effects that ripple through the rest of the sequence, this is taken as 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.

In future work, we aim to consider languages with a more complex voicing system such as Hindi. Hindi exhibits a four-way voicing contrast with voiced unaspirated, voiced aspirated, voiceless unaspirated and voiceless aspirated stops, as opposed to a two-way voicing contrast of German and Spanish. Moreover, Hindi admits complex onsets such as stop-rhotic and stop-lateral clusters. The four-way voicing contrast in Hindi will allow us to investigate the effects of a more comprehensive set of phonetic perturbations (not available in German or Spanish) on the phonetic indices which express syllabic organization. For example, let us consider the aspirated and unaspirated voiced stops in Hindi. The stop in a stop-rhotic Hindi cluster can be either voiced aspirated or voiced unaspirated, as in /bʱrəʃʈ/ 'corrupt' ∼ /brahmən/ 'priest'. According to Dutta (2007:113), aspirated stops have a shorter closure duration than unaspirated stops. We can thus investigate the effect of varying stop's duration on the phonetic indices expressing syllabic organization within the same voicing condition (voiced stop-rhotic clusters). Furthermore, assuming that the aspiration gesture in Hindi voiced stops has the same effect as the aspiration gesture in voiceless stops in simpler voicing systems such as those of English or German in that it keeps the stop apart from the following consonant (which as we will argue seems reasonable), this would mean that there will be some IPI variation between voiced aspirated and voiced unaspirated stop-rhotic clusters in Hindi. This would imply that not only closure duration but also IPI are varied within the relevant set of clusters in Hindi. The advantage of Hindi is thus that it allows us to examine perturbations such as varying C1 stop's duration and presumably varying degrees of IPI within voiced stop-rhotic clusters. Such a setting was not available in German or Spanish. In the case of German, perturbations related to C1 and IPI duration had to be elicited via prosodic variation. In the case of Spanish, the effect of varying C1 stop's duration was present, but substantial IPI variation between voiced and voiceless stop-lateral clusters was not present. Hindi thus offers both types of variation within one and the same language and presumably without resorting to extra conditions in the experimental design as we have in German in this work (so that we could induce the requisite variation in phonetic parameters). As such, Hindi would offer a highly appropriate new empirical ground for future work aiming to further test and sharpen the predictions of this dissertation.

Chapter 2

Global organization in Spanish onsets

2.1 Introduction

This Chapter addresses the relation between language-specific syllable structure and inter-segmental spatio-temporal coordination in Central Peninsular Spanish (henceforth, Spanish). This study is the first one to investigate temporal coordination patterns of word-initial clusters in Spanish using articulatory data. We examine word-initial voiced stop-lateral, voiceless stop-lateral, and voiceless stop-rhotic clusters all of which are reported to form complex onsets in Spanish (Harris 1983, Hualde 2005, Steriade 1999).

A main aim of the thesis is to compare the spatio-temporal organization of the same sequence of segments in two contrasting syllabic organizations. In the present Chapter, since clusters of different syllabic affiliations are not available in our design for Spanish, Moroccan Arabic will serve as a language with a different syllabic organization in that it does not admit clusters as part of the onset of the syllable. The comparison between similar word-initial clusters in Spanish and Moroccan Arabic allows us to expose the differing effects the two contrasting syllabic organizations have on the same or similar sequences of segments. We thus compare word-initial CCV sequences across two languages, Spanish and Moroccan Arabic, in which the segments involved have similar phonetic properties but different claimed phonological organizations, e.g., /kla/ is a single syllable in Spanish but not so in Arabic. The reason why Spanish and Arabic make a highly appropriate comparison is that in both languages, clusters show an open transition, a period of no constriction between the two clustered consonants. Moreover, the voicing contrast is also implemented in the same way across the two languages, that is, by lead voicing (negative VOT) for the voiced versus a short lag for the voiceless (Zeroual 2000 on Moroccan Arabic, Lisker & Abramson 1964 on Spanish). We did not use cross-word clusters C#CV in the design of our Spanish study (unlike in the German study in Chapter 3) as a context in which the syllabic organization is different from that of onset CCV clusters. This is due to the following reasons. First, the set of word-final consonants is limited in Spanish to sonorants and /s/ (at least for non-loanwords) and no word-final stop consonants are permitted (Harris 1983, Steriade 1999). It is such word-final stops that would be needed in the first C position of a C#CV so that we can compare for example a cross-word p#lV condition to a word-initial plV (which is attested). Hence, not all desirable contrasts to our multiple CCV word-initial onsets would be available. Second, it is unclear whether in Spanish there is resyllabification of the word-final consonant in the C#CV context so as to join the onset of the following syllable (see Shepherd 2003 for a literature review). For these reasons, we will thus make use of data from Moroccan Arabic which provide the needed appropriate comparison where the word-initial cluster is not a complex onset and crucially as explained above the phonetic properties of the segments across the two languages are the same. Conversely, using Moroccan Arabic as a contrasting syllabic organization case to that of German is not advisable. This is because German does not show robust open transitions in at least voiced stop-lateral cluster and its voicing implementation is expressed in short lag (voiced) versus long lag (voiceless) Voice-Onset-Time (VOT) opposition (Jessen 2002, Bombien & Hoole 2013).

In our results for Spanish, we demonstrate several sources of evidence for a global organization of the segments that are part of the hypothesized syllabic units. This evidence is reflected in a number of ways: shortening of the prevocalic consonant, reorganization of the relative timing of internal CV subsequence (in a CCV), early vowel initiation, and a strong compensatory relation between the duration of the open transition (in the obstruent-lateral juncture) and lateral duration. In other words, we find that the global organization that seems to hold over the segments partaking in these tautosyllabic CCVs is simultaneously expressed over a set of different phonetic parameters, a property we refer to as pleiotropy, rather than via a privileged metric such as c-center stability or any other such given single measure (employed in prior works). Our Spanish results offer a particularly striking illustration of this latter point. When we consider the stability of different intervals, the patterns are such that the interval whose stability has been taken to be indicative of a local organization is observed for voiced stop-laterals and the interval whose stability has been considered indicative of global organization is observed for voiceless stop-lateral clusters. Finally and crucially, the effects reflecting global organization in Spanish are not present or are less present in Moroccan Arabic.

2.2 Background

The relation between qualitative phonological organization and continuous phonetics is a fundamental problem in spoken language. We study a specific instance of this problem in the relation between syllable structure and phonetic indices for that structure.

Past analyses of syllable structure in articulation examine the variance of structurally relevant intervals, some spanning across and some spanning within the phonological constituents of interest, extracted from phonetic recordings of sound sequences with some hypothesized syllabic structure. The basic idea is that syllable structure corresponds to a form of temporal enquoteglue or temporal stability (Browman & Goldstein 1988). In these studies, patterns of relative stability of structurally relevant intervals are assessed and specifically these patterns are assessed over lexical forms of different shapes. Two patterns of stability have emerged, each characteristic of a particular qualitative syllabic organization. In languages that admit complex onsets like English, the most stable interval across CVC, CCVC and CCCVC utterances (where C is any consonant and V is any vowel) is an interval defined by the center of the prevocalic consonantal string and the end of the hypothesized syllable (Browman & Goldstein 1988, Honorof & Browman 1995, Byrd 1995, Marin & Pouplier 2010, Shaw & Gafos 2015). In contrast, it has been shown that in languages that do not admit complex onsets such as Arabic (Shaw *et al.* 2009), the most stable interval across CVC, CCVC and CCCVC utterances is defined by the immediately prevocalic consonant and the end of the hypothesized syllable (see also Goldstein *et al.* 2007, Hermes *et al.* 2015 on Berber). Thus, the stabilities of intervals seem to change depending on syllabic structure. These patterns of interval stabilities, local timing versus global timing stability, have thus been considered to be the phonetic correlates of different syllabic structures (simplex versus complex onsets). We henceforth refer to these phonetic, quantitative correlates as the stability-based indices of syllabic structure.

Nevertheless, subsequent experimental work has uncovered puzzling patterns which indicate that the phonetic manifestations of syllabic structure distinctions between simplex and complex onsets may be more involved (see Hermes *et al.* 2013 on Italian, Shaw *et al.* 2009 on Arabic, Marin & Pouplier 2010 on English, Pouplier 2012 and Brunner *et al.* 2014 on German, Marin 2013 on Romanian). Shaw *et al.* (2009, 2011) report speaker-specific patterns in Arabic data, some of which resemble patterns reported previously for English, even though the two languages (Arabic and English) are considered to be prototypical examples of the complex (English) versus simplex (Arabic) typological distinction. Staying within one language, let us consider German for which several recent studies based on articulatory data have been published: Pouplier (2012) examines the clusters /bl, pl, gm, km, sk/ in syllable onset as well as in coda position, Brunner *et al.* (2014) investigates /pl, kn, gl, kv, sk/ word-initially, word-medially, and under two accentuation conditions, and Mücke *et al.* (2014) reports a comparison of timing in /kl/ clusters (as in "klima") between a healthy control and three essential tremor patients (under deep brain stimulation). Each of these studies highlights a specific problem – how variability affects the phonetic indices of syllabic organization – in different ways. Pouplier (2012) finds some evidence for the expected pattern of c-center stability in onsets but also strong effects of cluster identity on the stability patterns. Brunner *et al.* (2014) finds evidence for the same result, reporting c-center stability in some clusters (e.g., /gl/) but also exceptions to it (e.g., /pl/) especially when the cluster occurs before a lax vowel (e.g., "Plätze" versus "plagen"); that is, not all clusters or cluster-vowel combinations show the expected c-center stability, even though for the clusters investigated (e.g., /gl, pl/) there is solid phonological evidence for their syllabic onset status (Wiese 1996). Mücke *et al.* (2014) find that unlike their control participant, their essential tremor patients exhibited increased variability in terms of individual consonant durations (suggested to be a symptom of dysarthria) which in turn precluded finding the expected interval stability pattern for complex onset organization as in the control participant. German is not the only language for which divergent results have been reported. Thus, Marin (2013), on Romanian, finds evidence for the expected c-center stability in sibilant-initial onset clusters /sk, sp, sm/ but also the unexpected right-edge stability for stop-initial onset clusters /ps, ks, kt, kn/ although native speakers classify the latter as well as the former as complex onsets.

Given this background, a study of Spanish consonant clusters and their spatio-temporal relation to the subsequent vowel both expands the empirical range of available data and promises to be informative in furthening our understanding of the relation between syllable structure and phonetic indices for that structure for the following reasons. Spanish is a language which shows certain phonetic characteristics, in the timing of consonant clusters, similar to languages like Moroccan Arabic but phonological properties, in terms of syllable organization, that are similar to English. Specifically, a recent articulatory study shows that stop-lateral and stop-tap clusters are timed with an open transition (Gibson *et al.* 2017), that is, a period of no vocal tract constriction between the release of the stop and the achievement of target of the lateral (for previous observations using acoustic data see Malmberg 1965, Colantoni & Steele 2005, Bradley 2006). This same timing pattern is also found in Moroccan Arabic (Gafos 2002, Gafos *et al.* 2010, Gafos *et al.* to appear). Moreover, the implementation of voicing is comparable between the two languages, with voiceless stops in particular being of the short lag Voice-Onset-Time (VOT) type. This means that, across Moroccan Arabic and Spanish, the phonetic profile of, for example, a /kla/ sequence is comparable. This would not be so if we were to refer to English or

German instead of Spanish. German and English differ from Moroccan Arabic in that they do not show robust open transitions in at least voiced stop-lateral clusters and in that their voicing systems involve a short lag (voiced) versus long lag (voiceless) VOT opposition. However, Spanish is a language which like English is assumed to admit clusters as onsets in syllables (Harris 1983:13–14, Hualde 2005:71) in contrast to Moroccan Arabic which is claimed not to admit clusters as complex onsets (Dell & Elmedlaoui 2002, Gafos *et al.* to appear). Thus, Spanish clusters exhibit timing patterns of the sort that characterize languages which do not admit clusters as syllable onsets but the language is prototypically of the complex onset type.

Our present study thus has two aims. First, we establish the basic timing patterns in Spanish stop-lateral and stop-tap clusters and their coordination with the subsequent vowel. Second, we draw implications of these findings for the issue of the relation between inter-segmental timing and syllabic organization by comparing Spanish with Moroccan Arabic. The Chapter is organized as follows. We begin in section 2.3 by outlining the method and stimuli used for the Spanish articulatory data. A detailed analysis of the registered data in section 2.4 then follows. As the Spanish clusters studied here have not been the subject of an articulatory study previously, we devote considerable attention, in section 2.4, to documenting their basic patterning using a variety of spatio-temporal measures: duration of the prevocalic sonorant across CV and CCV, relative timing of the lateral with the vowel across CV and CCV, relation between inter-gestural timing in the CC cluster (expressed by the interplateau interval between the two consonants) and sonorant duration, and finally vowel initiation with respect to the preceding segment or segmental combination. Section 2.5 summarizes the resulting patterns and argues that joint consideration of the various measures provides evidence for the presence of a global organization over the segmental sequences investigated herein. The key results from Spanish are then compared to patterns in corresponding clusters of Moroccan Arabic in section 2.6. We conclude in section 2.7 with implications of our results for the issue of the relation between syllable structure, an aspect of language-particular phonological organization, and spatio-temporal patterning in sequences of segments over which that syllable structure is hypothesized.

2.3 Method

2.3.1 Subjects

Articulatory data were collected at the University of Potsdam from six native speakers of Central Peninsular Spanish (vp01, vp02, vp03, vp04, vp05 and vp06). All speakers reported no speech or hearing problems, provided written informed consent prior to the investigation, and were reimbursed for their participation. All experimental procedures were approved by the Ethics Committee of the University of Potsdam.

2.3.2 Speech material

The stimuli consist of real disyllabic words in Spanish starting with consonant clusters or single consonants. Table 2.1 presents the list of Spanish CCV and CV paired words. CC-initial words begun with stop-lateral or stop-rhotic clusters. Their paired single consonant-initial words begun with a lateral or a rhotic such that in any CV∼CCV pair the prevocalic consonant remained the same across the CV- and CCV- initial words in that pair (e.g., /plato/ is paired with /lato/). Each stimulus word was embedded in the carrier phrase "Di _____ por favor," with the stimulus word location indicated by the '_____'. The clusters included in the analysis are /pl/, /bl/, /kl/, /gl/, /kɾ/, /tɾ/, /pɾ/[1]. There were 38 different CV- and CCV- initial words per speaker. For both clusters and singletons, the vowels following the consonant(s) of interest are the low vowel /a/ and the mid vowels /e/ and /o/. Furthermore, we also aimed to keep the postvocalic consonant the same within word pairs to the extent possible (e.g., as in our /plato/ paired with /lato/ example). For the cases where this was not possible – given that we opted for including real words as stimuli – the postvocalic consonant was different within the pair but either manner of articulation or place of articulation was maintained, as in /lomo/-/globo/. Each speaker produced ten or eleven repetitions of each item ($N = 38$) yielding a total of about 380 tokens per speaker.

2.3.3 Data acquisition

The data were acquired by means of Electromagnetic Articulometry (EMA) using the Carstens AG501 articulograph. The system tracks the three-dimensional movement of

[1] The voiced stop-rhotic clusters /bɾ, gɾ/ were not recorded because our corpus was quite large, consisting of both word-initial and word-medial clusters. The word-medial clusters are not part of the current study. Even if the stop-rhotic clusters are not balanced with respect to C1 voicing (since the voiced stop-rhotic clusters were excluded), we are still able to address the effect of voicing of the initial stop on inter-segmental temporal coordination using stop-lateral clusters with a voiced or voiceless stop as part of our corpus.

Table 2.1: Spanish stimuli.

	low vowel (/a/)		**mid vowel (/e/, /o/)**	
Cluster	**CCV**	**CV**	**CCV**	**CV**
pl	plato ‘plate’	lato ‘to whip’	plena ‘full’ plomo ‘lead’	Lena (proper name) lomo ‘loin’
bl	blata (model of motorcycle) blanda ‘soft’	lato ‘to whip’	bleque ‘tar’ bloque ‘block’	leco ‘nuts’ loco ‘crazy’
gl	glato (province in Italy) glana (surname)	lato ‘to whip’	gleba ‘mound of land’ globo ‘balloon, globe’	lema ‘slogan/motto’ lomo ‘loin’
kl	[k]lapas ‘stripping of unfertile land’	lapa ‘Lap (from Lapland)’	[k]lema ‘electrical connector’ [k]lono ‘clone’	lema ‘slogan/motto’ lomo ‘loin’
pr	prato (city in Italy)	rato ‘while, bit’	presa ‘pray’ promo ‘promotion’	Rena (proper name) romo ‘blunt’
kr	[k]rapa (name of Italian restaurant in Valencia)	rapa ‘flower of olive tree’	[k]rema ‘cream’ [k]lomo ‘chrome’	rema ‘to row’ romo ‘blunt’
tr	trapo ‘rag’	rape ‘monkfish’	trecho ‘stretch’ trono ‘throne’	Recho (fishing company in Galicia) roto ‘broken’

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 triggers the articulograph to start recording. The subject sat on a chair in a sound-proof booth and was instructed to read 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: a 'tongue tip' (TT) sensor attached 1 cm posterior to the apex of the tongue, a 'tongue mid' (TM) sensor attached 2 cm posterior to the TT and a '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.

2.3.4 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, r/) using the TT sensor, and labials (/p, b, m/) were measured using the lip aperture (LA). LA is a derived signal computed as the Euclidean distance between upper and lower lip sensors. The vowels following the consonant(s) of interest were measured using the TB sensor and 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 algorithm developed by Mark Tiede at Haskins Laboratories. The 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 2.1 illustrates an example parse of the velar and coronal gestures of the two initial consonants of the word /glato/ using the TB and TT sensors respectively. The panels from top to bottom illustrate the acoustic signal of the zoomed in portion of the utterance along with the TB and the TT movement trajectories. The black filled boxes correspond to the constriction phase, or plateau, of the gestures delimited by the target and release landmarks (left and right side of the black filled boxes). The left and right sides of the white boxes indicate the initiation and end of the gestures calculated as the timestamps at which velocity rises above and falls below a 20% threshold of local tangential velocity peaks.

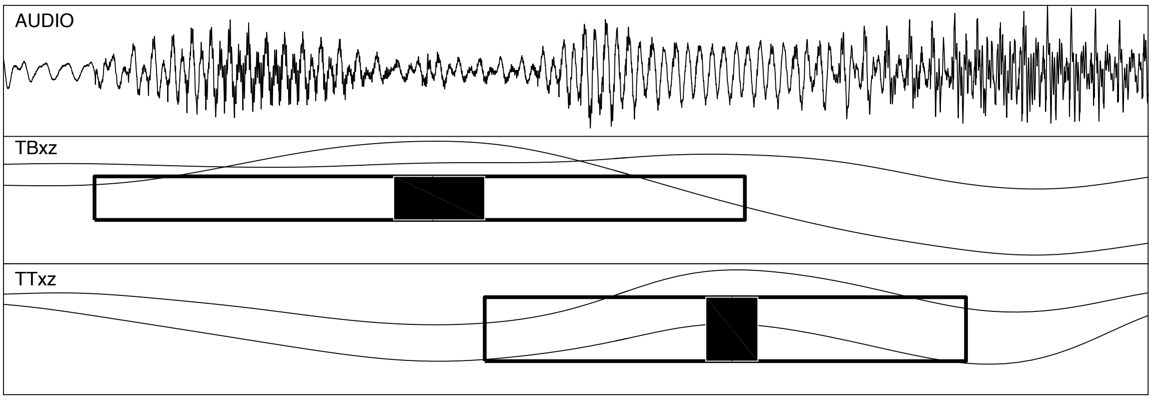

Figure 2.1: Parsing of the first two consonantal gestures of the word /glato/. Panels from top to bottom: acoustic signal of the region of interest, tongue back (TB) movement trajectory in the anterior-posterior and superior-inferior dimensions (x, z), tongue tip (TT) movement trajectory in the anterior-posterior and superior-inferior dimensions (x, z). The black filled rectangles indicate the constriction phases, or plateaus, of the TB and TT gestures. The left and right sides of the filled rectangles indicate the initiation and end of the gestures.

2.4 Spatio-temporal properties in Spanish onsets

2.4.1 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 ms. Positive IPIs indicate no temporal overlap between the plateaus of the two consonants, while negative IPIs indicate temporal overlap. Normalized IPI was also calculated to compensate for effects related to inter-speaker variability (cf. Bombien 2011). The IPI was normalized by dividing the raw measure by the total constriction duration of the cluster (i.e., IPI/(C2 release – C1 target)).

For the statistical analysis, linear mixed effects models were fitted with IPI (raw and normalized) as dependent variables, C1 voicing (voiced, voiceless), speaker and C1 place (labial, velar) as fixed effects. The variable "word", which corresponds to each unique token in the corpus, was used as a random factor. The results showed that there is a significant main effect of voicing on IPI ($p = 0.002$) with IPI being 10 ms greater in C1 voiceless stop-lateral clusters than in C1 voiced clusters across speakers. The effect of C1 place is at the limits of significance ($p = 0.05$) with IPI being 7 ms greater in C1 velar stop-lateral clusters than in labial stop-lateral clusters. There is no interaction between C1 place and C1 voicing. Figure 2.2 illustrates IPI values as a function of C1 voicing.

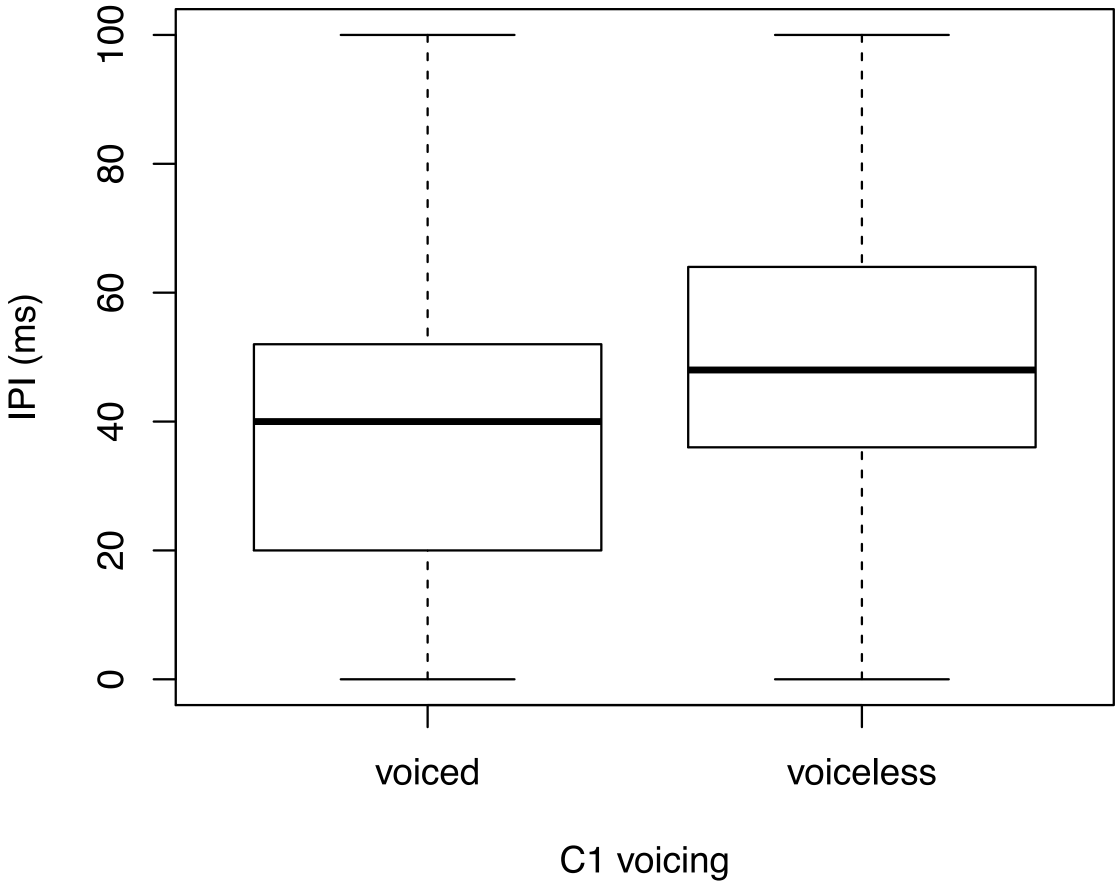

Figure 2.2: Interplateau interval (IPI) as a function of C1 voicing.

Raw IPI measures exhibit substantial variability as can be seen by their ranges in Figure 2.2. 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 durational extent of the cluster CC, the longer the IPI). Hence, it seems useful to also examine a normalized measure of IPI. The linear mixed effects model with the normalized IPI as dependent variable showed no main effect of C1 voicing and no main effect of C1 place. Figure 2.3 shows normalized IPI as a function of C1 place and voicing. It is evident that when IPI is normalized by taking into account total cluster duration, regardless of C1 voicing or C1 place of articulation, there is no difference in IPI.

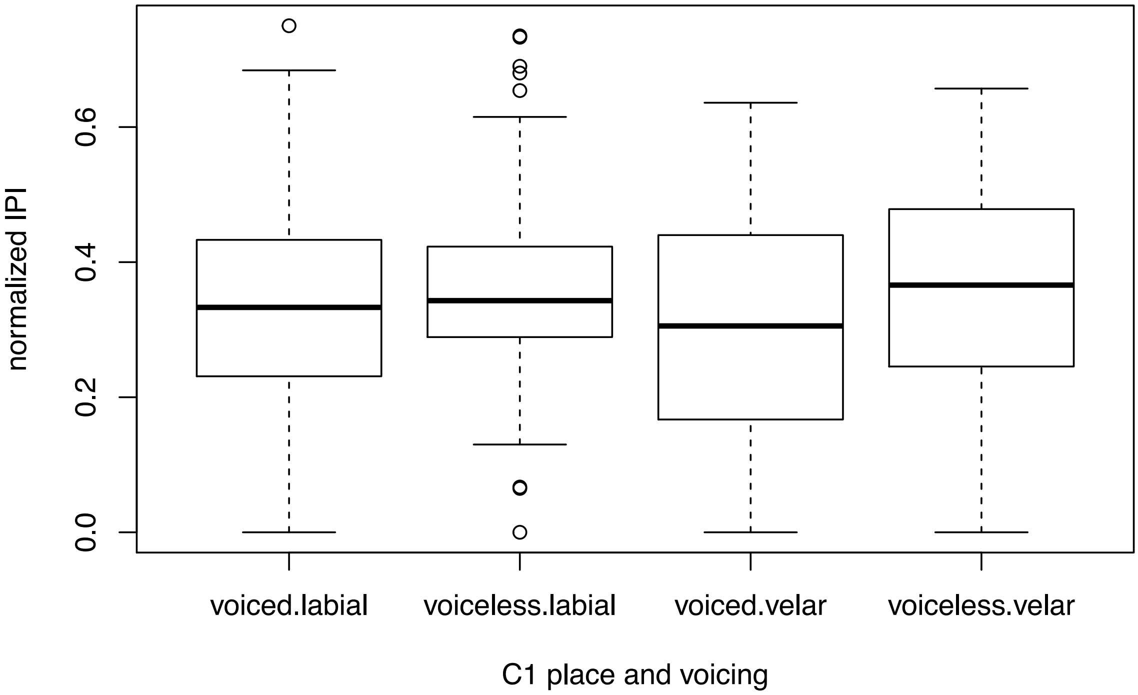

Figure 2.3: Normalized IPI as a function of C1 place and voicing.

Means and standard deviations for raw IPI and normalized IPI in C1 voiced and C1 voiceless stop-lateral clusters when C1 is a velar or a labial are shown in Table 2.2.

Table 2.2: Means and standard deviation of IPI as a function of voice and place.

Interplateau interval IPI (ms)			Normalized IPI	
	Mean	*sd*	Mean	*sd*
C1 voiced	37.9	21.1	0.32	0.16
Labial	35.9	18.2	0.33	0.15
Velar	40.2	23.7	0.31	0.17
C1 voiceless	48.2	20.8	0.35	0.13
Labial	46.9	16.8	0.35	0.11
Velar	49.7	24.9	0.34	0.15

2.4.2 C1 duration in stop-lateral clusters

In this subsection, we examine how C1 plateau duration changes as a function of C1 voicing and C1 place of articulation. Consonant plateau duration was calculated as the interval between C release and C target. C1 plateau duration was also normalized by dividing it by the total duration of the cluster: C plateau normalized = C plateau duration/(C2 release – C1 target). For the statistical analysis, linear mixed effects models were fitted with C1 plateau duration (raw and normalized) as dependent variables, C1 voicing (voiced, voiceless), speaker and C1 place (labial, velar) as fixed effects. The variable "word", which corresponds to each unique token in the corpus, was used as a random factor. The results showed that there is a main effect for C1 voicing with voiceless stops being 13 ms longer than voiced stops ($p = 0.0004$). This is of course consistent with the cross-linguistic generalization about voicing in stops (Lisker 1957). Furthermore, there is a main effect for C1 place with velars being 9 ms longer than labials ($p = 0.003$). With the normalized C1 plateau duration as dependent variable, the main effect of C1 voicing is maintained ($p = 0.01$). However, there is no main effect of C1 place. Figure 2.4 illustrates C1 plateau duration in ms as a function of C1 place and voicing. Figure 2.5 illustrates normalized C1 plateau duration as a function of C1 place and voicing. The boxplots for C1 voiceless clusters show longer C1 plateau duration than the C1 voiced clusters. No C1 duration difference is observed when looking at velar versus labial clusters.

Means and standard deviations for C1 plateau duration in ms and also normalized C1 duration in C1 voiced and C1 voiceless stop-lateral clusters when C1 is a velar or a labial are provided in Table 2.3.

Table 2.3: Means and standard deviations for C1 duration.

	Duration C1 (ms)				Normalized C1 duration			
	Labial	*sd*	Velar	*sd*	Labial	*sd*	Velar	*sd*
Voiced	31.7	13.1	44.2	23.6	0.30	0.09	0.33	0.11
Voiceless	49.5	17.2	52.4	21.5	0.36	0.08	0.37	0.12

2.4.3 C2 lateral duration and IPI-C2 compensatory relation in stop-lateral clusters

In this subsection, we examine C2 plateau duration as a function of C1 voicing and C1 place of articulation. The relation between IPI and C2 duration will also be investigated, in order to examine compensatory effects in the CCV string. Consonant plateau duration was calculated as the interval between C release and C target: C plateau = C release – C target. C2 plateau duration was also normalized by dividing it by the total constriction

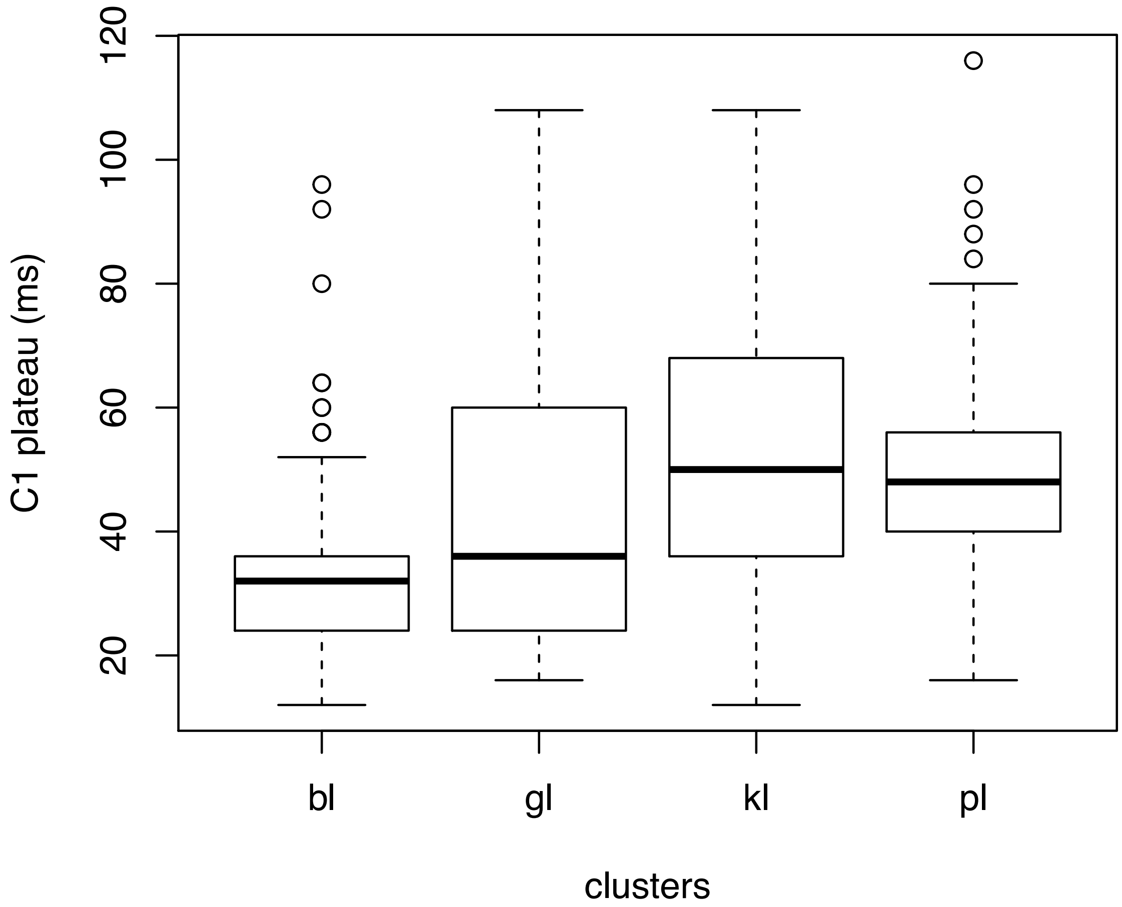

Figure 2.4: C1 plateau duration in ms as a function of cluster type. C1 plateau is greater for C1 voiceless stop clusters /pl, kl/ than for C1 voiced /bl, gl/ and greater for velar stop clusters /kl, gl/ than for labial stop clusters /bl, pl/.

duration of the cluster: C plateau normalized = C plateau duration/(C2 release – C1 target).

For the statistical analysis, linear mixed effects models were fitted with C2 plateau duration (raw and normalized) as dependent variables, C1 voicing (voiced, voiceless), speaker and C1 place (labial, velar) as fixed effects. The variable "word", which corresponds to each unique token in the corpus, was used as a random factor. The dependent variables C2 duration and normalized C2 duration were square rooted to approach a normal distribution. The results showed that the main effect of C1 voicing on the C2 plateau duration is at the limits of significance ($p = 0.06$) with the duration of the lateral C2 being 4 ms shorter when occurring after a C1 voiceless stop than after a C1 voiced stop. No main effect of C1 place was observed on the lateral's duration. The linear mixed effects models with the normalized C2 plateau duration as dependent variable showed a significant main effect on the lateral's duration for C1 voicing ($p < 0.0001$) with the lateral being shorter after a C1 voiceless stop (estimate = –0.06) and a significant main effect for C1 place (p

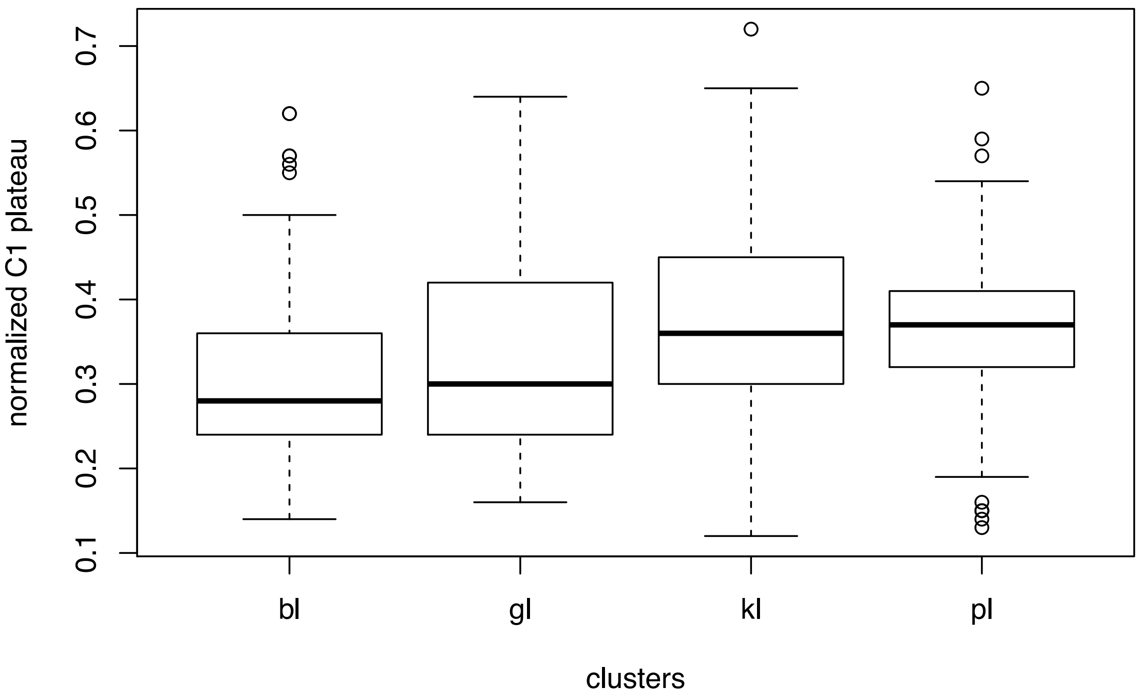

Figure 2.5: Normalized C1 plateau duration as a function of cluster type. C1 plateau is greater for C1 voiceless stop clusters /pl, kl/ than for C1 voiced /bl, gl/. No C1 duration difference is observed between velar stop clusters /kl, gl/ and labial stop clusters /bl, pl/.

= 0.02), with the lateral being shorter when following a velar (estimate = –0.02) than a labial. Figure 2.6 shows C2 plateau duration in ms as a function of cluster type. The duration of the lateral (C2) is slightly greater for /kl, pl/ than for /bl, pl/. The effect, however, becomes more obvious when looking at Figure 2.7 which shows normalized C2 plateau duration as a function of C1 voicing.

Furthermore, only with the normalized C2 lateral plateau duration as opposed to the raw values, there is an effect of C1 place of articulation with the lateral C2 being longer after velars than after labials. Means and standard deviations for C2 lateral duration in ms and normalized C2 duration in C1 voiced and C1 voiceless stop-lateral clusters when C1 is a velar or a labial are provided in Table 2.4.

Table 2.4: Means and standard deviations for C2 lateral duration in CCV.

	Duration C2 (ms)				Normalized C2 duration			
	Labial	*sd*	Velar	*sd*	Labial	*sd*	Velar	*sd*
Voiced	40	19.9	43.3	20.5	0.36	0.12	0.35	0.17
Voiceless	38.3	18.3	36.6	13.9	0.27	0.09	0.27	0.11

Segments in a cluster context have been found to be shorter than segments in the singleton context (Haggard 1973). Table 2.5 presents means and standard deviations

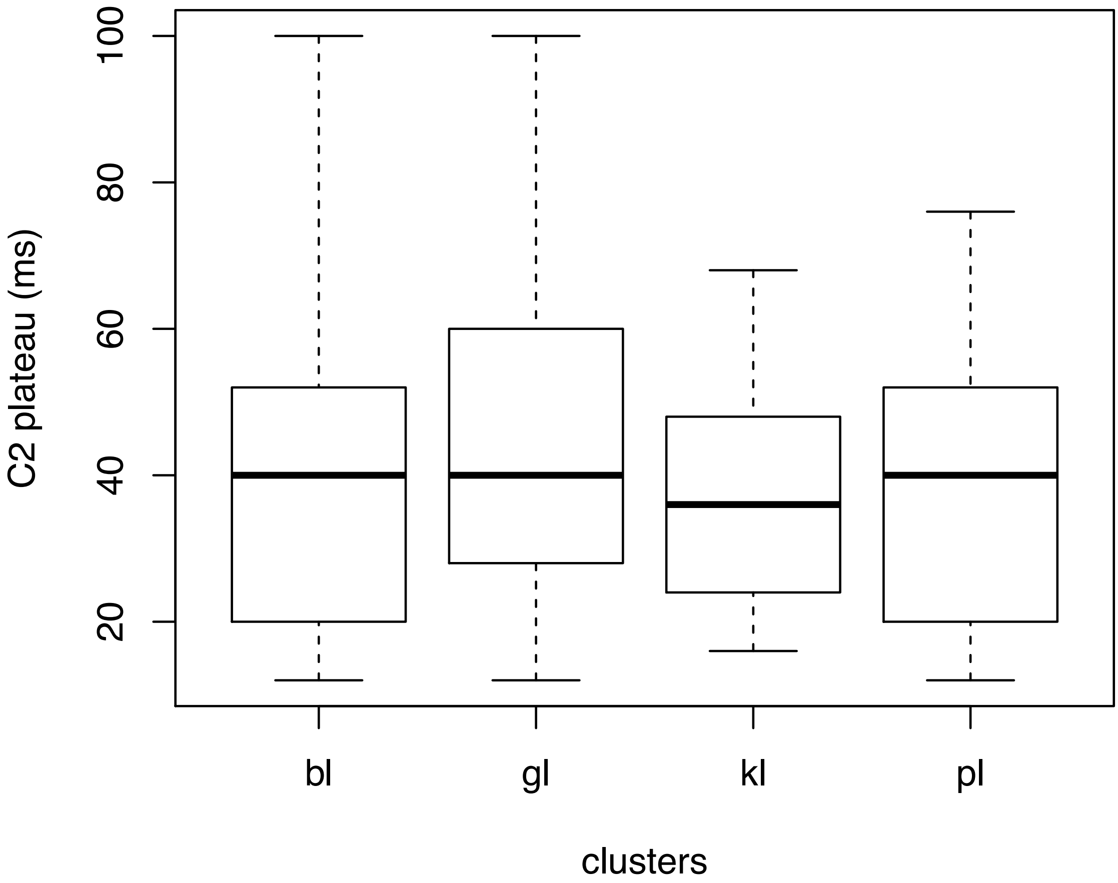

Figure 2.6: C2 plateau duration in ms as a function of cluster type.

for C2 lateral duration in CV and CCV. Our data confirm that the lateral in the CV context is longer than in the CCV context. This difference in duration is greater when the initial consonant is a voiceless stop. Thus, we find greater shortening of the lateral when a voiceless stop is added in the beginning of the CV string than when a voiced stop is added (15 versus 11 ms), a fact to which we will return to in the Discussion section. We argue that such a difference in the lateral duration between voiced and voiceless stop-lateral clusters is a result of global organization which imposes different degrees of shortening to the lateral from CV to CCV depending on the voicing of the initial stop and the different IPIs voiced and voiceless stop-lateral clusters exhibit.

Additionally, there is a strong negative correlation between normalized duration of the lateral C2 and normalized IPI ($r(587) = -0.70, p < 0.0001$), such that as IPI increases, C2 lateral duration decreases.[2] This compensatory relation between IPI and C2 duration can be seen in Figure 2.8. The relation is the same for both velars and labials, as can be

[2]When looking at the raw values of the two variables, there is no correlation ($r(587) = -0.16, p < 0.0001$), presumably due to the substantial variability across clusters and speakers.

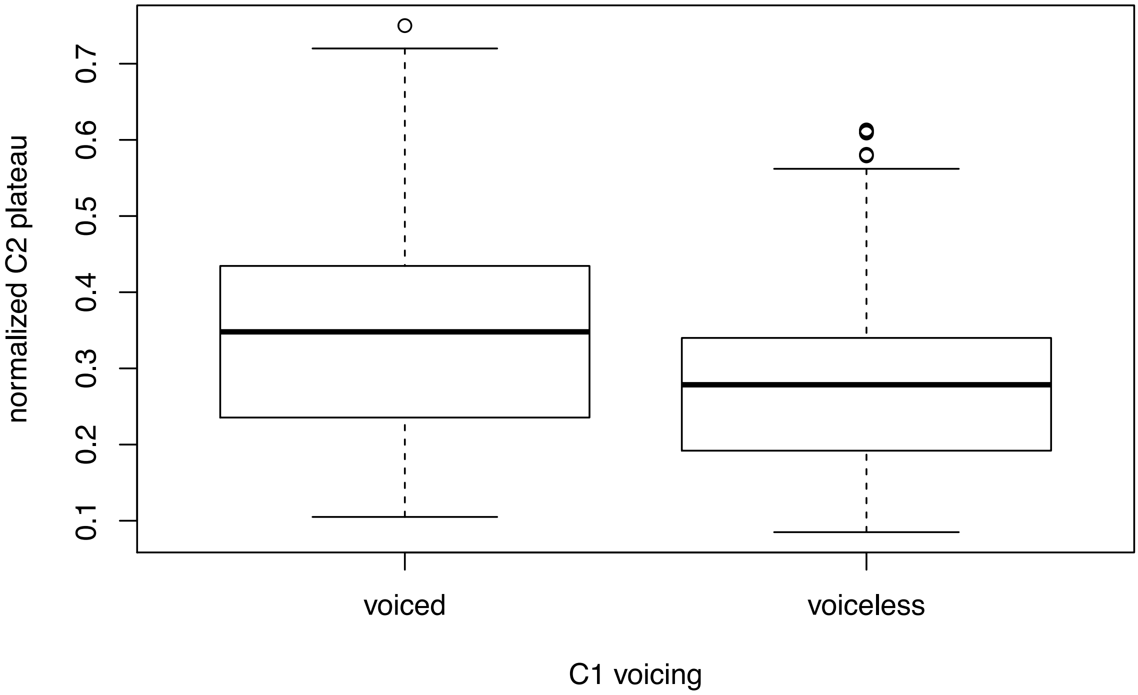

Figure 2.7: Normalized C2 plateau duration as a function of C1 voicing.

Table 2.5: Mean and standard deviations for C2 lateral duration in CV and CCV.

C2 lateral duration (ms)		
	Mean	*sd*
C1 voiceless		
CV	52.4	22.1
CCV	37.5	16.5
C1 voiced		
CV	52.9	23.5
CCV	41.6	20.2

seen in Figure 2.9. This relation points to a global organization which brings the vowel to overlap with the cluster as will be discussed in section 2.5.

2.4.4 Stop-rhotic clusters

The stop-rhotic clusters in our dataset consist of the voiceless stop-ɾclusters /pɾ, kɾ, tɾ/. Since no effect of voicing on the interplateau interval can be assessed in this case, we only report IPI values, rhotic plateau duration and their relation in CCV for a comparison with the stop-lateral clusters. Table 2.6 provides means and standard deviations for the interplateau intervals of stop-rhotic clusters. Overall, the mean IPI in voiceless stop-rhotic clusters is far greater than the IPI of voiceless stop-lateral clusters (74 versus 48 ms).

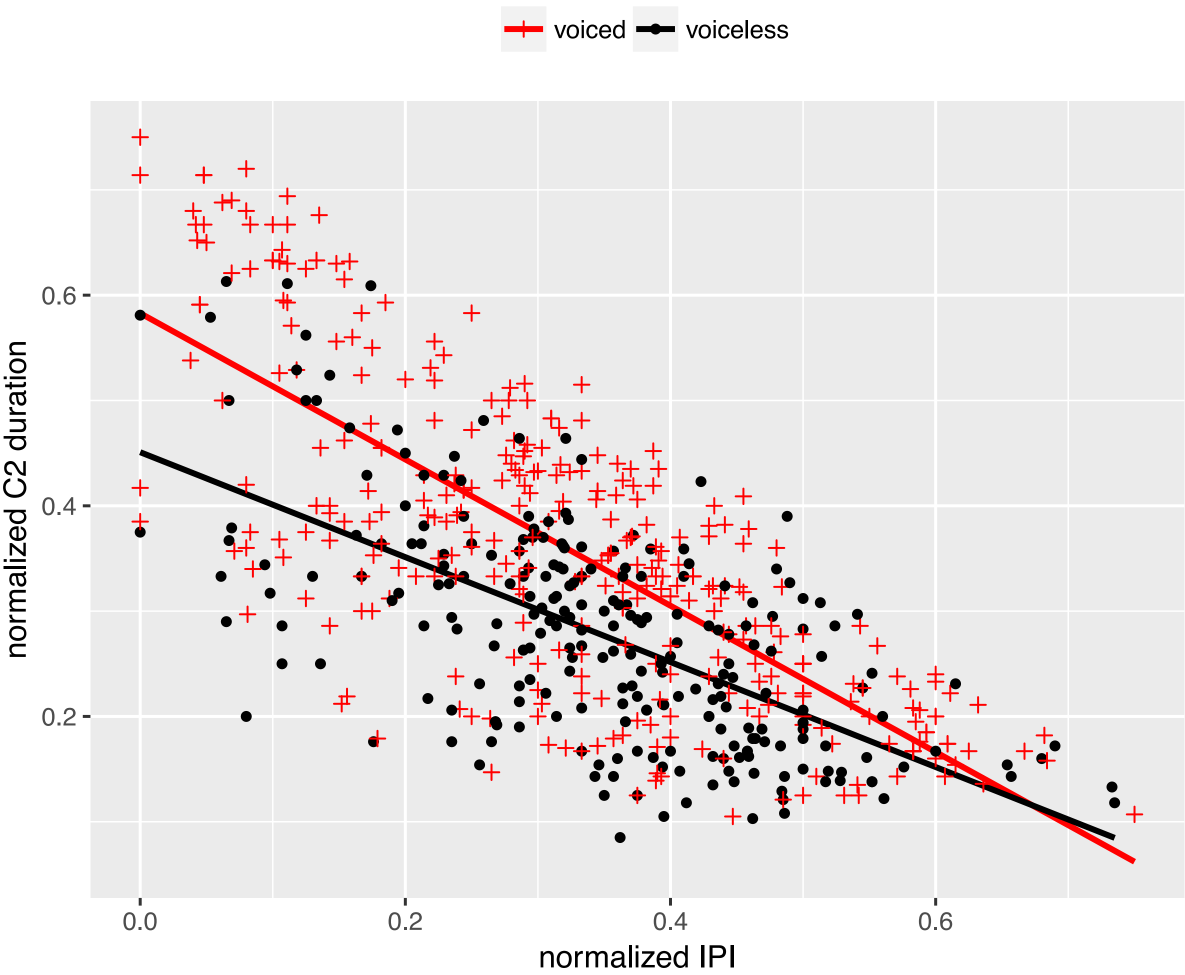

Figure 2.8: Scatterplot showing the compensatory relation between normalized C2 lateral duration and normalized IPI for voiced and voiceless stop-lateral clusters. As IPI increases, the lateral's (C2) duration decreases ($r(587) = -0.70, p < 0.0001$).

Table 2.6: Means and standard deviations for interplateau intervals (IPI) in stop-rhotics.

Interplateau interval IPI (ms)		
	Mean	*sd*
Across clusters	74.3	18.47
pɾ	69.8	18.1
kɾ	73.9	20.9
tɾ	79.5	14.2

Table 2.7 provides the duration of the C2 prevocalic rhotic in both singleton and cluster cases. The duration of the singleton has a mean value of 71 ms, while its duration in the cluster decreases substantially to 23 ms. In Spanish, the singleton rhotic is produced as a trill in the word-initial position, while in a cluster it is produced as a tap (Lipski 1990, Hualde 2005). These two segments have inherently different durations with the trill

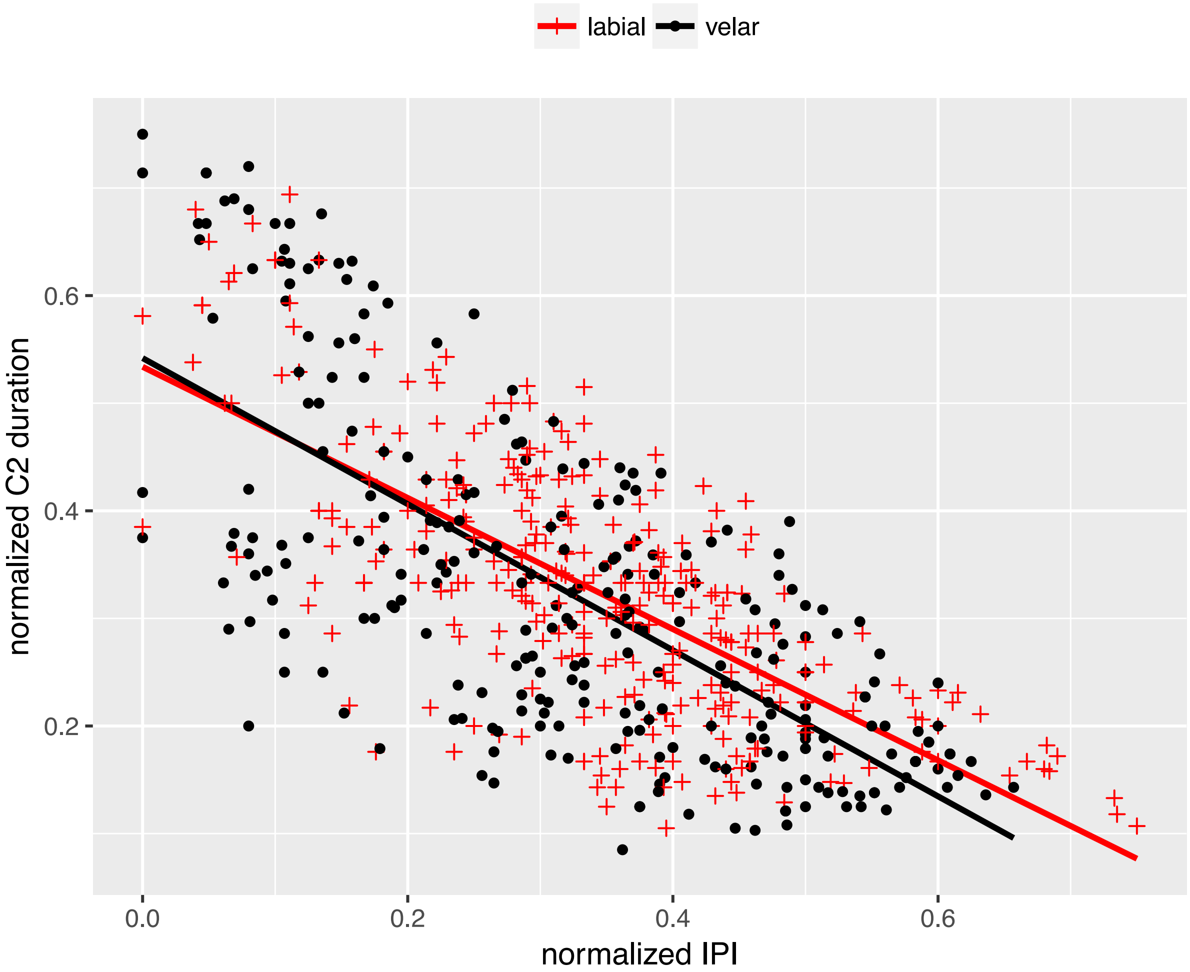

Figure 2.9: Scatterplot showing the relation between normalized C2 lateral duration and normalized IPI for labial and velar stop-lateral clusters. For both velar and labial C1 stops, as IPI increases, the lateral's (C2) duration decreases.

being longer than a tap. Regardless of whether a comparison between a singleton trill and a stop-tap cluster is meaningful, the point is that the duration of the rhotic decreases considerably from CV to CCV compared to the lateral. This difference in the degree of shortening between the rhotic and the lateral has an effect on the stability-based indices of syllabic organization (see Discussion).

Finally, there is no correlation between IPI and C2 tap duration in CCV using raw ($r(447) = 0.14, p = 0.002$) or normalized values ($r(447) = -0.25, p < 0.0001$).

2.4.5 Interim summary

To summarize, the place of articulation of C1 does not seem to have a strong effect on IPI, as there was only a 7 ms IPI difference between velar stop-lateral clusters and labial stop-

Table 2.7: Means and standard deviations for C2 rhotic duration.

C2 plateau duration (ms)		
	Mean	*sd*
CV	71.8	29.9
(p)r	72.3	29.7
(k)r	71.5	31.3
(t)r	64.8	27.5
CCV	23.05	11.1
pɾ	21.4	12.1
kɾ	22.5	10.6
tɾ	25.3	10.2

lateral which disappeared when looking at normalized IPI. Additionally, the results bore out predictions related to the duration of consonants in complex onsets and the effects of constriction location. Thus, C1 plateaus for voiceless consonants were significantly longer than for voiced as per the well-known cross-linguistic generalization (Lisker 1957) and slightly longer durations (without being statistically significant), independently of voice, for velars than for labials. The results of the effects that C1 voice and place have on the duration of the lateral are in line with Gibson *et al.*'s (2015) EPG study where it was found that laterals preceded by a tautosyllabic voiceless stop are systematically shorter than laterals preceded by a voiced stop.

Two novel results in our data concern the presence of a compensatory relation between lateral duration and interplateau interval and an asymmetry in lateral shortening depending on C1 voice. Specifically, we found that lateral duration and C1C2 interplateau interval (IPI) are linearly related such that the longer the IPI, the shorter the lateral. Furthermore, the voice specification of the C1 has a strong effect on the duration of the lateral, meaning that the lateral is shorter after a voiceless stop than after a voiced stop.

On the effect of voicing on IPI for stop-lateral clusters, the results are quantitatively different when looking at raw and normalized IPI. Voicing seems to have an effect on raw IPIs, with the IPI being 10 ms larger (less overlap) in voiceless stop-lateral clusters than in voiced stop-lateral clusters. The effect disappears when looking at normalized IPI. Although the results are quantitatively different, qualitatively they point to the same direction: IPI patterns in Spanish seem to be more similar to IPI patterns in French than in German (see Bombien & Hoole 2013 for a comparison between German and French and Kühnert & Hoole 2006 on French). This is because IPIs have overall large positive values in both Spanish and French (approx. 40 ms) and the effect of voicing on IPI is a difference of approx. 4 ms in French and 10 ms in Spanish. In German, however, as it has been reported in Bombien & Hoole (2013) and as we can verify with our own data in Chapter 3, the IPI patterns are quite different from French and Spanish in that voiced stop-lateral clusters show a very short, often negative IPI indicating temporal overlap between the plateaus of the consonants, while voiceless stop-lateral clusters show a larger IPI of approx. 30 ms. Furthermore, the IPI difference between voiced and voiceless stop-lateral clusters in German is approx. 22 ms, which is considerably greater than the IPI difference between voiced and voiceless stop-lateral clusters in Spanish and French. Overall, the IPI values observed in French and Spanish stop-lateral clusters regardless of voicing are similar to the IPI values of voiceless stop-lateral clusters in German. Our results, thus, support the hypothesis that language-particular voicing implementation (such as the difference between oral-laryngeal coordination in German/English versus Spanish/French) conditions intra-cluster timing.

For stop-rhotic clusters, we found a large IPI – even larger than for stop-lateral clusters. The overall large IPI values (little overlap) in stop-lateral and stop-rhotic clusters in Spanish regardless of C1 voicing point to the robustness of what has been referred to as an 'open transition' in consonant-sonorant clusters of Spanish (Catford 1988:118, Gafos 2002). This is in line with past acoustic studies dealing with Spanish complex onsets in which the existence a vocoid like element between the two consonants has been reported (see Quilis 1970, Quilis 1993, Levin 1987, Harms 1976, Hall 2006, Schmeiser 2009, Bradley 2006). Our study makes explicit the articulatory basis for such a phenomenon. Finally, we found no relation between IPI and C2 rhotic duration in CCV, unlike in stop-lateral clusters where a compensatory relation was observed between the two variables (we return to this asymmetry in the Discussion).

2.4.6 Stability analysis

In the previous subsections, we examined properties of individual consonants, such as the duration of the initial stop and the duration of the sonorant in CV and CCV contexts, as well as the inter-gestural timing, as expressed by IPI, of the two consonants before the vowel. We now turn to examine how the prevocalic material of a single consonant or consonant cluster relates to its subsequent vowel. We will do so, first in this subsection, by examining stability-based indices of syllabic structure, as have been employed in other work reviewed in section 2.2, on stop-lateral and stop-rhotic clusters in Spanish and their corresponding sonorant-vowel subsequences. The indices we quantify in this section refer to the interval stabilities of two intervals which have been claimed in prior work to be informative about syllabic organization. Subsequently, in quantifying the relation of the consonants to vowel by looking into vowel initiation measures, we will also consider measures of relative timing between the prevocalic consonant and the vowel as well as vowel initiation measures with respect to the preceding consonantism.

2.4.6.1 Stop-lateral clusters

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. For thoroughness, in assessing the robustness of results, we used three different anchors: the target of the constriction of the postvocalic consonant (C^{tar}), the maximum constriction of the postvocalic consonant (C^{max}), and the spatial extremum of the vowel (V^{max}) measured by the tangential velocity of the TB sensor. For the low vowel /a/, the spatial extremum corresponds to the timepoint where the tongue back reaches the lowest point. For the mid vowels /e/ and /o/, the spatial extremum corresponds to the timepoint where the tongue back reaches its highest position, which is midway between a high and a low vowel. 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 studies, e.g., Shaw *et al.* 2009, Shaw *et al.* 2011, Hermes *et al.* 2015, Shaw & Gafos 2015). The two intervals left-delimited by these landmarks were the c-center-to-anchor interval, extending between the temporal midpoint of the consonant(s) and the anchor, and the right-edge-to-anchor interval, extending between the constriction release of the (immediately) prevocalic consonant and the anchor. For ease of reference, we henceforth 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.

We begin with a first descriptive characterization of the data which consists in examining the relative standard deviation (RSD) of each interval calculated as the standard deviation of the interval divided by its mean duration and then multiplied by 100 to express the RSD value as percentage (Shaw *et al.* 2009). The interval with the lowest RSD value, if a lowest exists, indicates the most stable interval for any given word pair. For instance, for a given CV∼CCV word pair, if the RSD value of the global timing interval is lowest compared to the RSD value of the local timing interval, the global timing interval is more stable or shows less change (in comparison to the local timing interval) as the number of consonants increases from CV to CCV. RSD values of the two intervals global timing and local timing using three anchor landmarks (C^{tar}, C^{max}, V^{max}) for our word pairs (e.g., /lato/∼/plato/, /loco/∼/bloque/ and so on) are provided in Table 2.8.

Table 2.8: Relative Standard Deviation (RSD) of the intervals global timing and local timing using three anchors (C^{tar}, C^{max}, V^{max}) across pairs of words for six speakers (vp01-06).

		C^{tar}		**C^{max}**		**V^{max}**	
		global	**local**	**global**	**local**	**global**	**local**
	lato∼plato	8.46	12.66	10.83	17.38	12.37	18.24
	lato∼blanda	8.6	9.77	11.05	10.45	13.42	16.82
	lena∼plena	11.56	8.5	10.36	11.17	9.34	12.14
	leco∼bleque	13.93	13.15	18.06	14.12	23.76	48.18
	lomo∼plomo	5.36	14.11	5.99	13.39	6.88	20.1
	loco∼bloque	20.48	21.51	26.85	29.89	12.56	24.81
vp01	lato∼blata	7.38	10.73	9.34	13.07	11.16	13.12
	lato∼glato	13.76	7.38	14.3	7.51	22.62	8.16
	lato∼glana	10.4	8.09	12.39	8.61	20.24	5.38
	lema∼gleba	8.49	12.02	7.99	11.42	18.95	23.49
	lomo∼globo	40.07	47.49	37.18	43.67	27.57	26.73
	lapa∼clapas	5.44	13.96	3.92	11.92	4.52	20.73
	lema∼clema	5.64	11.96	4.66	10.84	12.28	22.87
		global	**local**	**global**	**local**	**global**	**local**
	lato∼plato	13.15	8.88	11.56	7.29	17.99	11.26
	lato∼blanda	12.19	5.67	15.67	9.61	15.51	10.26
	lena∼plena	11.23	7.14	10.97	6.51	12.71	26.29
	lomo∼plomo	12.19	7.22	11.47	8.73	16.27	15.96
	loco∼bloque	15.97	9.25	15.58	9.17	19.14	7.63
	lato∼glana	11.23	8.1	9.97	7.2	17.95	14.73
vp02	lato∼blata	13.76	5.07	11.42	5.87	18.04	10.5
	lato∼glato	7.88	5.85	5.27	6.22	11.76	8.21
	lema∼gleba	8.84	5.13	8.34	4.86	12.64	12.24
	lomo∼globo	13.77	5.49	11.82	6.61	10.96	19.42
	lapa∼clapas	10.5	12.78	9.08	13.48	18.73	8.27
	lema∼clema	9.8	10.44	9.43	9.01	18.51	12.32
	lomo∼clono	10.01	13.37	8.9	13.42	12.14	21.65
		global	**local**	**global**	**local**	**global**	**local**
	lato∼plato	10.02	8.69	13.13	15.08	13.46	9.33

	lato∼blanda	8.55	5.21	12.32	10.36	9.94	9.18
	lena∼plena	15.84	18.79	14.49	16.98	13.3	16.6
	loco∼bloque	8.88	7.17	11	12.91	14.9	16.8
	lato∼blata	8.21	6.5	11.39	12.93	11.06	6.1
	lato∼glato	10.96	11.03	14.73	16.58	37.95	69.24
vp03	lato∼glana	11.76	8.73	11.41	15.88	39.36	70.81
	lapa∼clapas	11.68	11	15.78	16.83	13.73	23.61
	lema∼clema	16.7	17.17	16.19	15.93	17.28	23.23
	lema∼gleba	15.97	14.47	15.23	13.99	42.5	71.69
	lomo∼clono	12.39	10.26	10.22	16.52	43.71	83.37
	lomo∼plomo	10.69	21.37	12.65	22.3	24.78	38.27

		global	**local**	**global**	**local**	**global**	**local**
	lato∼glato	11.66	7.61	10.94	7.49	19.8	35.84
	lato∼glana	11.79	10.37	10.8	5.42	36.54	45.7
	lato∼blanda	8.89	6.62	10.71	9.49	29.57	44.06
	lato∼blata	9.88	6.49	9.17	5.57	21.6	33.49
	lato∼plato	6.98	5.7	6.29	5.23	21.47	30.06
	lena∼plena	12.91	26	11.13	22.25	28.63	34.55
vp04	leco∼bleque	19.06	16.42	17.46	15.14	40.16	69.44
	loco∼bloque	15.25	12.37	15.33	12.48	21.11	26.92
	lapa∼clapas	13.57	10.14	12.32	9.46	30.71	51.56
	lema∼clema	13.82	11.82	13.15	10.94	38.9	59.59
	lema∼gleba	10.93	15.72	9.92	14.08	22.72	26.17
	lomo∼clono	12.58	8.82	11.74	9.81	48.49	98.45
	lomo∼plomo	10.99	15.76	10.69	16.42	49.08	61.81

		global	**local**	**global**	**local**	**global**	**local**
	lato∼glana	13.48	11.01	12.31	12.95	20.76	11.15
	lato∼blanda	10.27	9.16	13.88	8.62	13.45	29.45
	lato∼glato	13.61	8.61	12.94	8.86	20.13	13.93
	lato∼blata	13.46	6.31	12.02	6.5	23.45	36.21
	lato∼plato	13.62	7.81	9.78	10.98	17.55	17.47
vp05	lena∼plena	19.31	29.6	18.9	28.57	18.4	31.32
	leco∼bleque	17.44	23.24	18.22	22.55	18.5	61.35
	loco∼bloque	24.41	29.1	20.38	25.58	18.55	28.57
	lapa∼clapas	15.25	18.69	15.05	17.44	18.14	18.71

	lema∼clema	14.63	19.18	13.31	16.72	30.83	28.02
	lema∼gleba	16.73	15.64	15.07	14.25	30.13	30.74
	lomo∼plomo	9.86	22.94	11.39	24.46	31.07	32.54
	lomo∼globo	17.08	15.16	15.87	16.16	22.81	50.89
	lomo∼clono	17.2	9.35	17.29	12.72	21.29	18.3

		global	**local**	**global**	**local**	**global**	**local**
	lato∼blanda	6.61	11.24	7.76	10.47	25.49	38.97
	lato∼blata	11.58	7.59	10.61	9.17	14.41	12
	lato∼plato	8.96	9.27	8.56	10.48	12.77	14.22
	lato∼glato	12.55	6.61	12.11	7.82	25.21	39.4
	lato∼glana	13.44	6.89	11.54	8.03	22.19	12.99
	lena∼plena	10.1	7.37	9.28	7.75	31.9	31.54
vp06	leco∼bleque	11.11	7.35	14.43	6.19	41.83	78.55
	loco∼bloque	10.09	10.84	9.93	10.78	20.8	63.51
	lapa∼clapas	7.27	10.82	6.7	10.46	13.9	10.53
	lema∼clema	7.91	14.06	7.42	13.29	11.14	18.98
	lema∼gleba	10.09	9.47	8.74	9.25	13.56	13.3
	lomo∼plomo	8.89	10.52	7.51	10.27	17.61	18.31
	lomo∼clono	13.41	3.91	12.58	4.91	14.12	9.31

According to the stability-based indices of syllabic organization reviewed in section 2.2 (Background), we expect Spanish clusters to exhibit global timing interval stability (RSD minima should be consistently found for global timing). However, the interval stabilities, as seen in Table 2.8, seem quite variable. By simply consulting the RSDs in Table 2.8, as presented per speaker, per cluster and per vowel, it is not straightforward to discern overall consistent patterns. The patterns that are present in this data will become clear once we assess our intervals statistically, as we do below, by turning the relevant variables (speaker, cluster, vowel) into factors. However, we can still descriptively characterize the data by making the following observations. On the one hand, across speakers, there are some voiceless stop-lateral-vowel combinations with the RSD of the global timing interval being the lowest. On the other hand, again across speakers, for some voiced stop-lateral-vowel combinations, the RSD of the local timing interval remains the lowest. Regardless and in contrast to the expectation of global timing interval stability, it is clear that the RSD values of the global timing interval are not always the lowest. Across speakers, there is evidence for both global and local timing interval stability.

We now turn to evaluate statistically how the global timing and local timing intervals change as the number of consonants increases from CV to CCV (cluster size) across speakers. Furthermore, we investigate the effect of voicing on the interaction between cluster size and interval type. A series of linear mixed effects models were fitted with interval duration as a dependent variable. Interval duration was log transformed to approximate a normal distribution. Interval type, cluster size, C1 voicing and speaker were used as fixed effects. The variable "word", which corresponds to each word in the corpus, was introduced as a random factor. For stop-lateral clusters, there is a significant interaction between interval type (local versus global), cluster size and C1 voicing across all anchors (C^{tar}: $F(2) = 22.9$, $p < 0.0001$; C^{max}: $F(2) = 23.2$, $p < 0.0001$; V^{max}: $F(2) = 5.78$, $p = 0.003$). For voiceless stop-lateral clusters, using C^{tar} as anchor both intervals were found to change significantly with the global timing and local timing intervals showing the same degree of change to different directions. From CV to CCV, the global timing interval increases (C^{tar}: $p < 0.0001$, estimate = –0.13), while the local timing interval decreases (C^{tar}: $p < 0.0001$, estimate = 0.13). Using C^{max} and V^{max} as anchors, however, the stability pattern is clearly in favor of the global timing interval. Both intervals change significantly, but the global timing interval shows the smallest increase from CV to CCV (C^{max}: global timing $p = 0.0008$ estimate = 0.10, local timing $p < 0.0001$ estimate = 0.13; V^{max}: global timing $p = 0.0006$ estimate = –0.77, local timing $p < 0.0001$ estimate = 1.31). Vowel was not a significant factor in any of the statistical models. There was no interaction between interval type, cluster size and cluster type within voiceless stop-lateral clusters meaning that both /pl, kl/ clusters behave the same way with respect to changes of interval duration. Overall, voiceless stop-lateral clusters provide substantial evidence for the global timing interval stability.

For the voiced stop-lateral clusters, there is a three-way interaction between interval type, cluster size and cluster type (C^{tar}: $F(2) = 9.08$, $p = 0.0001$; C^{max}: $F(2) = 11.65$, $p < 0.0001$; V^{max}: $F(2) = 2.67$, $p = 0.06$), meaning that the way intervals change from CV to CCV depends on whether the cluster is a /bl/ or /gl/. For /bl/, the global timing interval increases significantly from CV to CCV (C^{tar}: global timing $p = 0.002$ estimate = –0.17; C^{max}: global timing $p = 0.003$ estimate = –0.19; V^{max}: global timing $p < 0.0001$ estimate = –1.16), while there is no effect on the local timing interval. For /gl/, both intervals change significantly with the local timing interval showing the smallest change. From CV to CCV, the global timing interval increases (C^{tar}: $p <.0001$, estimate = –0.16; C^{max}: $p = 0.003$, estimate = –0.13; V^{max}: $p < 0.0001$ estimate = –1.08), and the local timing interval decreases (C^{tar}: $p = 0.001$, estimate = 0.07; C^{max}: $p = 0.03$, estimate = 0.08; V^{max}: $p < 0.0001$, estimate = 0.88). Although the results between /bl/ and /gl/ are quantitatively different, they do not differ qualitatively. Across voiced stop-lateral clusters, local timing interval stability is observed across speakers and across the three different anchors. Figure 2.10 plots the duration of the two intervals global timing and local timing with C^{max} as anchor for voiced and voiceless stop-lateral clusters across speakers, as a function of the number of consonants (CV versus CCV).

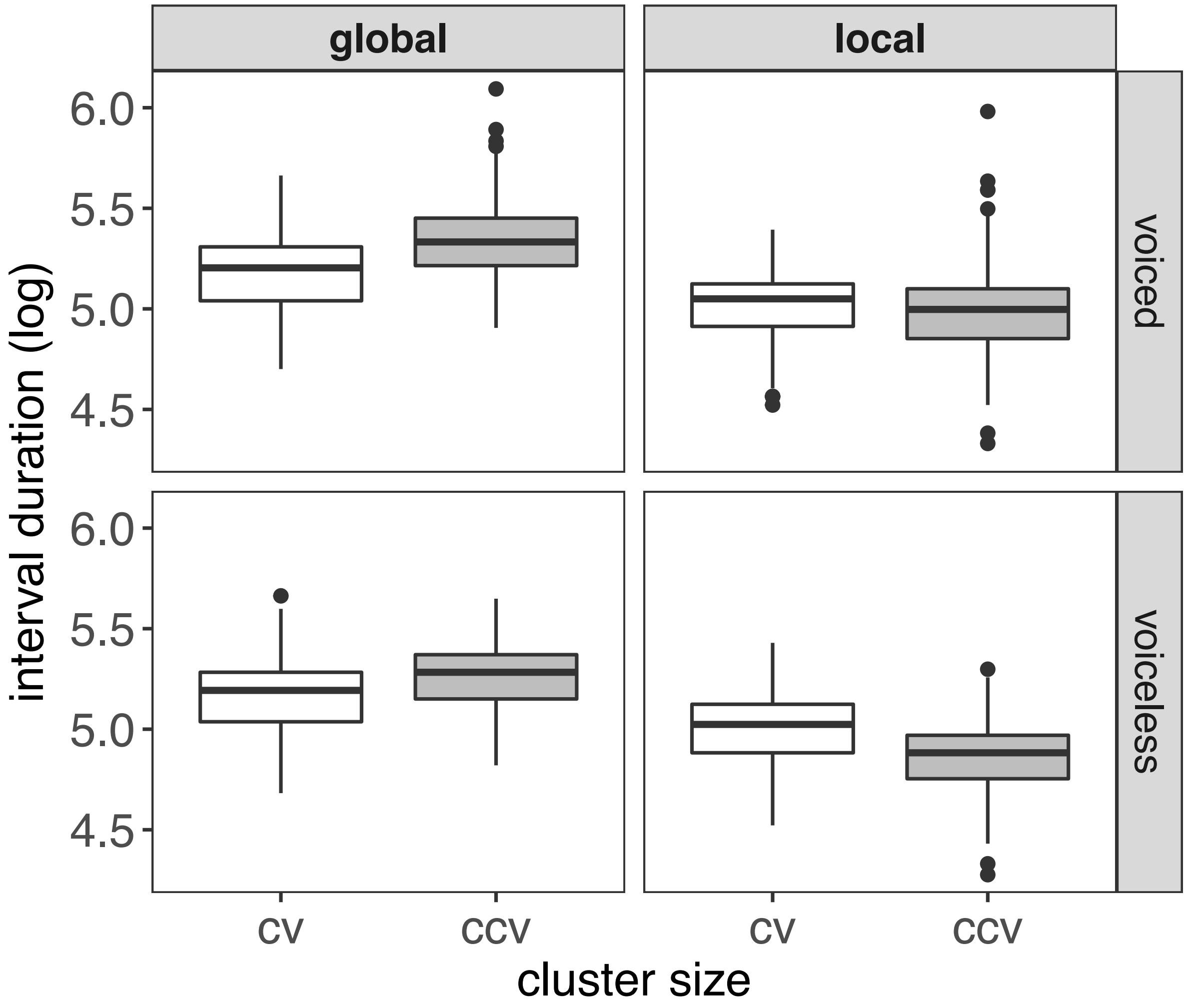

Figure 2.10: Duration (in log scale) of the two intervals global timing and local timing for CV (white) and CCV (grey) words. For voiced stop-lateral CCV words (top), the difference between CV and CCV for the global timing interval is greater than the difference for the local timing interval. For voiceless stop-lateral CCV words (bottom), the duration difference between CV and CCV for the global timing interval is smaller compared to the one for the local timing interval.

To summarize, C1 voicing affects the way intervals change as the number of consonants increases from CV to CCV. Voiceless stop-lateral clusters seem to exhibit global timing interval stability across speakers, while voiced stop-lateral clusters seem to show local timing interval stability. In the context of the profile of the patterns seen in other studies which assess interval stabilities so far, the global stability patterns seen for voiceless stop-lateral clusters in Spanish stand out. As it will be recalled, these clusters show robust open transitions. However, languages or clusters with robust open transitions or larger interplateau intervals between the two consonant constrictions seem to show strong evidence for local organization (Shaw *et al.* 2009, Shaw *et al.* 2011, Ridouane *et al.* 2014).

2.4.6.2 Addressing the spirantization issue in Spanish

A qualitative summary of the results above in 2.4.6.1 is that voiceless stop-lateral clusters exhibit global timing interval stability while voiced stop-lateral clusters exhibit local timing interval stability. In Spanish, voiced stops are generally realized as approximants except in word-initial position following a pause (Harris 1969). The target words in our experiments consist of clusters in the word-initial position. In the voiced Cl clusters (/bl/, /gl/), due to the spirantization of the initial stop, it is reasonable to hypothesize that there is more variability in the timestamp of the c-center landmark than in the case of the voiceless Cl cluster, because of variability in delimiting the C1 plateau which is the consonant that spirantizes (spirantization is not an issue in voiceless Cl clusters).

In this section, we distinguish voiced stops which are produced as such and those which are spirantized and thus produced as approximants. By doing so, we evaluate whether variability in C1 plateau duration contributes to the unexpected local timing interval stability for voiced stop-lateral clusters in Spanish. We distinguish voiced stops from approximants based on the duration of the segment. Typically, approximants are shorter in duration than stops, because their gesture has a similar nature but the approximants lack a constriction phase. Based on the distribution of the duration of C1 voiced stops across speakers and also on the fact that voiced stops in Spanish typically show duration of 70 ms (Martinez Celdrán 1993, Gibson *et al.* 2015), we decided to use the C1 plateau duration of 40 ms as a threshold. C1 voiced stops which are less than 40 ms are considered as approximants and thus excluded from the analysis for voiced stop-lateral clusters. By excluding the C1 voiced stop-l clusters whose C1 segment is produced as an approximant according to our heuristic, we wish to investigate whether improved global timing interval stability can be observed for voiced stop-lateral clusters since some of the variability linked to C1 plateau duration has been eliminated.

The same statistical procedure was followed as in the subsection above but now only for the voiced stop-liquid clusters. The analysis showed that no improved global timing interval stability is observed. Both intervals change significantly from CV to CCV with the local timing interval showing the smallest change (C^{tar}: global timing $p < 0.0001$ estimate = –0.19; local timing $p = 0.04$ estimate = 0.06; C^{max}: global timing $p < 0.0001$ estimate = –0.20, local timing p = non significant, estimate = 0.01; V^{max}: global timing $p < 0.0001$ estimate = –1.49; local timing $p = 0.005$ estimate = 0.73). Therefore, local timing interval stability is observed across voiced stop-lateral clusters.

To summarize, we attempted to eliminate some of the variability linked to C1 plateau duration which could be ascribed to spirantization of voiced stops in Spanish. This, however, did not improve the global timing interval stability as local timing interval stability was maintained throughout. We can thus exclude the possibility that the unexpected local timing interval stability for voiced stop-lateral clusters is a result of increased variability of C1 plateau duration linked to spirantization of voiced stops.

2.4.6.3 Stop-rhotic clusters

To evaluate interval stability patterns for the stop-rhotic clusters, we fitted linear mixed effects models with interval duration as a dependent variable (log transformed to better approach a normal distribution) and with interval type, cluster size, cluster type, vowel and speaker as fixed effects. The variable "word" which corresponds to each word in the corpus was used as a random factor. Results indicated a three-way interaction of interval type, cluster size and cluster type (C^{tar}: $F(4) = 21.3$, $p < 0.0001$; C^{max}: $F(4) = 16.6$, $p < 0.0001$; V^{max}: $F(4) = 3.02$, $p = 0.01$). This means that the way interval types change as the number of consonants increases from CV to CCV depends on the type of cluster (/pɾ, kɾ, tɾ/). Therefore, separate linear mixed effects models were run for each cluster type. Across anchors, for /pɾ/, there is no significant effect on the global timing and local timing intervals. From CV to CCV, the global timing interval changes slightly more than the local timing interval (C^{tar}: global timing estimate = –0.10, local timing estimate = 0.07; C^{max}: global timing estimate = –0.11, local timing estimate = 0.04), while when using a vowel anchor the global timing changes less than the local timing interval (V^{max}: global timing estimate = –0.12, local timing $p = 0.09$ estimate = 0.28). For the /kɾ/ cluster, both intervals change significantly (C^{tar}: global timing $p < 0.0001$ estimate = –0.13, local timing $p < 0.0001$ estimate = 0.10; C^{max}: global timing $p = 0.0001$ estimate = –0.12, local timing $p = 0.0007$ estimate = 0.08), with the global timing changing slightly more than the local timing interval. Using a vowel anchor, for /kɾ/, the local timing interval changes significantly (local timing $p = 0.04$ estimate = 0.25), while there is no effect on the global timing interval ($p = 0.08$ estimate = –0.20). For /tɾ/, the local timing interval changes significantly (C^{tar}: local timing $p = 0.004$ estimate = 0.26; C^{max}: local timing $p < 0.0001$ estimate = 0.25; V^{max}: local timing $p = 0.0004$ estimate = 0.63), while the global timing interval is at the limits of significance showing the smallest change (C^{tar}: $p = 0.05$ estimate = –0.12; C^{max}: $p = 0.01$ estimate = –0.06; V^{max}: p = non significant, estimate = –0.08). For none of the clusters was vowel identity (levels: low /a/, mid /e/, /o/) a significant factor. Figure 2.11 illustrates the global timing and local timing intervals with C^{max} as anchor for each of the stop-rhotic clusters across speakers.

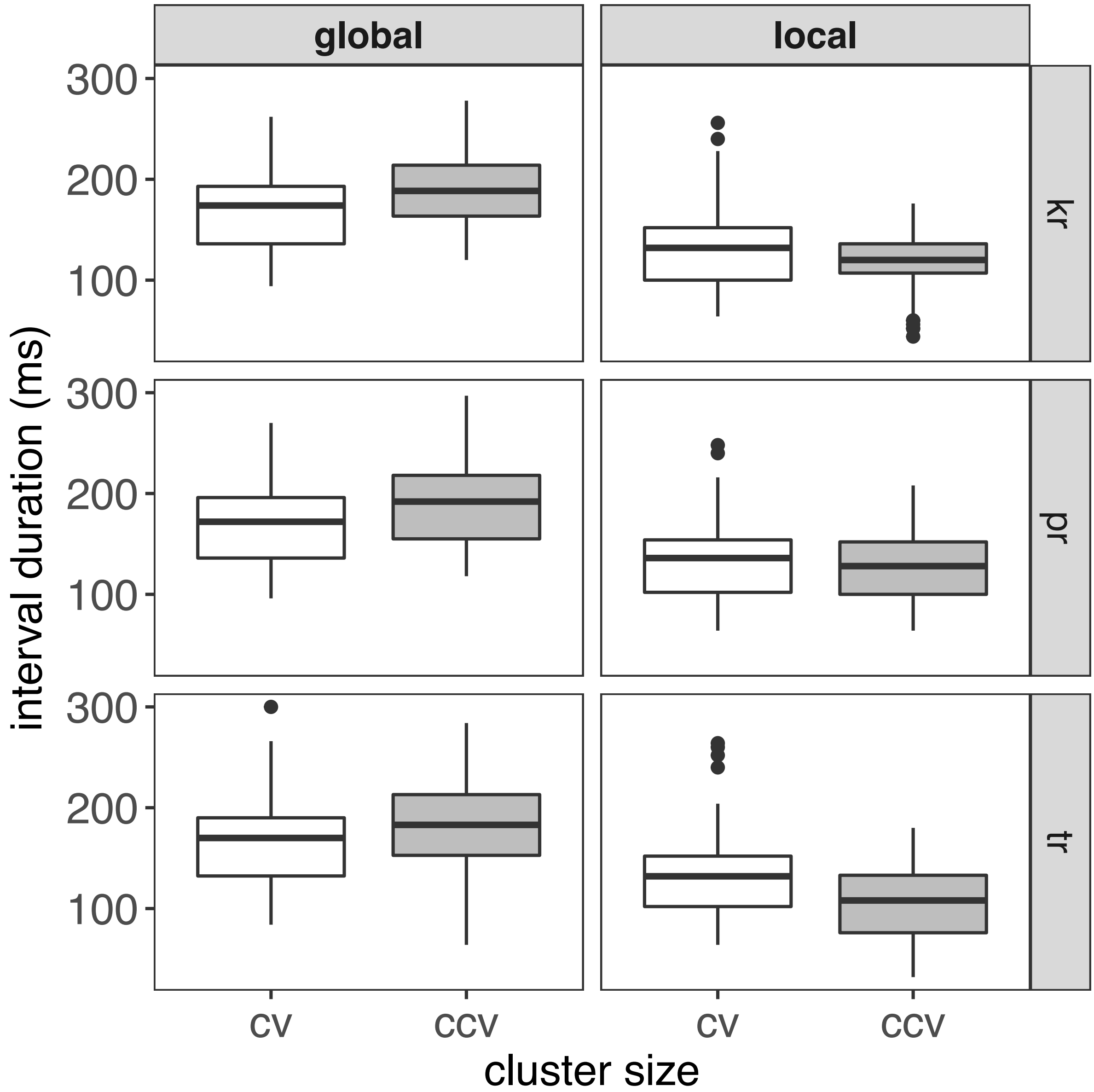

Figure 2.11: Duration (in ms) of the intervals global timing and local timing for CV (white) and CCV (grey) words of the /kɾ/, /pɾ/, /tɾ/ clusters.

Overall, across stop-rhotic clusters, the linear mixed effects model showed that both intervals change significantly with the global timing interval changing the least (C^{tar}: global timing $p = 0.001$ estimate = –0.12, local timing $p = 0.0002$ estimate = 0.14; C^{max}: global timing $p = 0.001$ estimate = –0.09, local timing $p < 0.0001$ estimate = 0.14; V^{max}: global timing $p = 0.04$ estimate = –0.16, local timing $p = 0.0001$ estimate = 0.36). In terms of stability indices of syllabic organization, the results seem to point to a global timing interval stability across speakers and across all anchors.

2.4.6.4 Relative timing in stop-lateral clusters

In this subsection, we ask how the relative timing of the lateral with the vowel changes from /lV/ to stop-/lV/ in our Spanish data. Relative timing is indexed by the lag between the release of the lateral and the maximum opening of the following vowel in the /lV/ and stop-/lV/ contexts (Goldstein *et al.* 2009, Marin 2013).

We fitted a linear mixed effects model with lag duration as a dependent variable, cluster size (CV, CCV) and speaker as fixed effects. "Item", which corresponds to each observation in our data, was treated as a random factor. Results show an interaction between cluster size and voicing of the initial stop ($X^2(2, N = 920) = 16.9$, $p = 0.0002$). When the initial stop is voiceless, the /lV/ lag decreases significantly in the stop-/lV/ context ($p < 0.0001$, estimate = 16.5 ms). When the initial consonant is voiced, the /lV/ lag also decreases significantly from CV to CCV but to a lesser extent ($p = 0.01$, estimate = 10 ms).

To sum up, the lag between the release of the lateral and the maximum opening of the vowel decreases when an initial stop – regardless of voicing – is added to the /lV/. This result is consistent with the complex onset organization. However, the extent to which the /lV/ lag decreases in the cluster context depends on the voicing of the initial stop: there is more shortening when the initial stop is voiceless than when it is voiced. Recall that voiceless stop-lateral clusters have larger IPIs than voiced stop-lateral clusters. We now see that the /lV/ lag decreases more in the former than in the latter. Such patterns of how different measures behave with respect to one another will be the focus of the forthcoming discussion section.

2.4.6.5 Vowel initiation

In this subsection, we examine vowel initiation in relation to the c-center of the cluster in different clusters of Spanish. Previous studies on the relation between syllabic organization and the temporal dimension of articulation have not quantified the vowel gesture explicitly. Vowel-specific articulatory movements tend to be longer than those of consonants and (thus) often lack clearly demarcated velocity peaks (as compared to consonants). This kinematic characteristic of vowels makes it hard for automatic algorithms to provide reliable parses of vowel gestures. Thus, previous studies have used indirect measures, such as consonant-to-consonant lag of the consonants surrounding the vowel in addressing consonant-vowel timing or vowel duration in CVC and CCVC sequences.

The current work seeks to quantify vowel initiation so as to gain an additional window into possible effects of syllabic organization on articulation. Given the qualifications above, our measurements will be based only on cases where the gestural onset of the vowel could be reliably parsed. Landmark identification for the vowel gestures was based on the tangential velocity of the tongue back (TB) positional signal. Individual cases where the parsing algorithm failed to provide a reliable parse of the vowel or where the landmarks demarcating the plateau did not correspond to the acoustic signal were discarded from the analysis. This means that a number of observations, 60 out of 977 for the stop-lateral cases, had to be discarded. However, since our corpus consists of seven different clusters /pl, bl, kl, gl, pɾ, kɾ, tɾ/ in three vowel contexts /a, e, o/ produced by six speakers, a substantial number of observations to perform an analysis was ensured even after discarding some of the data. Furthermore, we do not limit our presentation to vowel initiation analyses. Rather, we use analyses based on vowel initiation in addition to other measurements (such as the measurements of stability of intervals and CV relative timing) as a measure that has not been used before and which provides further insight on syllabic organization.

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 & Browman (1995: Figure 1, p. 552) shows the beginning of the vowel activation window to be at the c-center of the prevocalic consonantal cluster (made out of three consonants). Nam & 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 vowel initiation in any systematic way.

Figure 2.12 shows vowel initiation in C1 voiced and C1 voiceless stop-lateral clusters. Figure 2.13 shows vowel initiation in stop-rhotic clusters. The vertical lines delimited by black and white dots indicate the gestural plateaus of the stop and the lateral respectively. The triangle indicates vowel initiation and the horizontal dotted line represents the c-center landmark. As can be seen in Figure 2.12, the vowel starts 20 ms after the c-center landmark in voiced stop-lateral clusters. In voiceless stop-lateral clusters, the vowel starts 31 ms after the c-center landmark. Therefore, the vowel in voiceless stop-lateral clusters starts later relative to the c-center of the cluster than in voiced stop-lateral clusters.

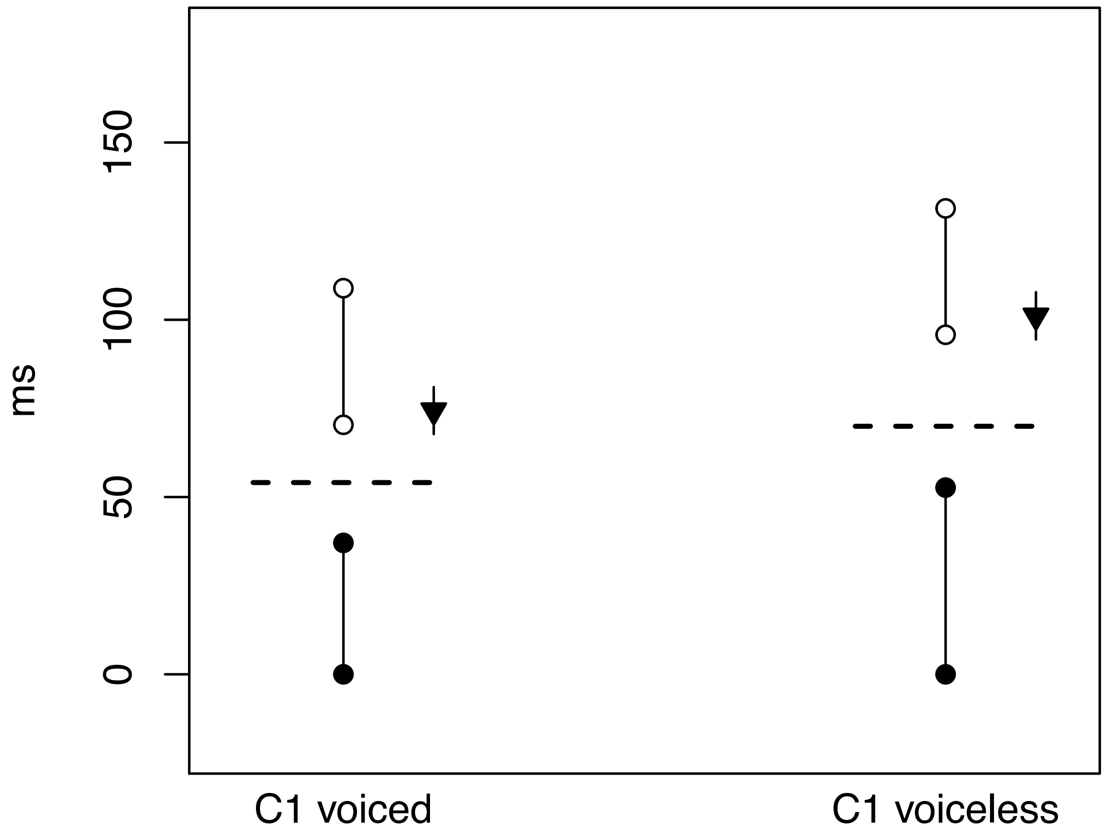

Figure 2.12: Vowel onset in relation to prevocalic consonants in C1 voiced and C1 voiceless stop-lateral clusters (x-axis). The vertical lines denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

For stop-rhotic clusters, as can be seen in Figure 2.13, the vowel starts 6 ms after the c-center of the /pɾ/ cluster, 22 ms after the c-center of the /tɾ/ cluster and 19 ms after the c-center of the /kɾ/ cluster.

In the following section, we discuss how the different vowel initiation patterns fit into a larger picture along with the stability patterns we observed for these groups of clusters. Furthermore, we discuss the different vowel initiation patterns observed for each cluster

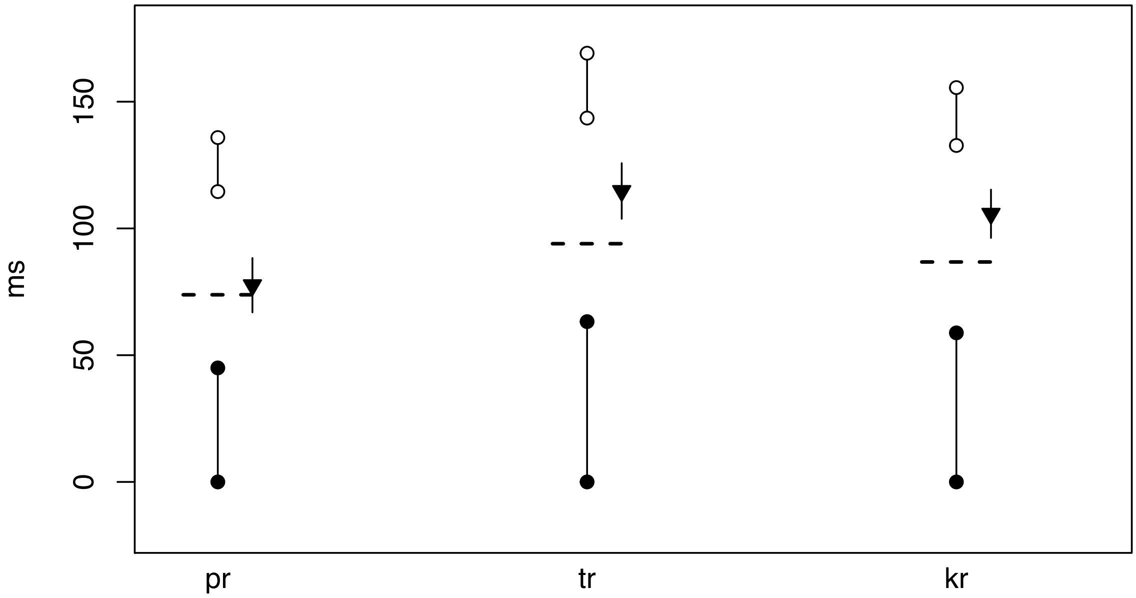

Figure 2.13: Vowel onset in relation to prevocalic consonants in stop-rhotic clusters /pɾ, tɾ, kɾ/ (x-axis). The vertical lines denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic tap consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

within the stop-rhotic category and how these can be rationalized on the basis of the degree to which the different (lingual or non-lingual) segments in these clusters allow the tongue to engage in vowel-specific movement.

2.5 Discussion

According to the stability-based indices of syllabic organization reviewed in section 2.2 (Background), complex onsets are expected to show global timing interval stability. Failure to verify this prediction in what are presumably complex onsets in Spanish may be interpreted as evidence that syllable structure does not have consistent phonetic manifestations in the articulatory record. An alternative is that the relation between syllable structure and phonetic properties is more nuanced. That alternative would predict that even in cases where the global timing interval stability cannot be found, there are other phonetic consequences that reflect a global organization of the whole sequence of segments that partake in the syllable. In what follows, we flesh out what is meant by global organization more specifically for each type of cluster we have examined, namely, voiced stop-lateral, voiceless stop-lateral and voiceless stop-rhotic clusters. Of particular interest is to identify the ways in which, in these different contexts, the individual segments combined in forming a presumed syllable onset – nucleus whole readjust their properties (as compared to their properties when these segments are not combined) so as to achieve overlap or closer proximity with the nucleus vowel. We will see that the presence as well as the degree to which these readjustments take place depend on the coincidental properties of the individual segments in subtle but systematic ways which in turn unveil the presence of an overarching global organization. For example, in accordance with cross-linguistic patterning, voiceless stops are longer than voiced stops, in Spanish. Because of the longer duration of the voiceless stop, when such stops are combined with laterals in forming complex stop-lateral onsets, the lateral shortens more in the voiceless stop context than in the voiced stop context and the relative timing between lateral and the subsequent vowel changes more in the voiceless stop-lateral context than in the voiced stop-lateral context. These readjustments and specifically the degree to which they are found seem to be motivated by constraints related to the global organization in the syllable.

2.5.1 Stop-lateral clusters

We begin by summarizing our results for the voiceless and voiced stop-lateral clusters and we discuss how cluster-specific properties account for the different temporal patterns observed across these two categories. For voiceless stop-lateral clusters, the duration of the prevocalic lateral is shorter, the IPI is greater, and the plateau duration of the initial consonant is greater than in voiced stop-lateral clusters. Under these conditions, global timing interval stability was observed for voiceless stop-lateral clusters, while local timing interval stability was observed for voiced stop-lateral clusters. The vowel in the former was found to start approx. 30 ms after the c-center of the cluster, while the vowel in the latter was found to start approx. 20 ms after the c-center of the cluster; i.e., in both groups of clusters, the vowel starts close to the c-center of the clusters to the extent permitted by properties of the individual segments in the cluster.

The schema in Figure 2.14 illustrates the temporal adjustments that occur when an initial voiced or voiceless stop is added to the lateral-vowel string according to the data. When a voiceless stop, with large duration and a large IPI, is added to the lateral-vowel string, the lateral shortens and it is "pushed" towards the vowel. Compare the single lateral /l/ in Figure 2.14a versus 2.14c when it is clustered. The lateral in Figure 2.14c is shorter and it occurs later than the single lateral in a. Across the CV and CCV (refer to Figure 2.14a and c), the c-center (indicated as cc in the Figure) landmarks in /lV/ and voiceless /ClV/ are better aligned (as highlighted by the red ellipsis) than the release (indicated as re in the Figure) landmarks. In effect, for the CV, CCV pairs where the first

C is voiceless, the local timing interval is shorter in CCV than in CV, while the global timing interval is comparable across CV, CCV. Compare this configuration to that for CV, CCV pairs when the first C is voiced. When a voiced stop with shorter duration and a shorter IPI than a voiceless stop is added to the lateral-vowel string, what can be observed throughout the data is that the lateral gets shorter but minimally "pushed" towards the vowel: compare the single lateral /l/ in Figure 2.14a versus 2.14b when it is clustered with a voiced stop. The lateral in Figure 2.14b shortens to some extent, but it does not shift later compared to the single lateral in Figure 2.14a. Thus, shortening of the prevocalic lateral, as seen in section 2.4.3, with some minimal rightward shift towards the vowel does not result in misalignment of the re landmarks between CV and CCV. Instead, the re landmarks remain better aligned (as highlighted by the red ellipsis) than the cc landmarks between CV and CCV, thus resulting in stability of the local timing interval.

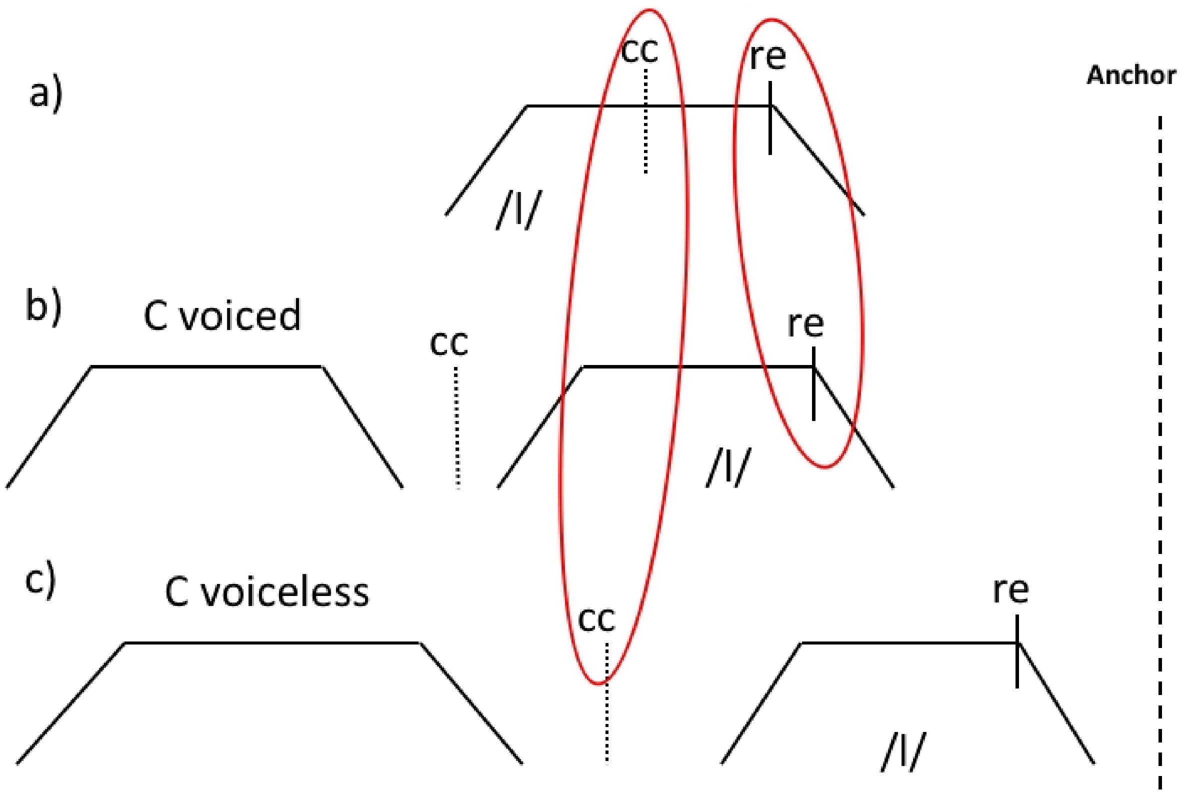

Figure 2.14: Temporal configurations of (a) singleton lateral /l/, (b) voiced stop-/l/ cluster and (c) voiceless stop-/l/ cluster. The trapezoids represent consonant gestures. The landmark marked as cc corresponds to the c-center and re to the right-edge of the singleton consonant (C) or the consonant cluster (CC). Across the singleton lateral-vowel sequence /lV/ and the cluster (stop-/lV/), the re landmarks are better aligned than the cc landmarks in /l/ - voiced C/l/ (a, b). The cc landmarks are better aligned than the re landmarks in /l/ - voiceless C/l/ (a, c). Ellipses highlight pairs of landmarks that remain relatively well aligned across the two contexts CV and CCV.

It is thus clear that phonetic properties, such as IPI and consonant duration, affect temporal patterns in ways that both local timing and global timing interval stability can be observed for complex onsets. That is, voiced stop-lateral clusters exhibit stability of the local timing interval while voiceless stop-lateral clusters exhibit stability of the global timing interval. Yet, these two cluster groups share a general property that seems to make reference to a global organizing principle: when the stop is added in front of the lateral to form a stop-lateral cluster, a number of readjustments, such as lateral shortening and change of relative timing between /l/ and the vowel, take place over the resulting CCV sequence. The nature of these readjustments indicate that the spatio-temporal structure of the individual segments adjusts, to the extent permitted by the segment-specific properties, so as to increase the overlap among the vowel and the prevocalic cluster. Furthermore, there is a compensatory relation between IPI and C2 lateral duration across stop-lateral clusters. This relation too seems to indicate the presence of global organization over the entire segmental sequence as we explain next.

Specifically, in the voiceless stop-lateral context, the phonetic properties of the segments involved – large C1 voiceless stop (51 ms), large IPI (48 ms) and a lateral, which is a segment with substantial duration (52 ms in a CV context as compared to a tap which is of the order of 25 ms) – conspire to create unfavorable conditions for establishing an organization where the vowel should overlap with the cluster. It is precisely in these conditions where we find that the lateral shortens substantially (compared to its singleton version) and the relative timing of the lateral with the vowel changes such that there is increasing overlap between the plateaus of the lateral and the vowel. These readjustments bring the vowel to start 30 ms after the c-center of the cluster in voiceless stop-lateral clusters. For the voiced stop-lateral clusters, both the lateral shortening and the change in the relative timing between the lateral and the vowel from CV to CCV are substantially smaller than in voiceless stop-lateral clusters. The minimal readjustments in voiced stop-lateral clusters already bring the vowel to start 20 ms after the c-center of the cluster.

There is no purely phonetic reason that would explain both why the lateral in the voiceless stop context shortens more than in the voiced stop context and why the relative timing between the lateral and the vowel changes more in the voiceless stop context than in the voiced stop context. These spatio-temporal readjustments in the voiceless stop-lateral context seem to be motivated by syllable structure, so that the vowel can overlap with the prevocalic two-consonant cluster. In the voiced stop-lateral context, the above readjustments are found to a lesser extent. In this context, the duration of the voiced stop is shorter than its corresponding voiceless counterpart, the IPI is shorter in voiced stop-lateral than in the voiceless stop-lateral clusters and the lateral shortens to some extent (but less than in voiceless stop-lateral clusters). Regardless of the different temporal readjustments that occur between voiced and voiceless stop-lateral clusters, there is also a compensatory relation between IPI and C2 lateral duration which holds true across the two cluster types. This relation too points to the presence of a global organization scheme that serves to bring the vowel in proximity with the cluster: increasing the lag between the constrictions of the two prevocalic consonants (those of the stop and the lateral) implies shortening of the second consonant (the lateral). If each of these, in principle, independent parts of the phonetic substance of a stop-lateral cluster were timed independently of one another, this compensatory relation would not be expected.

To sum up, we have seen that when looking at the commonly used stability-based heuristics, the results on Spanish stop-lateral clusters are rather confusing. The voiceless stop-lateral clusters showed the expected interval stability characteristic of complex onsets, while the voiced stop-lateral clusters showed the interval stability characteristic of simplex onsets. However, when looking beyond the stability-based heuristics, we see effects related to global organization for both voiced and voiceless stop-lateral clusters in a number of phonetic indices. Shortening of the prevocalic lateral from CV to CCV, change of the relative timing of the lateral and the vowel from CV to CCV, early vowel initiation to the extent permitted by segment-specific phonetic properties in the cluster, and the IPI-C2 compensatory relation all express syllabic organization in that they serve to bring the vowel to overlap with the cluster as prescribed by global organization.

2.5.2 Stop-rhotic clusters

In comparing voiceless stop-rhotic clusters to (voiceless) stop-lateral clusters, we find that IPI is considerably larger in the former than the latter (74 ms versus 48 msec) and yet earlier vowel initiation relative to the c-center of the cluster is observed in the former than the latter. This difference seems to be linked to the fact that the rhotic is substantially shorter than the lateral and thus allows for more overlap with the vowel without any segment being perceptually masked. Figure 2.15 illustrates the temporal readjustments that occur from /rV/ to /CɾV/. Compare the trill in the /rV/ context and the tap in the /CɾV/ context in Figure 2.15a and Figure 2.15b respectively. The tap is shorter and it occurs later with respect to the trill. Therefore, the prevocalic consonant is shorter in /CɾV/ than in /rV/ and is shifted towards the vowel in the cluster case. The re landmarks are thus misaligned and the cc landmarks are better aligned, as highlighted by the red ellipsis, giving rise to the patterns we have observed in the stability analysis (section 2.4.6.3) with the global timing interval remaining stable from CV to CCV.

For stop-rhotic clusters, we find no compensatory relation between IPI and prevocalic

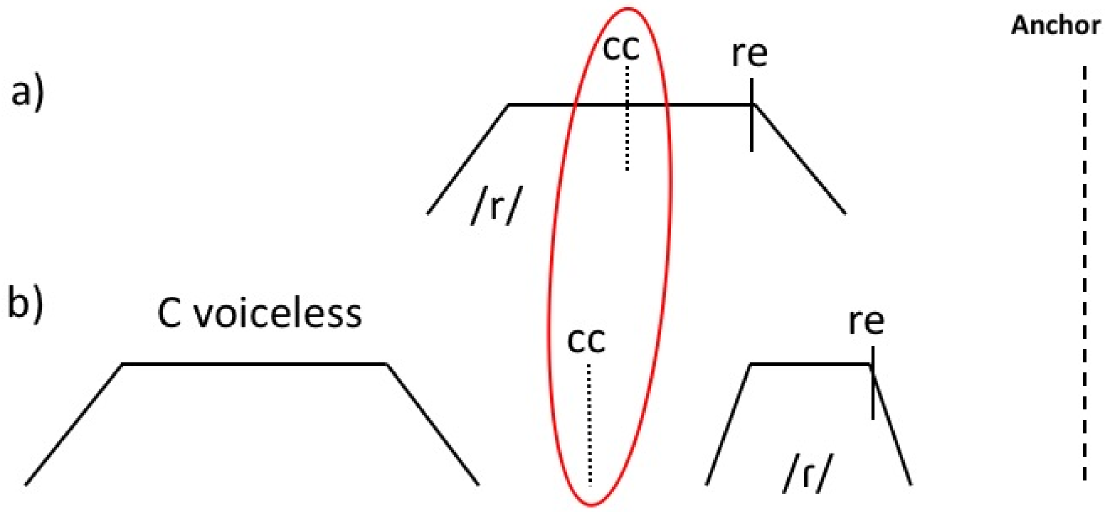

Figure 2.15: Temporal configurations of a singleton trill /r/ in (a) and a voiceless stop-/ɾ/ cluster in (b). The trapezoids represent consonant gestures. The mnemonic cc corresponds to the c-center landmark and re to the right-edge landmark. The cc landmarks are better aligned than the re landmarks across /rV/, shown in (a), and stop-/ɾV/, shown in (b). Ellipses highlight pairs of landmarks that remain relatively well aligned across the two contexts CV and CCV.

consonant duration in CCV. One reason for this may be that taps (the rhotics in second position of the relevant /CɾV/ clusters) are so-called momentary articulations (Catford 1977:130). The definitional characteristic of such momentary articulations (as opposed to maintainable articulations of, for example, fricatives) is that they have a very short duration (around 25 ms), which can neither be stretched in time (amounting to undoing their momentary nature) nor reduced further (amounting to eliminating the tap altogether). The latter property (neither lengthening, nor shortening) precludes or at least makes it very unlikely verifying a compensatory relation involving the duration of this element because this duration must exhibit a wide range of values for such a relation to be reliably detectable. We have called upon this compensatory relation as one indication for a global organization in Spanish complex onsets as it attests to an overarching constraint that the vowel overlaps with the cluster. However, in stop-rhotic clusters, we find other equally, if not more, telling indications of global organization.[3] In particular, as we have seen in

[3]Shortening of the prevocalic consonant from CV to CCV and change in the relative timing between rhotic and vowel from CV to CCV could be used as a further source of evidence for syllabic organization but will not be used so here. The reason is that the rhotic is a trill in CV but a tap in CCV. That is, the manner of articulation (Catford 1977) changes between the occurrences of the rhotic between these two different contexts. The phonetic consequences of this change in manner of articulation are such that, in accordance with cross-linguistic patterning, the tap is by its nature a much shorter segment than the trill. No such change in manner of articulation takes place for the /l/ in /lo/ - /glo/. The relevant point is that because of this change in manner, the shortening seen from trill to tap cannot be ascribed

Figure 2.13, the vowel is found to start quite close to the c-center of the cluster. Furthermore, any minor differences we observe among the stop-rhotic clusters with respect to vowel initiation or IPI, can be rationalized on the basis of properties of the segments in the different clusters. Thus, we have seen that the vowel starts earlier in /pɾ/ than in /kɾ, tɾ/. For the /pɾ/ cluster, the earlier vowel initiation compared to /kɾ, tɾ/ can be attributed to the fact that /pɾ/ cluster consists of a sequence of segments that implicate a labial (thus, non-lingual) and a lingual gesture with the former not placing any demands on the tongue, while all segments in /kɾ, tɾ/ implicate lingual gestures thus preventing vowel initiation as early as in the labial-rhotic case. The /pɾ/ cluster, in comparison to the /kɾ, tɾ/ clusters, constrains the tongue less, hence the earlier vowel initiation in /pɾ/ than in /kɾ, tɾ/. Furthermore, since /k, t/ both implicate the tongue but /p/ does not, vowel-specific movement for the /ɾ/ in /kɾ, tɾ/ cannot begin before the /k, t/ constrictions are completed whereas no such constraint is present in /pɾ/. This then explains why IPI in /pɾ/ is shorter than in /kɾ, tɾ/.

To sum up, voiceless stop-rhotic clusters in Spanish provide evidence for global organization in that the vowel starts close to the c-center of the cluster to the extent permitted by the properties of the segments in the cluster. Vowel initiation in stop-rhotic clusters constitutes the only reliable way to assess syllabic organization because it considers only the cluster context. Any other measure which may be used to assess syllabic organization by combining both the CV and the CCV context in stop-rhotic clusters, such as stability-based heuristics, CV relative timing across CV∼CCV, duration of the prevocalic consonant in CV and CCV, would be misleading because the prevocalic consonant in the CV and CCV contexts is a different segment – a trill in CV and a tap in CCV. These two segments, the trill and the tap, have inherently different duration and they thus affect the calculation of intervals which include the duration of the prevocalic consonant. For instance, although the stability-based heuristics (in section 2.4.6.3) indicate global timing interval stability for voiceless stop-rhotic clusters, we cannot rely on such a measure because the calculation of the global timing interval includes the duration of the prevocalic consonant. Furthermore, even though the IPI-C2 relation can be assessed only in the cluster context, such a relation was not present in stop-rhotic clusters. This is likely due to the nature of the tap, being a very short segment (around 25 ms). As such, it does not exhibit sufficient variability in duration which would allow for a compensatory relation to be detected. Due to all of the above points, the vowel initiation measure is the only reliable index for assessing syllabic organization. This measure is also in fact a rather

exclusively to the imposition of an overarching constraint on global organization; for example, there may be biomechanical and or aerodynamic reasons why the trill changes to a tap when clustered (see Rennicke 2015, Blecua Falgueras 2001, Barry 1997 for relevant discussion), although we also cannot exclude syllabic organization as a factor in the presence of this alternation in Spanish onsets.

direct way of assessing syllabic organization, because as the reader may recall from our literature review (Chapter 1) a key point in the landmark work of Browman & Goldstein (1988) concerned the vowel starting early in the consonant cluster (that is claimed to be a complex onset).

2.5.3 Summary

Overall, then, Spanish provides robust evidence that complex onsets do not always show the expected global timing interval stability as in the case of stop-lateral clusters. Moreover, it is clear that the degree to which segments change from CV to CCV affects the temporal coordination of the cluster with the vowel in ways that preclude a unique phonetic exponent of the relation between the hypothesized syllabic structure and timing patterns. For example, a small degree of lateral shortening is likely to result in local timing interval stability as in voiced stop-lateral clusters while more lateral shortening is likely to result in global timing interval stability as in voiceless stop-lateral clusters. However, across all clusters (voiced and voiceless stop-lateral clusters as well as voiceless stop-rhotic), in examining the different phonetic properties and their effects on the spatio-temporal organization of the whole CCV, we find evidence for spatio-temporal readjustments which seem to reflect demands of an overarching organization; an organization over the entire CCV which requires its constituent segments to be tightly bound together and specifically the vowel to overlap as much as possible with the cluster.

2.6 Comparison with Moroccan Arabic

In a general form, we can summarize the results of our preceding section by saying that adding a consonant to the left of a CV to obtain a CCV results in reorganization of the spatio-temporal structure of the entire CCV in Spanish. Our interpretation is that this reorganization is related to the fact that the examined CCV sequences are syllable onsets in Spanish. In turn, this interpretation predicts that in a language where the same consonant clusters are not admitted as syllable onsets, the signatures of global organization we have observed in Spanish should not be found. Here we pursue this prediction with data from Moroccan Arabic, a language which is claimed to disallow complex clusters as syllable onsets (Dell & Elmedlaoui 2002, pp. 252–253; Boudlal 2001, Kiparsky 2003).

Let us first establish why a comparison between Spanish and Moroccan Arabic would be particularly appropriate over and above the fact that the two languages are claimed to treat complex clusters differently in terms of syllabic organization. The main point is that the phonetic profile of consonant clusters in the two languages is very similar in terms of two crucial properties. The first property concerns inter-segmental coordination. The second concerns the implementation of the voicing contrast in the two languages. The combination of these two properties renders the phonetics of the relevant clusters very similar across the two languages. What makes the Spanish, Moroccan Arabic comparison especially appropriate is that the surface phonetics (in terms of these two properties) are very similar, but the claimed phonological organization of the relevant clusters when combined with a vowel is very different.

Consider first inter-segmental coordination which we quantify here by the duration of the interplateau interval between the two segments in CC sequence. In our sample from four Moroccan Arabic and six Spanish (Central Peninsular Spanish) speakers, for any given cluster C1C2, we normalize the interplateau interval (IPI) by the total duration of the cluster as measured from target of C1 to release of C2. The so normalized IPI durations for the two languages are 0.35 ($sd = 0.04$) for Moroccan Arabic and 0.34 ($sd = 0.15$) for Spanish.

Consider next voicing. The implementation of the voicing contrast in the two languages is comparable. In Spanish as well as in Moroccan Arabic, the /g/ or voiced element of the /g/-/k/ voicing opposition is physiologically implemented by the so-called prevoicing, that is, vocal fold vibration which is (at least partly) coextensive with the articulatory plateau of the velar constriction. The /k/ or voiceless part of the opposition in Spanish as well as in is Moroccan Arabic is produced with an open glottis and thus no vocal fold vibration during its constriction phase. Upon release of the velar constriction in /k/, a short burst of air escapes and is followed by a period of approximately 20 miliseconds of airflow or aspiration, after which modal vocal fold vibration can begin again. Voiceless stops, in other words, are of the short lag VOT type in both languages.

For concreteness, consider how the above two phonetic properties would have been if we had used English or German, for which relevant articulatory recordings are also available, instead of Spanish. The voiced /b/ in German shows a similar physiological behavior to the voiceless /p/ of Spanish, in that German /b/ is characterized by a short burst of air followed a period of aspiration (airflow escape). German /p/, however, is characterized by a longer period of aspiration that lasts approximately 60 miliseconds (Bombien & Hoole 2013) following the release of the bilabial constriction. Aspiration in a typical German voiceless /p/ is thus substantially longer (hence, long lag VOT) than in a typical Spanish /p/. German and English differ from Moroccan Arabic not only in that their voicing systems are different but also in that they do not show robust open transitions in at least voiced stop-lateral clusters.[4] Thus, in German voiced stop-lateral

[4]See Katz (2012) for the potentially confounding role of VOT regarding certain clusters in previous

clusters, the constriction plateaus of the two gestures either overlap in time or are very close to overlapping, a temporal configuration that does not exist in Spanish which instead exhibits a robust interplateau interval between the two consonants as shown earlier. On the surface, then, across Moroccan Arabic and Spanish, the profile of, for example, a stop-lateral sequence is comparable.

After having established why it is appropriate to compare Spanish and Moroccan Arabic, we now turn to the spatio-temporal reorganization effects we have observed in Spanish in order to assess their presence in Moroccan Arabic. As we have seen, in Spanish even though the global timing interval is not consistently the most stable interval, we find evidence for adjustments in the segments that partake in the Spanish onset which make reference to a global organization. These readjustments consist of shortening of the prevocalic lateral, relocation of the lateral with respect to the vowel, and a compensatory relation between IPI and lateral duration. The above readjustments and the extent to which they occur in different clusters (e.g., across voiced stop-lateral versus voiceless stop-lateral) serve to bring the vowel to overlap with the cluster to the extent permitted by properties of the individual segments in the relevant clusters. For each of these effects, we now turn to consider the evidence from Moroccan Arabic.

One first indication of the reorganization that takes place in Spanish when a stop is added in front of a lateral is the shortening of the lateral as observed in our data. Specifically, we have seen that the lateral shortens 12 ms in the cluster overall regardless of the initial stop's voicing. In assessing lateral adjustments in Moroccan Arabic, the duration of the single lateral in the /lam/ ($N = 36$) was compared with the lateral duration in the /blat/ ($N = 16$), /klat/ ($N = 18$) and /klam/ ($N = 26$) in our available data from four speakers. The duration of the lateral increases by 3 ms in the cluster compared to the singleton context. This change in duration is minimal and in the opposite direction from that of Spanish. Table 2.9 provides measures of lateral duration in CV and CCV in the available Moroccan Arabic contexts. Thus, there is no shortening of the lateral from CV to CCV in Moroccan Arabic.

Consider now measures that reflect the reorganization of the temporal structure of the CV internal to a CCV. Here effectively the only appropriate comparison between the two languages must be restricted to /lV/∼/klV/. This is so because of two reasons. First, Spanish does not show evidence for global timing interval stability for voiced consonant-lateral clusters such as /lV/∼/glV/ or /lV/∼/blV/. Second, Moroccan Arabic's phoneme inventory does not include the voiceless bilabial stop and hence there are no /pl/ clusters to be compared to corresponding clusters in Spanish. Our available data for Moroccan

works on the relation between syllabic organization and timing in English. These two properties, intergestural timing and VOT, are related (see Bombien& Hoole 2013).

Table 2.9: Lateral duration in CV and in second position of a CCV cluster in Moroccan Arabic.

C2 lateral duration (ms) in CV and CCV contexts		
	Mean	*sd*
CV	36.1	11.8
CCV	39.4	9.6
bl	44.6	9.2
kl	37.5	9.1

Arabic consist in 36 /lam/ words and 26 /klam/ words from two speakers, given that this is an earlier corpus not designed for the purpose of this study. In Moroccan Arabic, when /k/ is added to a word-initial /lV/, the local timing interval (with anchor being the maximum opening of the following vowel, as in our Spanish measurements), decreases from 100 ms to 93 ms, while the global timing interval increases from 119 to 165 ms. Thus, local timing changes to a lesser extent than global timing. This is, again, the opposite effect of what was observed in Spanish. In Spanish, from /lV/ to /klV/, the local timing interval was found to change to a greater extent (from 84 to 54 ms) than the global timing interval (from 111 to 119 ms). In terms of interval changes, then, Moroccan Arabic and Spanish are mirror images of one another: from CV to CCV, the global timing interval changes more than the local timing interval in Moroccan Arabic but in Spanish the opposite is true. Figure 2.16 illustrates the intervals global timing and local timing for /lV/∼/klV/ in Moroccan Arabic (MA) and Spanish (SP).

In terms of vowel initiation, the vowel in Moroccan Arabic /bl/ and /kl/ is found to start at 0.3 ms and 6 ms after the release of the lateral respectively. Figure 2.17 shows vowel initiation in /bl/ and /kl/ clusters in Moroccan Arabic. The vertical lines delimited by black and white dots indicate the gestural plateaus of the stop and the lateral respectively. The triangle indicates vowel initiation. It can be seen that, in contrast to the Spanish patterns, the vowel in Moroccan Arabic starts close to the release of its immediately preceding consonant. This is of course consistent with other studies on Moroccan Arabic reporting the vowel being tightly organized with respect to its immediately preceding consonant only (Shaw *et al.* 2009, 2011, Shaw & Gafos 2015, Gafos *et al.* 2018).

With respect to the relation between IPI and lateral duration, in Spanish, a compensatory effect was observed such that as IPI increases, the duration of the lateral decreases. For the Moroccan Arabic stop-lateral tokens /blat/ ($N = 16$), /klat/ ($N = 18$), /klam/ ($N = 26$) of four speakers, we find no linear relationship between IPI and the duration of the lateral when using raw values of the variables ($r(58) = 0.07$, $p = 0.59$). When using normalized values, there is a weak negative correlation ($r(58) = –0.37$, $p = 0.003$), such that as IPI increases, the duration of the lateral tends to decrease. At any rate,

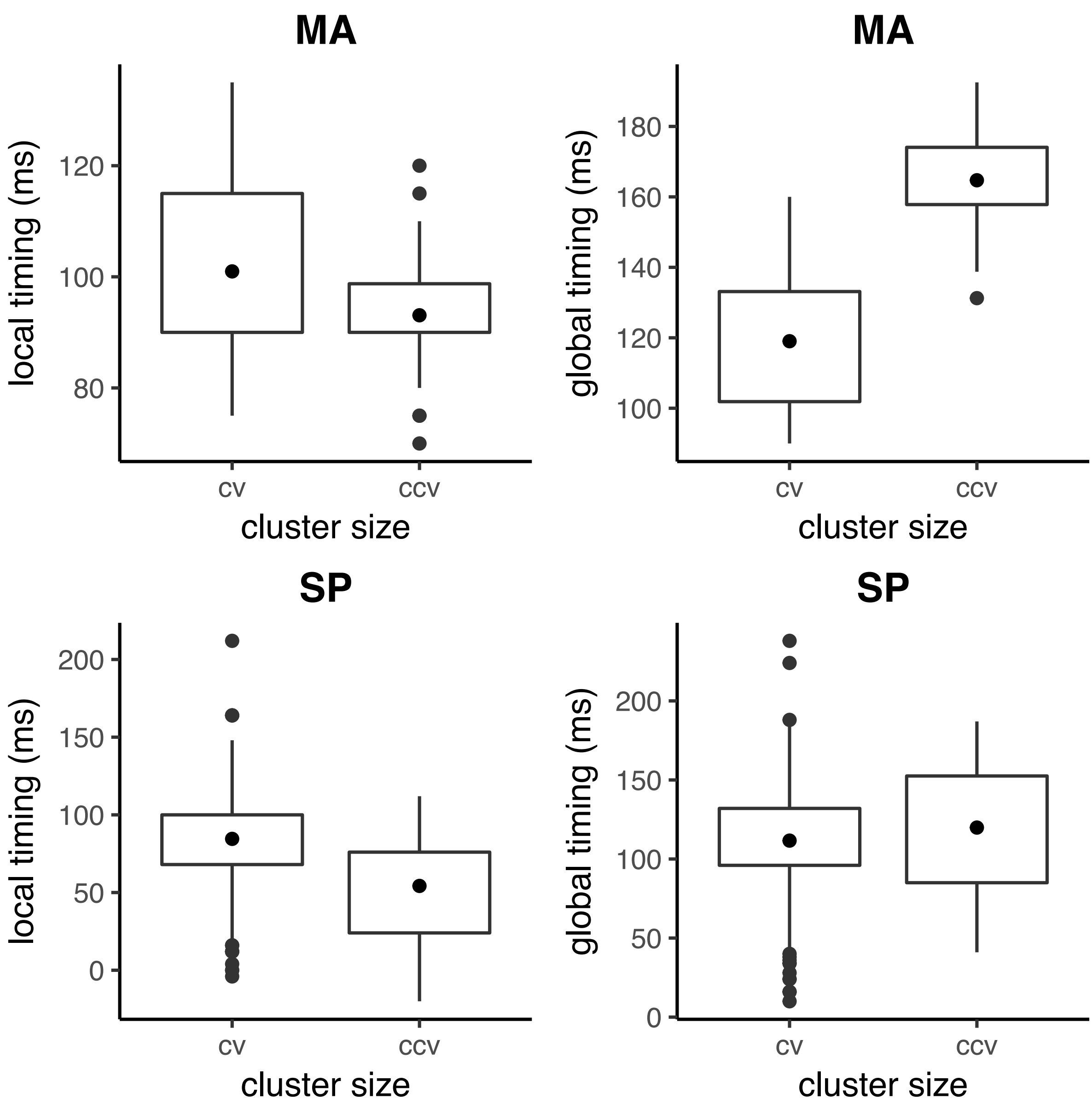

Figure 2.16: Duration (in ms) of the intervals global timing, local timing for CV (/lV/) and CCV (/klV/) words in Moroccan Arabic (MA) upper row and in Spanish (SP) lower row. The difference between CV and CCV of the local timing interval is minimal compared to the difference of the global timing interval in MA, while in SP the opposite pattern is observed.

the negative correlation is strong in Spanish while it is at best weak in Moroccan Arabic. Figure 2.18 and Figure 2.19 show the relation between IPI and C2 lateral duration (raw and normalized values respectively).

Overall, adding a segment to a sequence of segments in Spanish produces effects that ripple through the inner sequence in Spanish but not so or less so in Moroccan Arabic. We can express this patterning by using a notion of span of organization. A schema

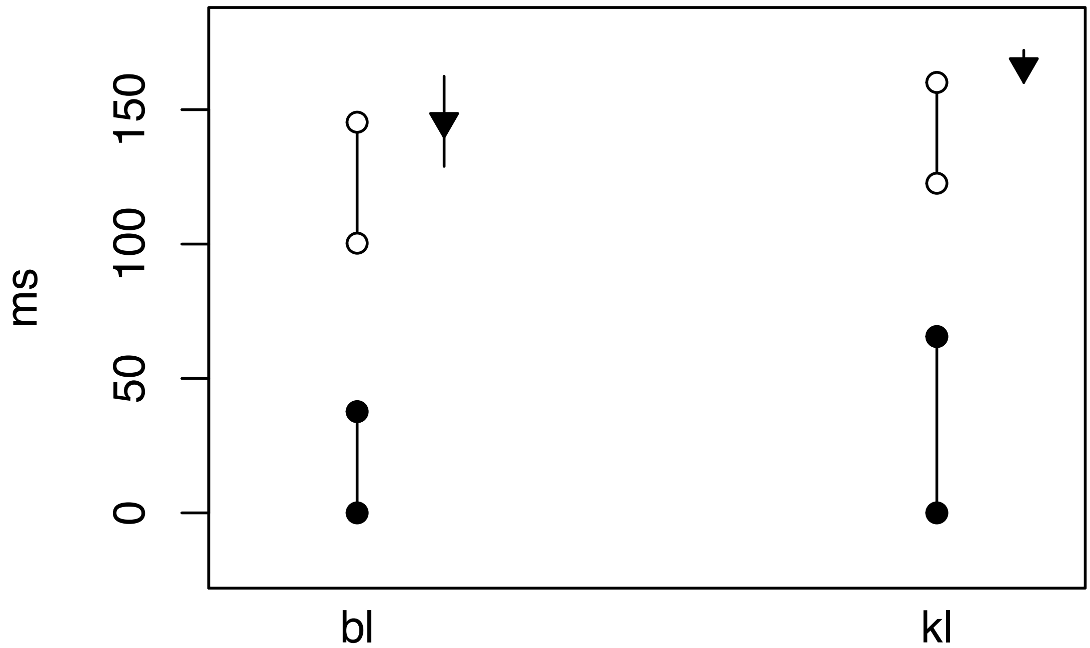

Figure 2.17: Vowel onset in relation to prevocalic consonants in /bl/ and /kl/ clusters (x-axis) in Moroccan Arabic. The vertical lines shown denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant (here, /b/ and /k/). Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants.

is in Figure 2.20 (left). In a CCV sequence, global organization ranges over the entire segmental complex whereas local organization ranges over the inner CV subsequence only. Metrification is global in the former case but local in the latter. Spanish represents the former, global case whereas Moroccan Arabic represents the latter local case. A partial indication of the ripple effects that adding a segment to a sequence of segments in Spanish is shown in Figure 2.20 (right, top half). Therein, schematized are the changes in the CV subsequence of a CCV that take place when an extra consonant is added in front of that CV. The duration of the second C shortens and the timing relation between that C and the vowel changes to accommodate the addition of the new consonant. Both these effects are absent in Figure 2.20 (right, bottom half) which schematizes the case of Moroccan Arabic. Finally, we have characterized the effects shown in Figure 2.20 (right) as a partial, as far as effects that derive from global organization, because, as we have argued elsewhere, the precise effects and how they interact with one another depend on the properties of the segments involved in each case. The presence of the different effects and the relative extent to which they are attested in the different clusters are the telling indicators of the global mode of organization in our Spanish data.

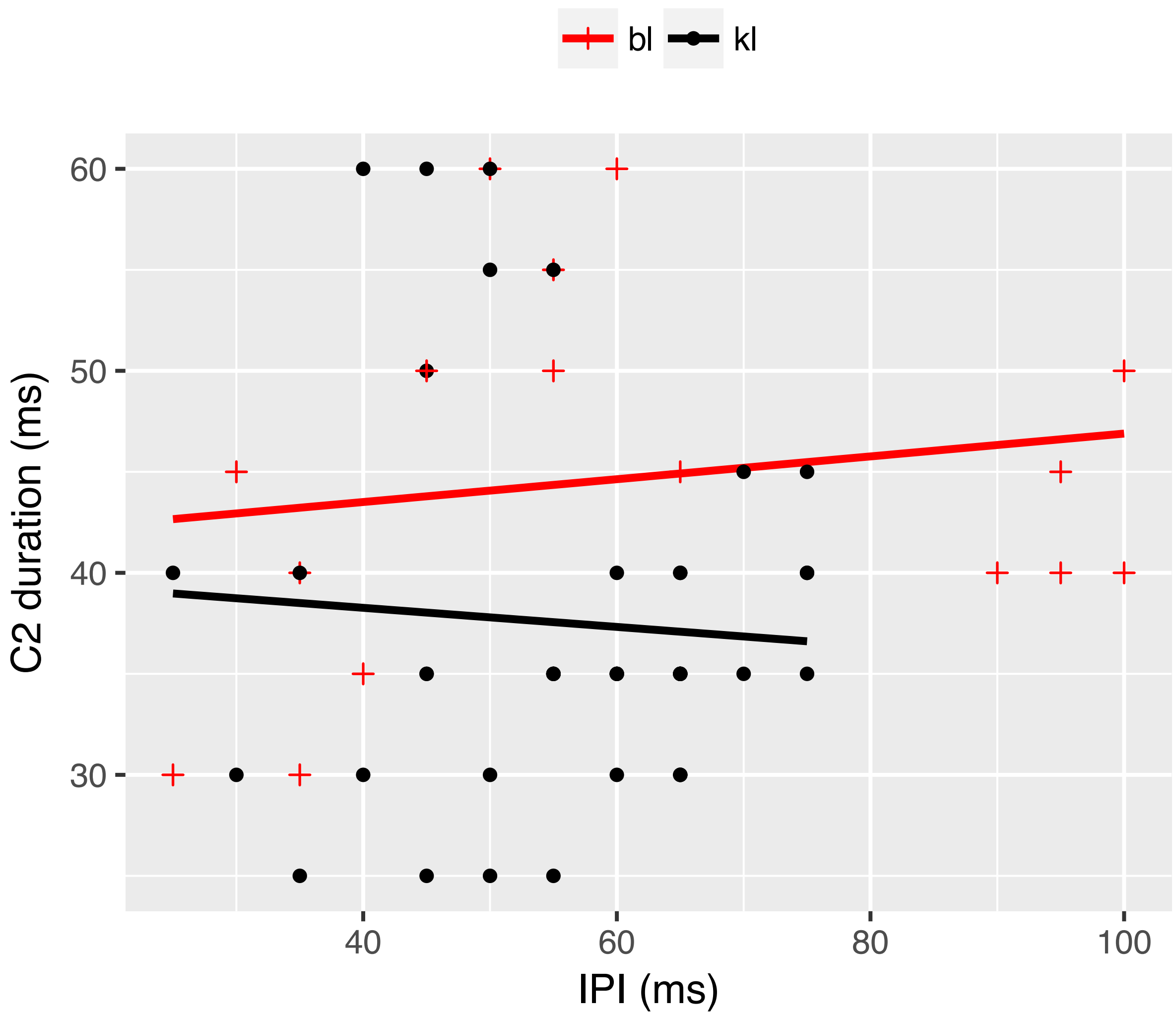

Figure 2.18: Scatterplot showing no relation between C2 lateral duration and IPI for the /bl/ and /kl/ clusters in Moroccan Arabic ($r(58) = 0.07$, $p = 0.59$).

2.7 Conclusion

This Chapter has addressed the relation between language-specific syllable structure and inter-segmental spatio-temporal coordination. Specifically, we studied the spatio-temporal properties of the clusters /pl, bl, kl, gl, pɾ, kɾ, tɾ/ in three vowel contexts /a, e, o/ produced by six speakers of Central Peninsular Spanish. We furthermore compared spatio-temporal properties of selected Central Peninsular Spanish syllable onset clusters with similar clusters of Moroccan Arabic, a language where similar clusters are found but are claimed not to be syllabified as syllable onsets.

In its most succinct and general form, our main result can be stated by saying that adding a consonant to the left of a CV to obtain a CCV results in a reorganization of the temporal structure of the internal CV in the onset clusters investigated in Spanish but not

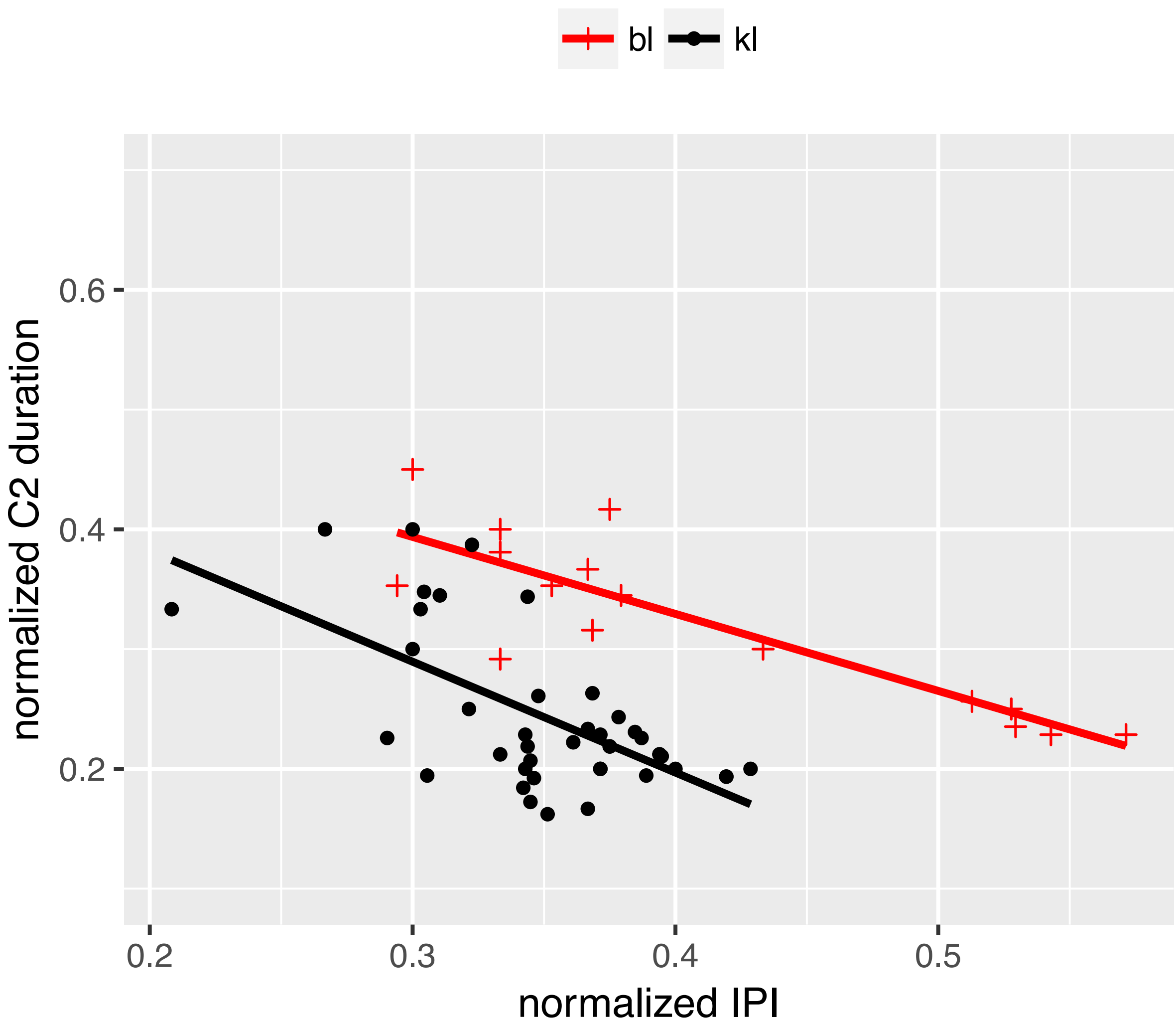

Figure 2.19: Scatterplot showing a weak negative correlation between normalized C2 lateral duration and normalized IPI for the /bl/ and /kl/ clusters in Moroccan Arabic ($r(58) = -0.37$, $p = 0.003$). As IPI increases, the lateral's C2 duration tends to decrease.

in Moroccan Arabic. We will review the specifics of this result below, after highlighting those aspects of our study that seem crucial in demonstrating this result.

Our study goes beyond previous studies in that it quantifies several different spatio-temporal measures: shortening of the prevocalic sonorant from CV to CCV, change of the relative timing of the lateral with the vowel from CV to CCV, relation between inter-gestural timing in the CC cluster (expressed by the interplateau interval or IPI) and C2 sonorant duration, and finally vowel initiation with respect to the preceding segment or segmental combination. With respect to the latter measure, employed in our study for the first time, we have seen that knowledge of vowel initiation helps identify connections with the other observed spatio-temporal readjustments in a CCV. With respect to the totality

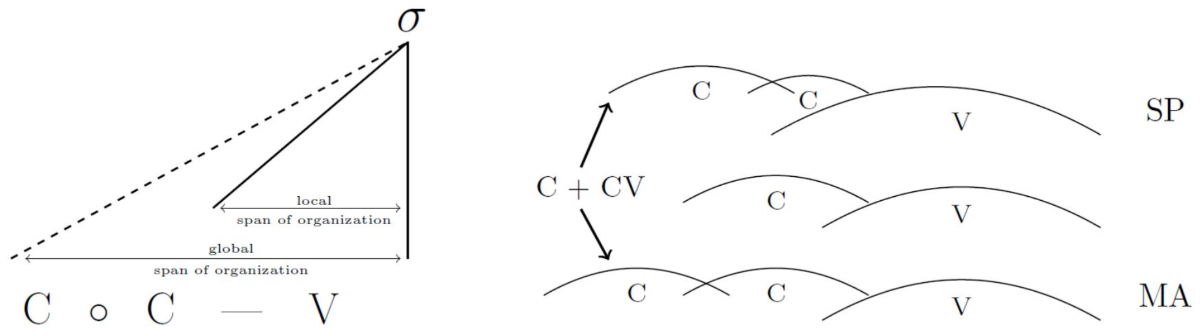

Figure 2.20: Span of organization, local versus global, over a CCV sequence (left) and schematic of concomitant adjustments from /la/ to /kla/ in Spanish versus Moroccan Arabic (right). Adding /k/ to the sequence /la/ (CV schema in right, middle panel) results in a number of readjustments which include substantial shortening of the /l/ in Spanish but not in Moroccan Arabic and an increase in overlap between the /l/ and the vowel in Spanish but not in Moroccan Arabic.

of the measures, also combined in our study for the first time, their joint consideration enables us to see patterns of how different measures relate to one another. It is such joint consideration, as we have argued and review below, that provides crucial clues on the presence of a global versus local organization over the segmental sequences investigated herein.

In our results, we find evidence for global organization in the form of a number of adjustments that effectively bring the vowel to overlap as much as possible with the cluster. Conceptually, perhaps the most important implication of our results is that there is no unique phonetic exponent of global organization. The exponent varies as a function of cluster, as in for example stop-rhotic versus stop-lateral clusters, and sometimes the degree to which the exponent is expressed varies as a function of within-segment properties as in the voicing contrast between voiced stop-lateral versus voiceless stop-lateral clusters. One reason why pleiotropy or this multiplicity of phonetic exponents expressing global organization is conceptually important is that previous studies, as reviewed in Chapter 1 and in this Chapter, use a small number of (one or two) measures to assess syllabic organization and even use somewhat different measures across different studies, with concomitant results so far failing to offer a unifying picture across studies (e.g., see Tilsen *et al.* 2012 who use just one measure; see also Brunner *et al.* 2014:449 for suggestions that the endeavor may be futile). If prosodic and more specifically syllabic organization is pleiotropic, as we argue in this thesis, lack of uniformity of results across different studies cease to be surprising. Any given measure represents some unique phonetic exponent and as we have seen the phonetic exponent expressing global organization cannot be assumed to be fixed. Let us illustrate this pleiotropy of global organization with some examples.

For stop-rhotic clusters which show quite large IPI (74 ms), we found global timing stability across clusters (to be treated with caution due to the different segment, trill in the /rV/ context and the tap in the /CrV/) and vowel initiation close to the c-center of the cluster (10-20 ms). For stop-lateral clusters, global organization is achieved in a number of ways: the prevocalic lateral duration decreases from CV to CCV, the relative timing of the internal /lV/ string changes and there is a compensatory relation between IPI and C2 lateral duration such that as IPI increases, C2 lateral duration decreases. The degree to which these changes are found depends on the coincidental phonetic properties of the segments before the vowel, such as their consonant duration or voicing and their consequences for the local inter-segmental timing between these consonants (such as IPI, which is longer in voiceless stop-lateral than voiced stop-lateral). To illustrate this latter point, let us contrast the effects observed in voiced stop-lateral versus voiceless stop-lateral clusters. Recall that in the latter cluster type, the initial stop is of longer duration than in the former case (i.e., /k/ is longer than /g/) and the IPI is longer than in the former than in the latter cluster type. Hence, for the vowel to overlap with the cluster to the extent possible, reorganization in the prevocalic material must be more extensive in voiceless stop-lateral than in voiced stop-lateral clusters. Indeed, as we observed, more lateral shortening and a greater change in the relative timing of the /lV/ string is found in the former than in the latter cluster type. Thus, given the phonetic properties of the segments and their local sequencing, syllabic organization imposes different degrees of lateral shortening, different degrees of relocation of the lateral in the cluster, a compensatory relation between IPI and lateral duration, and vowel initiation around or close to the c-center of the cluster to the extent possible. Overall, we find that even though some indices that have been used in prior work to diagnose syllabic structure do not show the expected patterns (for example, the usually employed global timing interval stability pattern which fails to show the expected result for the voiced stop-lateral clusters), there are other indices which do reflect a global organization. Crucially, these adjustments are not present or are less present in Moroccan Arabic, a language which we have argued provides a highly apt comparison to Spanish. This is so because, unlike Spanish, Moroccan Arabic (freely admits clusters but) does not permit complex onsets but shares with Spanish the same phonetic properties at the level of individual segments at least as far voicing is concerned and thus allows for a better comparison pair than say English and Arabic or German and Arabic as used by, for example, Shaw & Gafos (2015) in previous work. Thus, in /kl/, Moroccan Arabic shows no shortening of the lateral, no change in the relative timing of the /lV/ from /lV/ to /klV/, vowel onset initiation tied to the immediately preceding segment /l/ (as opposed to near the c-center of the prevocalic cluster as seen in Spanish) and a weak compensatory relation between IPI and lateral duration. These properties express an organization according to which the vowel is locally timed with the prevocalic consonant of the cluster.

Chapter 3

Global organization in German onsets

3.1 Introduction

This Chapter investigates the relation between syllabic organization and inter-segmental temporal coordination in German stop-lateral clusters. Specifically, it examines the effect of syllabic organization on the temporal coordination patterns in these clusters under different prosodic boundary conditions and in word-initial versus cross-word contexts.

Articulatory data were used to register the spatio-temporal form of stop-lateral clusters composed of the same segments parsed in two contrasting syllabic organizations, e.g., word-initial /kl/ as in /klɑːgə/ versus cross-word /k#l/ as in /pɑk#lɑːgə/. There have been previous studies investigating temporal coordination patterns of various consonant clusters in German but not under prosodic variation (Pouplier 2012, Brunner *et al.* 2014). These studies looked at the temporal predictions that derive from the span of syllabic organization. Furthermore, there have been also studies investigating intra- and inter-segmental timing patterns under prosodic variation in a subset of clusters in German (Bombien *et al.* 2013, Bombien *et al.* 2010). The latter do not include voiced stop-lateral clusters (/bl/, /gl/) and only consider the effect of prosodic variation and cluster composition on the duration of the consonants and on intra-cluster timing (overlap) without examining how prosodic variation affects the temporal coordination patterns in clusters with a different syllabic affiliation. The present study, for the first time, combines the scope of investigation of the previous studies mentioned on German. Namely, it examines 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 and it investigates the effect that these properties have on the inter-segmental temporal coordination patterns in clusters with different syllabic affiliation.

Why is prosodic variation a crucial key to understanding the nature of the link between (phonological) syllabic organization and the phonetic spatio-temporal manifestation of that organization? The effects of prosodic variation on segmental properties and on the overlap between the segments, we argue, offer the right pathway to discover patterns related to syllabic organization. This is so because, to uncover evidence for global organization, the sequence of segments partaking in that organization as well as properties of these segments or their relations with one another must be somehow locally varied. The consequences of such variation on the rest of the segmental sequence can then be used to unveil the span of organization. When local perturbations to segments or relations between adjacent segments have effects that ripple through the rest of the sequence, this is an indication 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.

The current study first examines the effect of varying prosodic boundary strength on the duration of the consonants in the cluster and on intra-cluster timing. This examination is required to ensure that the experimental design was successful in eliciting differences in consonant duration and in the degrees of overlap between the consonants in the cluster. Such differences are crucial because, as recounted above, it is the presence of consequences deriving from such differences that reveals the span of organization presiding over the segmental sequences studied here. Furthermore, the present study investigates the effect that these varying segmental parameters have on inter-segmental temporal coordination depending on syllabic organization. By inter-segmental temporal coordination, we refer to the usual stability-based heuristics and CV relative timing that previous studies have investigated before. However, as we have seen with the previous Chapter, in our work we extend examination to novel measures, such as the relation between intra-segmental timing and C2 lateral plateau duration as well as vowel initiation with respect to the prevocalic consonants. We argue that such a joint consideration of measures provides evidence for syllabic organization even in cases where one of the measures, such as the often relied on stability-based heuristics, fails to do so. It is only when prosodic and more specifically syllabic organization is examined by using a variety of phonetic indices and their relations that such organization can be reliable diagnosed in the phonetic record.

3.2 Background

The relation between syllabic organization and temporal coordination patterns constitutes the core of investigation in the present dissertation. 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 the temporal organization in 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, Hermes *et al.* 2013 on Italian, Pastätter & Pouplier 2015 and Hermes *et al.* 2017 on Polish). The working hypotheses 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 that has been said to reflect this global organization is an interval that spans from the center of the whole consonantism (hence, again, global) 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 the entire 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 prediction 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 *et al.* 2018). 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 *et al.* 2007 and Hermes *et al.* 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 immediately prevocalic 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 pattern of stabilities, 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, most of the studies on the relation between syllabic organization and temporal coordination patterns have shown that the stability-based heuristics related to syllabic organization often break down due to various reasons related to variability or cluster-specific phonetics.

One of the aims of the current Chapter is to better understand how directly the proposed stability-based heuristics for syllabic structure reflect syllabic organization in articulatory data from German. 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 *et al.* (2014). The Shaw *et al.* (2011) study is particularly relevant to our current purposes. 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 or 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). In their simulations, they varied these two phonetic parameters and considered the effect of that variation on the stability-based heuristics. The 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 simulation 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 be as simple as in statements that 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 in the literature it meets challenges in the data reported so far: there are clear cases where the expected stability-based stability 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, as promoted in Shaw *et al.* (2011), 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 (as previously done in 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 when the variability of certain parameter increases. The results showed that complex onset organization allowed for variability in the data to some extent but after some threshold the pattern for complex onset organization ceased to be stable.

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 Shaw *et al.* (2011) approach as well as our present experimental study, because it is then when the predictions of syllabic organization may be revealed. Shaw *et al.* (2011) investigated shortening of the prevocalic consonant and vowel compression as phonetic parameters. We investigate the lag between the two consonants, what we call interplateau interval or IPI, and lengthening of the initial stop consonant and the prevocalic consonant duration in word-initial CCV and C#CV where #indicates word boundary in German. 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 following the experimental design of Byrd & Choi (2010) on English. Byrd & Choi (2010) found that as the prosodic boundary strength increases, the lag between the two consonants increases and the effect is stronger for heterosyllabic consonants than onsets. For C1 duration, they found that both in onsets and in heterosyllabic clusters, C1 lengthens with increasing boundary strength. For C2 prevocalic consonant duration, they found that in onsets it does not seem to lengthen consistently with increasing boundary strength, while it does so in heterosyllabic clusters under prosodic strengthening. Byrd & Choi's (2010) prior work on English, thus provided us with reasonable evidence that an experimental design along their lines, if implemented in our German experiments, may provide analogous variability in the parameters we seek to introduce variability to. In turn, if this would turn out to be the case, this means that we would have generated the appropriate experimental test to pursue the simulation predictions of Shaw *et al.* (2011) in German.

Therefore, the present study has two aims. The first aim 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. The above three constitute the phonetic parameters that we aim to change in the present study as a result of the experimental manipulation. The second aim is to examine the way these effects of prosody on intra- and inter-segmental properties affect the stability-based heuristics in clusters of different syllabic affiliation. 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 lessons learned are the same as those in our previous Chapter on Spanish. First, joint consideration of measures (which include but are not limited to the usual stability-based heuristics) proves highly informative in diagnosing syllabic structure. That is one lesson. There is another one: evidence for complex organization emerges when phonetic parameters are scaled. Moreover, this evidence emerges even in cases where consideration of the stability-based heuristics in prior work on German (Brunner *et al.* 2014, Pouplier 2012) have indicated that German does not exhibit the expected patterns for global organization across clusters.

The Chapter is organized as follows. First, an outline of the methodology and the German stimuli is provided in section 3.3. A detailed analysis of the data including the changing of phonetic parameters under prosodic strengthening and the temporal coordination patterns that emerge in clusters of different syllabic affiliation follows in section 3.4. Section 3.5 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 3.6 with implications of our results for the issue of the relation between syllable structure and phonetic indices related to temporal 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 clusters) makes different predictions about the way phonetic indices respond to perturbations of phonetic parameters such as IPI or C1 lengthening.

3.3 Methods

3.3.1 Subjects

Articulatory data using the Carstens AG501 Articulograph were collected from five native German speakers (vp01, vp02, vp03, vp04, vp05), one male and four females, between 20 and 35 years old. All speakers reported no speech or hearing problem. They provided a 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.

3.3.2 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., /lɑːgə/∼/plɑːgə/). The cluster occurred word-initially as in the CCV sequences and across words as in C#CV where # indicates word boundary. Thus, for the cluster /pl/, the word /plage/ corresponds to a CCV sequence and the words /knɑp#lɑːgə/ 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., /lɑːgə/ 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 clusters consist of only the voiceless stop-lateral clusters /pl, kl/ because German has been reported to have final devoicing (Wiese 1996). This means that voiced stops in the word-final position are produced as voiceless. 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 three prosodic conditions with varying boundary strength preceding the stimulus word. For the cross-word clusters, 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 intonational phrase boundary to utterance phrase boundary. The word boundary condition was elicited by embedding the stimulus word in the carrier phrase "Ich sah ____ an" with the stimulus word location indicated by the "____". The intonational phrase boundary condition was elicited by embedding the stimulus word in the carrier phrase "Als ich Tom sah, ____ sagte er sofort".[1] The utterance phrase boundary condition was elicited by embedding

[1]The sentence "Als ich Tom sah, ____ sagte er sofort" as part of the intonational phrase boundary condition may be considered non-standard or even perhaps ungrammatical and as such could potentially affect the way participants produced the target word embedded in this sentence. This sentence was, however, chosen, because it fulfilled most of the conditions which play a role in successfully collecting and analyzing articulatory data (one of which being that the movement data have to be reliably segmented and thus the consonant(s) of interest should be preceded by a low vowel). The data collected as part of the above sentence do not shake the validity of our results in any way, because the utterance phrase boundary condition expressed through our other grammatically correct sentence agrees with the results of this

the stimulus word in the carrier phrase “Zunächst sah ich Anna. _____ sagte sie”. The stimulus word is always preceded by a low vowel. The preceding low vowel context allows for a clear distinction between a gesture of a lingual consonant and a vowel gesture as the two gestures are incongruent. Each speaker produced ten repetitions of each item ($N = 10$) in three prosodic conditions yielding a total of 300 tokens per speaker. Table 3.1 presents the list of complex onsets (CCV), simplex onsets (C#CV) and the respective singleton (CV) words.

Table 3.1: German stimuli.

Cluster	**CCV**	**C#CV**	**CV**
pl	Plage ‘plage’	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 ‘to grab a position’	Lage ‘location, position’

3.3.3 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

sentence which is used to elicit the intonational phrase boundary condition. Moreover, the participants familiarized themselves with the utterances prior to the experiment.

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.

3.3.4 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 algorithm developed by Mark Tiede at Haskins Laboratories. The 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 /plɑːgə/ 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 tangential 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 and end of the gesture calculated as the timestamps at which velocity rises above and falls below a 20% threshold of local tangential velocity peaks.

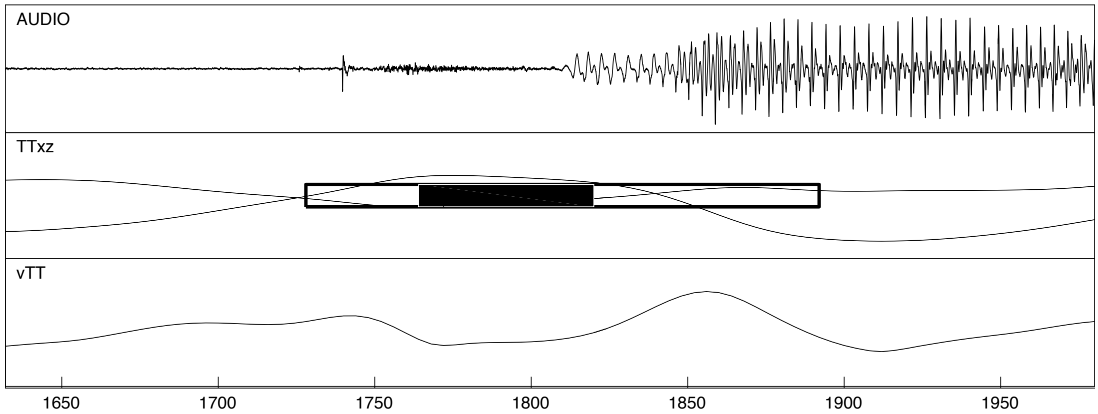

Figure 3.1: Parsing of a gesture using Mview. Panels from top to bottom: Acoustic signal of the zoomed in portion of the word "Plage", tongue tip (TT) movement trajectory in the anterior-posterior and superior inferior dimensions (x, z), tongue tip (TT) tangential velocity signal based on the movement in both x and z. 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.

3.4 Results

3.4.1 Interplateau internal

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 ms. Positive IPIs indicate no temporal overlap between the plateaus of the two consonants, while negative IPIs indicate temporal overlap.

In the current study, we elicited degrees of IPI by embedding the stimulus word in different prosodic conditions as in Byrd & Choi (2010). The first condition is the word boundary condition or control condition, where there is no phrase boundary or pause preceding the target word as in the carrier phrase "Ich sah [target word] an". The second condition corresponds to the intonational phrase boundary, where a short pause precedes the target word as in the carrier phrase "Als ich Tom sah, [target word] sagte er sofort". The third condition corresponds to the utterance phrase boundary, where a long pause precedes the target word as in the carrier phrase "Zunächst sah ich Anna. [Target word] sagte sie". Based on previous findings in Byrd & Choi (2010), it is predicted that IPI of the word-initial cluster of the target word increases with increasing boundary strength; i.e., the longer the pause preceding the target word, the longer the IPI of the cluster.

In this section, we quantify IPI as a function of prosodic condition to examine whether our experimental design successfully elicited different degrees of IPI. For word-initial clusters CCV, we fitted a linear mixed effects model with IPI as a dependent variable and prosodic condition, voicing of the initial stop (C1 voicing) and speaker as fixed effects. The variable "word" which corresponds to each unique word in the corpus was used as a random factor. The results of the model showed that IPI increases significantly from the word boundary condition (henceforth, wb) to the intonational phrase boundary (henceforth, ip) ($p < 0.0001$, estimate = 15.5 ms), to the utterance phrase boundary condition (henceforth, ut) ($p = 0.003$, estimate = 6.6 ms). Figure 3.2 illustrates IPI in CCV as a function of prosodic condition. IPI increases with increasing boundary strength, from wb to ip to ut. Furthermore, IPI is 28 ms larger in C1 voiceless stop-lateral clusters than in C1 voiced stop-lateral clusters ($p = 0.03$) across prosodic conditions. Figure 3.3 shows IPI as a function of C1 voicing. Means and standard deviations for IPI in C1 voiced and C1 voiceless stop-lateral clusters when C1 is a velar or a labial are shown in Table 3.2.

Table 3.2: Means and standard deviations of IPI in CCV as a function of prosodic condition, voice and place.

IPI (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
C1 voiceless	35.9	21	57.7	27.3	61.5	28.3
labial	44	19.5	66.1	23.5	64.7	29.1
velar	27.8	19.5	49.3	28.5	58.1	27.4
C1 voiced	15.5	19.4	25.7	23.9	35.1	26.5
labial	22.2	20.9	35.9	23.1	33	22.8
velar	9.9	16.1	17.1	21.2	36.8	29.5

For cross-word clusters C#CV, the IPI or lag between consonantal plateaus is obscured by the pause duration which separates the two consonants in the intonational and utterance phrase boundary conditions. The notion of IPI is thus not identical across word-initial and cross-word clusters (because it essentially reflects the pause duration in the latter context). However, we do report results on IPI or the lag between consonantal plateaus for cross-word clusters as well and we examine how variation in their IPI affects some of the phonetic indices in the C#CV string. This is because, even though the notion of IPI is not the same across word-initial and cross-word clusters, the reporting of IPI for cross-word clusters illustrates rather clearly how different syllabic organizations respond differently to perturbations of local properties such as IPI.

The results for cross-word clusters showed that IPI increases as a function of prosodic condition just as we saw for word-initial clusters. The IPI differences between our prosodic conditions word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut) are greater in C#CVs than in CCVs. For example, compare the values in

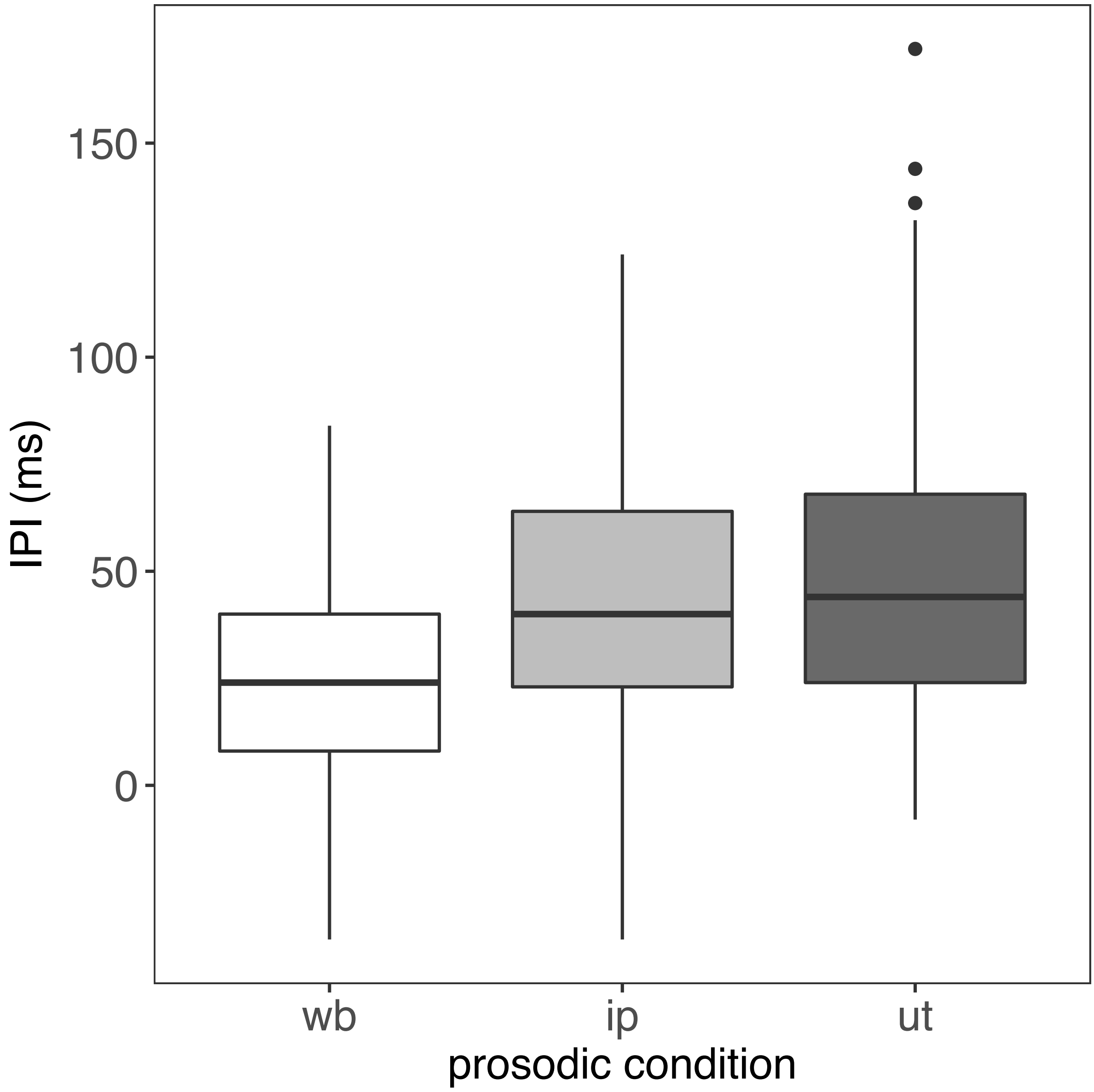

Figure 3.2: Interplateau interval (IPI) in CCV as a function of prosodic condition.

row "C1 voiceless" between wb, ip and ut in Table 3.2 and Table 3.3. 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 pauses of increasing duration). It is then reasonable that in the intonational phrase boundary condition (short pause), the IPI of the two consonants in C#CV is greater than in the word boundary condition with no pause. Similarly, IPI is greater in the utterance phrase boundary condition than in the intonational phrase boundary condition, because the pause in the latter is longer than in the former. Compare IPI values between wb, ip and

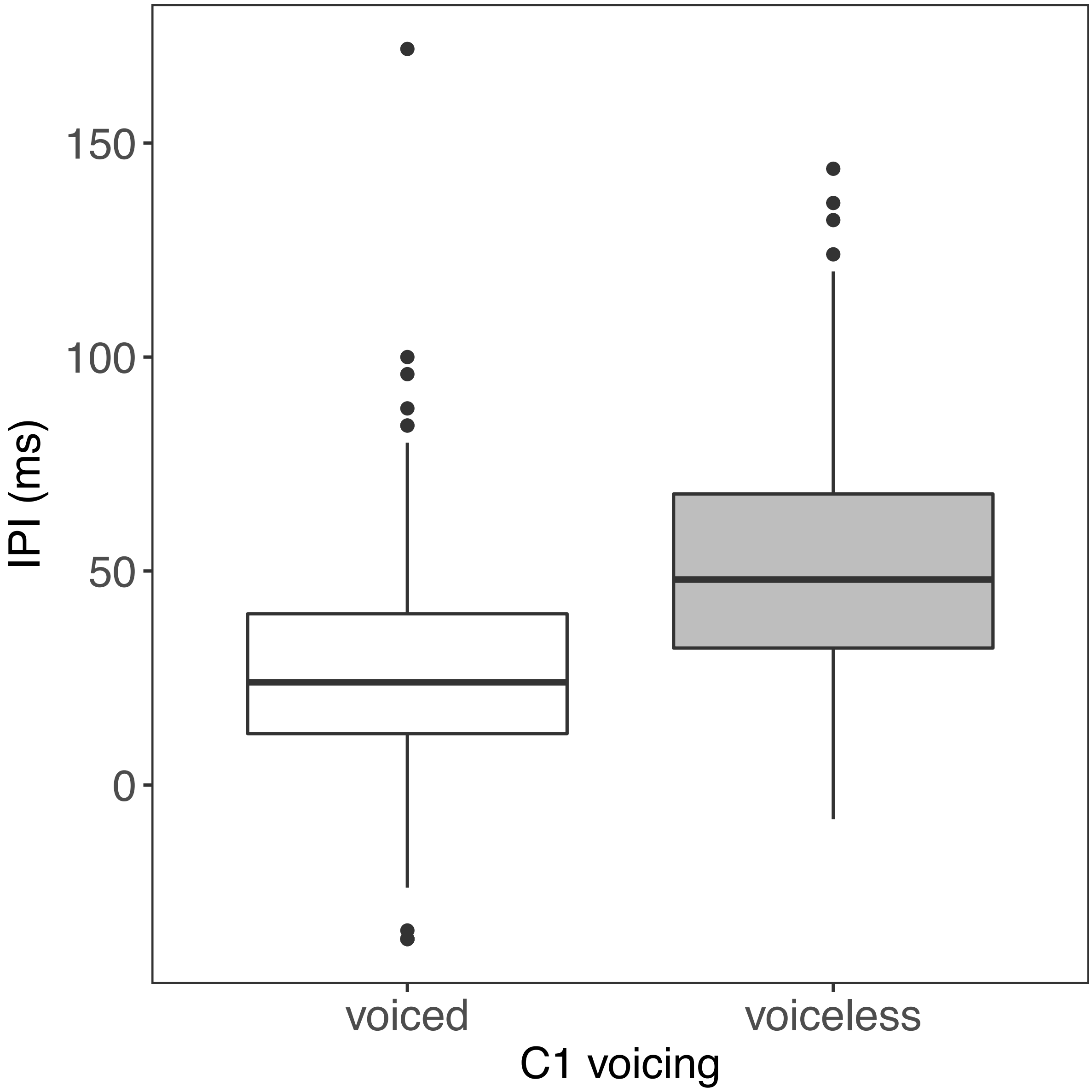

Figure 3.3: Interplateau interval (IPI) in CCV as a function of C1 voicing.

ut in Table 3.3. 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 3.4 illustrates IPI in C#CVs as a function of prosodic condition. It can be seen that with increasing boundary strength, from wb to ip to ut, IPI increases.

Table 3.3 presents means and standard deviations for IPI in voiceless C#CV sequences when the initial stop is a labial /p/ or a velar /k/.

To sum up, IPI in both word-initial complex onsets and cross-word clusters increases with increasing boundary strength corroborating the results in Byrd & Choi (2010) on

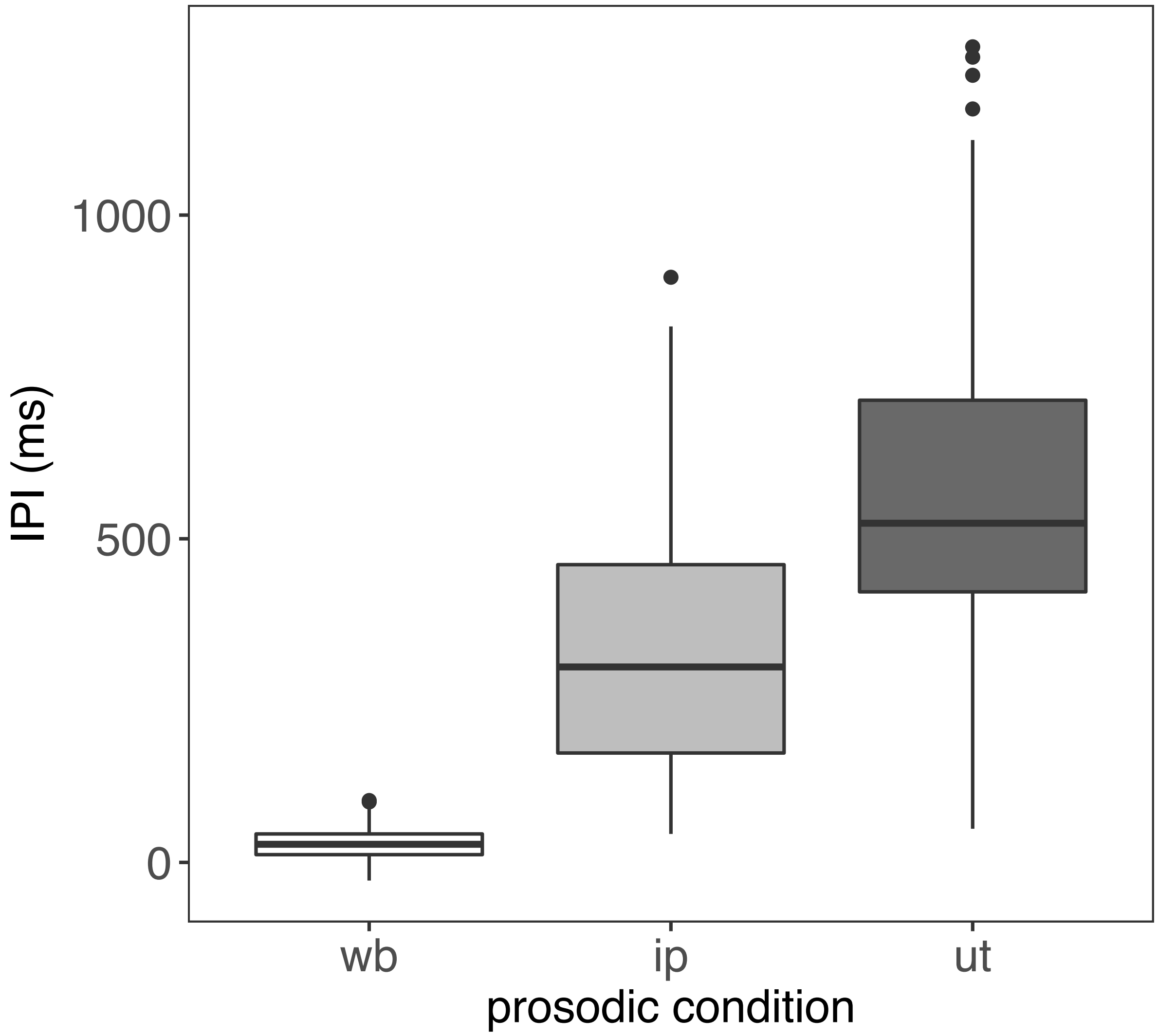

Figure 3.4: Interplateau interval (IPI) in C#CV as a function of prosodic condition.

Table 3.3: Means and standard deviations of IPI in CCV as a function of prosodic condition, voice and place.

IPI (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
C1 voiceless	28	23.9	330	200	590	272
labial	29.8	21.4	341.2	210	580	262
velar	26.1	26.3	319	192	600	285

English. Thus, the crucial take away for now is that our experimental design successfully elicited different degrees of IPI. This is a crucial prerequisite as noted in the introduction as we aim to use such differences in IPI to assess how syllabic organization adapts or responds to such perturbations or changes in the phonetic parameters of the segments that take part in that organization. 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 syllables in which these segments are included. Recall that the effect of C1 voicing on IPI cannot be examined in cross-word clusters due to final devoicing reported in German.

3.4.2 C2 lateral duration

In this subsection, we examine C2 lateral duration when it occurs as a single consonant in simple CV sequences and when it occurs in a cluster as a prevocalic consonant in a hypothesized complex onset CCV and in a cross-word cluster C#CV. C2 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.

Our results show that the duration of the lateral decreases 15 ms from CV to CCV across prosodic conditions (M = 58.22 ms, *sd* = 28.7 versus M = 43 ms, *sd* = 15.8). However, within prosodic condition, this result holds true only for the intonational phrase (ip) and utterance phrase (ut) boundary conditions. For the word boundary case (wb), the lateral's duration slightly increases (3 ms) from CV to CCV. Table 3.4 provides means and standard deviations of the lateral's duration from CV to CCV in each prosodic condition.

Table 3.4: Means and standard deviations for C2 lateral duration in CV and CCV in three prosodic conditions.

C2 lateral duration (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
CV	38.8	11.4	62.1	28.1	73.9	30.2
CCV	41.7	17.1	43.6	14.9	43.8	15.4

The differences in C2 lateral plateau duration from CV to CCV are statistically evaluated next. A linear mixed effects model was fitted with C2 lateral duration (log transformed) as a dependent variable. Speaker and the interaction of cluster size, C1 voicing and prosodic condition were used as fixed effects. "Item" which corresponds to each unique observation in the data was used as a random factor. The interaction between cluster size, C1 voicing and prosodic condition is at the limits of significance $F(2) = 3.028$, $p = 0.05$. The Tukey post-hoc test 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 = –0.02, p = non significant; C1 voiceless: estimate = –0.02, p = non significant), but it does decrease in the intonational (C1 voiced: estimate = 0.41, $p < 0.0001$; C1 voiceless: estimate = 0.23, $p = .003$) and utterance phrase boundary conditions (C1 voiced: estimate = 0.63, $p < 0.0001$; C1 voiceless: estimate = 0.38, $p < 0.0001$). Figure 3.5 illustrates C2 lateral duration from CV to CCV per voicing and per prosodic condition. The duration of the lateral decreases from CV to CCV in the intonational (ip) and utterance (ut) phrase boundary condition for both voiced and voiceless stop-lateral clusters, but it remains the same in the word boundary (wb) condition.

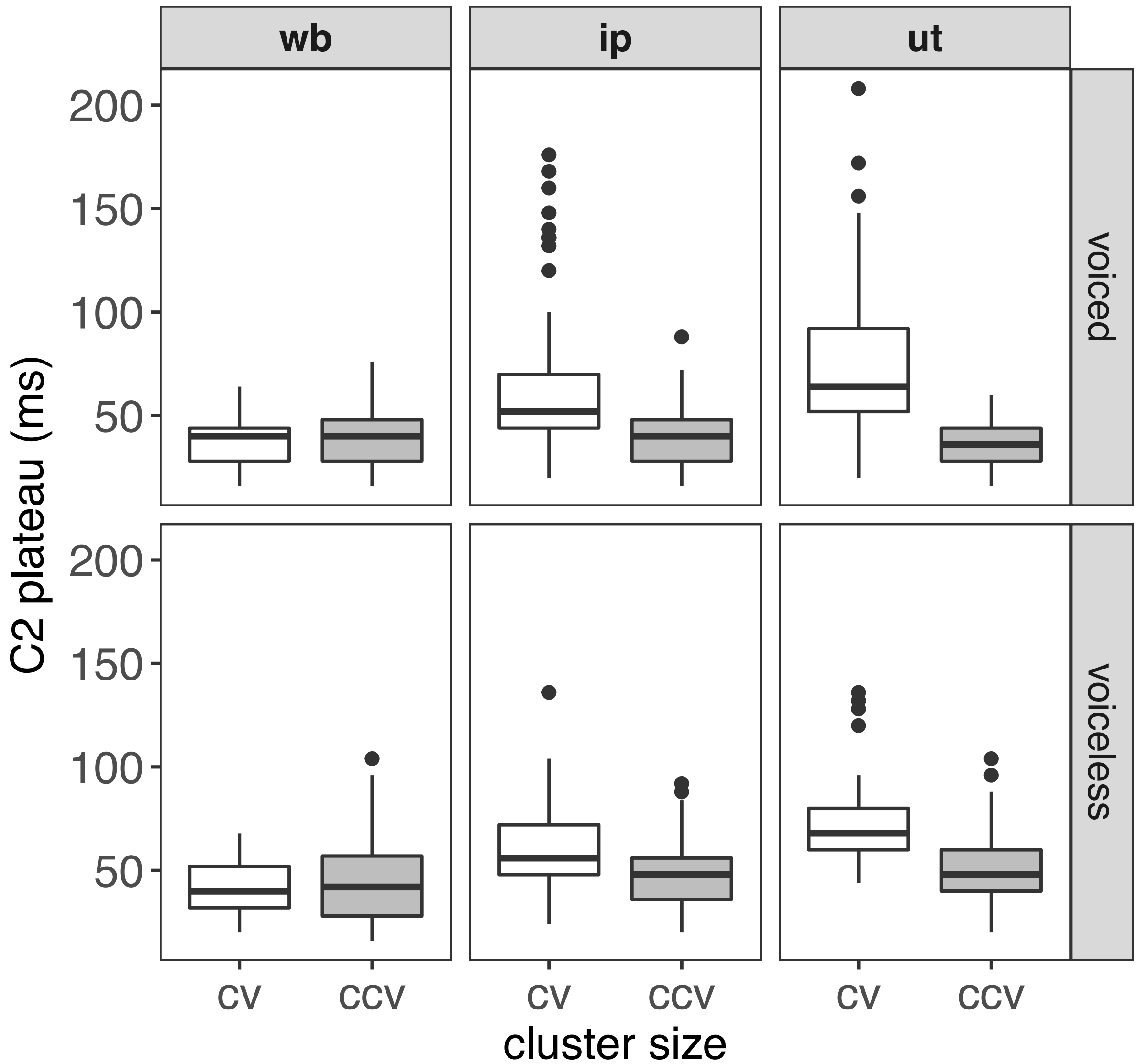

Figure 3.5: C2 lateral plateau duration as a function of cluster size, C1 voicing and prosodic condition.

We now turn to examine the effect of C1 voicing and prosodic condition and their interaction on C2 lateral duration within stop-lateral CCV clusters. A linear mixed effects model was fitted with C2 plateau duration as a dependent variable. The dependent variable was log transformed in order to better approximate a normal distribution. The variables C1 voicing, prosodic condition and speaker were modeled as fixed effects. “Item” which corresponds to each unique observation in our data ($N = 622$) was used as a random factor. There is an interaction between prosodic condition and C1 voicing ($F(2) = 4.907$, $p = 0.008$) which means that across CCVs the effect of C1 voicing on lateral's duration depends on the prosodic condition. The post-hoc Tukey test showed that for the word boundary position there is no effect of C1 voicing on lateral duration. For the intonational phrase boundary and the utterance phrase boundary conditions, the lateral's duration is shorter when it is preceded by a voiced than a voiceless stop (ip: estimate = 0.22, $p = 0.0002$, ut: estimate = 0.32, $p < 0.0001$). Figure 3.6 shows C2 lateral plateau duration as a function of C1 voicing and prosodic condition in CCV sequences.

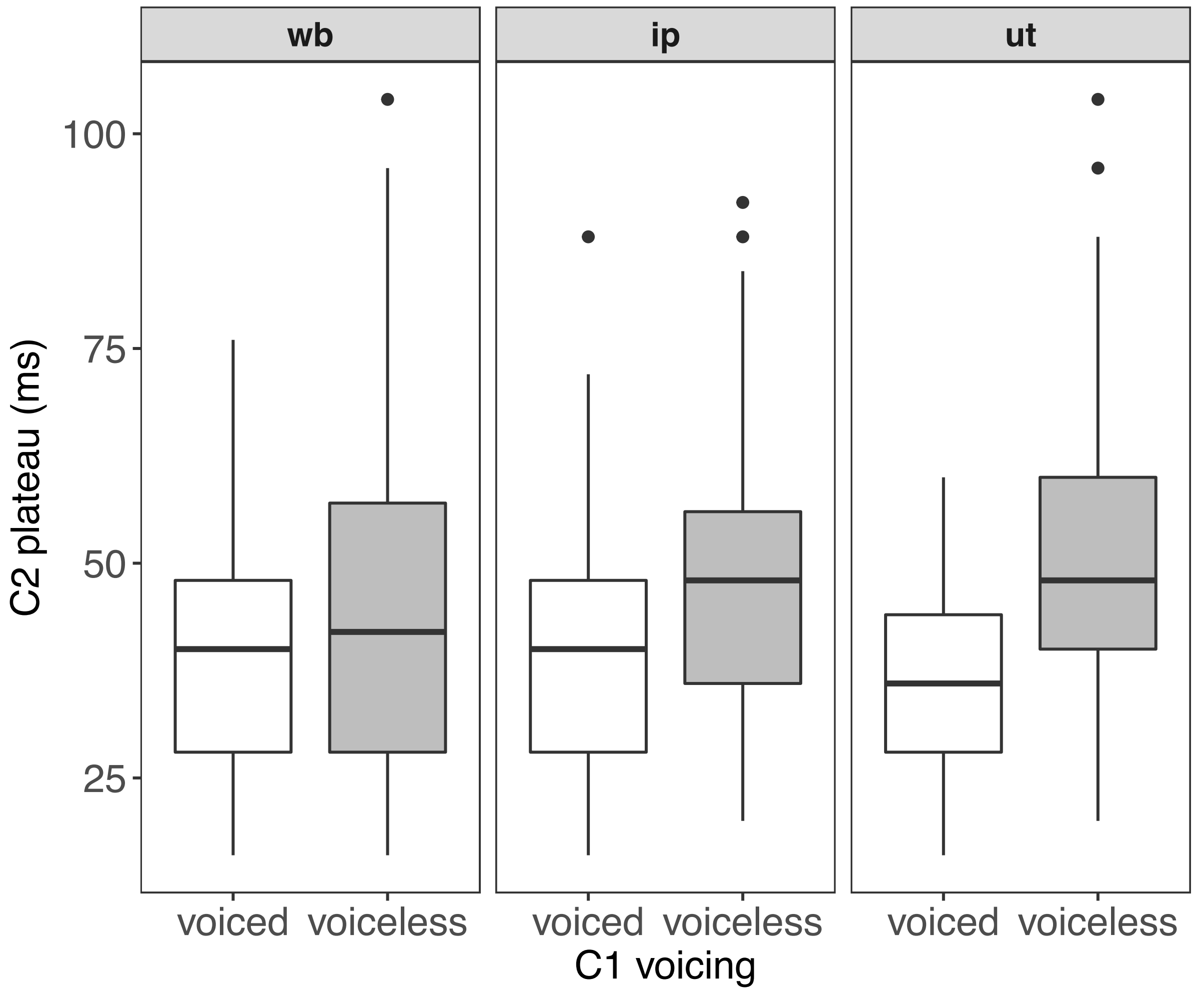

Figure 3.6: C2 lateral plateau duration as a function of C1 voicing and prosodic condition in CCV sequences. The C2 lateral in CCV is shorter when the initial stop is voiced than when it is voiceless only in the intonational phrase (ip) and in the utterance phrase (ut) boundary conditions. C2 lateral duration is the same regardless of C1 voicing in the word boundary (wb) condition.

Table 3.5 presents means and standard deviations for C2 lateral plateau duration in CV and CCV when the initial stop in the cluster is voiced or voiceless in three prosodic conditions.

Table 3.5: Means and standard deviations for the C2 lateral plateau duration from CV to CCV when the initial consonant in CCV is a voiced or a voiceless stop in three prosodic conditions.

C2 lateral duration (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
C1 voiceless						
CV	41.5	11.9	61.9	20.7	74.1	20.7
CCV	44.2	19.3	48.2	14.9	50.6	16.2
C1 voiced						
CV	37.2	10.9	62.1	31.5	73.8	34.5
CCV	39.3	14.4	39.3	13.7	37.1	11.2

Consider now C2 lateral duration from CV to C#CV. The duration of the lateral increases 8 ms from CV to C#CV across prosodic conditions (M = 59.1, *sd* = 22.6 versus M = 67 ms, *sd* = 27.3). Increasing lateral duration from CV to C#CV holds true also within prosodic condition but to different degrees. Specifically, in the word boundary condition the lateral lengthens 12 ms, in the intonational phrase boundary condition the lateral lengthens 9 ms and in the utterance phrase boundary condition the lateral lengthens 3 ms. Thus, with increasing boundary strength the lateral duration increases from CV to C#CV but the extent of this increase diminishes with increasing boundary strength (from wb to ip to ut). Table 3.6 presents means and standard deviations from C2 lateral plateau duration in CV and C#CV in the word boundary (wb), intonational phrase (ip) boundary and utterance phrase boundary (ut) conditions.

Table 3.6: Means and standard deviation for C2 lateral plateau duration in the CV and C#CV contexts in three prosodic conditions.

C2 lateral duration (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
CV	41.5	11.9	61.9	20.7	74.1	20.7
C#CV	53.9	16	70.3	25.3	77	32.8

In what follows, we evaluate statistically the change in C2 lateral plateau duration from CV to C#CV. For the cross-word clusters and their respective singletons ($N = 479$), we fitted a linear mixed effects model with C2 lateral duration (log transformed) as a dependent variable. Speaker and the interaction of cluster size, prosodic condition were used as fixed effects. "Item" which corresponds to each unique observation in the data was used as a random factor. There is an interaction between cluster size and prosodic condition ($F(2) = 6.936$, $p = 0.001$), which means that the way the lateral duration changes from CV to C#CV depends on the prosodic condition. The post-hoc Tukey test showed that only in the word boundary condition does the lateral duration increase significantly from CV to C#CV (estimate = –0.24, $p < 0.0001$). In the intonational and utterance phrase boundary conditions, the lateral duration does not change significantly from CV to C#CV (estimate = –0.11 and estimate = 0.01 respectively). Figure 3.7 illustrates C2 lateral duration from CV to C#CV in three prosodic conditions, word boundary, intonational phrase boundary and utterance phrase boundary.

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 intonational and utterance phrase boundary conditions, however, the lateral in the cluster CCV is shorter than the single lateral in CV and its shortening is greater (albeit not significant) when the initial stop is voiced than when it is voiceless. Now we turn to the lateral within CCV only (no CV context). Within the stop-lateral CCV clusters, there is no effect of C1 voicing in the lateral's duration in the word boundary condition. However, in the intonational and utterance phrase boundary conditions, the lateral in CCV is shorter when the initial stop is voiced than when it is voiceless. The above differences regarding lateral duration as a function of C1 voicing in contexts of prosodic strengthening for word-initial clusters enable us to bring out evidence for global organization. It is in the contexts of segmental lengthening and in phonetic contexts related to the voicing implementation (short IPI for voiced clusters versus long IPI for voiceless clusters) where fine-grained changes occur to the lateral. These changes, as will be discussed in section 3.5.1, will be argued to be motivated by the imposition of a global organization on the whole CCV sequence. For cross-word clusters, the lateral in C#CV is longer than the single lateral in CV across prosodic conditions. This last result, however, holds true only in the control condition and not in the other two contexts of prosodic strengthening (intonational phrase and utterance phrase boundary conditions) where the duration of the lateral in C#CV does not differ from that of the lateral in CV. Regardless, lengthening or no change in lateral duration from CV to C#CV across prosodic conditions points to local organization in cross-word clusters.

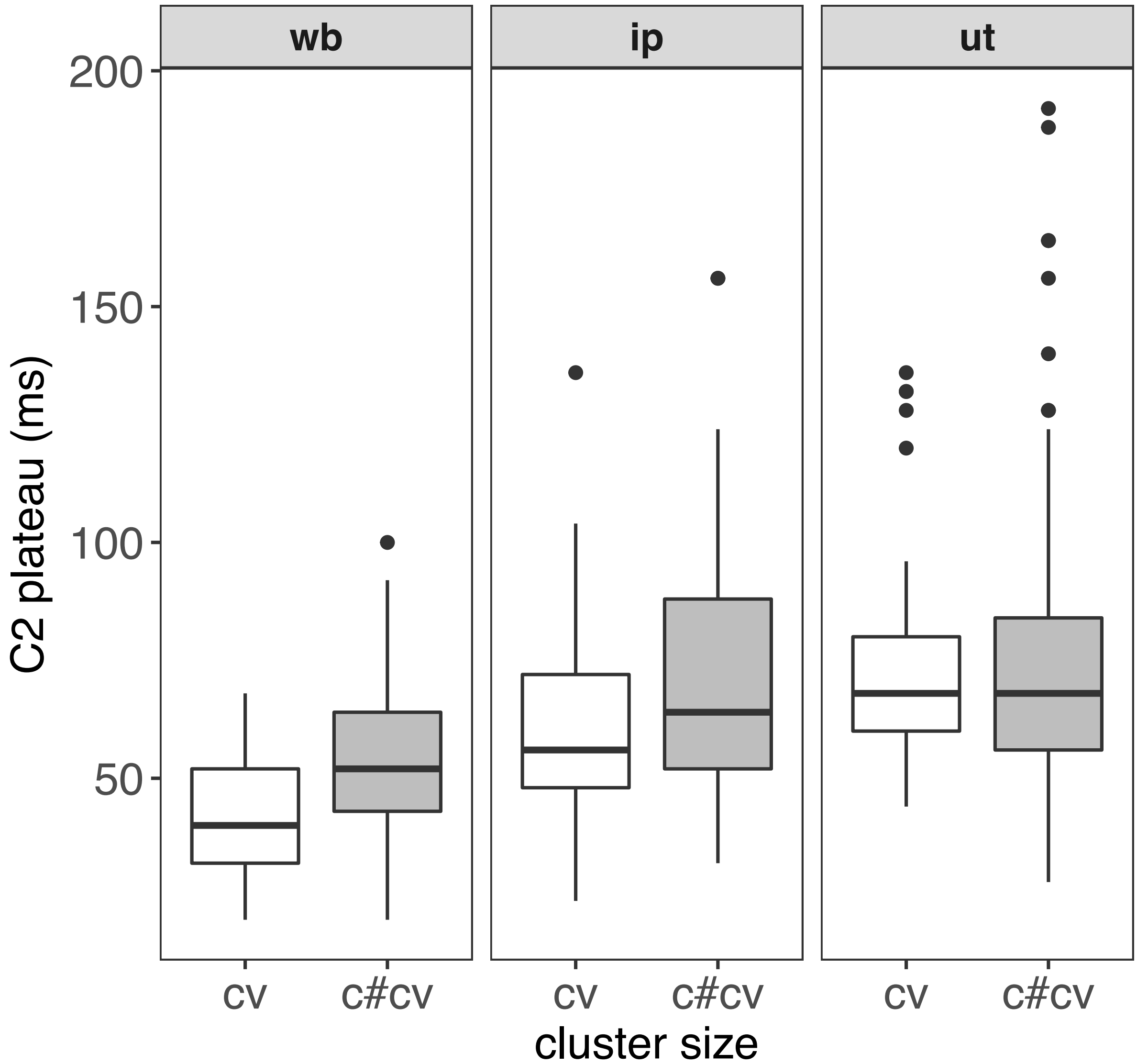

Figure 3.7: C2 lateral plateau duration from CV to C#CV in the word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut) conditions. The lateral in the cluster C#CV is longer than in CV only in the word boundary condition. In the intonational and utterance phrase boundary conditions, there is no significant difference with respect to C2 lateral duration from CV to C#CV.

3.4.3 IPI – C2 lateral compensatory relation

This subsection examines the relation between IPI and C2 lateral plateau duration in word-initial stop-lateral complex onsets across and within prosodic conditions. Our aim is to assess the presence of compensatory effects in the CCV string, that is, effects where spatio-temporal 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 offer evidence for global organization.

We first define the two relevant dependent variables. IPI was defined in section 3.4.1 as the lag between release of the initial consonant C1 and target of the second consonant C2 (C2 target – C1 release). The other relevant dependent variable is consonant duration. Consonant plateau duration was calculated as the interval between C release and C target: C plateau = C release – C target. C2 plateau duration was also normalized by dividing it by the total constriction duration of the cluster: C plateau normalized = C plateau duration/(C2 release – C1 target). The relation between IPI and C2 lateral plateau will not be investigated in the cross-word clusters, because the notion of IPI is not meaningful in that condition as it is obscured by the pause duration which separates the two consonants in the intonational and utterance phrase boundary conditions. Indeed, pause duration in cross-word clusters was produced with high variability within and across speakers and it can by no means be compared to the IPI in word-initial complex onsets. The obscured IPI in cross-word clusters is particularly relevant for the IPI-C2 lateral duration relation, because we rely on normalized IPI and C2 duration values which are normalized based on the total constriction duration and thus include the IPI duration. Normalizing IPI and C2 lateral duration in cross-word clusters, based on an interval which consists of the pause duration with high variability, is thus not meaningful.

For the word-initial stop-lateral complex onsets CCV ($N = 620$), there is no correlation between IPI and C2 lateral duration across prosodic conditions when looking at raw values $r(618) = -0.003$, $p = 0.93$. 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 intonational phrase boundary condition, $r(202) = -0.08$, $p = 0.23$; for the utterance phrase boundary, $r(205) = 0.07$, $p = 0.29$. Figure 3.8 shows no relation between raw IPI and C2 lateral plateau duration across CCV. 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 3.3 and Figure 3.6. 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 values, there is a weak negative correlation across prosodic conditions $r(618) = -0.36$, $p < 0.0001$, such that as IPI increases, C2 lateral duration tends to decrease. Figure 3.9 illustrates this weak compensatory relation between normalized IPI and C2 lateral plateau duration across CCV. However, there are inconsistencies within prosodic condition. For the word boundary condition, there is a moderate negative correlation $r(207) = -0.58$, $p < 0.0001$, for the intonational phrase boundary, there is a weak negative correlation $r(202) = -0.40$, $p < 0.0001$, and for the utterance phrase boundary, there is no correlation $r(205) = 0.10$, $p = 0.13$. Figure 3.10 illustrates the relation between normalized IPI and C2 lateral plateau across CCV in each prosodic condition (wb, ip, ut).

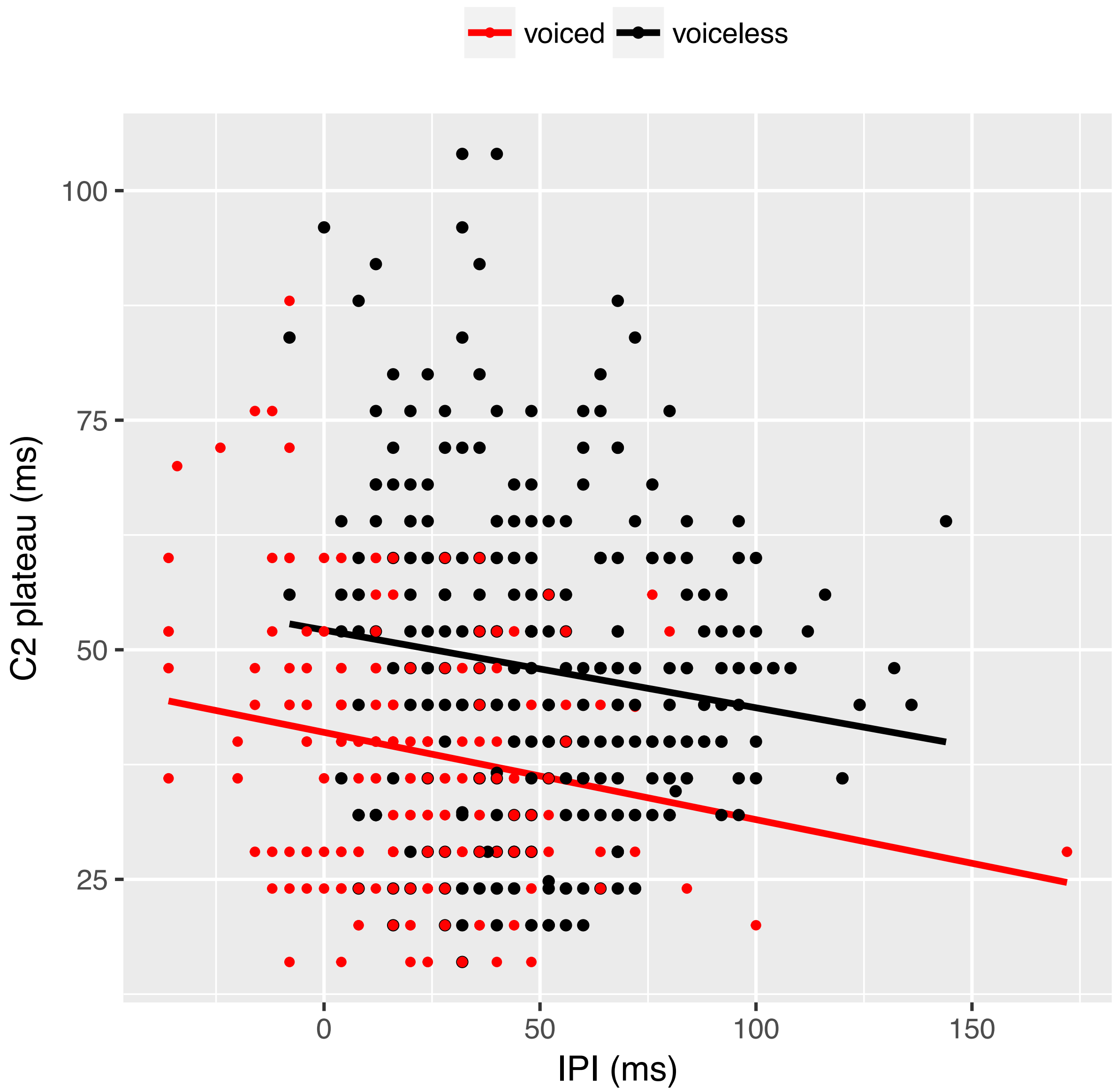

Figure 3.8: Scatterplot showing no relation between C2 lateral duration and IPI (in ms) for voiced and voiceless stop-lateral clusters CCV ($r(618) = –0.003$, $p = 0.93$).

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

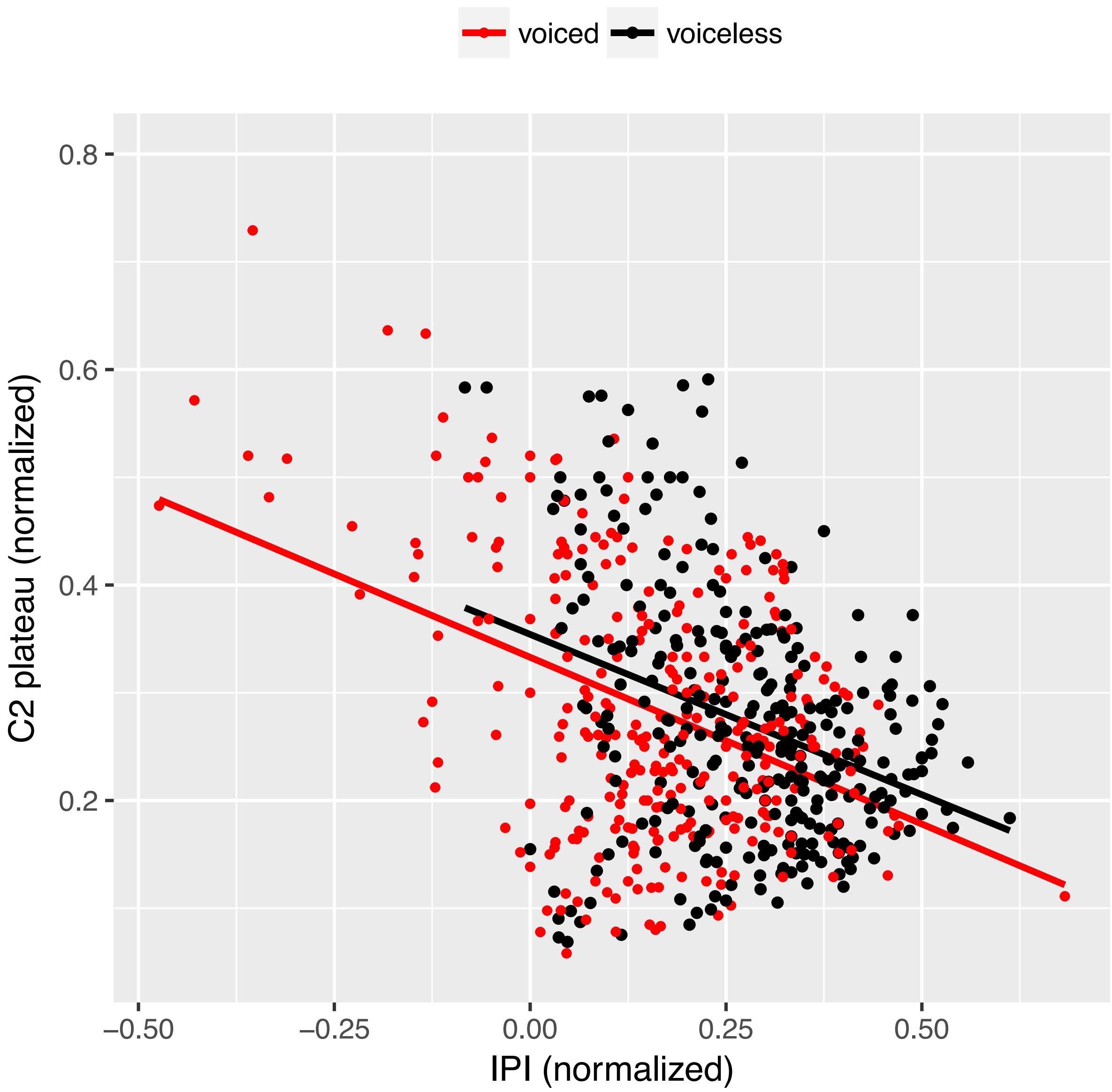

Figure 3.9: Scatterplot showing a weak relation between normalized C2 lateral duration and normalized IPI for voiced and voiceless stop-lateral clusters CCV ($r(618) = –0.36$, $p < 0.0001$).

decrease. This compensatory relation is stronger in the word boundary condition and it weakens with increasing boundary strength. In order to find a compensatory relation between two variables, variability is required. From the distribution of the data seen in Figure 3.10, 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 increasing boundary strength as can be seen in the progressive shrinking of the y-axis values from wb to ip to ut. Thus, the apparent lack of a compensatory relation in the other two contexts is not surprising. The compensatory IPI-C2 lateral plateau duration relation,

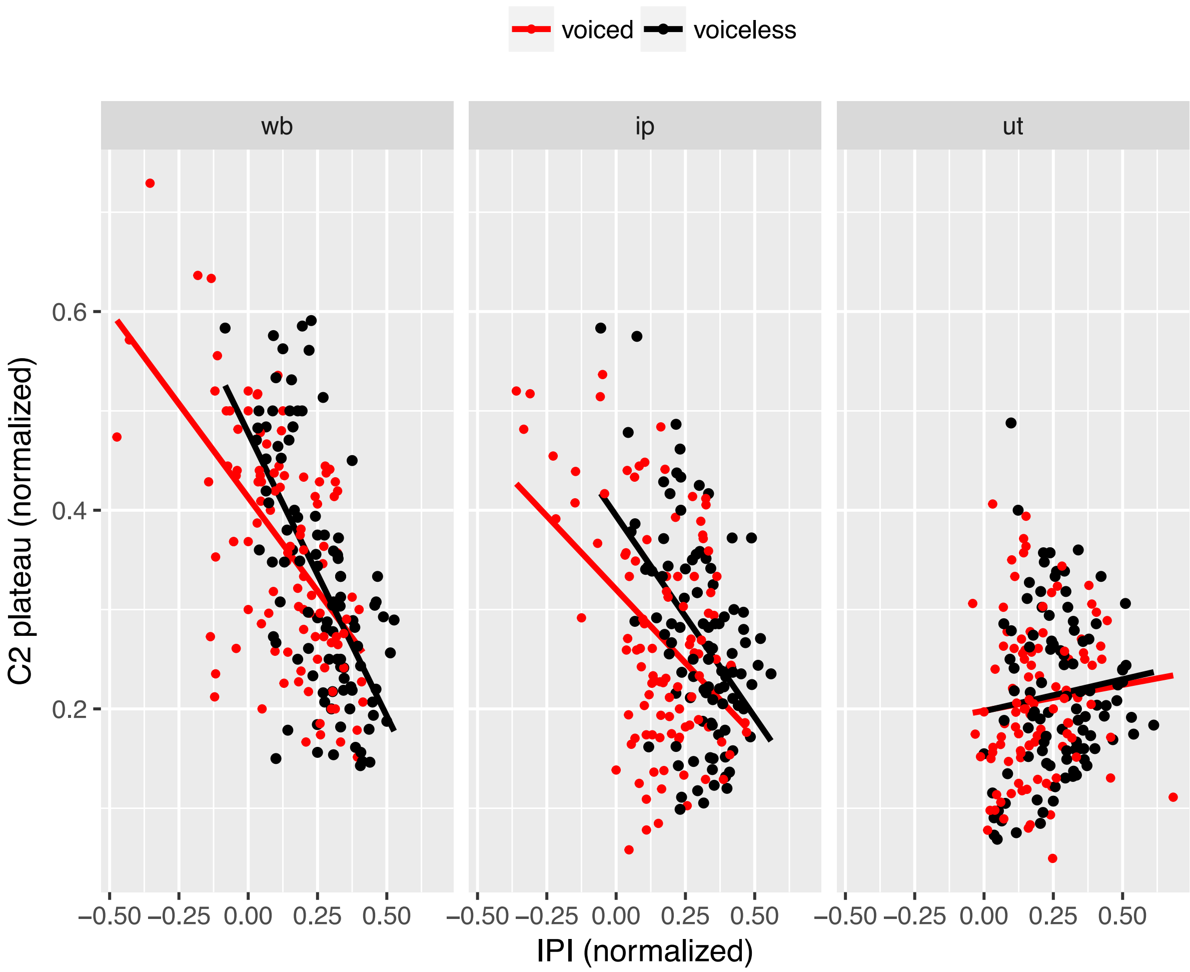

Figure 3.10: Scatterplots showing the relation between normalized C2 lateral duration and normalized IPI for voiced and voiceless stop-lateral clusters CCV in three prosodic conditions. In the word boundary condition (wb), there is a moderate negative correlation ($r(207) = –0.58$, $p < 0.0001$) such that as IPI increases, C2 lateral plateau decreases. In the intonational phrase boundary condition (ip), there is a weak negative correlation ($r(202) = –0.40$, $p < 0.0001$). In the utterance phrase boundary condition, there is no correlation ($r(205) = 0.10$, $p = 0.13$).

in contexts where it is found, as will be argued in section 3.5.2, serves as a hallmark of global organization.

3.4.4 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, prosodic condition and speaker were modeled as fixed effects. “Word” which corresponds to each token in the data was used a random factor. There is no interaction between C1 voicing and prosodic condition. There is no main effect of C1 voicing on the duration of C1 plateau. Figure 3.11 plots C1 plateau duration as a function of C1 voicing. No difference in C1 stop's duration is observed. Across prosodic conditions C1 voiced stops are 88 ms and C1 voiceless stops are 82 ms. There is a main effect of prosodic condition on the duration of C1 plateau with C1 plateau being shorter in word boundary than in intonational phrase boundary condition (estimate = 0.45, $p < 0.0001$) and C1 plateau being shorter in intonational phrase than in utterance phrase boundary condition (estimate = 0.26, $p < 0.0001$). Figure 3.12 plots C1 plateau duration as a function of prosodic condition. With increasing boundary strength from wb to ip to ut, C1 plateau lengthening can be observed. Across voicing conditions, stops are 53.4 ms in word boundary condition, 87 ms in the intonational phrase boundary condition and 116 ms in the utterance phrase boundary condition. Means and standard deviations for C1 stop plateau duration in ms in C1 voiced and C1 voiceless stop-lateral CCV clusters are provided in Table 3.7.

Table 3.7: Means and standard deviations for C1 plateau duration as a function of C1's voicing in each prosodic condition.

C1 stop duration (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
C1 voiceless	53.5	14.8	83.1	43	112.3	58.1
C1 voiced	53.4	14.4	92.3	53.9	119.8	62.2

For the cross-word stop-lateral clusters C#CV, the C1 stop consonant in C1#C2V slightly shortens with increasing boundary strength. 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 and speaker were used as fixed effects. The linear mixed effects model showed that the C1 plateau duration is slightly shorter in the intonational phrase boundary than in the word boundary condition (estimate = –0.12, $p = 0.05$). The intonational phrase boundary and utterance boundary conditions do not change significantly with respect to C1 plateau duration (estimate = –0.07). Figure 3.13 illustrates C1 plateau duration across C#CV as

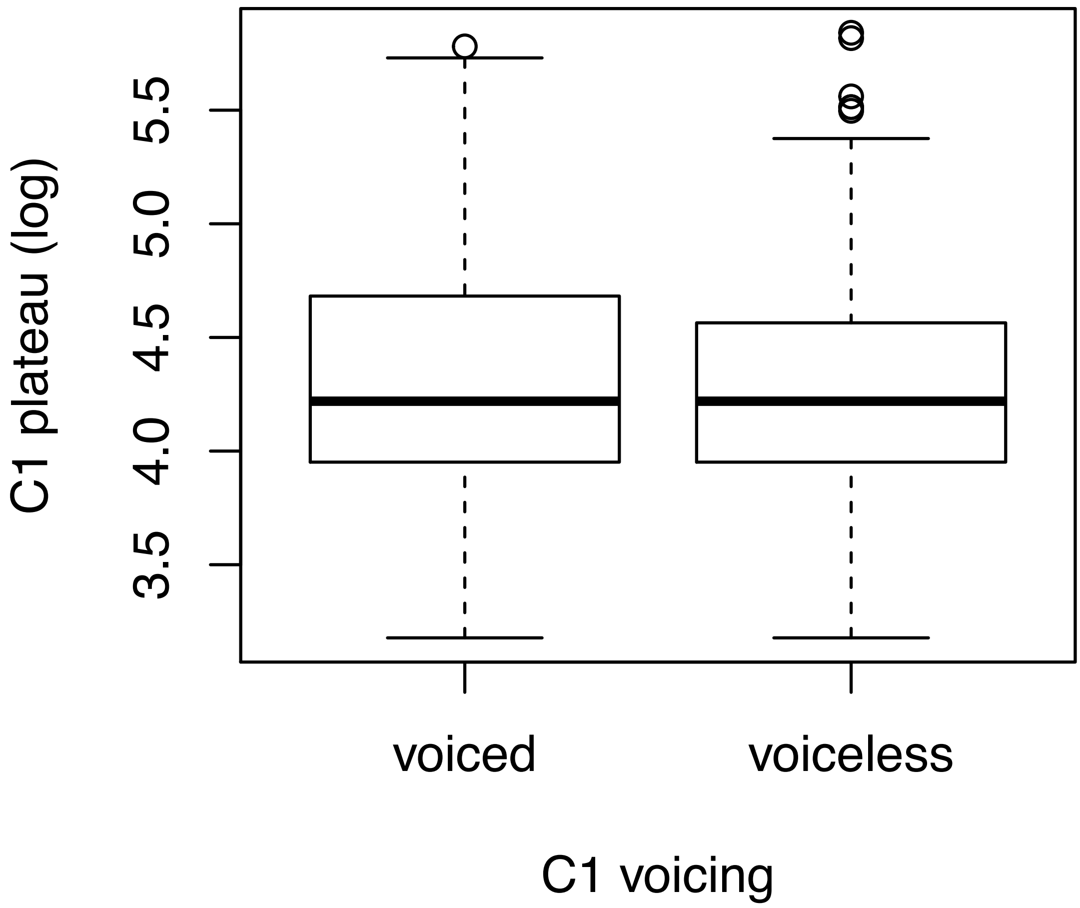

Figure 3.11: C1 plateau duration (log transformed) as a function of C1 voicing. C1 plateau duration remains the same regardless of C1's voicing.

a function of prosodic condition. With increasing boundary strength from wb to ip to ut, the stops' duration remains approximately the same.

Means and standard deviations for C1 plateau duration in ms in voiceless C#CV clusters when C1 is velar or labial are provided in Table 3.8.

Table 3.8: Means and standard deviations 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.

C1 stop duration (ms)	wb		ip		ut	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
C1 voiceless	57	17.2	52.2	20.1	47.8	16.3
p#l	56.9	16	43.6	16.1	40.1	13.2
k#l	57	18.5	60.7	20.3	56	15.4

Summarizing, there is no effect of C1 voicing on the duration of the initial stop in word-initial stop-lateral complex onsets. There is an effect of prosodic condition on the duration

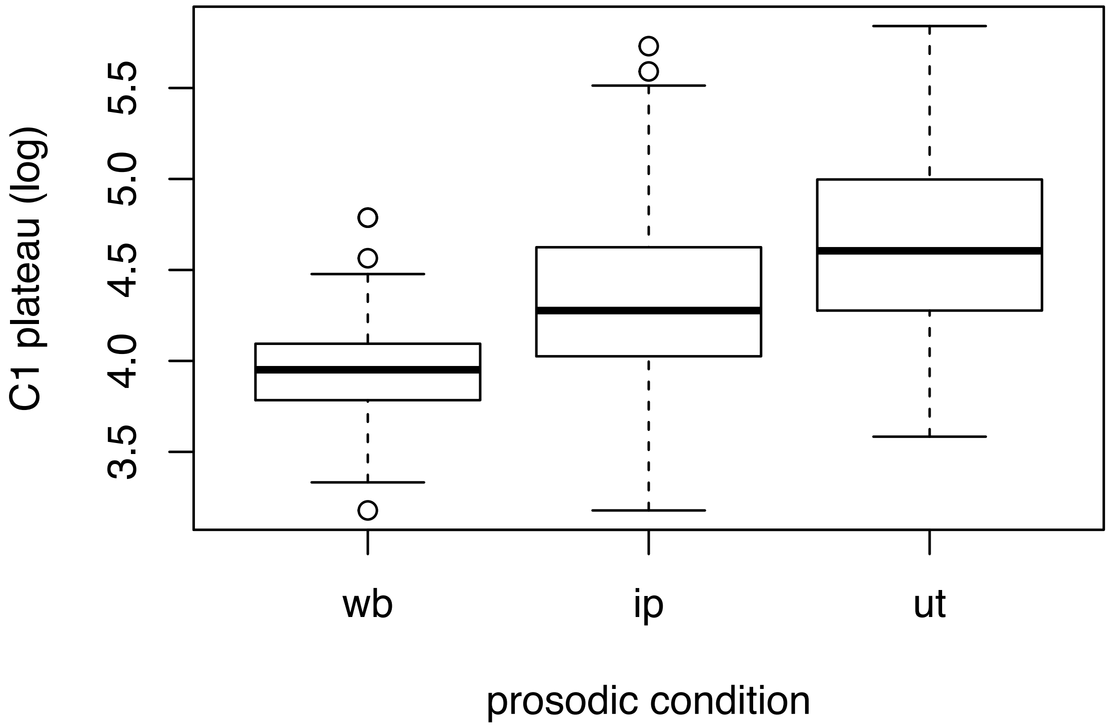

Figure 3.12: C1 plateau duration (log transformed) as a function of prosodic condition: word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut). C1 plateau duration increases with increasing boundary strength from wb to ip to ut.

of the initial stop across voicing. This means that the duration of the initial stop increases with increasing boundary strength 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 clusters. Thus, there is no effect of prosodic condition on the duration of the initial stop in cross-word clusters. 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 phonetic perturbations under which we can assess whether, as a result of these perturbations, temporal readjustments in (other parts of) the CCV string occur.

3.4.5 Vowel initiation

In this subsection, we examine (for the first time in the relevant literature) vowel initiation in stop-lateral clusters in German. We will do so by first examining vowel initiation in word-initial stop-lateral CCV clusters where the initial stop is voiced or voiceless in each of the following prosodic conditions: word boundary (wb), intonational phrase boundary (ip)

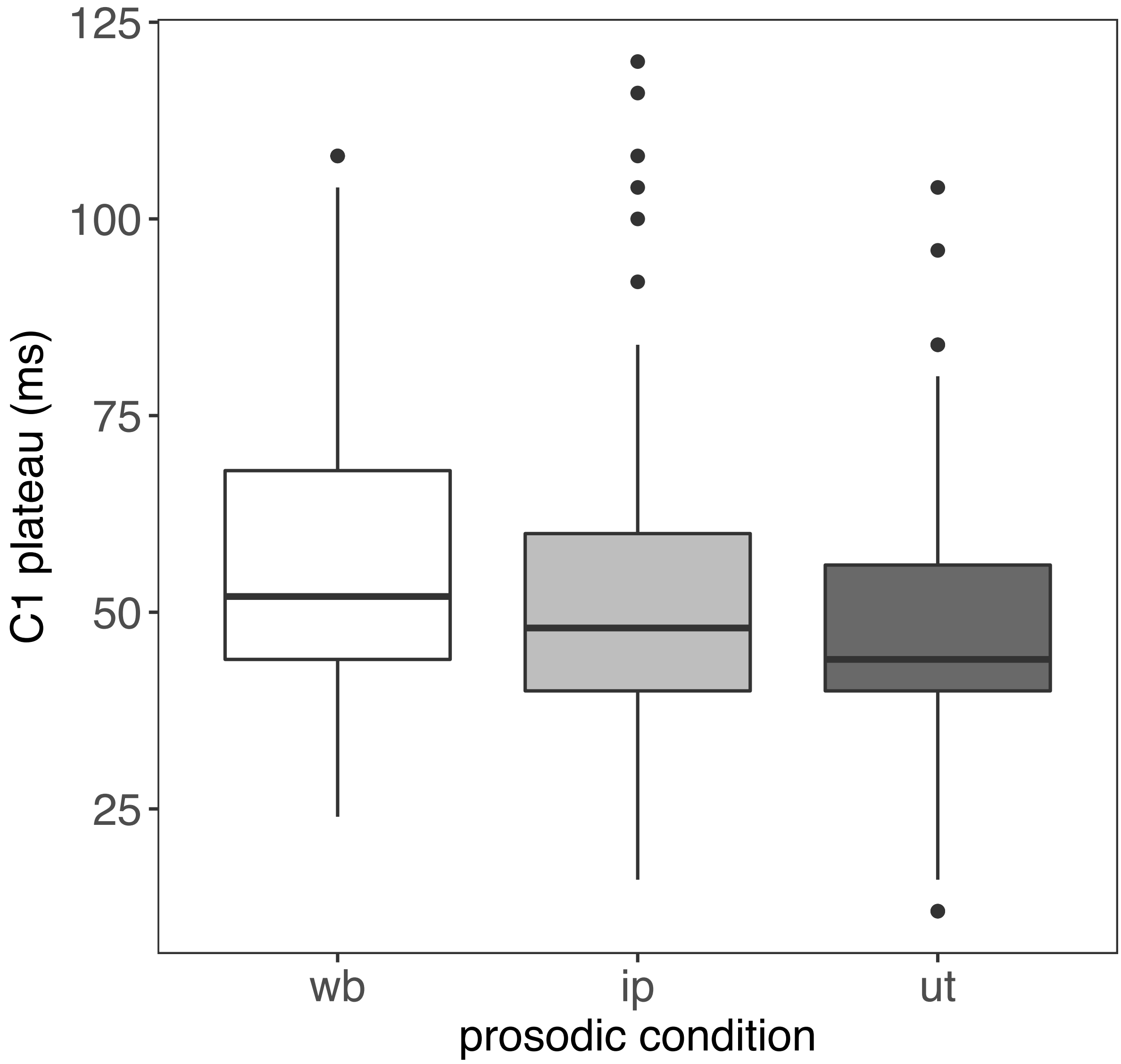

Figure 3.13: C1 plateau duration in ms for C#CV sequences as a function of prosodic conditions, in word boundary, intonational phrase boundary and utterance phrase boundary conditions.

and utterance phrase boundary (ut). Figure 3.14, Figure 3.15, and Figure 3.16 illustrate vowel initiation in word-initial stop-lateral CCV clusters where the initial consonant is voiced and voiceless in three prosodic conditions (word boundary, intonational phrase boundary and utterance phrase boundary conditions respectively). The vowel was parsed by using the tongue back sensor based on the tangential velocity of the signal. Where the vowel gesture failed to be reliably identified using the tangential velocity of the tongue back sensor, the vertical velocity was used instead. In cases where the tongue back sensor overall did not provide a reliable parse for the vowel, the tangential velocity of the tongue mid sensor was used.

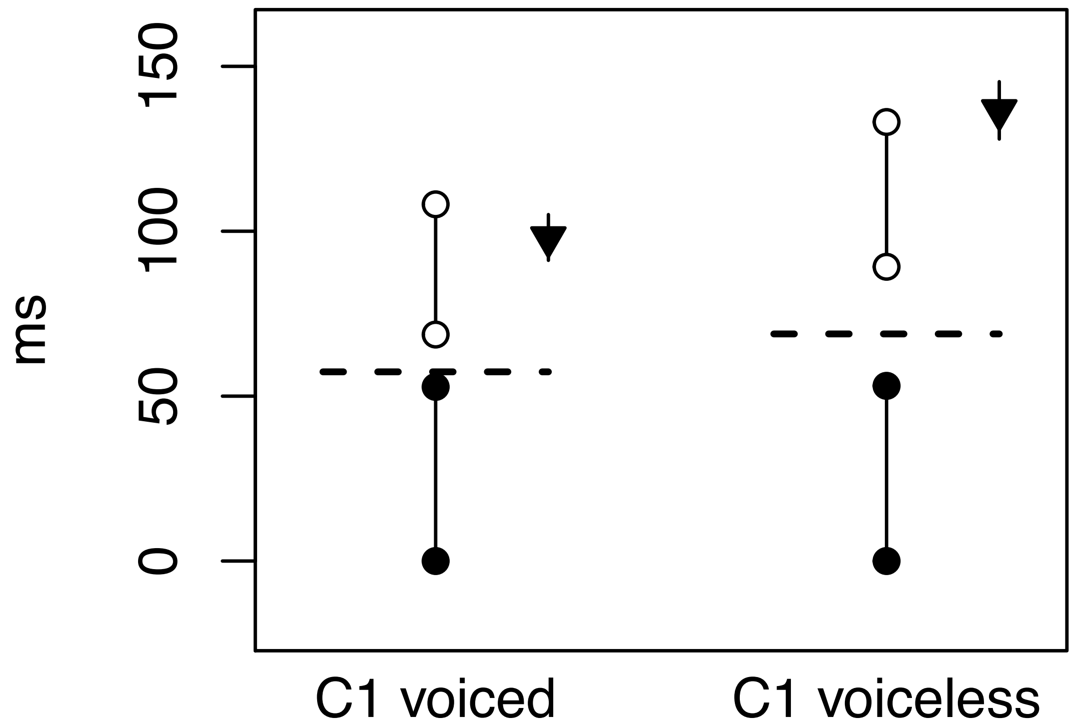

Figure 3.14: Vowel onset in relation to prevocalic consonants in C1 voiced and C1 voiceless stop-lateral CCV clusters (x-axis) in the word boundary condition. The vertical lines denote intervals corresponding to gestural plateaus of the two consonants in the stop-lateral cluster. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

Consider the interval between the target of the lateral (denoted by the bottom open circle of the horizontal line that denotes the plateau of the lateral) and vowel initiation (CV lag) in Figure 3.14, Figure 3.15 and Figure 3.16. This interval is progressively decreasing with increasing boundary strength. Across all CCVs, it decreases from 38.1 ms in the word boundary condition to 36 ms (p = non significant) in the intonational phrase and 27 ms ($p = .03$) in the utterance phrase boundary condition. In what follows, we examine how the CV lag, the interval between the target of the prevocalic lateral and the vowel initiation, changes from CV to CCV.

In the preceding, we examined vowel initiation in CCV where the initial consonant is voiced and voiceless in three prosodic conditions. We now turn to vowel initiation differences between CV and CCV in three prosodic conditions. For the CV#CCV pair ($N = 1105$), we fitted a linear mixed effects model with CV lag as a dependent variable. Speaker and the interaction of cluster size, C1 voicing and prosodic condition were used as fixed effects. "Item" was used as a random factor. The model showed no three-way

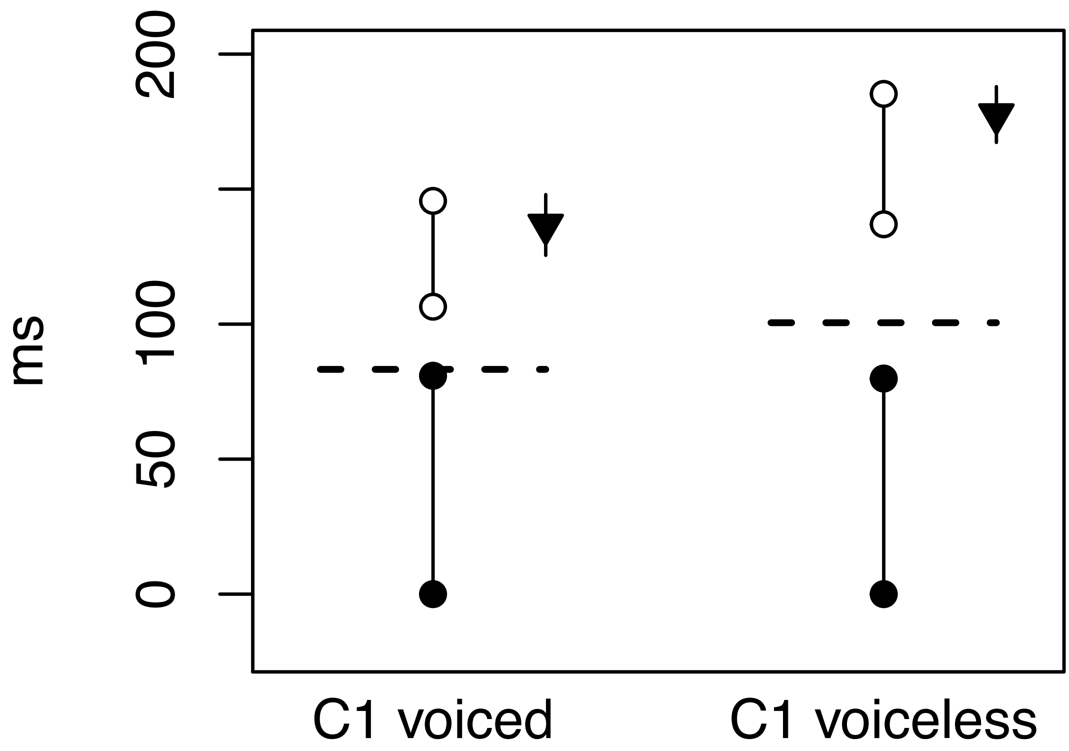

Figure 3.15: Vowel onset in relation to prevocalic consonants in C1 voiced and C1 voiceless stop-lateral CCV clusters (x-axis) in the intonational phrase boundary condition. The vertical lines denote intervals corresponding to gestural plateaus of the two consonants in the stop-lateral cluster. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

interaction of cluster size, C1 voicing and prosodic condition, but a significant interaction between cluster size and prosodic condition $F(2) = 18.9$, $p < 0.0001$. The Tukey post-hoc test showed that for the word boundary condition the CV lag does not change significantly from CV to CCV (estimate = –7.7 ms). In the intonational phrase boundary condition, the CV lag decreases by 15.6 ms ($p = 0.005$) from CV to CCV, and in the utterance phrase boundary condition, the CV lag decreases more by 31.8 ms ($p < 0.0001$) from CV to CCV. Figure 3.17 illustrates the CV lag across CV#CCV when the initial stop in the cluster is voiced and voiceless in the three prosodic conditions (wb, ip, ut).

With increasing boundary strength from wb to ip to ut, the vowel is found to start earlier with respect to the target of the prevocalic lateral in the word-initial CCV cluster than in CV. The exact opposite pattern is observed in cross-word clusters with increasing boundary strength as we will see below. The vowel is found to start later with respect to

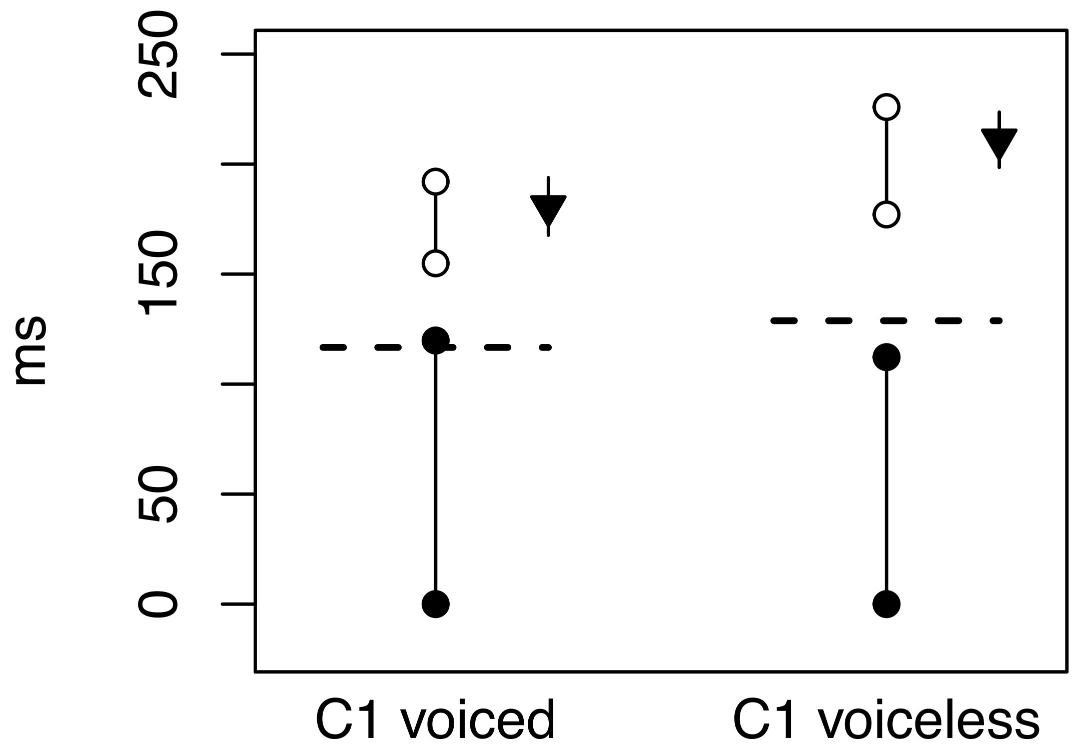

Figure 3.16: Vowel onset in relation to prevocalic consonants in C1 voiced and C1 voiceless stop-lateral CCV clusters (x-axis) in the utterance phrase boundary condition. The vertical lines denote intervals corresponding to gestural plateaus of the two consonants in the stop-lateral cluster. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

the target of the prevocalic lateral in cross-word C#CV clusters with increasing boundary strength from wb to ip to ut. Figure 3.18 shows vowel initiation with respect to the prevocalic consonants in cross-word stop-lateral clusters C#CV where the initial consonant is voiceless. The interval between the target of the prevocalic lateral and the vowel initiation is progressively increasing with increasing boundary strength. Across C#CVs (N = 299), the CV lag increases from 58.3 ms in word boundary to 64.3 ms (p = non significant) in intonational phrase boundary and 74.8 (p = 0.001) in the utterance phrase boundary condition. Thus, across C#CVs, the vowel starts later with increasing boundary strength from wb to ip to ut. Recall that the exact opposite pattern is observed across word-initial CCV clusters, where the vowel starts earlier with increasing boundary strength (refer to Figure 3.14, Figure 3.15, Figure 3.16).

We have heretofore considered vowel initiation with increasing boundary strength in

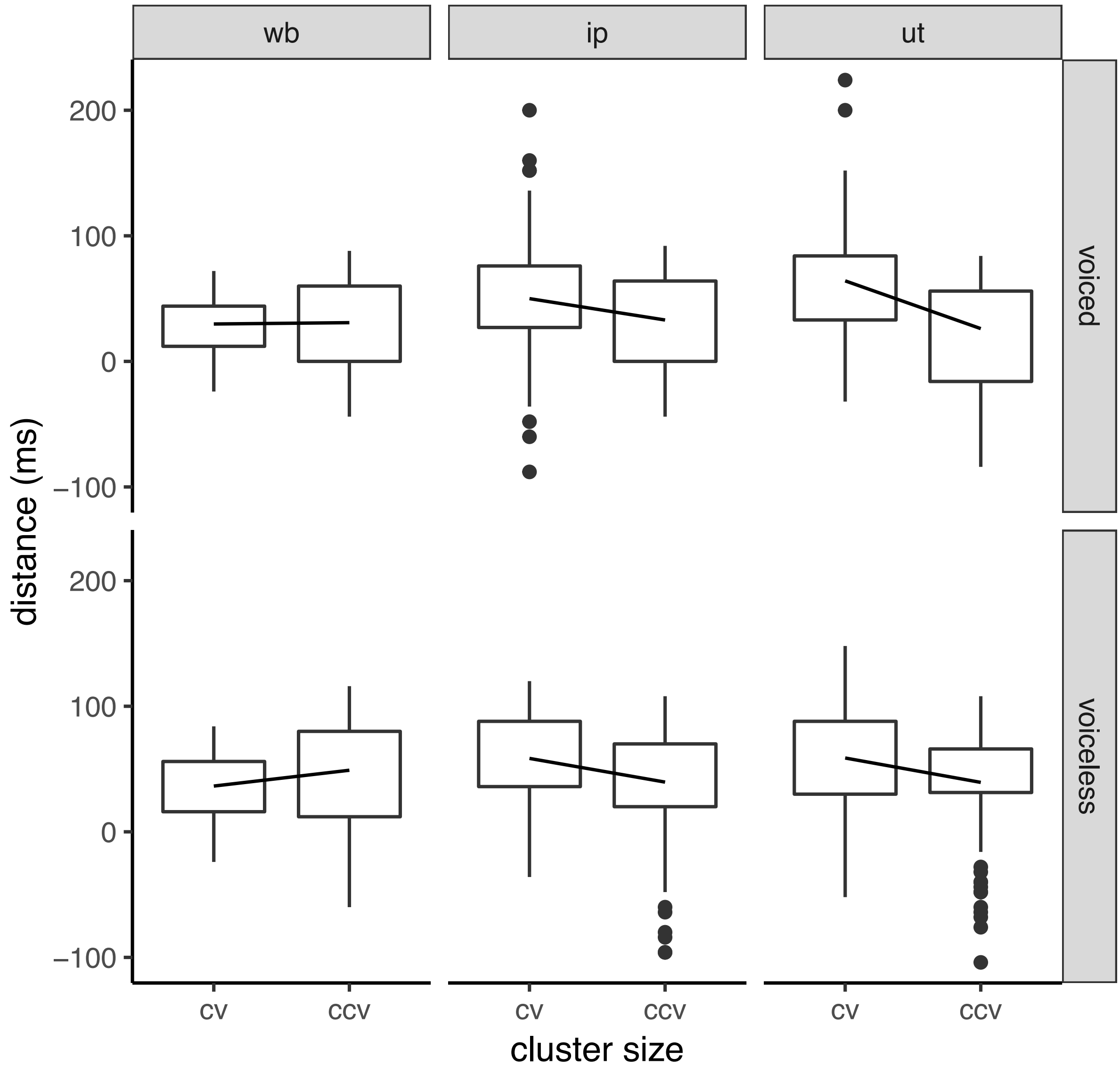

Figure 3.17: CV lag in ms between the target of the lateral and vowel initiation in lateral-vowel sequences (CV) and stop-lateral sequences (CCV) when the initial stop is voiced and voiceless in three prosodic conditions, word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut). The CV lag remains the same across CV∼CCV in wb, but it decreases from CV to CCV in the intonational and utterance phrase boundary conditions.

cross-word C#CV clusters only. We now turn to vowel initiation differences between CV and C#CV with increasing boundary strength. For the CV∼C#CV pairs (N = 481), we fitted a linear mixed effects model with CV lag as a dependent variable. Speaker and the interaction between cluster size and prosodic condition were used as fixed effects. Item was used as a random factor. There is an interaction between cluster size and prosodic condition $X^2(4,\ N = 481) = 23.01$, $p = 0.0001$. According to the post-hoc Tukey test, for the word boundary condition, the CV lag increases by 21.4 ms from CV to C#CV

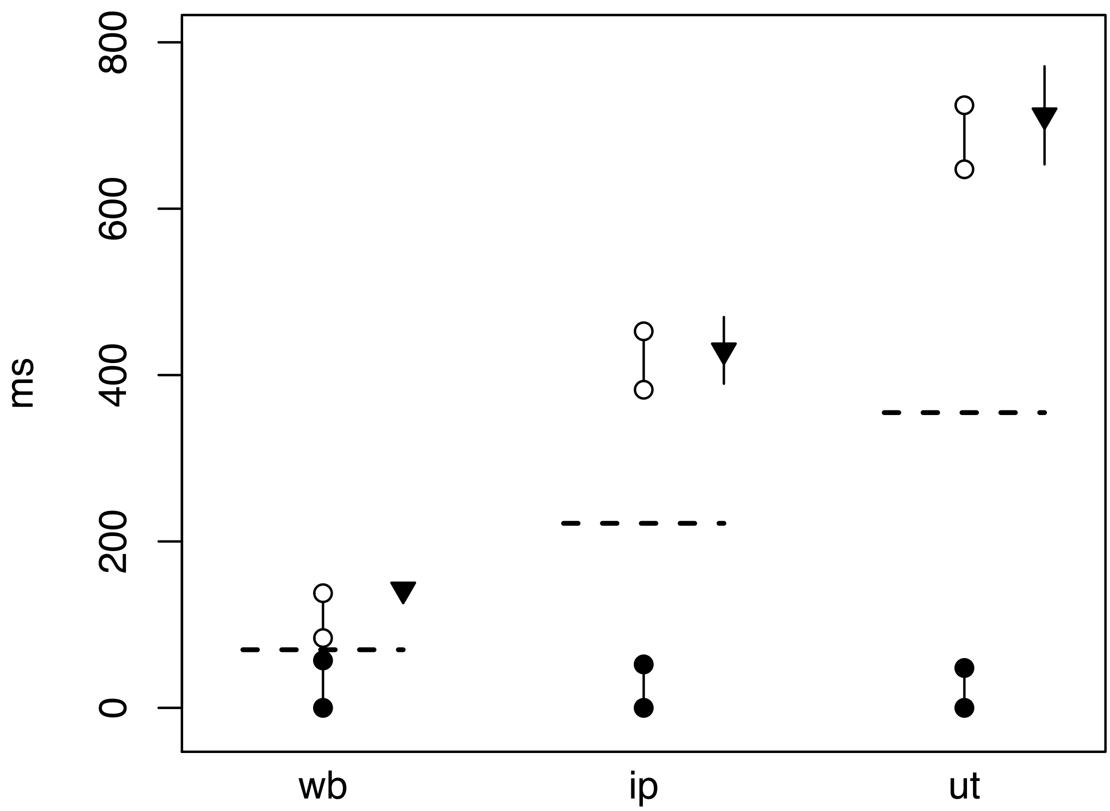

Figure 3.18: Vowel onset in relation to prevocalic consonants in C1 voiceless stop-lateral C#CV clusters in three prosodic conditions (x-axis): word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut). The vertical lines denote intervals corresponding to gestural plateaus of the two consonants in the stop-lateral cluster. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial voiceless stop. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

(p = 0.004), for the intonational phrase boundary condition it increases by 9 ms from CV to C#CV (p = non significant), and for the utterance phrase boundary condition the CV lag increases by 17 ms from CV to C#CV (p = 0.03). Figure 3.19 plots the CV lag in lateral-vowel (CV) and voiceless stop-lateral (C#CV) sequences in three prosodic conditions, word boundary, intonational phrase boundary and utterance phrase boundary.

To sum up, the CV lag, which describes the proximity of vowel initiation to the target of the prevocalic lateral, is progressively decreasing with increasing boundary strength across the stop-lateral word-initial complex onsets CCV. The exact opposite pattern is observed across cross-word clusters, where the CV lag progressively increases with increasing boundary strength. Thus, with increasing boundary strength the vowel is found to start earlier in word-initial complex onsets but later in cross-word clusters composed

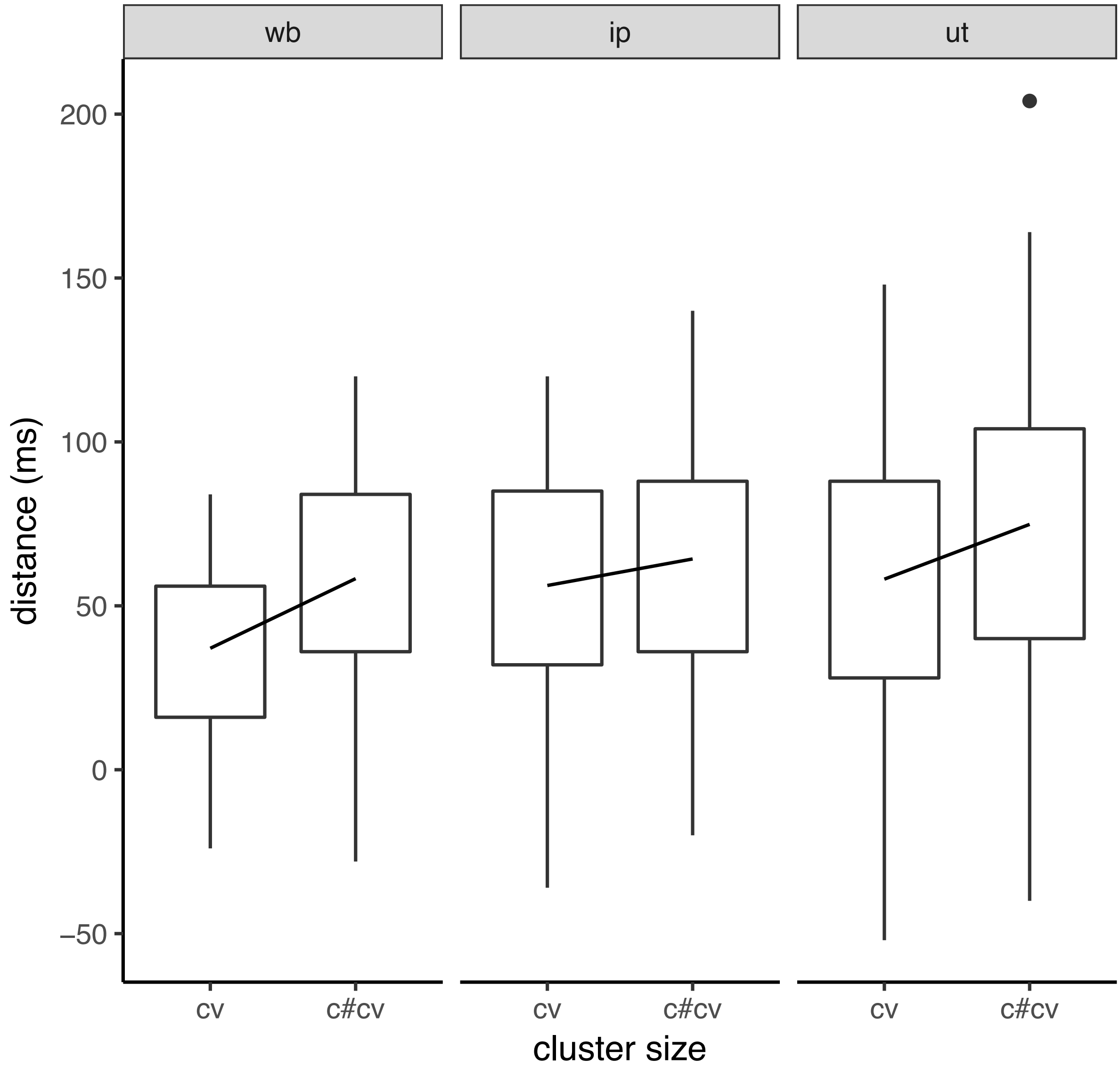

Figure 3.19: CV lag in ms between the target of the lateral and vowel initiation in lateral-vowel sequences (CV) and voiceless stop-lateral sequences (C#CV) in three prosodic conditions, word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut). The CV lag increases from CV to C#CV across prosodic conditions.

of the same consonants as in complex onsets. Such a vowel initiation difference between word-initial and cross-word clusters is crucial in revealing the two different organizing principles between the two as will be discussed in section 3.5.3; an early vowel initiation is a hallmark of a global organizing principle where the vowel overlaps with the entire cluster or starts as early in the cluster as possible, while a later vowel initiation is indicative of a more local organizing principle where the vowel is timed only with the immediate prevocalic consonant.

Let us also summarize the results regarding vowel initiation from CV to CCV and CV to C#CV, but these results should be treated with caution as they do not only reflect syllabic organization but also effects of prosodic strengthening on segmental properties which are different for a CV than a cluster. From CV to CCV, the CV lag remains the same in the word boundary condition, but it progressively shortens with increasing boundary strength. On the other hand, from CV to C#CV, the CV lag was found to increase within each prosodic condition.

3.4.6 Stability-based heuristics

For the interval stability analysis, the durations of different intervals and their stability are calculated for each stimulus observation. These intervals are defined on the basis of certain left- and right-delimiting landmarks. 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^{tar}), the maximum constriction of the postvocalic consonant (C^{max}), and the temporal midpoint of the vowel plateau (V^{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 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 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 (global, local) changes as the number of consonants increases from CV to a cluster with different syllabic affiliation CCV and C#CV.

For word-initial clusters, from CV to CCV ($N = 1118$), we fitted a linear mixed effects model with interval duration as a dependent variable. Speaker and a four-way interaction of interval type, cluster size, C1 voicing and prosodic condition were used as fixed effects. "Item" was used as a random factor. The Anova showed that there is a three-way interaction of interval type, cluster size and C1 voicing (C^{tar}: $F(1) = 23.28$, $p < 0.0001$; C^{max}: $F(1) = 16.74$, $p < 0.0001$; V^{mid}: $F(1) = 28.1$, $p < 0.0001$) and a three-way interaction of interval type, cluster size and prosodic condition which is at the limits of significance (C^{tar}: $F(2) = 2.77$, $p = 0.06$; C^{max}: non significant; V^{mid}: $F(2)$

= 3.4, p = 0.03). The post-hoc Tukey test showed that, for voiced stop-liquid clusters, the global timing interval increases significantly from CV to CCV (C^{tar}: estimate = –24 ms, p = 0.0002; C^{max}: estimate = –23, p = 0.01; V^{mid}: estimate = –29.6, p < 0.0001), while the local timing interval decreases from CV to CCV (C^{tar}: estimate = 10 ms; C^{max}: estimate = 10 ms; V^{mid}: estimate = 5.4 ms) but not to a significant extent. For voiceless stop-liquid clusters, the global timing interval increases from CV to CCV (C^{tar}: estimate = –36.7 ms, p < 0.0001; C^{max}: estimate = –36.3, p = 0.007; V^{mid}: estimate = –38.7, p < 0.0001), while there is no effect on the local timing interval which decreases from CV to CCV but reaches significance only when using the vowel anchor (C^{tar}: estimate = 15 ms; C^{max}: estimate = 16 ms; V^{mid}: estimate = 11.6 ms, p = 0.03). Figure 3.20 plots the duration of the two intervals global timing and local timing with a vowel anchor for voiced and voiceless stop-lateral clusters across speakers, as a function of the number of consonants (CV, CCV).

The post-hoc Tukey test for the three-way interaction of interval type, cluster size and prosodic condition showed that in the word boundary condition, the global timing interval increases significantly from CV to CCV (C^{tar}: estimate = –35.3, p = 0.0005; V^{mid}: estimate = –40 ms, p < 0.0001), while there is no effect on the local timing interval (C^{tar}: estimate = 2 ms; V^{mid}: estimate = –2.9 ms). In the intonational phrase boundary condition, the global timing interval increases significantly from CV to CCV (C^{tar}: estimate = –29 ms, p = 0.01; V^{mid}: estimate = –33 ms, p < 0.0001), while the local timing interval decreases from CV to CCV (C^{tar}: estimate = 15 ms; V^{mid}: estimate = 10.4 ms) but not to a significant extent. In the utterance phrase boundary condition, the global timing interval increases from CV to CCV (C^{tar}: estimate = –26 ms, p = 0.05; V^{mid}: estimate = –28.6 ms, p < 0.0001) and the local timing interval decreases from CV to CCV but reaches significance only when using a vowel anchor (C^{tar}: estimate = 22 ms; V^{mid}: estimate = 18 ms, p = 0.0004). Figure 3.21 plots the duration of the two intervals global timing and local timing with a vowel anchor for the three prosodic conditions, word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut) across speakers, as a function of the number of consonants (CV, CCV).

For cross-word clusters, across CV∼C#CVs (N = 458), we fitted a linear mixed effects model with interval duration as dependent variable. Speaker and the interaction of interval type, cluster size and prosodic condition were used as fixed effects. “Item” was introduced as a random factor. The Anova model showed that there is a three-way interaction of interval type, cluster size and prosodic condition (C^{tar}: $F(2)$ = 125.46, p < 0.0001; C^{max}: $F(2)$ = 121.86, p < 0.0001; V^{mid}: $F(2)$ = 152.50, p < 0.0001). The post-hoc Tukey test showed that in the word boundary condition, the global timing interval increases significantly (C^{tar}: estimate = –47 ms, p < 0.0001; C^{max}: estimate = –45.5 ms, p <

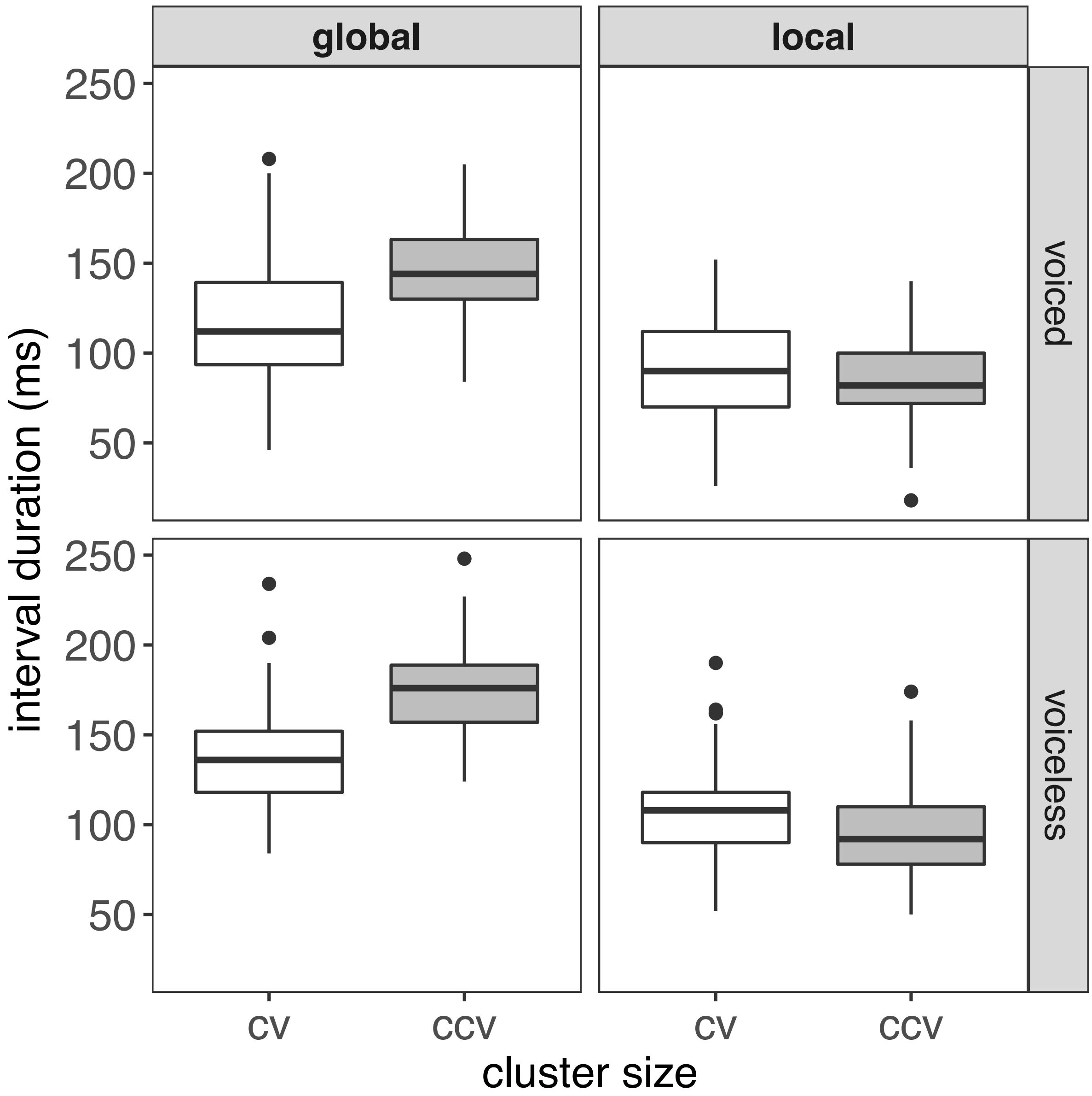

Figure 3.20: Duration (in ms) of the two intervals, global timing and local timing, for CV (white) and CCV (grey) words. For voiced stop-lateral CCV words (top), the difference in duration between CV and CCV for the global timing interval is greater than the difference for the local timing interval. 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.

0.0001; V^{mid}: estimate = –48.7 ms, $p < 0.0001$), while there is no effect on the local timing interval (C^{tar}: estimate = 0.3 ms; C^{max}: estimate = 1.8 ms; V^{mid}: estimate = –1.12 ms). In the intonational phrase boundary condition, the global timing interval increases significantly (C^{tar}: estimate = –196 ms, $p < 0.0001$; C^{max}: estimate = –196 ms, $p < 0.0001$; V^{mid}: estimate = –170 ms, $p < 0.0001$), while there is no effect on the local

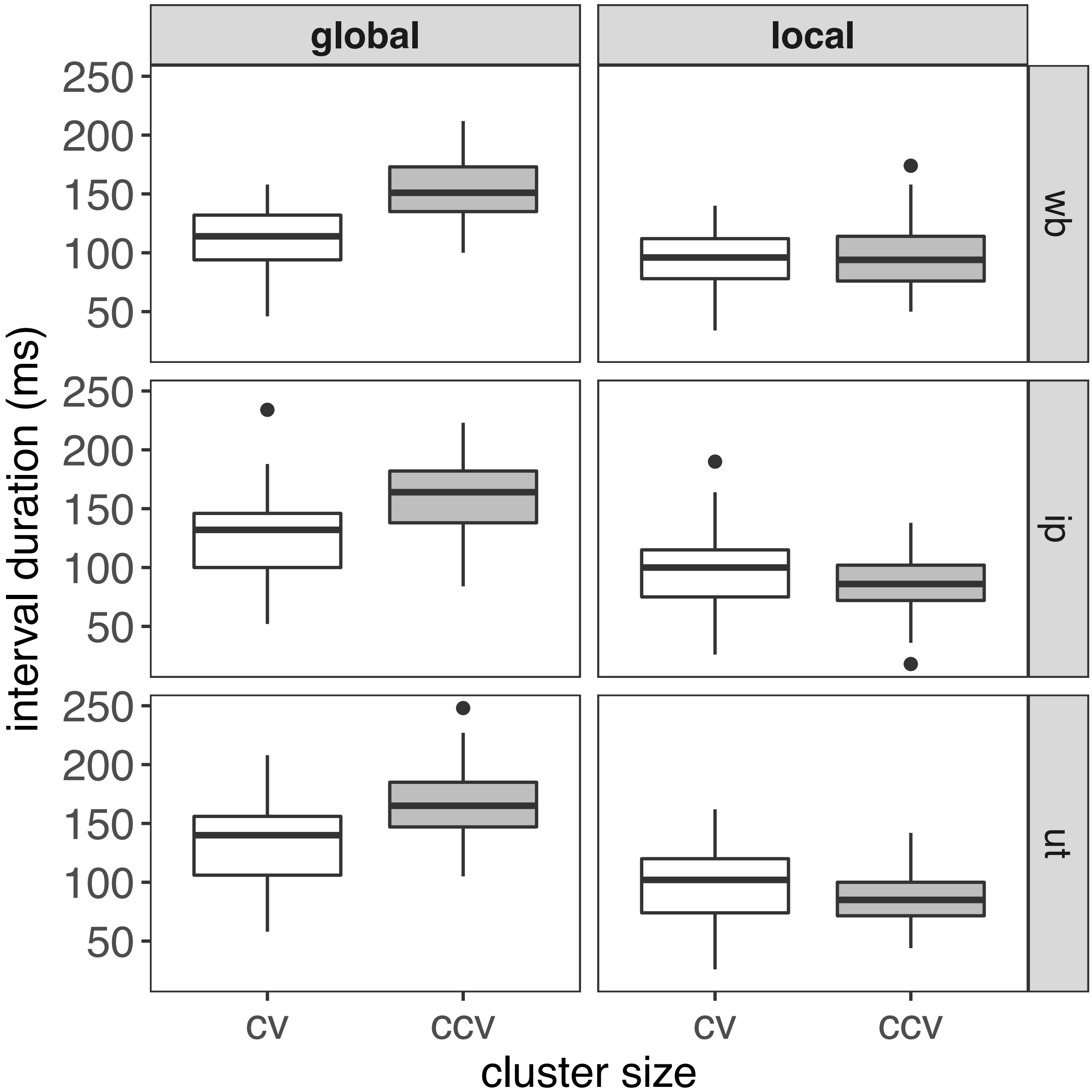

Figure 3.21: Duration (in ms) of the two intervals, global timing and local timing, for CV (white) and CCV (grey) words in three prosodic conditions. Across prosodic conditions, word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut), the difference between CV and CCV for the global timing interval is greater than the difference for the local timing interval.

timing interval (C^{tar}: estimate = 0.65 ms; C^{max}: estimate = 0.38 ms; V^{mid}: estimate = –0.54 ms). In the utterance phrase boundary condition, the global timing interval increases significantly (C^{tar}: estimate = –264 ms, $p < 0.0001$; C^{max}: estimate = –266 ms, $p < 0.0001$; V^{mid}: estimate = –247 ms, $p < 0.0001$), while there is no effect on the local timing interval (C^{tar}: estimate = –0.2 ms; C^{max}: estimate = –2 ms; V^{mid}: estimate = 2.87 ms). Figure 3.22 plots the duration of the two intervals global timing and local timing with a vowel anchor for the three prosodic conditions, word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut) across speakers, as a function of the number of consonants (CV, C#CV).

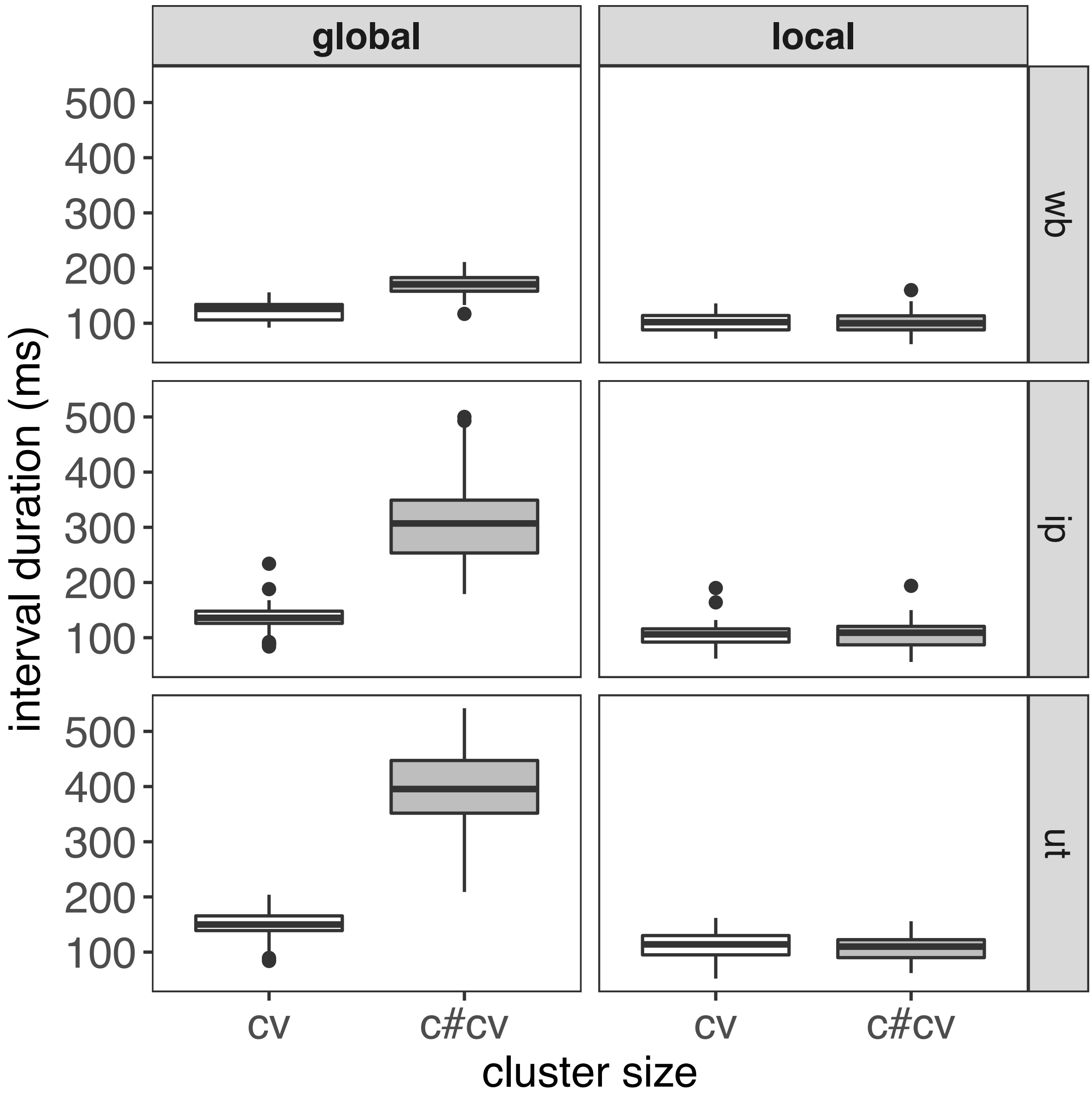

Figure 3.22: Duration (in ms) of the two intervals global timing and local timing for CV (white) and voiceless C#CV (grey) words in three prosodic conditions. Across prosodic conditions, word boundary (wb), intonational phrase boundary (ip) and utterance phrase boundary (ut), the difference between CV and C#CV for the global timing interval is far greater than the difference for the local timing interval.

To summarize, for word-initial stop-lateral clusters across CV∼CCV, local timing interval stability was observed across three different anchors and across prosodic conditions and regardless of the initial stop's voicing. When adding a voiceless stop to 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 stop-lateral clusters across CV∼CCV, local timing interval stability was found across three different anchors within each prosodic condition, word boundary, intonational phrase boundary and utterance phrase boundary. However, with increasing boundary strength, the global and local timing intervals gradually change to a similar extent as a consonant is added to the CV string. This means that with increasing boundary strength, the local timing interval stability begins to weaken and the global timing interval stability tends to improve. On the other hand, for stop-lateral cross-word clusters 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 increasing boundary strength, 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 measures such as vowel initiation, C2 lateral plateau duration, IPI-C2 lateral relation, in the previous sections will be addressed in section 3.5.

3.5 Putting it all together

The key empirical contribution of our experimental design is the registration of the spatio-temporal form of clusters composed of the same segments parsed in two contrasting syllabic organizations, e.g., word-initial /kl/ as in /klɑːgə/ versus cross-word /k#l/ as in /pɑk#lɑːgə/, and the quantification of this form in terms of several measures. The word-initial stop-lateral clusters are prototypical complex onsets in German and other Germanic languages, while 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). The results indicate clearly that this experimental design of embedding target sequences in contexts of increasing boundary strength successfully elicited different degrees in IPI in word-initial and cross-word stop-lateral clusters in German (as Byrd & Choi 2010 have done for English) and different degrees of C1 lengthening for word-initial clusters (see Byrd & Choi 2010 and Fougeron & Keating 1997 for similar effects of prosodic boundaries in English). The elicitation of different degrees of IPI and C1 duration is the crucial prerequisite that needed to be satisfied in order to investigate the effect of such phonetic perturbations on the temporal coordination of segments with different syllabic affiliations. It will be our main preoccupation in this section to consider and contrast the consequences of the elicited phonetic perturbations for segmental sequences under the two contrasting organizations. Our discussion is organized by looking at the effects as enumerated in our earlier results section, namely: C2 lateral duration, the IPI-C2 lateral plateau duration relation, vowel initiation and stability-based heuristics, each addressed in individual subsections below. For each of these, we will contrast the patterning of effects for word-initial clusters (which are parsed phonologically as complex onsets) versus cross-word clusters (which are not parsed as complex onsets), e.g., word-initial /kl/ in /klɑːgə/ versus cross-word /k#l/ as in /pɑk#lɑːgə/. What we will learn from this comparison is that syllabic organization (word-initial versus cross-word clusters) makes different predictions about the way phonetic indices respond to perturbations of phonetic parameters such as IPI or C1 lengthening. Furthermore, global organization is expressed in the form of compensatory relations between phonetic indices, while in local organization such relations are either absent or weak.

3.5.1 C2 lateral duration

This subsection is divided in two parts. Subsection 3.5.1.1 discusses C2 lateral duration in the single consonant-vowel or CV context and in clusters with different syllabic affiliation such as word-initial CCV and cross-word clusters C#CV in contexts of increasing prosodic boundary strength. In this subsection, we disentangle effects on C2 lateral duration associated with higher level prosody versus syllabic structure. It is only effects in the second case which are relevant for the purposes of this dissertation. Next, in subsection 3.5.1.2, we discuss C2 lateral duration differences as a function of stop's voicing in stop-lateral CCV clusters in contexts of prosodic boundary strength and how such differences indicated the presence of global organization.

3.5.1.1 C2 lateral duration and syllabic organization

We begin by discussing effects related to C2 lateral duration in the single consonant-vowel or CV context and in clusters with different syllabic affiliation, that is, word-initial clusters in the CCV context and cross-word clusters in the C#CV context in three different prosodic conditions. It is these different prosodic conditions that were used to induce scaling of phonetic parameters under the two different phonological organizations, CCV versus C#CV.

For word-initial stop-lateral CCV clusters, the lateral's duration does not change across CV and CCV, e.g., across /lɑːgə/, /plɑːgə/ in the control condition (word boundary condition). In contrast, for cross-word clusters, the lateral's duration is greater in C#CV than in CV in the word boundary condition, e.g., across /lɑːgə/, /knɑp#lɑːgə/. This is a first indication that the duration of the lateral in the cluster compared to the single CV context depends on syllabic affiliation, CCV (both consonants within the same syllable) versus C#CV (each consonant in a different syllable). Although our result of no lateral duration change between CV and word-initial CCV does not agree with studies on English which have shown that segments in word-initial clusters tend to be shorter than segments in isolation (Haggard 1973, Klatt 1976), in the cross-word clusters a different pattern is observed, where the lateral lengthens in the cross-word cluster compared to CV. Lengthening of the lateral from CV to C#CV is in line with the local organization pattern observed where the immediate prevocalic consonant maintains a stable relation with the vowel; a pattern which is not perturbed if the lateral lengthens in the C#CV cluster content compared to the single consonant CV context. Shortening of the lateral from CV to CCV for word-initial clusters could be seen as a result of global organization in order to bring the vowel to overlap with the cluster and as we have seen this is what was found for Spanish in Chapter 2. Even though this is not what we find in our German data, it is still the case that we do not find lengthening of the lateral in the cluster as we find for cross-word clusters in German. This means that the lateral behaves different in the cluster compared to the single CV depending on the syllabic affiliation of the cluster.

After discussing lateral duration differences between CV and clusters of different syllabic affiliation, that is, CCV versus C#CV, in the word boundary or control condition, we now turn to consider contexts of increasing prosodic boundary strength. In doing so, our aim is to disentangle effects of prosodic strengthening from effects of syllabic structure on lateral duration. In contexts of increasing prosodic boundary strength, the lateral is shorter in word-initial CCV clusters than in CV but it remains approximately the same in CV and cross-word C#CV clusters. In the contexts of increasing prosodic boundary strength, the shortening in the CCV compared to CV for word-initial clusters can be ascribed to higher prosody and not necessarily to syllabic structure. Segments which are closer to word or phrase edges lengthen more than segments found word- or phrase-internally (Pierrehumbert & Talkin 1992, Byrd 1996, Fougeron & Keating 1997, Byrd & Saltzman 1998, Cho & Keating 2001, Fougeron 2001, Keating *et al.* 2003, Byrd *et al.* 2005, Bombien *et al.* 2006, Byrd & Choi 2010). This is the case for the lateral in CV and CCV respectively. The lateral in CV occurs immediately after the prosodic boundary and thus lengthens more than the lateral in the second position of CCV as in the latter case the lateral is farther from the prosodic boundary preceding the word-initial cluster. In a similar fashion, for the cross-word clusters, the fact that the lateral does not change in duration between CV and C#CV in contexts of increasing boundary strength could also be ascribed to higher level prosody. This is because the prosodic boundary in both CV and C#CV contexts precedes the prevocalic lateral, i.e., the CV string, regardless of whether it is preceded by a C or not, is produced after a long pause due to the presence of a strong prosodic boundary. The long pause which expresses the prosodic boundary strength preceding the prevocalic lateral makes the C#CV similar to a CV. That is, the lateral's duration in "Lage" does not change in contexts of prosodic boundary strengthening regardless of whether the lateral is part of a CV "Lage" or a C#CV "knapp Lage". Therefore, the effect of syllabic affiliation on lateral duration in CV and in clusters with different syllabic affiliation can only be investigated in the control condition (word boundary condition) as seen in the beginning of this subsection, where higher prosody does not affect the lateral's duration as it does in contexts of prosodic boundary strengthening (intonational phrase and utterance phrase boundary conditions). In this context (control condition or word boundary), where we can safely exclude that any effects found are due to higher prosody, as we have seen the lateral does not change in duration between CV and CCV but it does increase in duration from CV to C#CV. This difference in the behavior of the lateral between word-initial and cross-word clusters reveals the different syllabic organization across these clusters.

Next, let us consider lateral duration within clusters of different syllabic affiliation in contexts of increasing boundary strength. It is in these contexts, as opposed to across CCV and CV, that effects of syllabic organization will become evident. Our results here indicate that the C2 lateral's duration in word-initial CCV clusters remains the same with increasing boundary strength, while the C1 stop's duration gradually lengthens. These results are to a large extent in line with Byrd & Choi (2010) who also report that the duration of the initial or C1 consonant in a CCV cluster lengthens with increasing prosodic boundary strength while the duration of the C2 consonant is not affected as much as C1's in these contexts. For cross-word clusters, we saw in our data that the lateral in the cluster lengthens in contexts of prosodic strengthening compared to the control condition (word boundary) lacking a phrase boundary, since the lateral immediately follows a stronger prosodic boundary in the former than in the latter. The lateral lengthening result, however, is not graded with increasing boundary strength, meaning that the lateral lengthens between the control condition and the individual conditions of prosodic boundary strength, but there is no difference in lateral duration between the prosodic conditions as a function of prosodic boundary strength. The fact that the lateral duration does not change between the two contexts of increasing boundary strength (intonational phrase and utterance phrase boundary conditions) is desirable for our purposes because it allows us to examine effects related to syllabic organization and not to higher level prosody. The above information on lateral duration in CCV and C#CV in contexts of increasing boundary strength is crucial, because it shows that within CCV the C2 lateral duration is not affected by prosodic strengthening and within C#CV there are contexts (intonational phrase and utterance phrase boundary) where C2 lateral duration is also not affected by prosodic strengthening. Therefore, we can argue that the temporal readjustments (related to C2 lateral duration) that will be addressed in the following subsections 3.5.2, 3.5.3, 3.5.4 reflect syllabic organization and are not a result of any differences due to prosodic strengthening.

In summary, modulations of lateral duration in our data do not have a unique cause. Sometimes these modulations are due to prosodic strengthening which affects the duration of the lateral based on its proximity to the prosodic boundary. Other times, these modulations seem to be due to syllabic structure. In this latter case, the lateral duration modulation effects reveal the presence of syllabic organization superimposed over the relevant sequences of segments.

3.5.1.2 C2 lateral duration as a function of stop's voicing in CCV

We have so far seen the effect of phonetic perturbations (IPI and C1 lengthening) on the duration of the lateral in clusters of different syllabic affiliation. We now turn to consider fine-grained changes that occur to the lateral in contexts of increasing boundary strength depending on the voicing of the initial consonant. Recall that it was the voicing context in Spanish in Chapter 2 that offered the right environment for evidence for global organization to emerge. For German, both the voicing context (short IPI for voiced stop clusters versus long IPI for voiceless stop clusters) and lengthening effects due to prosodic strengthening available in our dataset will show evidence for global organization as will be explained next. An effect of C1 voicing on the lateral duration has been found in our results on German only for word-initial clusters and not for cross-word clusters since in the latter no voiced-initial clusters exist due to final devoicing in German. In word-initial clusters, laterals following a voiced stop are shorter than laterals following a voiceless stop but only in contexts of prosodic boundary strength. The lateral duration difference as a function of stop's voicing only in contexts of increasing boundary strength and not in the control condition (word boundary) is a key fact to be addressed next. Such differences on lateral duration as a function of stop's voicing in contexts where the phonetic parameters are scaled (IPI and C1 plateau lengthen as a function of prosodic strengthening) reveals effects related to the global organization of the syllable. An alternative account would aim to link the lateral duration differences as a function of the stop's voicing to issues of perceptual recoverability of the lateral following aspirated stops versus non-aspirated stops. In what follows, we present both accounts and we argue that the perceptual recoverability account alone cannot explain the lateral duration differences observed as a function of the stop's voicing in contexts of prosodic strengthening.

Voiceless stops in German are aspirated. In the cluster context, voiceless aspirated stops do not overlap much with the following consonant, as seen by the long IPI in clusters beginning with a voiceless stop. This long IPI ensures that the aspiration is maintained and that the following segment will not be perceptually masked by it (Bombien & Hoole 2013). However, laterals following aspirated stops have been reported to be partially devoiced in German (Bombien & Hoole 2013, Hoole 2006). Since the prevocalic lateral in voiceless stop-lateral clusters is threatened by the stop's aspiration, shortening of the lateral is not possible because it would result in the lateral being perceptually masked. For this reason, the lateral in voiceless stop-lateral clusters is longer than the lateral in voiced stop-lateral clusters. In the latter, the perceptual recoverability of the lateral is not threatened (as it is not adjacent to an aspiration gesture) and thus there is more freedom for shortening of that segment as a result of the global organization in which it is part of. Recall that in Spanish, in Chapter 2, we found the opposite pattern; the lateral is shorter after a voiceless than after voiced stop. The languages appear thus to pattern in opposing ways, but as we will see in Chapter 4 order and coherence can be brought to such apparently odd patterning when we consider the differing language-particular voicing implementation systems involved. A phonological account explaining the longer lateral in voiceless than in voiced stop-lateral clusters considers aspiration as a gesture which, just like any other gesture, has a temporal and spatial dimension. This glottal gesture intervenes between the stop and the lateral and therefore it is closer to the boundary before the initial stop than the lateral. As is known (Byrd & Saltzman 2003, Byrd & Choi 2010), larger changes in the duration of gestures occur for segments that are closer to the boundary as in the case of voiced stop-lateral clusters where IPI is short and no extra aspiration gesture intervenes between the stop and the lateral. It is in this context of close proximity where we do find that the lateral is shorter than in voiceless stop-laterals. It appears thus that the case of **voiceless** stop-laterals in German is similar to the case of **voiced** stop-laterals in Spanish in that the IPI between stop and lateral is large (and of similar duration across the two languages; compare IPI of 36 ms in German and IPI of 38 ms in Spanish) and we find that the lateral is longer than in its occurrence in the respective opposite value of voicing clusters, that is, voiced stop-laterals in German and voiceless stop-laterals in Spanish. In sum, what appear to be divergent patterns of how lateral duration is modulated in the two languages turn out to be effects related to global organization which is manifested differently across these two languages.

We now come to the crux of the argument for why the lateral duration effects cannot be seen as motivated exclusively by perceptual recoverability factors and why additional constraints related to the imposition of a global organization seem to be involved. As mentioned already, the lateral duration difference between voiced and voiceless stop-lateral clusters emerges only when C1 and IPI lengthen under prosodic strengthening. The fact that it emerges only in cases of increasing boundary strength (intonational phrase and utterance phrase boundary) and not in the word boundary condition (no phrase boundary) excludes the perceptual recoverability account as the only force responsible for this effect. This is so because the aspiration gesture of the voiceless stop is present in all three conditions (word boundary, intonational phrase and utterance phrase boundary) and it stretches to the following lateral regardless of prosodic condition. Therefore, even though the perceptual recoverability of the lateral is presumably threatened also in the word boundary condition, we do not find any difference in the lateral duration as a function of the stop's voicing. This indicates that the effects we see (longer lateral after a voiceless than after a voiced stop) in contexts of prosodic strengthening cannot be the result of exclusively recoverability considerations.

Let us now see how this difference in lateral duration as a function of stop's voicing in contexts of prosodic strengthening connects to other effects such as the shift of the lateral to the vowel. It is in such contexts of prosodic strengthening that we also observe the lateral shifting towards the vowel and such shifting is found to a greater extent in the voiceless than in the voiced cluster context (hence the greater local timing interval shortening in voiceless than in voiced stop-lateral clusters to be addressed in 3.5.4). Since the lateral is shifted towards the vowel in the voiceless stop-lateral context, it is no longer necessary to shorten extensively as a result of a global organization, because shifting the lateral towards the vowel brings the vowel already closer to the cluster. It is, however, necessary to shorten the lateral in the voiced stop-lateral context where the lateral is not shifted extensively towards the vowel and thus shortening its duration serves as a readjustment that again brings the vowel close to the cluster as required by global organization. The above is an account of why the lateral shortens in the voiced but not in the voiceless stop contexts that crucially relies on the degree of lateral shift to the vowel as a result of a global organization imposition.

3.5.2 IPI – C2 lateral compensatory relation

We now turn to see when and how a dependency between IPI and C2 lateral duration reveals effects of syllabic organization. Only word-initial CCV clusters can be investigated with respect to this relation because, as explained in section 3.4.3, cross-word clusters do not offer a context where IPI can be meaningfully defined. This is so because the presence of the pause between the two consonants in C#CV does not make it possible to define a notion of IPI that can be measured independently of the pause.

Looking at the relation between IPI and C2 lateral plateau duration in word-initial stop-lateral CCV clusters in the control condition (word boundary), a compensatory relation is observed such that as IPI increases C2 lateral duration decreases. This compensatory relation is a first indication of global organization that serves to bring the vowel to overlap with the cluster. Specifically, when the lag between the plateaus of the two consonants, what we call IPI, increases, shortening of the lateral 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 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 control condition. However, in the other two prosodic conditions, 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. It seems that the prosodic strengthening (that is, in terms of its temporal consequences, lengthening) induced on the duration of the first C in CCV as well as in the duration of the IPI between the two consonants freeze the extent of variability in these two variables, thus precluding the manifestation of compensatory relation. Emphatically, however, this does not mean that there are no indications for global organization in the other non-control prosodic conditions. To the contrary, we find strong and perhaps even stronger indications of global organization in these contexts, to be addressed next in subsections 3.5.3 and 3.5.4.

3.5.3 Vowel initiation

We first discuss 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, in three prosodic contexts. For word-initial clusters, the CV lag which describes 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); i.e., there is no difference in vowel initiation from CV to CCV in the control condition. However, with increasing boundary strength (intonational phrase and utterance phrase boundary), the vowel starts progressively earlier with respect to the target of the prevocalic lateral in CCV compared to CV. Consider now cross-word C#CV clusters. With respect to vowel initiation in cross-word C#CV clusters compared to single CV sequences, the CV lag is longer in C#CV than in CV across and within prosodic conditions. This means that the vowel occurs later with respect to the prevocalic lateral in the cluster C#CV than in CV. This is exactly the opposite of what happens between CV and CCV where the vowel occurs earlier with respect to the prevocalic lateral in the cluster CCV than in CV across prosodic conditions. The earlier vowel initiation from CV to CCV in contexts with increasing boundary strength can be attributed to effects of segmental duration varying due to higher prosody but this does not seem to be the case for the vowel initiation pattern observed in CV and C#CV as will be explained next. As mentioned in 3.5.1, the lateral in CCV is shorter than the lateral in CV in contexts of increasing boundary strength possibly due to the proximity of the lateral to the prosodic boundary. Such a lateral shortening in CCV compared to CV can result in an earlier vowel initiation in CCV without this early vowel initiation being related to a global organization of the syllable. For cross-word clusters, the fact that vowel initiation occurs later in the cluster than in the single CV in contexts of prosodic strengthening does not seem to relate to lateral duration as a result of prosodic boundary strength, because the lateral is only slightly but not significantly longer in the cluster than in the single CV context. Furthermore, the lateral in C#CV is longer than the lateral in CV also in the control condition where no phrase boundary or pause precedes the CV string. This shows that lateral duration is not affected by prosodic strengthening across CV and C#CV and thus later vowel initiation in C#CV than in CV can be attributed to the syllabic structure. According to this syllabic structure, the onset consists of only one consonant in both CV and C#CV and either no vowel initiation difference across CV∼C#CV or later vowel initiation in the cluster (which is what we find) is in line with the local organization where the vowel is timed only with the immediate prevocalic consonant. The fact that the vowel initiation pattern cannot be examined across CV and CCV in contexts of prosodic strengthening due to the lateral being affected in different ways as a result of prosodic strengthening does not mean that there are no indications for global organization. We do find such indications when looking at within CCV clusters as will be explained immediately below.

Let us now consider vowel initiation with respect to the target of the prevocalic lateral within clusters with different syllabic affiliation in contexts of increasing boundary strength. For word-initial clusters, the CV lag which describes the lag between vowel initiation and target of the prevocalic consonant, decreases with increasing boundary strength.[2] For cross-word clusters, the CV lag slightly increases or remains approximately the same with increasing boundary strength.[3] Thus, the vowel initiation pattern

[2]Refer to Figure 3.14, Figure 3.15, Figure 3.16 and specifically to the distance between the triangle indicating vowel initiation and the bottom white dot indicating the target of the prevocalic lateral.

[3]Refer to Figure 3.18 and the distance between the triangle indicating vowel initiation and the bottom white dot indicating the target of the prevocalic consonant in three prosodic conditions (wb, ip, ut) on the x-axis.

is markedly different between clusters of different syllabic affiliation in contexts of prosodic strengthening. The reason for the vowel initiation differences seems to directly relate to syllabic structure and not to higher prosody. For word-initial clusters, we do not see any difference in lateral duration as a function of prosodic condition, as discussed in 3.5.1, and thus the earlier vowel initiation with increasing boundary strength cannot be ascribed to any effect related to changes in the lateral duration due to higher prosody. For cross-word clusters, across C#CV, there is a slight increase in the lateral duration from word boundary to intonational phrase boundary condition but no change in the utterance phrase boundary condition. Thus, although there is an increase in the lateral duration, the effect is not gradient as a function of prosodic boundary strength. The fact that the lateral lengthens in the intonational phrase boundary condition could be taken as an argument that any difference in the vowel initiation pattern between the control condition and the intonational phrase boundary condition could be a result of the lateral's lengthening in the latter. However, no vowel initiation difference was observed between these two conditions. Therefore, we can exclude the possibility that differences in vowel initiation patterns between clusters of different syllabic structure can be related to segmental lengthening effects due to higher prosody. Instead, it seems that these vowel initiation differences relate to syllabic structure and specifically the two different global versus local organizations with the former corresponding to the CCV and the latter to the C#C context.

For word-initial clusters, vowel initiation occurs earlier with increasing boundary strength in order for the vowel to overlap with the cluster. 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. As a result of this, 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 where planned independently of the other segments, such compensatory adjustments are not expected. In other words, the fact that the lag between vowel initiation and target of the prevocalic lateral decreases with increasing boundary strength 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 such as in the cross-word clusters). For cross-word clusters, where no organizing principle prescribes that the vowel overlaps with the cluster as a unit, there is no difference in vowel initiation patterns even when expansion occurs (IPI lengthening) due to prosodic boundary strength. Thus, the vowel initiation patterns differ as a function of syllabic structure and these differences become apparent as phonetic parameters are scaled as in contexts of increasing IPI.

3.5.4 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 contexts of increasing boundary strength.

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, but with increasing boundary strength the local timing interval becomes less stable while the global timing interval becomes more stable. Let us first disentangle effects of prosodic strengthening from effects of syllabic structure that pertain to this result. As mentioned in 3.5.1, the lateral is shorter in CCV than in CV in contexts of prosodic boundary strength presumably due to different lengthening effects of the lateral based on its proximity to the prosodic boundary; the lateral is shorter in CCV than in CV just because the former is farther from the prosodic boundary than the latter. This lateral shortening from CV to CCV affects potentially 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 increasing boundary strength. Thus, the improved global timing interval stability in contexts of increasing boundary strength may not be driven exclusively by syllabic organization. However, the worsening of the local timing interval stability in contexts of increasing boundary strength (intonational phrase and utterance phrase boundary conditions) is driven by syllabic organization in contexts of increasing boundary strength. This is so because as opposed to the global timing interval, the local timing interval and its (in)stability is not affected by any change in duration of the prevocalic lateral as will be explained next. In our results, the local timing interval shortens from CV to CCV progressively to a larger extent with increasing boundary strength. Lateral shortening from CV to CCV in contexts of prosodic boundary strength 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 (all else being equal). Just shortening the lateral without shifting it towards the vowel would result in lengthening of the local timing interval. Therefore, the crucial parameter which relates to the instability of the local timing interval is shifting the lateral towards the vowel which occurs with increasing boundary strength, i.e., as IPI and C1 lengthen. It is when we introduce variability via the prosodic boundary manipulation that the local timing interval ceases to be stable. The latter must be considered as an effect of syllabic organization, because a markedly different pattern is observed for cross-word clusters. Specifically, for cross-word clusters, 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) with increasing boundary strength, while the global timing interval changes to a great extent. As we have reviewed in 3.5.1, there are no lateral duration differences between CV and C#CV in contexts of prosodic boundary strength 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 clusters.

Now we turn to consider the temporal readjustments mentioned above for word-initial clusters only (no cross-word clusters) and see how these readjustments result in the observed temporal coordination patterns across prosodic conditions. We will consider the voiced and voiceless stop-lateral clusters separately, because their phonetic properties exhibit similarities but also differences and it is these differences that reveal effects related to the global organization of word-initial complex onsets in German. Across prosodic conditions, the voiced and voiceless stops are of similar duration, the IPI is larger in the voiceless than in the voiced stop context, and the lateral is shorter in the voiced than in the voiceless stop context. The schema in Figure 3.23 illustrates the temporal adjustments that occur when an initial voiced or voiceless stop is added to the lateral-vowel string according to the data. Across prosodic conditions, when adding a voiced stop with a short IPI to the lateral-vowel string, the lateral shortens to a large extent but it is only minimally "pushed" towards the anchor. Compare the single lateral /l/ in Figure 3.23a versus the prevocalic lateral in the cluster in 3.23b. The lateral in Figure 3.23b is substantially shorter presumably due to effects related to higher prosody but it is not relocated towards the anchor to a great extent compared to the single lateral in Figure 3.23a. Across the CV and CCV (refer to Figure 3.23a and b), the release landmarks (indicated as re in the Figure) in /lV/ and voiced /ClV/ are better aligned than the c-center (indicated as cc in the Figure) landmarks. In effect, for the CV, CCV pairs where the first C is voiced, the local timing interval is comparable across CV, CCV, while the global timing interval is substantially greater in CCV than in CV. Compare this configuration to that for CV, CCV pairs when the first C is voiceless. When adding a voiceless stop with a large IPI to the lateral-vowel string, the following temporal readjustments occur: the lateral in CCV is "pushed" towards the anchor and it shortens compared to the CV context but the shortening is of a lesser extent than in the voiced stop-lateral context. Compare the single lateral /l/ in Figure 3.23a versus 3.23c in the cluster. The lateral in Figure 3.23c is shorter presumably due to effects related higher prosody and it occurs later than the single lateral in Figure 3.23a. Therefore, shortening of the prevocalic lateral and shift of the lateral towards the anchor in CCV compared to CV results in misalignment of the release landmarks (indicated as re in Figure 3.23). At the same time, however, the c-center landmarks (indicated as cc in the Figure) are also misaligned because the degree of IPI in voiceless stop-lateral clusters along with the degree of lateral shortening do not suffice to align the c-center landmarks. However, the important point is that it is in the voiceless stop-lateral context that we observe improved global timing interval stability and worse local timing interval stability with increasing boundary strength because the large IPI shifts the lateral towards the vowel leading to substantial shortening of the local timing interval from CV to CCV (and hence worse stability of that interval). Such a lateral shift does not occur in the voiced stop-lateral context and thus clear local timing interval stability is observed with increasing boundary strength. The degree of shortening of the local timing interval depending on the stop's voicing relates to the duration of the lateral which is shorter after voiced than after voiceless stops, as will be addressed in the following paragraph.

The fact that local timing interval stability is consistently found for voiced stop-lateral clusters does not mean that these clusters do not show evidence for global organization. There are reasons why the local timing interval does not decrease substantially in the voiced stop-lateral clusters (hence, why local timing interval stability is consistently found) while it does decrease in voiceless stop-lateral clusters. These reasons are related with the variability that we introduced in scaling phonetic parameters and the coincidental phonetic properties of the clusters which are different between voiced and voiceless stop-laterals. Specifically, the variability that we introduced by increasing IPI and C1 duration in the cluster does not affect voiced and voiceless stop-lateral clusters the same way. Increasing IPI and C1 duration for voiced stop-lateral clusters which originally had a short IPI, was not sufficient to result in a substantial decrease of the local timing interval. It was, however, sufficient to result in a substantial decrease of the local timing interval for voiceless stop-lateral cluster which originally had a long IPI. Still, even voiced

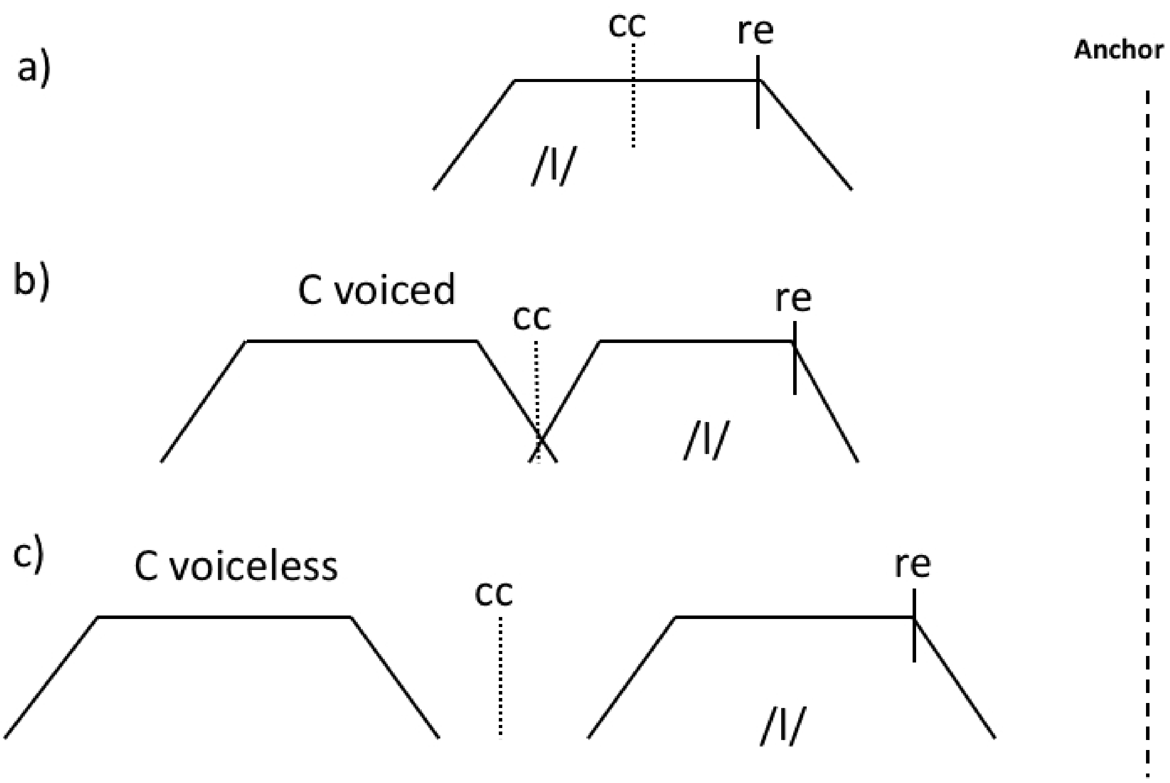

Figure 3.23: Temporal configurations of (a) singleton lateral /l/, (b) voiced stop-/l/ cluster and (c) voiceless stop-/l/ cluster. The trapezoids represent consonant gestures. The landmark marked as cc corresponds to the c-center and re to the right-edge of the singleton consonant (C) or the consonant cluster (CC). Across the singleton lateral-vowel sequence /lV/ and the cluster (stop-/lV/), the re landmarks are better aligned than the cc landmarks in /l/ - voiced C/l/ (a, b). The re landmarks are also slightly better aligned than the cc landmarks in /l/ - voiceless C/l/ (a, c), but both cc and re landmarks are more misaligned in /l/ - voiceless C/l/ (a, c) than in /l/ - voiced C/l/ (a, b).

stop-lateral clusters show evidence for a global organization which can be seen in that the lateral is shorter in voiced than in voiceless stop-lateral clusters and in that the lateral is shifted more to the vowel in voiceless than in voiced stop-lateral clusters. 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 with the clusters. Such shortening is one 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 to a great extent; this is also a signature of global organization that brings the vowel to overlap with the cluster in the case of voiceless stop-lateral clusters. The above fine-grained differences between voiced and voiceless stop-laterals serve as an example of what the current dissertation is contributing: considering other measures, other than just stability-based heuristics, proves highly informative. It is via such joint consideration of measures when evidence for global organization emerges. If diagnosis of syllabic organization was based exclusively on the stability-based heuristics, the results on voiced and voiceless stop-lateral clusters would indicate very little or even contradictory evidence for syllabic organization in German due to the exceptional stability pattern of some clusters (cf. Brunner *et al.* 2014, Pouplier 2012). Moreover, consideration of different measures (such as prevocalic consonant duration and lateral shift towards the vowel) does not only bring out evidence for global organization; it also enables one to understand why the interval stabilities pattern the way they do. For instance, the local timing interval does not change across lateral-vowel and voiced stop-lateral clusters to the same extent as in voiceless stop-laterals because the lateral is shorter in voiced than in voiceless stop-lateral clusters thus enabling the vowel to establish proximity to the cluster due to this reduced lateral duration.

The next subsection summarizes all the effects addressed so far and how these when considered together provide evidence for a given syllabic organization.

3.5.5 Overall summary

We now summarize the effects for clusters of different syllabic affiliations across prosodic conditions. As phonetic parameters are scaled, e.g., with increasing IPI or C1 stop duration (the latter only in CCV, not in C#CV clusters), our results indicate that the temporal readjustments occurring in the same sequences of segments differ depending on the syllabic organization of the sequence, CCV (word-initial complex onset) versus C#CVs (word-initial simplex onset preceded by a word-final consonant) where the two consonants involved are the same across the two sequences. Within word-initial CCV clusters, vowel initiation occurs progressively earlier as a function of increasing prosodic boundary strength, while for C#CVs, vowel initiation either does not change or occurs progressively later as a function of increasing prosodic boundary strength. The earlier vowel initiation in CCV is a response to the expansion of the syllable due to the prosodic boundary strengthening, which causes C1 and IPI to lengthen. The greater the expansion of the syllable, the greater the pressure in the CV subpart of the CCV to compensate, as seen by the earlier vowel initiation, for that increased C1 and IPI duration. In other words, this 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). In contrast, for cross-word C#CV clusters, which instantiate the other syllabic organization, no earlier vowel initiation is observed in any of the prosodic conditions, indicating no interdependency between local CV timing and the lag between the plateaus of the two consonants.

As for the stability-based heuristics, as IPI increases, our results show that the local timing interval is perturbed to a large extent, thus ceasing to be the most stable interval for word-initial clusters, while it remains remarkably stable for cross-word clusters. Within word-initial clusters, the compensatory effects such as earlier vowel initiation due to pressure in CV substring and shortening of C2 in response to IPI lengthening (due to the IPI-C2 compensatory relation) ripple through to other dependent measures as well. As we have seen, the compression of the CV substring in CCV due to expansion of IPI and C1 plateau duration also results in compression or shortening of the local timing interval. For this reason, the local timing interval becomes less stable as it decreases from CV to CCV. Furthermore, local timing interval shortening is greater for voiceless than for voiced stop-lateral clusters, meaning that the lateral is shifted more towards the vowel when the initial stop is voiceless than voiced. This finding seems to relate to the lateral's duration; namely, that the lateral is shorter in voiced than in voiceless stop-lateral clusters. Therefore, a shorter lateral co-occurs with less lateral shift to the vowel, while a longer lateral co-occurs with more lateral shift to the vowel. A shorter lateral does not need to shift to a large extent towards the vowel, while a longer lateral needs to shift more towards the vowel in order to bring the vowel closer to the cluster. Such a relation between lateral duration and degree of shift in word-initial clusters, along with the earlier vowel initiation, the IPI-C2 lateral relation and the reduced local timing interval stability, are signatures of a global organization bringing the vowel to overlap with the cluster. Crucially, a cross-word C#CV cluster is not fleshed out in the same way as a word-initial CCV cluster. Across C#CVs, with increasing boundary strength, our results show that vowel initiation either does not change or occurs progressively later and the local timing interval stability remains remarkably stable.

In the context of prior work which usually considers a limited number of measures (e.g., see Tilsen *et al.* 2012 who use just one measure or Marin 2013 who uses two measures), revealing the systematic but at times subtle effects that syllabic organization has in our data requires the joint consideration of several measures and the disclosure of compensatory relations which offer telling indicators of global organization. Crucially, the strength of the demonstration for a global organization stands clear even if some of the methods used in prior work to diagnose syllabic organization (namely, stability-based heuristics) fail to show the expected patterns in some of our conditions, e.g., as in the word boundary condition where the 'wrong' stability pattern is met.

3.6 Conclusion

The present Chapter addressed the relation between language-specific syllable structure and spatio-temporal coordination between segments in German. Specifically, we investigated the effect of phonetic perturbations related to prosodic effects on consonant clusters with different syllabic affiliations. We did so by effectively eliciting different degrees of IPI and C1 stop lengthening in stop-lateral clusters under various degrees of prosodic strengthening. We studied stop-lateral word-initial clusters /bl, pl, gl, kl/ and cross-word clusters /pl, kl/ in a tense vowel context produced by five native speakers of German. The main finding can be summarized by saying that the effect of phonetic perturbations on the temporal coordination patterns depends on the syllabic organization superimposed on these clusters. Changing phonetic parameters in consonant clusters leads to different inter-segmental temporal coordination patterns depending on whether these clusters are word-initial complex onsets or cross-word clusters which do not combine to form a syllable onset. In short, syllabic organization is reflected in the spatio-temporal patterns of the segments partaking in this organization. However, to reveal this dependence, 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) and as we have seen in our results, such stability-based heuristics are unreliable in consistently diagnosing the syllabic organization in our datasets. The crucial novelty in our study is that the Shaw *et al.* (2011) conclusions on stability-based heuristics form complex onsets were based on simulated data, whereas our results here are based on experimental data (Shaw *et al.* 2011 had experimental data for simplex but not for complex onsets). We have, in other words, demonstrated for the first time experimentally that the distinctions between simplex and complex onsets which Shaw *et al.* (2011) predicted are attested in our German datasets. The basic difference between our current study and Shaw *et al.* (2011) is that the phonetic parameters that we manipulated relate to lengthening effects due to prosodic strengthening, while the phonetic parameters manipulated in Shaw *et al.* (2011) relate to compression effects such as shortening of the prevocalic consonant. Regardless of the different phonetic parameters manipulated, the conclusion is the same: as phonetic parameters are scaled, syllabic organization makes different predictions about how stability-based heuristics respond to such perturbations. Moreover, apart from looking at stability-based heuristics, we extended our investigation to other measures and specifically to how these other measures as well their relation to one another are affected as phonetic parameters are scaled (again, as a function of syllabic organization). As we have seen, joint consideration of several measures points to the conclusion that syllabic organization is not reflected only in the stability-based heuristics (indeed, exclusive attention to these heuristics is unreliable or points to the wrong conclusions) but also in other measures such as the duration of the prevocalic consonant depending on the voicing of the initial stop, vowel initiation in the cluster, compensatory effects between IPI and duration of the prevocalic consonant. Specifically, as we have seen in our German data, increasing the lag between two consonants (IPI) and C1 stop's duration in word-initial stop-lateral clusters leads to earlier vowel initiation within CCV, and reduces 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 in contexts where 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. Other effects observed in CCV stop-laterals when IPI and C1 stop duration lengthen and which point to global organization are the lateral duration differences as a function of C1 stop's voicing along with the degree of lateral shift towards the vowel in these contexts. A shorter lateral in voiced stop-lateral clusters co-occurs with a minimal shift of the lateral towards the vowel, while a longer lateral co-occurs with a greater shift of the lateral towards the vowel. All the above are temporal readjustments that occur in word-initial clusters and serve to bring the vowel to overlap with the cluster in order to establish a relation with it as prescribed by a global organization. The effects observed are markedly different when we turn to cross-word stop-lateral clusters. In the latter, increasing the lag between two consonants in cross-word stop-lateral clusters leads to no change in vowel initiation across C#CVs and robust local timing interval stability across CV, C#CV. 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) crucially modulate the spatio-temporal coordination of the same segmental material.

To conclude, our study further informs research on the relation between syllabic organization and phonetic indices. 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. 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, IPI variability. It is rather that as IPI varies, different syllabic organizations respond to that variability differently. In the complex onset organization, as IPI increases we have found that the local timing interval ceases to be 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 IPI or the lag between consonantal plateaus increases, we have found that 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 (pleiotropy) and the disclosure of compensatory relations which offer telling indicators of global organization.

Chapter 4

Conclusion

In this final Chapter of the dissertation, we begin by a summary of the main findings from the different languages. After this summary, a synthesis of our findings across the range of data we have considered is presented. This synthesis leads to new perspective on the relation between syllabic organization and phonetics. In presenting this new perspective, we highlight the two key parts of our main claim: the pleiotropy of phonetic indices expressing syllabic organization and the expression of global organization via compensatory relations between phonetic parameters.

One of the consequence of the new perspective is that we can pinpoint, we believe for the first time, more precisely the role of language-particular phonetics in the expression of the phonological concept of syllabic organization. We will specifically argue that the language-specificity of the phonetics (that is, the fact that a [p] is a different phonetic entity in Spanish versus German) does not entail that all aspects of the temporal organization can be reduced to properties of those language-specific phonetics. Rather, phonological organization and specifically in our case syllabic organization is crucially involved in how the phonetic substance is expressed for any given combination of segments within any given language. Phonetics is a key part of the picture but not the whole picture. The same or similar sequences of segments across two languages with different syllabic organizations or within the same language but parsed under different syllabic structures are phonetically expressed differently. The Chapter closes by pointing to some directions for future work that should be helpful in furthening the proposed perspective and advancing our understanding of the relation between phonological organization and phonetics.

The Chapter is organized as follows. Subsection 4.1 provides a summary and the main findings from Chapter 2 on Spanish and Moroccan Arabic. Subsection 4.2 provides a summary and the main findings from Chapter 3 on German. The relation between language-specific phonetics and syllabic structure is taken up in subsection 4.3.

Subsections 4.4 and 4.5 present more specifically the manifestations of global and local organization respectively. Subsection 4.6 concludes with directions for further research.

4.1 Summary and findings from Chapter 2

For Spanish, articulatory data of six native speakers of Central Peninsular Spanish were collected using the Carstens AG501 articulograph. Our stimuli words consisted of word-initial stop-lateral clusters where the initial stop is voiced and voiceless /bl, pl, gl, kl/ and word-initial stop-rhotic clusters where the initial stop is voiceless /pɾ, kɾ, tɾ/. The paired one-consonant sonorant-initial words begun with a lateral or a rhotic such that in any CV∼CCV pair the prevocalic consonant remained the same across CV∼CCV words (e.g., /lato/∼/plato/, /ɾato/∼/pɾato/). Using voiced and voiceless stop-lateral clusters, we examined how the modulations in the CCV string caused by the implementation of the voicing contrast in Spanish affect the spatio-temporal organization of the CCV sequence. In addition to stop-lateral clusters, we also considered stop-rhotic clusters. In Spanish, the latter clusters provide a context where the lag between the two consonants is even greater than for stop-lateral clusters. We thus also examined how global organization is achieved in clusters with very large interplateau interval, i.e., when intra-cluster timing hinders the vowel to approximate the c-center of the clusters as prescribed in a global organization. For Moroccan Arabic, articulatory data of four native speakers of Moroccan Arabic were collected using the Carstens AG500 articulograph. We compared Spanish to Moroccan Arabic because the latter imposes a different syllabic organization from that of Spanish for the same clusters but has the same voicing system as SP. For example, clusters like /bl, kl/ exist as word-initial clusters in both Spanish and Moroccan Arabic, share the same implementation of voicing and they are characterized by an open transition. But the phonological parse of these clusters is different across these two languages – they are licit syllable onsets in Spanish but cannot form syllable onsets in Moroccan Arabic where syllables maximally have one consonant as an onset. A comparison between Moroccan Arabic and Spanish word-initial clusters thus allows us to examine the effects of syllabic organization on the temporal coordination between segments which share similar phonetic properties.

The overall profile of our results can be summarized by saying that adding a consonant to the left of a CV to obtain a CCV results in a reorganization of the temporal structure of the internal CV in the onset clusters investigated in Spanish but not in Moroccan Arabic. Given the phonetic properties of the segments and their local sequencing in Spanish, syllabic organization imposes different degrees of lateral shortening, different degrees of relocation of the lateral in the cluster, a compensatory relation between IPI and lateral duration, and vowel initiation around or close to the c-center of the cluster to the extent possible. Overall, we find that even though some indices that have been used in prior work to diagnose syllabic structure do not show the expected patterns (for example, the usually employed global timing interval stability pattern which fails to show the expected result for the voiced stop-lateral clusters in Spanish), there are other indices which do reflect a global organization. Crucially, these adjustments are not present or are less present in Moroccan Arabic, a language which (freely admits clusters but) does not permit complex onsets. Thus, in /kl, bl/, Moroccan Arabic shows no shortening of the lateral, vowel onset initiation tied to the immediately preceding segment /l/, a weak compensatory relation between IPI and lateral duration, and no change in the relative timing of the /lV/ from /lV/ to /klV/. These properties express an organization according to which the vowel is locally timed with the prevocalic consonant of the cluster.

With the comparison between Spanish and Moroccan Arabic, we have demonstrated that syllabic organization is reflected in a number of phonetic indices, which corresponds to the pleiotropy part of our main claim. The second part of our main claim in this dissertation concerns the way in which syllabic organization is expressed. Our results demonstrate that syllabic organization is expressed via compensatory relations between phonetic parameters. In the general form of these relations, one phonetic parameter compensates as a response to a change of another phonetic parameter. An example is seen in the IPI-C2 compensatory relation which was one indication of a global organization in Spanish stop-lateral clusters: in this relation, increasing the lag between the two consonants in a C1C2V is compensated by a reduction in the duration of the prevocalic lateral so as to maintain the proximity of the vowel to the entire cluster. In Moroccan Arabic, in contrast, the local perturbation of increasing the lag between two consonants in the cluster remains local without affecting the duration of the prevocalic consonant. Such a lack of compensatory relations between phonetic parameters of the segments that partake in a given syllabic organization indicates that these segments are locally organized.

4.2 Summary and findings from Chapter 3

For German, articulatory data of five native speakers of German were collected using the Carstens AG501 articulograph. Word-initial CCV clusters /bl, pl, gl, kl/ and cross-word C#CV /pl, kl/ stop-lateral clusters followed by a low tense vowel were recorded. The cross-word voiced stop-lateral /bl, gl/ clusters were not registered as voiced stops do not occur word-finally in German due to final devoicing (Wiese 1996). The paired single consonant-initial words begun with a lateral such that in a CV∼CCV and CV∼C#CV pair the prevocalic consonant remained the same across CV∼CCV and across CV∼C#CV words (e.g., /lɑːgə/∼/plɑːgə/, /lɑːgə/∼/knɑp#lɑːgə/). Each stimulus word was recorded in three prosodic conditions with varying boundary strength preceding the stimulus word. For the cross-word clusters, the prosodic boundary is between the two consonants of the cluster and thus it precedes the second word. The prosodic boundary strength in our experimental design is such that it increases progressively from no phrase boundary (this is what we refer to interchangeably as simply the word boundary or control condition), to intonational phrase boundary to utterance phrase boundary. The word boundary condition was elicited by embedding any given stimulus word in the carrier phrase "Ich sah ____ an." with the stimulus word location indicated by the "____." The intonational phrase boundary condition was elicited by embedding any given stimulus word in the carrier phrase "Als ich Tom sah, ____ sagte er sofort." The utterance phrase boundary condition was elicited by embedding the stimulus word in the carrier phrase "Zunächst sah ich Anna. ____ sagte sie." Via the different prosodic conditions, we successfully elicited different degrees in interplateau interval in word-initial and cross-word stop-lateral clusters and different degrees of C1 lengthening for word-initial clusters. The elicitation of different degrees of IPI and C1 duration was the crucial prerequisite that needed to be satisfied in order to investigate the effect of such phonetic perturbations on the temporal coordination of segments with different syllabic affiliations.

The main finding can be summarized by saying that the effect of phonetic perturbations on the temporal coordination patterns depends on the syllabic organization superimposed on these clusters. Changing phonetic parameters in consonant clusters leads to different inter-segmental temporal coordination patterns depending on whether these clusters are word-initial complex onsets or cross-word clusters which do not combine to form a syllable onset. For word-initial complex onsets, increasing the lag between two consonants (IPI) and C1 stop's duration leads to earlier vowel initiation within CCV and reduced local timing interval stability. Furthermore, a compensatory relation between IPI and C2 lateral plateau is also observed within CCV in the no phrase boundary (or word boundary) condition where the variability of each of these two parameters was sufficient in order for such a relation to emerge. This relation is such that as IPI increases, C2 lateral duration decreases. Moreover, fine-grained differences on lateral duration and lateral shift to the vowel as a function of C1 stop's voicing emerged. A shorter lateral in the voiced stop-lateral clusters co-occurs with a minimal shift of the lateral towards the vowel, while a longer lateral co-occurs with a greater shift of the lateral towards the vowel. All the above effects are readjustments that occur in word-initial clusters and serve to bring the vowel to overlap with cluster in order to establish a relation with it as prescribed by a global organization.

In contrast, for cross-word clusters, increasing the lag between two consonants leads to no change in vowel initiation across C#CVs. Unlike in the case of word-initial complex onsets, for these cross-word clusters local perturbations remained local without affecting inter-segmental temporal coordination in the rest of the CV string. Moreover, robust local timing interval stability across CV∼C#CV was observed even when the lag between the two consonants in C#CV increased. As for the prevocalic lateral's duration, it did not change with increasing the lag between the two consonants. All the results above for the cross-word clusters indicate that local perturbations, such as increasing the lag between two consonants, remained local without affecting inter-segmental temporal coordination in the rest of the CV string. This is exactly what we expect. In these two sequences (CV, C#CV) the syllable consists of only the prevocalic consonant and the vowel to the exclusion of the consonant preceding the prevocalic consonant.

In sum, our results clearly demonstrate that syllabic organization crucially modulates the effects that scaling of phonetic parameters (as in increasing the lag between the two consonants of the cluster) has on inter-segmental temporal coordination.

4.3 Language-specific phonetics and syllabic structure

We now turn to a synthesis of our results, aimed at demonstrating that, as a consequence of our work, we can better characterize the role of language-specific phonetics in the relation between syllabic structure and its manifestation in the speech signal. To illustrate this point, we begin by a brief review of the key driving ideas on the topic syllable structure and its phonetics, starting with the origins of the research program on the relation between syllabic organization and spatio-temporal dimensions and continuing up to the most recent studies. The latter studies highlight the cluster-specificity of the results and, at times, hint that it is conceivable that finding consistent manifestations of syllabic organization in the speech signal may not be feasible. This naturally leads to asking:

does phonological syllabic organization play any role at all or can all observed effects be simply derived (as has been suggested at times) by phonetic considerations? Our work offers clear evidence that the latter part of this question can be answered in the negative.

Key early studies (Browman & Goldstein 1988, Byrd 1995) expressed the idea that syllabic organization is reflected in the way segments are temporally coordinated with each other. The notion of timing has been considered crucial in determining whether consonants combine to form an onset as in CCV or not as in C.CV where the '.' indicates a syllabic boundary. The two types of syllabic organizations, complex onset organization for CCV and simplex onset organization for C.CV, have initially been described in early studies using relations between interval stabilities of the c-center-to-anchor and right-edge-to-anchor intervals (Browman & Goldstein 1988, Byrd 1995). The c-center landmark is a property of the entire cluster (complex onset) and thus highlights a global organization of the syllable. In such an organization, complex onset organization or global organization, it is the c-center landmark that maintains a stable relation with the nucleus vowel. This organization was thought to be reflected in the stability of the c-center-to-anchor interval. In contrast, the right-edge landmark in a C.CV sequence constitutes a property of a single consonant, the immediate prevocalic consonant, to the exclusion of the initial consonant in the cluster. As a property of the immediate prevocalic consonant, the right-edge landmark highlights a local organization in the syllable where the right-edge landmark maintains a stable relation with the vowel, with the initial consonant not being syllabically affiliated with the rest of the CV sequence. This organization was thought to be reflected in the stability of the right-edge-to-anchor interval.

As we have reviewed in Chapter 1, subsequent studies have deepened the empirical terrain on English and have looked beyond English by investigating various clusters in a number of languages in an effort to better understand the relation between syllabic structure and phonetic indices (Hermes *et al.* 2013 on Italian, Shaw *et al.* 2009 on Arabic, Marin & Pouplier 2010 on English, Pouplier 2012 and Brunner *et al.* 2014 on German, Marin 2013 on Romanian, among others). Stability-based heuristics (such as c-center-to-anchor interval stability being indicative of complex onset organization and right-edge-to-anchor interval stability being indicative of simplex onset organization), along with the temporal predictions that arise from a given syllabic organization (such as leftward and rightward shift for complex onsets, but no rightward shift for simplex onset) have been widely used in these studies in order to diagnose syllabic organization. The results have been puzzling as the stability-based heuristics did not always diagnose syllabic structure (see Hermes *et al.* 2013 on Italian, Shaw *et al.* 2009 on Arabic, Marin & Pouplier 2010 on English, Pouplier 2012 and Brunner *et al.* 2014 on German, Marin 2013 on Romanian). In a similar fashion, the expected leftward or rightward shift associated with a given syllabic organization has not reliably been found (Goldstein *et al.* 2009 and Marin & Pouplier 2010 on English, Pouplier 2012 on German, Marin 2013 on Romanian, Hermes *et al.* 2017 and Pastätter & Pouplier 2015 on Polish). Instead, there is by now substantial evidence for cluster-specific patterns even though there is no doubt about the phonological status of these clusters being onsets. Intra-cluster timing (Pouplier 2012), coarticulatory resistance of the segments (Marin & Pouplier 2014, Pastätter & Pouplier 2015), and voicing implementation (Brunner *et al.* 2014) have been among the parameters implicated in these cluster-specific patterns in that they have been found to affect temporal coordination in the CCV context.

In sum, what appeared at first as a rather promising, encompassing hypothesis has in more recent work been broken down into 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, then what is the role, if any, of syllabic structure? Do perhaps all temporal aspects of organization derive from the cluster-specific phonetics regardless of the syllabic, phonological organization of the language?

The current dissertation answers this last question in the negative. The dissertation is engaged in demonstrating that syllabic organization is in fact involved in the phonetics and more specifically is expressed in a pleiotropy of phonetic indices. This was shown to be the case even in a language where prior evidence for the role of that organization has looked weak in previous studies, as is the case with German (Brunner *et al.* 2014, Pouplier 2012), and in a language where the phonetic indices used in prior work offer puzzling, apparently uninterpretable evidence, as is the case with Spanish where voiceless stop-lateral clusters show the expected global timing interval stability but voiced stop-lateral clusters show the unexpected local timing interval stability.

At the most general level, the main way in which syllabic organization is expressed is via compensatory relations between phonetic parameters. A compensatory relation between two parameters describes how one parameter responds to changes (in the values) of another parameter. Both individual parameters vary, but their ways of variation are not independent of one another. The presence of such relations indicates that the organization of the different parts of CCV is not independently planned and produced, hence global organization. Specifically, if each segment in a CCV is planned independently of the other segments, then an increase or decrease in the duration of that segment 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. A joint consideration of several phonetic indices (duration of the prevocalic consonant, relative timing of the prevocalic consonant and the vowel in CV and CCV, vowel initiation with respect to the prevocalic consonant, stability-based heuristics) and the presence or absence of compensatory effects among these indices has been shown in our work to reliably diagnose syllabic structure. None of the previous studies have jointly considered several phonetic indices and how their combination may reveal syllabic organization. Furthermore, none of the previous studies have looked at vowel initiation, even though temporal coordination patterns between consonants and their subsequent vowel have been at the center of the research program since the early key studies of Browman & Goldstein (1988) and Honorof & Browman (1995). Finally, none of the previous studies have considered the hypothesis that compensatory effects between phonetic indices manifest syllabic organization.

It is important to highlight that the expression of syllabic organization is not via any single measure or any single effect. In fact, when one considers sets of phonetic effects in clusters across different languages, these effects may appear to be contradictory (across the languages considered) and yet they may, in all cases, indicate the same structural uniformity: namely, global organization. Let us illustrate with an example from Spanish and German. Spanish and German word-initial stop-lateral clusters exhibit different properties mainly due to the different implementation of voicing which affects inter-segmental temporal coordination different in these two languages. Specifically, the interplateau interval or IPI between the consonantal plateaus of the two consonants in Spanish CC clusters is large and overall larger than in German. Furthermore, the duration of the initial stop changes in Spanish as a function of voicing, but not so in German (in our clusters data). For German voiceless stop-lateral clusters, the aspiration gesture associated with the long VOT for the voiceless stop is an extra gesture which has to be timed between the stop's oral gesture and the following lateral. Such an aspiration gesture is not present in that position in voiced stop-lateral clusters in German and is also not present in any stop-lateral cluster in Spanish. These language-specific phonetic properties affect two of the phonetic indices we investigated: the duration of the prevocalic lateral and the stability-based heuristics or more specifically the local timing interval or the degree of shift of the lateral towards the vowel. For German, the lateral was shorter after a voiced than after a voiceless stop. As for the local timing interval or the shift of the prevocalic lateral to the vowel, the lateral was shifted more towards the vowel in the latter than in the former. Thus, for German, a shorter lateral co-occurs with less lateral shift to the vowel, while a longer lateral co-occurs with more lateral shift to the vowel. In Spanish, the opposite was found, i.e., the lateral was shorter after a voiceless than after a voiced stop and the lateral was shifted more towards the vowel in the former than in the latter. That is, for Spanish, shorter lateral co-occurs with more lateral shift to the vowel, while longer lateral co-occurs with less lateral shift to the vowel.

Even though the listing of the above effects shows markedly different profiles between the two languages and in fact seemingly the effects go in opposite directions, these effects in fact serve as hallmarks of the one and the same concept: namely, that in both languages, the clusters are parsed as complex onsets. We now turn to see how these different effects of lateral shortening and degree of lateral shift to the vowel based on voicing of the initial stop in German and Spanish both express a global organization, by bringing the vowel in proximity with its tautosyllabic cluster to the extent permitted by the coincidental (language-particular) phonetic properties in the cluster.

The basic point is that global organization is achieved in different ways in German and Spanish because the language-particular phonetics lead to different temporal readjustments. It is precisely because of the latter fact, that language-particular phonetics exists, that the way in which global organization is expressed must be different in different languages. In other words, if one seeks to find manifestations of complex onset organization in different languages, then these will likely not be in the form of any given effect or any given value of some index (as in the case of stability indices) found across the languages examined, but rather in relational properties of how the effects pattern with respect to one another within any given language. Let us make this more precise in what follows.

For Spanish, because of the longer duration of the voiceless stop (compared to the voiced stop) and the substantially large lag between the consonants in the cluster, when such stops are combined with laterals in forming complex stop-lateral onsets, the lateral shortens more in the voiceless stop context than in the voiced stop context and the relative timing between lateral and the subsequent vowel changes more in the voiceless stop-lateral context than in the voiced stop-lateral context. This is because in the voiceless stop-lateral clusters, due to the longer duration of the stop and the longer lag between the stop and the lateral, more compression in the rest of the CCV must be achieved to bring the vowel in closer proximity to the cluster. The readjustments and specifically the degree to which they are found (more readjustments in voiceless stop-lateral context than in the voiced stop-lateral context) are thus motivated by constraints related to the global organization. For German, as mentioned above the opposite pattern is observed: shorter lateral co-occurs with less lateral shift to the vowel in voiced stop-lateral clusters, while longer lateral co-occurs with more lateral shift to the vowel in voiceless stop-lateral clusters. Shortening of the lateral is not 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 to a great extent; this shifting in voiceless stop-lateral clusters in German brings the vowel to overlap with the cluster in the case of voiceless stop-lateral clusters. In both languages, German and Spanish, the readjustments help bring the vowel in proximity with its tautosyllabic cluster. This is what a global organization prescribes: not just the immediately prevocalic consonant but the entire prevocalic cluster along with the vowel make up a unit that is organized globally as a whole. As we have seen, both Spanish and German have complex onsets and demonstrate the presence of such a global organization over what appear to be similar sequences of segments (e.g., /klV, glV/ appear in both languages as onsets), but the ways in which this global organization is achieved depends on the coincidental (language-particular) phonetic properties of the clusters in question.

It is also a consequence of the main claim of this dissertation that uncovering the effects of syllabic organization is not via a single universally applicable experimental design. The design must be tailored to the language. Let us see how this is so by considering the methodological aspects which allowed us to unveil effects related to global organization of complex onsets in Spanish versus German. The phonetic parameters present in Spanish clusters, such as open transition or large lag between consonantal plateaus, and varying stop duration as a function of the stop's voicing provided sufficient perturbations in phonetic parameter values so that global organization effects could emerge, that is, for us, as seen in the presence of compensatory relations among these phonetic parameters. For German, however, the overall short lag between consonantal plateaus compared to Spanish (for German there is a very short lag for voiced stop-laterals versus a longer lag for voiceless stop-laterals but still shorter than that for Spanish) did not offer the right environment for effects related to global organization to emerge. In German, effects such as shortening of the lateral and degree of lateral shift as a function of the C1 stop's voicing only emerged when C1 plateau and lag between consonantal plateaus (IPI) lengthened as a function of increasing boundary strength. It is these modulations on the CCV string via prosodic variation which allowed us to unveil effects related to global organization of complex onsets in German. Recall that previous researchers, faced with the lack of evidence for c-center-to-anchor stability in complex onsets in languages like German and English, have already hinted at possible relations between stability-based heuristics and phonetic parameters in the clusters under study. Thus, Pouplier (2012) and Brunner *et al.* (2014) report results from German which show that c-center-to-anchor stability seems to be met more in clusters with short interplateau lags than in clusters with large lags between the consonants. This dissertation challenged this view in the following way. Using prosodic variation, we lengthen C1 stop's duration and the lag between consonantal plateaus in German clusters (Chapter 3). We find that such a manipulation actually improves the stability of the c-center-to-anchor interval or global timing interval (in our terms). Furthermore, in Spanish (Chapter 2), the lag between consonantal plateaus is larger than in German clusters, but c-center-to-anchor stability (global timing interval stability, in our terms) is observed. Clearly, the extent of lag between the two consonantal plateaus in a CCV is unrelated to whether stability is observed for the c-center-to-anchor or the global timing interval.

Let us consider how the German results would have looked like had we not used the prosodic modulations in the design of our core German study. Without prosodic modulations, we would have recourse to only one out of our current three prosodic conditions available, which would correspond to data from only our control condition (word boundary condition) where no phrase boundary precedes the target word. This is the context used in Pouplier (2012) and Brunner *et al.* (2014). Our results would then be the following. Across stop-lateral clusters in the control condition, local timing interval stability was observed using three different anchor landmarks. This can be interpreted as no lateral shift towards the vowel across CV∼CCV. Furthermore, there was no difference in lateral duration between voiced and voiceless stop-lateral clusters, no lateral shortening from CV to CCV, and no earlier vowel initiation with respect to the prevocalic lateral between CV and CCV. Therefore, regardless of whether the sequence is a CV lateral-vowel or a CCV stop-lateral-vowel, there is no effect pointing to a global organization. Had we not included conditions that elicit prosodic modulations in the design of our German study, all the above listed results would have provided substantial evidence against a global organization of word-initial stop-lateral clusters in German (again, even though from the phonological side, these clusters constitute prototypical complex onsets in German).

The only phonetic index in the control prosodic condition which would point to a global organization is the compensatory effect between IPI and C2 lateral duration across stop-lateral clusters. However, this is not a relation that is noted in previous studies which tend to use stability-based heuristics or lag measures (as seen in our literature review) aimed at quantifying intervals. As we have seen in our German results, in the control prosodic condition, a negative correlation was observed between IPI and C2 lateral duration, such that as IPI increases C2 lateral duration decreases. This compensatory relation is one indication of global organization that serves to bring the vowel to overlap with the cluster. Specifically, when the lag between the plateaus of the two consonants, what we call IPI, increases, shortening the lateral compensates by bringing the vowel to overlap more with its tautosyllabic cluster. This and other such relations we have brought out in our data are signatures of a global organization across the CCV string.

For German, when C1 stop duration and IPI lengthened via manipulations of prosodic boundary strength, a number of phonetic indices revealed effects related to a global organization. Such effects consist of variation in the duration of the prevocalic lateral as a function of the stop's voicing in stop-lateral clusters, decrease of the local timing interval across CV, CCV to different degrees as a function of the stop's voicing or, in other words, degree of lateral shift towards the vowel as a function of the stop's voicing, reduced local timing interval stability, and earlier vowel initiation with respect to the prevocalic lateral within stop-lateral clusters with increasing boundary strength. All the above effects point to a global organization of stop-lateral complex onsets in German and none of these effects would have emerged had we not not used scaling of phonetic parameters via prosodic variation. Therefore, with the example of our German study we have recognized the necessity of two things in diagnosing syllabic structure: first, looking at various phonetic indices and, second, scaling phonetic parameters using prosodic variation and unveiling in this way compensatory effects in relations between phonetic parameters.

4.4 Manifestations of global organization

Let us summarize the phonetic indices that reveal effects related to global organization in German and Spanish. For Spanish, the stability-based heuristics provided confusing results, as we found global timing interval stability for voiceless stop-lateral clusters and (voiceless) stop-rhotic clusters but local timing interval stability for voiced stop-lateral clusters. All these clusters are complex onsets in Spanish and yet the often-used expectation that they should thus exhibit global timing interval stability is not borne out in our data. In our approach, several other measures that aimed to quantify degrees of readjustments across CV and CCV showed that larger effects occurred in voiceless stop-lateral clusters than in voiced stop-laterals. For instance, the duration of the prevocalic lateral was shorter in CCV than in CV and this lateral shortening was greater in voiceless stop-laterals than in voiced stop-lateral clusters. That is, the relative timing of lateral with the vowel changed more with the addition of a voiceless than a voiced stop in front of the lateral. An explanation for this asymmetry in the magnitude of effects between voiced and voiceless stop-lateral clusters can be found in the vowel initiation patterns. The vowel in voiced stop-lateral clusters is found to start closer to the c-center of the cluster than in voiceless stop-lateral clusters. Hence, extensive temporal readjustments are not necessary in voiced stop-lateral clusters since given the phonetic context (shorter C1 stop's duration and shorter IPI than in voiced stop-laterals) the vowel already starts quite close to the c-center of the cluster. For voiceless stop-laterals, the magnitude of effects is larger because the phonetic environment is more unfavorable for the vowel to establish a relation with the cluster than in voiced stop-laterals. Recall that C1 stop and IPI are longer in voiceless stop-laterals than in voiced stop-laterals. The temporal readjustments that occur serve the main organizing principle of global organization which is to bring the vowel to overlap with the onset cluster to the extent possible. For stop-rhotic clusters, using several measures as we did for stop-lateral clusters was not possible or not meaningful because the prevocalic rhotic in CCV and CV is a different segment; it is a trill in /rV/ versus a tap in stop-/ɾV/. These two segments have inherently different durations and phonetic properties and thus measures such as shortening of the prevocalic rhotic from CV to CCV or relative timing between rhotic and vowel in CV and CCV would not be meaningful. For stop-rhotic clusters, we used the stability-based heuristics and vowel initiation to assess syllabic organization. We found global timing interval stability which, however, as we acknowledged, can also be related to the tap in CCV being shorter than the trill in CV and to the very large IPI in stop-rhotic clusters. The large IPI in stop-rhotic clusters is shifting the prevocalic tap towards the vowel and thus the local timing interval ceases to be stable from CV to CCV rendering the global timing interval more stable. The stability-based heuristics in this case thus, although indicating at face value global organization, should be treated with caution as they are computed by quantifying intervals across CV and CCV where the prevocalic C is not the same segment across these sequences. For this reason, it was crucial to look again at the vowel initiation patterns across stop-rhotic clusters /pɾ, tɾ, kɾ/ which can help us assess syllabic organization in a more direct way. A clear result indicating global organization emerged here. The vowel was found to start quite close to the c-center of the cluster across stop-rhotic clusters.

For German, in the control prosodic condition where no phrase boundary preceded the target word, most of the results did not point to a global organization. We found local timing interval stability, which in past studies is taken as evidence for simplex onset or local organization, we found no effects related to prevocalic lateral shortening from CV to CCV, no difference in lateral duration as a function of C1 stop's voicing, and finally no vowel initiation differences between CV and CCV. The only result that pointed to a global organization in the control condition was the compensatory relation between IPI and C2 plateau duration, such that as IPI increases, C2 lateral duration decreases. Crucially, however, a greater range of effects related to a global organization became evident as phonetic parameters were scaled, i.e., as C1 stop duration and IPI lengthened using prosodic boundary of increasing strength. With increasing boundary strength, we found the following two effects which point to global organization: a difference in lateral duration as a function of the C1 stop's voicing and a reduction in local timing interval stability seen to a greater extent for voiceless than for voiced stop-lateral clusters. The latter means that the local timing interval decreased more from CV to CCV when the initial consonant was voiceless than when it was voiced. Furthermore, we found a decrease in the lag between target of the prevocalic lateral and vowel initiation which means earlier vowel initiation with respect to the prevocalic lateral with increasing boundary strength. The above results, reduced local timing interval stability and earlier vowel initiation with respect to the prevocalic lateral are indices of global organization because they serve to bring the vowel to overlap with the cluster to the extent possible.

Based on the above review of the phonetic indices that express global organization in Spanish and German, it has become clear that there is no unique phonetic exponent which expresses syllabic organization. The commonly employed stability-based heuristics have failed in our core studies in this dissertation to diagnose syllabic structure in both German and Spanish when used exclusively as they also failed in several languages in previous studies. Any single phonetic index does not seem to reliably diagnose syllabic organization across languages and even within a language. Consider duration of the prevocalic lateral in stop-lateral clusters. For Spanish, a short lateral co-occurs with global timing interval stability, while in German a short lateral co-occurs with local timing interval stability. The lateral's duration does not necessarily predict how the stability-based heuristics will pattern. Sometimes the relation between lateral duration and stability-based heuristics is complementary, meaning that the effect of one enhances the effect of the other index as in Spanish, with the short lateral and the large shift of the lateral to the vowel resulting in reduced local timing interval stability and thus in global timing interval stability. But other times, the relation between lateral duration and stability-based heuristics is compensatory as in the case of German, where the short lateral and the small shift of the lateral to the vowel co-occur with (and perhaps cause in this case) local timing interval stability. This demonstrates that the specific form of the relation between phonetic indices is not fixed and cannot be necessarily predicted beforehand. Rather, the specific form of that relation reflects the necessary readjustments that need to occur in order to achieve global organization given the coincidental phonetic properties of the segments that partake in this organization. Phonological organization finds its way to get by with whatever the phonetic material it is given to deal with.

4.4.1 Compensatory effects

Phonetic indices reflect syllabic organization individually, as seen in the previous subsection, but also in terms of compensatory effects. Such compensatory effects are common across German and Spanish CCV sequences and they provide evidence for a global organization that corresponds to the phonological notion of the whole CC being a complex onset of the syllable whose nucleus is the V. In stop-lateral CCV clusters, a compensatory relation between the duration of the interval between the two consonants (what we refer to as IPI) and C2 lateral duration is observed in Spanish and in German (only in the control condition, for German) such that as IPI increases, C2 lateral duration decreases. As we have seen, for German, the two variables that enter into the compensatory relation, IPI and C2 lateral duration, individually show sufficient variability to express such a relation only in the control condition of our experimental study. In the other two prosodic conditions, 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. It seems that the prosodic strengthening (that is, in terms of its temporal consequences, lengthening) induced on the duration of the first C in CCV as well as in the duration of the IPI between the two consonants freeze the extent of variability in these two variables, thus precluding the manifestation of a compensatory relation between IPI and C2 lateral plateau duration. Emphatically, however, this does not mean that there are no indications for global organization in the other non-control prosodic conditions. To the contrary, we find strong and perhaps even stronger indications of global organization in these contexts in the relation between vowel initiation and the prevocalic consonant in stop-lateral clusters. With increasing boundary strength, the lag between target of the prevocalic lateral and vowel initiation is found to decrease. This is because prosodic strengthening causes some parts of the hypothesized syllable to expand (C1 or IPI lengthening). As a result of this, the inner CV substring in CCV gets compressed and earlier vowel initiation with respect to the prevocalic lateral is observed: when C1 and IPI in CCV lengthen due to prosodic strengthening, the rest of the string (the inner CV) has to shorten to compensate for the C1's and IPI's increased length. In other words, earlier vowel initiation in CCV with increasing boundary strength is another species of compensatory effect, just like the IPI-C2 lateral duration relation (observed in the control condition), 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).

Compensatory effects of the sort just exemplified are hallmarks of global organization holding over the CCV string: the presence of such 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.

A main theme running throughout the thesis with respect to methodology has been that diagnosis of phonological and specifically syllabic organization is best achieved when one considers how the string of segments (over which the nature of the phonological organization is assessed) responds to perturbations (scaling of phonetic variables) of localized properties (such as durations) within that string. Specifically, variation in phonetic variables and more specifically prosodic variation is a crucial key to understanding the nature of the link between (phonological) syllabic organization and the phonetic spatio-temporal manifestation of that organization. The effects of prosodic variation on segmental properties and on the overlap between the segments, we argue, offer the right pathway to discover patterns related to syllabic organization. This is so because, to uncover evidence for global organization, the sequence of segments partaking in that organization as well as properties of these segments or their relations with one another must be somehow locally varied. The consequences of such variation on the rest of the sequence can then be used to unveil the span of organization. When local perturbations to segments or relations between adjacent segments have effects that ripple through the rest of the sequence, this is taken as 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.

It is important to point out that there is no guarantee that any given phonetic parameter will exhibit comparable degrees of variability across two languages even if the profile of the contexts over which variability is assessed appears to be the same. An illustrative case is that of IPI variability observed in Spanish versus German stop-lateral word-initial clusters. While in Spanish the phonetic properties in a CCV string offer the right environment for effects related to a global organization to emerge, for German there are contexts where these effects are not evident (intonational phrase and utterance phrase boundary conditions). Why is this so? In Spanish, there is sufficient variability in the interval between the two consonants (what we refer to as IPI). As IPI varies and specifically stretches out over a comparably wide range, the lateral must compensate by shortening in duration. For clusters following a strong prosodic boundary in German (as in the intonational and utterance phrase boundary conditions), C2 lateral duration does not vary as much. The lengthening of C1 duration and IPI duration (and the presence of the aspiration gesture in between the stop and the lateral in German) freeze the extent to which C2 lateral duration can vary. Consequently, the compensatory relation seen in Spanish is not found in this context in German. However, by inducing perturbation in phonetic parameters for German, other effects related to global organization similar to those seen in Spanish such as early vowel initiation and reduced local timing interval stability become evident also in German as phonetic parameters are scaled. In other words, for both languages, the nature of the phonetic parameter which offers the right variability so that the emergence of compensatory relations can be examined (on the basis of this variability) cannot be known *a priori*. What we can say is that the nature of global organization is the same across the two languages and the methodological approach is in its broad lines entirely parallel between the two languages. That is, as we have seen, just like in Spanish, the scaling of phonetic parameters via prosodic variation in German has been 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) crucially modulate the spatio-temporal coordination of the same segmental material. The researcher, thus, must seek to induce the right perturbations in the segmental sequence on a language to language basis, informed by previous research and or intuition, as we did for German.

Once again, generalization across languages is found in how segmental sequences respond to perturbations, not in some fixed parameter or relation which remains the same across the languages. The presence of a uniform organization is common across (complex onset admitting) languages. But the phonetics, that is, the substance that fits into any given segmental sequence is language-particular. Thus, the expression of syllabic organization in terms of specific phonetic effects across languages cannot be the same. There is no unique index of syllabic organization across contexts and languages. But conceptually the way in which such organization is expressed, in terms of compensatory relations, is entirely general.

4.5 Manifestations of local organization

We have so far summarized effects related to a global organization across Spanish and German. We now turn to consider effects related to a local organization in Moroccan Arabic word-initial stop-lateral clusters and in cross-word voiceless stop-lateral /pl, kl/ clusters in German. Both in Moroccan Arabic word-initial clusters and in cross-word German clusters, the two consonants of the cluster do not form an onset. Rather, they are separated by a syllabic boundary. Thus, these clusters exemplify simplex onset organization which should be expressed in terms of local organization of only the prevocalic consonant with the vowel to the exclusion of consonants preceding the prevocalic consonant. In past studies, the phonetic manifestation of this local organization is taken to be the stability of an interval spanning just the prevocalic consonant and the vowel (the local timing interval or as termed in other work the right-edge-to-anchor interval). However, as emphasized throughout in this dissertation, interval stability is just one index of syllabic organization which must be used along with other indices.

We now summarize the phonetic indices that express simplex onset organization or local organization in Moroccan Arabic word-initial clusters in contrast to Spanish word-initial complex onsets. In Moroccan Arabic /kl, bl/ clusters, the lateral's duration did not change from CV to C.CV (again, within words the '.' indicates syllabic boundary). In Spanish, however, the prevocalic lateral was found to shorten substantially in the CCV cluster compared to CV. In Moroccan Arabic, vowel initiation occurs at the release of the prevocalic lateral in both clusters; i.e., vowel onset is tied to the immediately preceding segment /l/ as prescribed by a simplex onset or local organization. In Spanish, however, the vowel is found to start as close to the c-center of the cluster as possible across stop-lateral and stop-rhotic clusters. In Moroccan Arabic, there is a weak compensatory relation between IPI and C2 lateral duration as opposed to a strong compensatory relation between the two variables observed in Spanish. This means that an increase in the lag between the two consonants does not affect the duration of the prevocalic consonant in Moroccan Arabic. That is, there are no consequences of IPI variation on the rest of the CV sequence. The local perturbation between the two consonants does not have effects that ripple through the rest of the sequence as is the case with globally organized stop-lateral clusters in Spanish. Furthermore, there is no change in the relative timing of the lateral with the vowel from /lV/ to /k.lV/ in Moroccan Arabic, while in Spanish the relative timing of the lateral with the vowel changes from /lV/ to stop-/lV/. In Moroccan Arabic, robust local timing interval stability has been observed across /lV/∼/k.lV/, while global timing interval stability has been observed across /lV/∼/klV/ in Spanish. Each of the above listed properties for Moroccan Arabic expresses an organization according to which the vowel is locally timed with the prevocalic consonant of the cluster. The comparison between Spanish and Moroccan Arabic demonstrates that syllabic organization is reflected in a number of phonetic indices. This is the pleiotropy part of our main claim.

For German, word-initial complex onsets were compared to cross-word simplex onsets, because the two groups of clusters exhibit a different syllabic organization. Again, a central idea of this dissertation is that syllabic organization makes different predictions about the way phonetic indices respond to changes in phonetic parameters. German cross-word clusters offer a clear demonstration of this idea. For cross-word clusters, when increasing boundary strength, there was robust local timing interval stability and no change in vowel initiation across C#CV. This is the opposite pattern of what it was observed for German complex onsets, that is, the same clusters but in word-initial position (CCV), where with increasing boundary strength in German, there was reduced local timing interval stability and earlier vowel initiation.

In sum, comparing word-initial complex onsets as in CCV versus the same clusters at the cross-word context as in C#CV, where only the second C is part of the onset of the syllable with the following vowel and hence simplex onsets in German, demonstrates that syllabic organization makes different predictions about the way phonetic indices respond to perturbations of phonetic parameters.

4.6 Directions for further research

It would not be an exaggeration to say that research with relevant articulatory data on the topic of the relation between syllabic organization and its phonetics has been so far pursued on a relatively limited number of languages (Browman & Goldstein 1988, Honorof & Browman 1995, Byrd 1995, and Marin & Pouplier 2010 on English, Shaw *et al.* 2009 and Shaw *et al.* 2011 on Moroccan Arabic, Pouplier 2012 and Brunner *et al.* 2014 on German, Marin 2013 and Marin & Pouplier 2014 on Romanian, Hermes *et al.* 2013 on Italian, Pastätter & Pouplier 2015 and Hermes *et al.* 2017 on Polish, Goldstein *et al.* 2007, Hermes *et al.* 2011 and Hermes *et al.* 2017 on Berber, Pouplier & Benus 2011 on Slovak, Goldstein *et al.* 2007 on Georgian, Kühnert *et al.* 2006 on Continental French). Given the limited number of languages, there is ample choice on what other languages one may consider next. It is thus reasonable to seek choice criteria which would at once enable one to sharpen the main proposal of this dissertation and to broaden the empirical domain. With such an aim in mind, we will propose in this section that future work would profit by investigating the relation between syllabic structure and phonetic indices in languages with a more elaborate voicing system than German and Spanish. An appropriate case is offered by Hindi. Whereas German and Spanish have a two-category voicing contrast between voiced and voiceless stops, Hindi has a four-way voicing contrast: voiced unaspirated, voiced aspirated (also known as breathy), voiceless unaspirated, and voiceless aspirated stops. This four-way voicing contrast system implies the presence of variation in terms of the phonetic parameters of stop duration and interplateau interval (or IPI) within the same language and without having to elicit such variation using prosodic modulations as had to be done for German in this dissertation. Furthermore, while Spanish does offer varying stop durations as a function of voicing, IPI did not vary across the two voicing categories. That is, the IPI was comparable across voiced stop-lateral and voiceless stop-lateral clusters. Hindi, however, offers both varying stop durations and varying IPI not only within the same language, but also within the same voicing condition (within the voiced or within the voiceless category) as will be explained below.

Let us first introduce some relevant aspects of syllabic structure and phonetic properties of clusters in Hindi. Hindi admits word-initial clusters which include stop-lateral /bl, pl, kl, gl/, stop-rhotic /pr, br, gr, kr/, stop-fricative /kv, k^{h}v, dv, d^{ɦ}v/, and nasal-liquid /ml, mr/ clusters (Kumar 2005). As mentioned, unlike German and Spanish, stops in Hindi show a four-way voicing contrast. For instance, let us illustrate this four-way contrast using labial stops. Within the voiced category, there is the voiced unaspirated stop /b/, which is produced with vocal fold vibration during closure, and the voiced aspirated /b^{ɦ}/, which is produced with vocal fold vibration during closure followed by a breathy release. Within the voiceless category, there is the voiceless unaspirated stop /p/, which is produced with an open glottis and no aspiration, and the voiceless aspirated stop /p^{h}/ produced with an open glottis and followed by an aspiration phase. In Hindi, there are word-initial clusters consisting of voiced aspirated stop-r /b^{ɦ}r, g^{ɦ}r/ as in / b^{ɦ}rəʃʈ / ‘corrupt’ and /g^{ɦ}rat/ ‘a smell’, voiced unaspirated stop-liquid sequences /br, gl/ as in /brahmən/ ‘priest’ and /glan/ ‘diseased’, and voiceless unaspirated stop-liquid sequences /kl, pr/ as in /klib/ ‘impotent’ and /prətha/ ‘custom’. Voiceless aspirated stop-liquid clusters do not seem to occur in Hindi but voiceless stop aspirated-fricative clusters such as /k^{h}v/ do occur word-initially (Kumar 2005).

Spanish and German exhibit a two-way voicing contrast which means that these languages distinguish just voiced from voiceless stops. In the expression of this binary contrast, one crucial phonetic parameter is the well-known Voice-Onset-Time (VOT), defined as the lag between the release of the stop and the initiation of voicing (vocal fold vibration). The German voicing contrast is mainly expressed in terms of a short versus long lag VOT difference, while Spanish voicing contrast is mainly expressed in terms of a lead voice versus short lag VOT difference. Voiceless word-initial stops in German are aspirated with a VOT in the range of 45–120 ms (depending on place of articulation and speaker, Hullebus *et al.* 2018) whereas German voiced stops word-initially are not truly voiced, in that there is no robustly present vocal fold vibration during the closure period of the stop (Jessen 2002) and instead show the so-called short lag VOTs of about 10–20 ms (Bombien & Hoole 2013). Spanish voiced stops show negative VOT values in the range of 108–138 ms and voiceless stops have VOT values in the range of 4–25 ms (Lisker & Abramson 1964). In Hindi, VOT is not by itself sufficient to encode the needed contrasts – it cannot distinguish between voiced aspirated and voiced unaspirated stops, as both exhibit negative VOT (Dutta 2007). Dutta (2007) finds that other cues like closure duration are relevant for expressing the contrast between the stop types. Closure duration has been found to be shorter for aspirated than unaspirated stops in Hindi (Dutta 2007:113) and voiceless stops have been found to be longer than fully voiced stops in Hindi (Shrotriya *et al.* 1996). The latter observation is seen also for Spanish voiced and voiceless stops (as we have seen in our data) with the fully voiced stops being shorter than voiceless unaspirated stops. Recall that in our main studies in this dissertation, such differences in closure duration were crucial in bringing out evidence for how clusters and following vowels are organized into bigger units or complex onsets. Because of its more elaborate voicing system, Hindi adds more relevant opportunities in investigating how voicing affects intra-cluster timing and how that timing expresses complex onset organization.

The voiced aspirated category in Hindi does not have an equivalent in Spanish or German. It thus provides new empirical ground wherein the effect of the aspiration gesture on inter-segmental temporal coordination patterns can be examined in a voiced context. The voiced unaspirated counterpart is also available in Hindi and will serve as a comparison with the aspirated context. Dutta (2007), in his dissertation on the four-way contrast in Hindi, reports that the standard view claiming that the main distinction between voiced aspirated and unaspirated stops is the breathy release does not seem to find much support in the data. Other properties seem to contribute to the distinction between these two categories such as closure duration. As mentioned in the previous paragraph, Dutta (2007:113) states that aspirated stops have shorter closure duration than unaspirated stops and this seems to hold true within voiced and voiceless stops. Therefore, C1 stop's duration varies as a function of aspiration. Furthermore, the presence of aspiration in voiceless aspirated stop-initial clusters or breathy phonation in voiced aspirated stop-initial clusters will likely cause some IPI variation compared to their (both voiced and voiceless) unaspirated counterparts. Specifically, the aspiration in voiceless aspirated stops may yield some delay in modal voicing and thus cause the following segment to occur later compared to the case where aspiration is absent. This prediction is based on reasonable extrapolations from English articulatory data collected in our lab where the effects of an intervening aspiration gesture on the timing of its surrounding oral gestures is studied. Let us review this evidence as it is both novel and related to the point being made here. The data derive from an experimental study with four native speakers of American English producing simple CV sequences ($N = 207$), where C was either the voiced or voiceless labial /b, p/ and V was either the low vowel /ɑ/ (as in /bɑt/, /pɑt/ or the high vowel /i/ (as in /biːʧ/, /piːʧ/). In this data, we examined vowel initiation with respect to the prevocalic labial in /pV/ versus /bV/ in order to investigate the effect of aspiration on vowel initiation. Figure 4.1 shows vowel initiation in /bV/ versus /pV/ in English. The vertical lines delimited by black dots indicate the gestural plateau of the bilabial stop /b, p/ and the triangle indicates vowel initiation. As can be seen in Figure 4.1, the vowel starts 24 ms before the target of /b/ and 14 ms before the target of /p/. Thus, the vowel starts later when it is preceded by a voiceless aspirated /p/ as opposed to a voiced /b/.

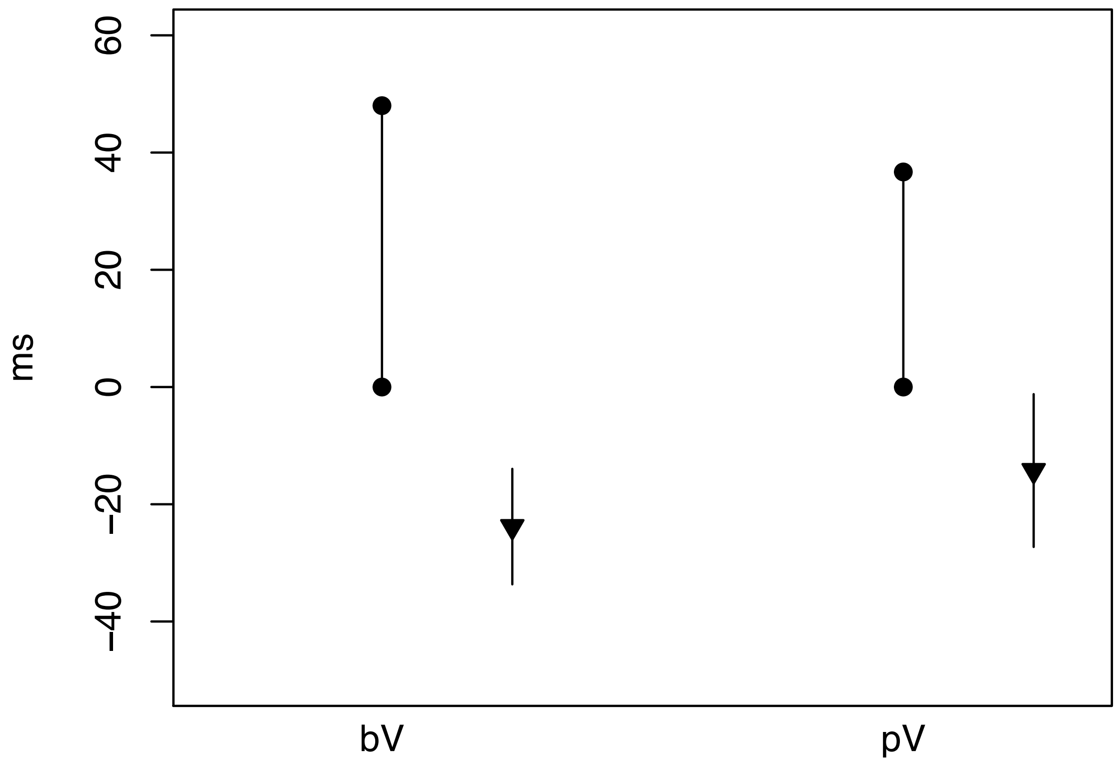

Figure 4.1: Vowel onset in relation to prevocalic consonants /b/ and /p/ (x-axis). The vertical lines denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonants /b/ and /p/. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateau of the prevocalic labial stop.

The results from English provide evidence that aspiration may delay initiation of the following segment. Turning now back to Hindi, we can say that if the presence of aspiration affects (or delays) timing of the following segment, we would expect to see greater IPI between the consonants in the voiced aspirated stop-rhotic clusters as opposed to the voiced unaspirated stop-rhotic clusters. The above prediction rests on the assumption that the aspiration gesture in voiceless stops affects the timing of the subsequent consonant in the same way as the breathy release in voiced aspirated stops. Dutta (2007) provides some evidence about the relation between aspiration (or breathy release) in voiced aspirates and aspiration in voiceless aspirates. Specifically, Dutta (2007) found that the duration of aspiration in voiced aspirated stops is longer than in voiceless aspirated stops. Furthermore, Dutta (2007) suggests that aspiration as produced following voiceless aspirated stops is acoustically and aerodynamically different from the breathy release following voiced aspirated stops. Even though the aerodynamic and acoustic nature of the glottal gesture in voiced aspirated versus voiceless aspirated stops is different, they are both glottal gestures and they both intervene between the two consonants in the cluster (as in /g$^{\text{ɦ}}$rat/ 'a smell' and in /k$^{\text{h}}$vab/ 'dream'). It thus seems reasonable to expect that these two different glottal gestures still provide a source of IPI variation compared to their unaspirated counterparts (voiced unaspirated and voiceless unaspirated stops) where such a glottal gesture is absent. Therefore, Hindi offers sequences such as /b$^{\text{ɦ}}$rəʃʈ/ 'corrupt'∼/brahmən/ 'priest' within the voiced condition and sequences such as /k$^{\text{h}}$vab/ 'dream'∼/kvat$^{\text{h}}$/ 'decoction' within the voiceless condition with varying stop duration and presumably varying IPI based on the presence of aspiration or not. This offers new and highly relevant empirical territory in further examining the effect of such phonetic variation (IPI, C1 stop duration) on the phonetic indices which express syllabic organization. Furthermore, a comparison within voiced stop-initial clusters and within voiceless stop-initial clusters will allows us to see whether effects related to syllabic organization are common across these two contexts or whether voicing of the initial stop is yet another factor that necessitates different kinds of temporal readjustments based on the presence or absence of aspiration.

Let us now focus on Spanish and Hindi to show how the voicing profile of these languages enables one to test a prediction of our main thesis that we could not test so far. Spanish, as opposed to German, shows a distinction in stop duration in that voiced stops are shorter than voiceless stops. This also seems to be the case in Hindi voiced versus voiceless stops (Shrotriya *et al.* 1996). This is the expected and indeed attested pattern in voicing systems which express voicing contrasts based on a prevoicing versus short lag VOT difference. The prevoicing versus short lag VOT constrast in Spanish is a subset of the voicing system in Hindi which describes voiced unaspirated stops (prevoiced) versus voiceless unaspirated stops (short lag). Thus, Spanish and Hindi are similar in two ways: in that voiced stops are shorter than voiceless stops and in that they share the same two-way voicing contrast in terms of prevoicing versus short lag. Furthermore, it is reasonable to expect that the prevoicing-short lag contrast in Spanish and Hindi yields similar timing patterns between consonants in these two languages if intra-cluster timing is partly determined by voicing conditions; i.e., we may find the same overall large IPI or little overlap across voiced unaspirated and voiceless unaspirated stops in Hindi as in voiced and voiceless stops in Spanish. Given that language-particular phonetics are highly similar across these languages, if syllabic organization is expressed in a similar way in similar phonetic contexts, then we expect to find the same phonetic indices expressing complex onset organization across Spanish voiced and voiceless stop-laterals and Hindi voiced unaspirated and voiceless unaspirated stop-initial clusters. For instance, the clusters /bl, pl/ in Spanish and Hindi share similar phonetic properties in that stop duration varies as a function of the stop's voicing and IPI patterns could prove similar based on the same voicing contrast these clusters exhibit. Furthermore, the /bl, pl/ clusters form complex onsets across these two languages. Thus, a comparison between Spanish and Hindi allows us to test a prediction of our approach. This prediction is that syllabic organization should be expressed uniformly across languages which share similar phonetic properties and the same syllabic organization. We have had so far no opportunity to test this prediction. German and Spanish share the relevant phonology, in that for both /kl/ is a complex onset, but the phonetic properties of the /k/ across the two languages are markedly different.

We now consider similarities across the three languages German, Spanish and Hindi which could help us understand the effects of laryngeal settings (which underlie the voicing contrasts in these languages) on intra-cluster timing. Voiceless unaspirated stops as initial consonants in a cluster are common in all three languages German, Spanish and Hindi. In the case of German, the voiceless unaspirated category corresponds to word-initial voiced stops /b, g/, because word-initially German voiced stops are not produced as fully voiced but as voiceless unaspirated stops. The voiceless unaspirated stops across Hindi, Spanish and German all show short positive VOT. If intra-cluster timing or CC overlap is partly determined by the laryngeal properties of the consonants (Bombien & Hoole 2013), we would expect to find the same degree of IPI across these types of clusters: /bl/ in German, /pl/ in Spanish, /pl/ in Hindi. This, however, is not the case according to our data, as shorter IPI is observed in German voiced stop-laterals than in Spanish voiceless stop-laterals. Hindi presents us with new territory here as it is a language which exhibits both the short-long lag contrast (within the voiceless category) and the prevoicing-short lag contrast (between voiced and voiceless stops). Hindi could then either exhibit little overlap between a voiceless unaspirated stop and the following sonorant (thus, following the pattern of voiceless stop-lateral clusters in French and Spanish) or it could exhibit high overlap following the pattern of voiced stop-lateral clusters in German. Examining Hindi will thus elucidate how a language with a more complex voicing system than English and

German responds to the short lag VOT with respect to the overlap pattern it exhibits. The laryngeal specifications related to the implementation of the voicing contrast do not necessarily yield the same overlap pattern across languages; i.e., short lag or short VOT in German voiced stop-laterals co-occurs with high overlap, but short lag or short VOT in French and Spanish voiceless stop-laterals co-occurs with low overlap. However, long lag VOT may generally not co-occur with high overlap. Evidence for large IPI (or little overlap) can be found in voiceless stop-lateral clusters in German (long lag VOT) in our study and in Bombien & Hoole (2013). We can thus discern a potential for a conjecture in that short lag VOT may co-exist with high or low overlap, but long lag VOT seems to co-exist only with low overlap. Such a typology provides one example of how laryngeal settings and specifically the aspiration gesture influences oral timing patterns. Another example of how laryngeal settings affect oral timing patterns is evident in the low overlap (large IPI) observed in clusters consisting of fully voiced consonants in French and in Moroccan Arabic. Voiced stop-lateral /bl, gl/ clusters, where the stop is fully voiced, have been found to exhibit low overlap in French (Bombien & Hoole 2013). Low overlap has also been found in stop-stop clusters in Moroccan Arabic (Gafos *et al.* 2010, Shaw *et al.* 2009). Voiced stops in Moroccan Arabic are prevoiced as in French. The low overlap setting between the two fully voiced segments facilitates the aerodynamic requirements for initiating and maintaining voicing during closure (Bombien & Hoole 2013). Thus, the overlap patterns seem to be language-specific but there still may be some role for the hypothesis that the voicing system also contributes to the patterns in that the configuration of long lag VOT with high overlap is avoided (as per the conjecture above).

The rich cluster inventory in Hindi in combination with its four-way voicing system would allow us to examine intra-cluster timing to better understand the role of laryngeal settings in oral timing patterns. Moreover, further investigation of how intra-cluster timing affects the organization of the cluster with its subsequent vowel will shed light on whether syllable structure contributes constraints which are manifested in different ways depending on language-specific realization of the voicing contrast. For instance, our results from German and Spanish show that there is a global organization pattern (where the vowel begins close to the c-center of the cluster) which in a language with a short lag versus long lag VOT (as in German) requires different degrees of CC overlap between voiced and voiceless stop-initial clusters. However, in a language with a prevoicing versus short lag VOT contrast (as in Spanish), global organization does not require different degrees of CC overlap between voiced and voiceless stop-initial clusters. Hindi could help us sharpen the above prediction, because its voicing system includes both types of voicing contrasts, that of German and that of Spanish. Therefore, in Hindi, clusters like /k^hv, kv/, whose stops exhibit a short lag-long lag VOT contrast, would show different degrees of CC overlap as observed in German /pl, bl/ clusters. However, in Hindi, clusters such as /bl, pl/ whose stops exhibit a prevoicing versus short lag should show similar degree of CC overlap as observed in Spanish /bl, pl/ clusters.

To conclude, a language with a four-way voicing contrast like Hindi which also admits complex onsets seems highly apt in further extending our understand of the role of laryngeal settings in intra-cluster timing, beyond what has been achieved so far. A comparison of the appropriate clusters which share similar phonetic properties with respect to stop duration and voicing in Hindi and Spanish will further elucidate how syllabic organization is expressed across languages that exhibit similar phonetic properties. Moreover, the case of voiced aspirated stop-initial clusters in Hindi is not met in Spanish or German. Such clusters will elucidate the effect of voicing and of the aspiration gesture on intra-cluster timing which, in turn, may potentially reveal new ways in which complex onset organization is achieved in such a phonetic context. Finally, Hindi complex onset clusters offer degrees of variation of phonetic parameters which are not met in the languages we have investigated, such as variation in stop duration and IPI within the same language and even within the voiced or voiceless context. Such variation is present in Hindi without having to elicit it in the experimental design as had to be done in our German study. It is such variation of phonetic parameters which is crucial to the methodological approach promoted in our work and would offer new contexts to further test and sharpen the predictions of this dissertation.

Bibliography

Barry, W. (1997). Another R-tickle. *Journal of the International Phonetic Association, 27*(1–2), 35–45.

Bertinetto, P. M. (2004). On the undecidable syllabification of /sC/ clusters in Italian: converging experimental evidence. *Italian Journal of Linguistics/Rivista di Linguistica, 16*, 349–372.

Blecua Falgueras, B. (2001). *Las vibrantes del español: manifestaciones acústicas y procesos fonéticos* (Doctoral dissertation). Universitat Autònoma de Barcelona, Spain.

Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *Handbook of phonological theory Vol. 1* (pp. 206–244). London: Blackwell.

Bombien, L., Mooshammer, C., Hoole, P., Kühnert, B. & Schneeberg, J. (2006). An EPG study of initial /kl/ clusters in varying prosodic conditions in German, *Proceedings of the 7th International Seminar on Speech Production*. Ubatuba, Brasil.

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.

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.

Boudlal, A. (2001). *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic* (Doctoral dissertation). Universite Mohammed V, Marocco.

Bradley, T. (2006). Spanish complex onsets and the phonetics-phonology interface. In F. Martinez Gil & S. Colina (Eds.), *Optimality-Theoretic Studies in Spanish Phonology* (pp. 15–38). Amsterdam, Netherlands: John Benjamins.

Browman, C. P. & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, 45(2–4), 140–155.

Browman, C. P. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, 49, 155–180.

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.

Byrd, D. (1995). C-Centers revisited. *Phonetica*, 52, 263–282.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24*, 209–244.

Byrd, D. & Saltzman, E. (1998). Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics, 26*, 173–199.

Byrd, D. & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary adjacent lengthening. *Journal of Phonetics, 31*(2), 149–180.

Byrd, D., Lee, S., Riggs, D. & Adams, J. (2005). Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America, 118*, 3860–3873.

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.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Bloomington, USA: Indiana University Press.

Catford, J. C. (1988). *A practical introduction to phonetics*. Oxford: Clarendon Press.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 419–447). Berlin, New York: Mouton de Gruyter.

Cho, T. & Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*, 155–190.

Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 283–333). Cambridge, UK: Cambridge University Press.

Colantoni, L. & Steele, J. (2005). Phonetically-driven epenthesis asymmetries in French and Spanish obstruent-liquid clusters. In R. Gess & E. Rubin (Eds.), *Experimental and Theoretical Approaches to Romance Linguistics* (pp. 77–96). Amsterdam, Netherlands: John Benjamins.

Crawley, Michael J. (2013). The R book. Chichester, West Sussex, United Kingdom: Wiley.

Dell, F. & Elmedlaoui, M. (1996). On consonant releases in Imdlawn Tashlhiyt Berber. *Linguistics, 34*, 357–395.

Dell, F. & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic.* Dordrecht, Netherlands: Kluwer.

Dutta, I. (2007). *Four-way stop contrasts in Hindi: An acoustic study of voicing, fundamental frequency and spectral tilt* (Doctoral dissertation). University of Illinois at Urbana-Champaign Program, USA.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America, 101*, 3728–3740.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29,109–135.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory*, 20(2), 269–337.

Gafos, A., Hoole, P., Roon, K. & Zeroual, C. (2010). Variation in timing and phonological grammar in Moroccan Arabic clusters. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallée (Eds.), *Laboratory Phonology 10* (pp. 657–698). Berlin: Mouton de Gruyter.

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.

Gafos, A. (2018). Stems and paradigms in Modern Standard Arabic and dialects. In E. Benmamoun & R. Bassiouney (Eds.), *The Routledge Handbook of Arabic Linguistics* (pp. 62–86). Routledge.

Gafos, A., Roeser, J., Sotiropoulou, S., Hoole, P. & Zeroual, C. to appear. Structure in mind, structure in vocal tract. *Natural Language and Linguistic Theory.*

Gibson, M., Fernández Planas, A. M., Gafos, A. & Remirez, E. (2015). Consonant duration and VOT as a function of syllable complexity and voicing in a subset of Spanish clusters, *Proceedings of the International Speech Communication Association (Interspeech)* (pp. 1690–1994). Dresden, Germany.

Gibson, M., Sotiropoulou, S., Tobin, S. & Gafos, A. (2017). On some temporal properties of Spanish consonant-liquid and consonant-rhotic clusters. In M. Belz, S. Fuchs, S. Jannedy, C. Mooshammer, O. Rasskazova & M. Zygis (Eds.), *Proceedings of the 13th Tagung Phonetik und Phonologie im deutschsprachigen Raum (PP13)* (pp. 73–76). Berlin, Germany.

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. *Action to Language via the Mirror Neuron System*, 215–249.

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.

Goldstein, L., Nam, H., Saltzman, E. & Chitoran, I. (2009). Coupled oscillator planning model of speech timing and syllable structure. In G. Fant, H. Fujisaki & J. Shen (Eds.), *Frontiers in phonetics and speech science* (pp. 239–250). Beijing: The Commercial Press.

Haggard, M. (1973). Correlations between successive segment durations: values in clusters. *Journal of Phonetics, 1*, 111–116.

Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. Phonology, 23(3), 387–429.

Harms, R. T. (1976). The segmentalization of Finnish 'nonrules', *Papers from the Transatlantic Finnish Conference* (pp. 73–88). Texas Linguistic Forum 5, Austin.

Harris, J. W. (1969). *Spanish phonology.* Cambridge, MA: MIT Press.

Harris, J. W. (1983). *Syllable structure and stress in Spanish.* Cambridge, MA: MIT Press.

Hermes, A., Ridouane, R., & Mücke, D. (2011). Gestural coordination in Tashlhiyt syllables, *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS).* Hong Kong, China.

Hermes, A., Mücke, D. & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology, 30*, 1–25. DOI:10.1017/S095267571300002X.

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 Tashlhyit Berber and Polish. *Journal of Phonetics, 64*, 127–144.

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.

Hoole, P. (2006). *Experimental studies of laryngeal articulation* (Unpublished habilitation thesis). Ludwig-Maximilians-Universität, Germany.

Hualde, J. I. (2005). *The sounds of Spanish*. New York City: Cambridge University Press.

Hullebus, M. A., Tobin, S. J., & Gafos, A. (2018). Speaker-specific structure in German voiceless stop voice onset times, *Interspeech 2018* (pp. 1403–1407). Hyderabad, India.

Jessen, M. & Ringen, C. (2002). Laryngeal features in German. *Phonology*, 19, 189–218.

Katz, J. (2012). Compression effects in English. *Journal of Phonetics, 40*(3), 390–402.

Keating, P., Cho, T., Fougeron, C. & Hsu, C. S. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden & R. Temple (Eds.), *Papers in Laboratory Phonology 6* (pp. 143–161). London: Cambridge University Press.

Kiparsky, P. (2003). Syllables and moras in Arabic. In C. Fery & R. van der Vijver (Eds.), *The syllable in optimality theory* (pp. 147–182). Cambridge, UK: Cambridge University Press.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208–1221.

Kubozono, H. (1989). The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech*, 32(3), 249–278.

Kühnert, B. & Hoole, P. (2006). Cohésion temporelle dans les groupes C1/l/ initiaux en français. *Actes des XXVIes Journées d'Etude sur la Parole*, 545–548.

Kumar, A. (2005). *Aspects of Hindi syllabic structure* (Doctoral dissertation). University of Michigan, USA.

Levin, J. (1987). Between epenthetic and excrescent vowels. In M. Crowhurst (Ed.), *Proceedings of the Sixth West Coast Conference on Formal Linguistics (WCCFL 6)* (pp. 187–201). Tuscon, USA.

Lipski, J. (1990). Spanish taps and trills: phonological structure of an isolated opposition. *Folia Linguistica*, 24, 153–174.

Lisker, L. (1957). Closure duration and intervocalic voiced-voiceless distinction in English. *Language, 33*, 42–49.

Lisker, L. & Abramson, A. (1964). A cross-language study of voicing in initial stops. *Word, 20*, 284–422.

Löfqvist, A. & Gracco, V. (1999). Interarticulator programming in VCV sequences: Lip and tongue movements. *Journal of the Acoustical Society of America, 105*, 1854–1876.

Malmberg, B. (1965). *Estudios de fonética hispánica (Collectanea Phonetica, I.).* Madrid: Consejo Superior de Investigaciones Científicas.

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.

Marin, S. (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics, 41*, 211–227.

Marin, S., & Pouplier, M. (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. *Journal of Phonetics, 42*, 24–36. http://dx.doi.org/10.1016/j.wocn.2013.11.001.

Martinez Celdrán, E. (1993). La percepción categorial de /b-p/ en español basada en las diferencias de duración. *Estudios de Fonética Experimental, 5*, 224–239.

McCarthy, J. J., & Prince, A. S. (1993). Generalized alignment. In *Yearbook of Morphology* (pp. 79–153). Dordrecht, The Netherlands: Kluwer.

Mücke, D., Becker, J., Barbe, M. T., Roettger, T. B., Meister, I., Liebhart, L., Timmermann, L. & Grice, M. (2014). The effect of deep brain stimulation on the speech motor system in essential tremor patients. *Journal of Speech, Language, and Hearing Research, 57*(4), 1206–1218.

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.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer Verlag.

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.

Pierrehumbert, J. & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 90–117). Cambridge: Cambridge University Press.

Pouplier, M. & Benus, S. (2011). On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory Phonology*, 2(2), 243–273.

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.

Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report 2.

Quilis, A. (1970). El elemento esvarabático en los grupos [pr, br, tr], *Phonetique et Linguistique Romaines: Melanges offerts a M. Georges Straka* (pp. 99–104). Lyon-Strasbourg: Societe de Linguistique Roman.

Quilis, A. (1993). *Tratando de fonologia y fonética españolas*. Madrid: Gredos.

Rennicke, I. (2015). *Variation and change in the rhotics of Brazilian Portuguese* (Doctoral dissertation). University of Helsinki in cotutelle with Universidade Federal de Minas Gerais, Brazil.

Ridouane, R., Hermes, A. & Hallé, P. (2014). Tashlhiyt's ban of complex syllable onsets: phonetic and perceptual evidence. *STUF-Language Typology and Universals, 67*(1), 7–20.

Schmeiser, B. (2009). An acoustic analysis of intrusive vowels in Guatemalan Spanish /rC/ clusters. *University of Pennsylvania Working Papers in Linguistics, 15*(1), 193–202.

Shaw, J., Gafos, A., Hoole, P. & Zeroual, C. (2009). Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability. *Phonology, 26*(1), 187–215.

Shaw, J., Gafos, A., Hoole, P. & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure. *Phonology, 28*, 455–490.

Shaw, J. & Gafos, A. (2015). Stochastic time models of syllable structure. *PloS One, 10*(5). DOI: 10.1371/journal.pone.0124714. eCollection 2015.

Shepherd, M. (2003). *Constraint interactions in Spanish phonotactics* (M.A. Thesis). Cal State Northridge (ROA-639).

Shrotriya, N., Verma, R., Gupta, S. K., Agrawal, S. S. (1996). Durational characteristics of Hindi consonant clusters, *International Conference on Spoken Language Processing (ICSLP-1996)* (pp. 2427–2430). Philadelphia, USA.

Steriade, D. (1999). Alternatives to syllable-based accounts of consonantal phonotactics. In O. Fujimura, B. Joseph & B. Palek (Eds.), *Proceedings of LP* (pp. 205–245). Prague: The Karolinum Press.

Tilsen, S., Zec, D., Bjorndahl, C., Butler, B., L'Esperance, M. J., Fisher, A., Heimsdottir, L., Renwick, M. & Sanker, C. (2012). A cross-linguistic investigation of articulatory coordination in word-initial consonant clusters. *Cornell Working Papers in Phonetics and Phonology.*

Westbury, J. R. (1994). *X-ray microbeam speech production database user's handbook.* Madison, WI: University of Wisconsin.

Wiese, R. (1996). *The phonology of German.* Oxford, UK: Oxford University Press.

Zeroual, C. (2000). *Propos controversés sur la phonétique et la phonologie de l'arabe marocain* (Doctoral dissertation). Université de Paris8, France.

Appendix

The information provided in this Appendix addresses comments of the two reviewers of my dissertation.

1. Comment on kinematic record showing vowel initiation

   The reviewer suggested adding a figure of the kinematic record with the vowel onset identified especially on sequences such as /kɾV/ or /klV/ and how the velar gesture was separated from the vocalic gesture (since the both use the same articulator, tongue back).

   An example of parsing the velar gesture for C1 /k/, the coronal gesture for C2 /l/ and the vocalic gesture for V /a/ for the item /klapas/ is provided below. In most of the cases there were more than one velocity peaks of the tongue back (TB) sensor after the release of the velar stop. Therefore, it was usually the case that the V initiation landmark did not coincide with the release landmark of the velar.

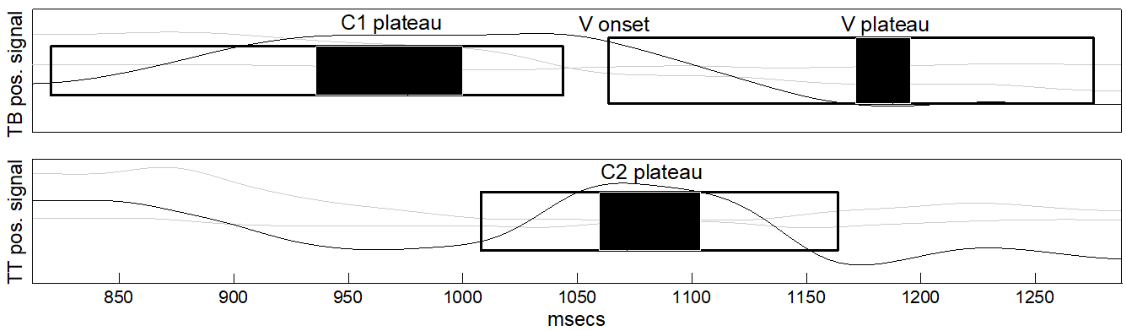

Figure A1: Parsing of the first two consonantal gestures (C1, C2) and the vocalic gesture (V) of the word /klapas/. Panels from top to bottom: tongue back (TB) positional signal in the superior-inferior (dark line), anterior-posterior and lateral dimensions (faint lines), tongue tip (TT) positional signal in the superior-inferior (dark line), anterior-posterior and lateral dimensions (faint lines). The black filled rectangles indicate the constriction phases, or plateaus, of the C1 /k/, C2 /l/ and V /a/. The left and right sides of the white filled rectangles indicate the initiation and end of the gestures.

2. Comment on statistics and model comparisons

The reviewer asked for some clarification as to why I did not perform model comparisons and how I decided to use the variable C1 place as a fixed factor in one analysis (p. 46) but not in another one (p. 61).

Performing model comparisons is a common practice when one has lots of variables and one does not necessarily have specific predictions about them; thus, it is not clear which one should be included in the model without overfitting it. In the studies of this dissertation, however, we did not have lots of variables of interest; additionally, there were specific predictions regarding those variables of interest. In this case, it makes more sense to fit the linear mixed effects model with the variables of interest to an Omnibus Anova which examines the significance of the model. One can still do model comparisons in this case as well, but there is in principle nothing against using an Omnibus Anova the way it was done in this dissertation.

Regarding C1 place of articulation as a fixed effect, for the stop-laterals (p. 61) where I evaluate statistically the interval-based indices, the analysis is done in the following way. First, I fit a model with interval duration as a dependent variable. Fixed effects are: C1 voicing, cluster size, interval type. The three-way interaction of C1 voicing, cluster size and interval type is significant which means the way interval type changes as the number of consonants increases from CV to CCV depends on C1 voicing. Note that a four-way interaction of interval type, cluster size, C1 voicing and C1 place was not possible to be examined because the model was too complex for the amount of data available. Then (p. 61-62), I looked at voiced and voiceless stop-laterals separately and fitted two separate models. For each of these models, interval duration was the dependent variable. Fixed effects were interval type, cluster size and cluster type (which is /bl, gl/ for the voiced stop-laterals, and /pl, kl/ for the voiceless stop-laterals). This means that with the factor cluster type within the voiced (levels: /bl/, /gl/) and voiceless stop-lateral subset of data (levels: /pl/, /kl/), I do examine C1 place. Cluster type (or C1 place in other words) is examined within the voiced and voiceless stop-lateral subset, in order to see whether the result for voiced (and voiceless) stop-laterals holds true for each of the two cluster types.

Regarding the effect of C1 place of articulation on the vowel initiation patterns, the reviewer comments that she/he would expect that there would be a C1 place of articulation effect on the vowel initiation patterns since the tongue dorsum is far less constrained in labial stop-laterals than in velar stop-laterals. The figure below illustrates vowel initiation patterns for each word-initial CCV cluster /bl, pl, gl, kl/ across speakers in German in the word boundary condition (control condition). There is clearly an effect of C1 voicing on vowel initiation patterns with the V starting later in C1 voiceless stop-laterals /pl, kl/ than in C1 voiced stop-laterals /bl, gl/. There is no effect of C1 place of articulation with the V starting earlier in labial stop-laterals /bl, pl/ than in velar stop-laterals /gl, kl/ due to articulatory constraints as the reviewer would expect.

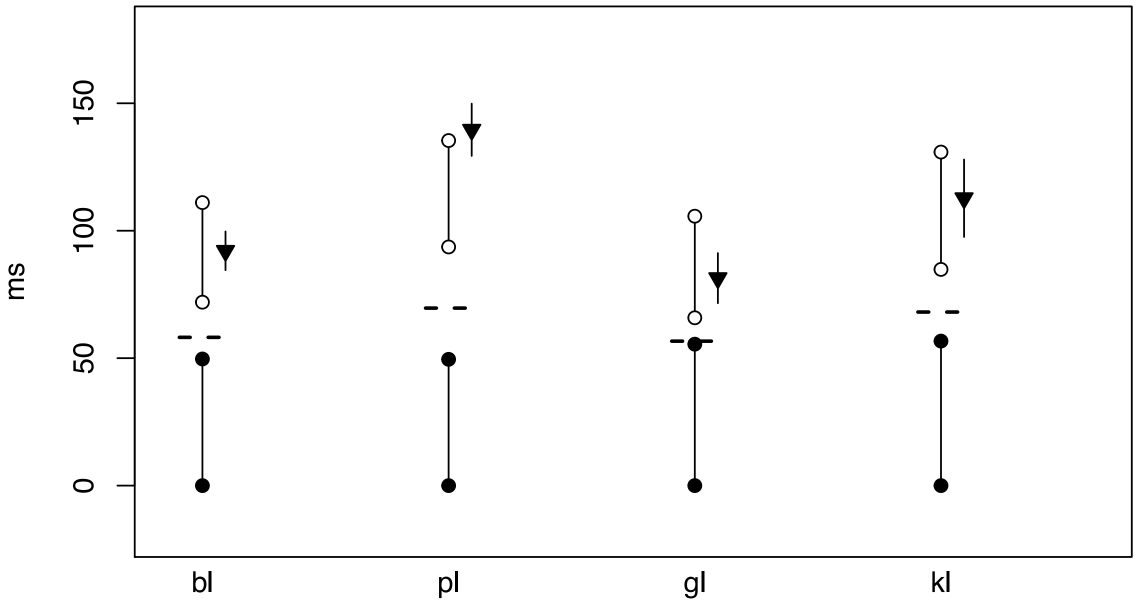

Figure A2: Vowel onset in relation to prevocalic consonants in the word-initial CCV clusters /bl, pl, gl, kl/ (x-axis) across speakers in the word boundary condition in German. The vertical lines denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals delimited by white dots indicate the plateau onset and offset timestamps of the prevocalic /l/ consonant. The black triangle indicates the vowel start, plus-minus SE of mean, in relation to the plateaus of the two consonants. The horizontal dotted line indicates the c-center landmark of the cluster.

3. Comment on using "speaker" as a fixed effect

According to Crawley (2013) a prerequisite for including a factor as random is the large number of levels, otherwise it is better to be included as a fixed effect because the model will not perform a proper estimation of random effects. In the studies of this dissertation on Spanish and German, there were 6 and 5 speakers respectively, which are not considered large numbers. Furthermore, an Anova comparison (not reported in the dissertation) comparing the model with speaker as a fixed effect and the model with speaker as random factor showed a preference of the former over the latter. However, in many cases, I did try models with speaker as a random factor (not reported in the dissertation) and the results between the model with speaker as a fixed effect and speaker as a random factor were not qualitatively different, sometimes not even quantitatively. Note that when including “speaker” as a fixed effect, the model provides an intercept for each speaker and thus the model does account for variation between speakers. Perhaps it is not ideal to include “speaker” as a fixed effect because one does not get one intercept across speakers which shows how great the variance among speakers is (which is what the reviewer asks), but it is, however, better than failing to do a proper estimation of random effects due to a small number of levels for “speaker”.

4. Comment on the test sentences on the German study (Chapter 3)

   Regarding the experimental description and the test sentences for the C#CV words, the carrier phrases are the following: “Zunächst sagte Anna „pack“, „Lage“ war das nächste Wort“ for the intonational phrase boundary condition and “Zunächst sagte Anna „pack“. „Lage“ war das nächste Wort“ for the utterance phrase boundary condition.