Journal of Phonetics 82 (2020) 100995

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

Global organization in Spanish onsets

Stavroula Sotiropoulou [a,*], Mark Gibson [b], Adamantios Gafos [a,c]

[a] University of Potsdam, Haus 14, Karl-Liebknecht-Str 24/25, 14476 Potsdam, Germany
[b] Universidad de Navarra, Campus Universitario, Edificio Amigos, 31009 Pamplona, Spain
[c] Haskins Laboratories, 300 George St. 900, New Haven, CT 06511, USA

ARTICLE INFO

Article history:
Received 23 August 2018
Received in revised form 18 April 2020
Accepted 18 April 2020
Available online 25 August 2020

Keywords:
Syllables
Inter-segmental coordination
Central Peninsular Spanish
Obstruent-lateral clusters
Obstruent-rhotic clusters

ABSTRACT

This paper addresses the relation between syllable structure and inter-segmental temporal coordination. The data examined are Electromagnetic Articulometry recordings from six speakers of Central Peninsular Spanish (henceforth, Spanish), producing words beginning with the clusters /pl, bl, kl, gl, pɾ, kɾ, tɾ/ as well as corresponding unclustered sonorant-initial words in three vowel contexts /a, e, o/. In our results, we find evidence for a global organization of the segments involved in these combinations. This is reflected in a number of ways: shortening of the prevocalic sonorant in the cluster-initial case compared to the unclustered case, reorganization of the relative timing of the internal CV subsequence (in a CCV) in the obstruent-lateral context, early vowel initiation, and a strong compensatory relation between the duration of the obstruent-to-lateral transition and the duration of the lateral. In other words, we find that the global organization presiding over the segments partaking in these tautosyllabic CCVs is pleiotropic, that is, simultaneously expressed over a set of different phonetic parameters rather than via a privileged metric such as c-center stability or any other such given single measure (employed in prior works).

1. Introduction

A basic idea about syllable structure is that it corresponds to a form of “glue” binding the group of segments over which that structure is imposed. A corresponding well-known hypothesis about how this binding can be assessed with phonetic data is by measures of temporal stability over intervals coextensive with the respective phonetic strings whose organization is at issue (Browman & Goldstein, 1988). Two patterns of stability have emerged, each thought to be characteristic of a particular syllabic organization.

In languages that admit complex onsets like English, it is often reported that 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; Byrd, 1995; Honorof & Browman, 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, Gafos, Hoole, & Zeroual, 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, Chitoran, & Selkirk, 2007; Hermes, Auris, & Mücke, 2015 on Berber). Thus, the stabilities of intervals seem to change depending on syllabic structure. These patterns of interval stabilities, global timing versus local 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 interval-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, Mücke, & Grice, 2013 on Italian, Shaw et al., 2009 on Arabic, Marin & Pouplier, 2010 on English, Pouplier, 2012; Brunner, Geng, Sotiropoulou, & Gafos, 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

* Corresponding author.
E-mail address: stsotiro@uni-potsdam.de (S. Sotiropoulou).

https://doi.org/10.1016/j.wocn.2020.100995
0095-4470/

languages (Arabic and English) are considered to be prototypical examples of the complex (English) versus simplex (Arabic) onset 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) investigate /pl, kn, gl, kv, sk/ word-initially, word-medially, and under two accentuation conditions. Pouplier (2012) finds some evidence for the expected stability pattern in onsets but also strong effects of cluster identity on the stability patterns. Brunner et al. (2014) find evidence for the same result, reporting the expected stability pattern 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 global timing stability, even though for the clusters investigated (e.g., /gl, pl/) there is solid phonological evidence for their syllabic onset status (Wiese, 1996). German is not the only language for which divergent results have been reported. Thus, Marin (2013), on Romanian, finds evidence for the expected stability pattern in sibilant-initial onset clusters /sk, sp, sm/ but also the unexpected local timing stability for stop-initial onset clusters /ps, ks, kt, kn/ although native speakers classify the latter as well as the former as complex onsets. Hermes, Mücke, and Auris (2017) on Polish report the expected stability pattern only for the /pl/ cluster with the clusters /pr, kr, kl/ deviating from it.

Given this background, we turn to Spanish which promises to be particularly informative in furthering our understanding of the relation between syllable structure and its phonetic indices for the following reasons. Spanish is a language which shows certain phonetic characteristics 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 in Spanish are timed with an open transition (Gibson, Sotiropoulou, Tobin, & Gafos, 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 for clusters in general and not only for those in which C2 is a rhotic or a lateral (Gafos, 2002; Gafos, Hoole, Roon, & Zeroual, 2010; Gafos, Roeser, Sotiropoulou, Hoole, & Zeroual, 2020). 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, Roeser, Sotiropoulou, Hoole, & Zeroual, 2020). 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.

In the context of past work on evidence for syllabic organization in phonetic data which tends to consider one or two (presumed) privileged measures, we quantified several different spatio-temporal measures for the first time: specifically, 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 main oral C constrictions) and sonorant duration, and finally vowel initiation with respect to the preceding segment or segmental combination. We argue that joint consideration of these measures enables one to see patterns of how different measures relate to one another and provides crucial evidence for global organization of the segments that are part of the hypothesized syllabic units. This evidence would not otherwise (i.e., when looking at isolated measures) be discernible.

The paper is organized as follows. We begin in Section 2 by outlining the method and stimuli used for the Spanish articulatory data. A detailed analysis of the articulatory data collected then follows in Section 3. As the Spanish clusters studied here have not been the subject of an articulatory study previously, we devote considerable attention, in Section 3, 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 tautosyllabic 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 4 pulls together 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. We conclude in Section 5 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. In particular, the main implication of our study is that global organization, holding over the segments in tautosyllabic CCVs, is simultaneously expressed via a set of different phonetic parameters and relations between these parameters rather than via a privileged metric such as c-center stability or any other such given single measure employed in prior works.

2. Method

2.1. 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.

2.2. Speech material

The stimuli consist of real disyllabic words in Spanish starting with consonant clusters or single consonants. Table 1 presents the list of Spanish CCV and CV paired words. CC-initial words began with stop-lateral or stop-rhotic clusters. Their paired single consonant-initial words began 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 word-initial /pl/, /bl/, /kl/, /gl/, /pɾ/, /kɾ/, /tɾ/.[1] There were 38 different CV- and CCV- initial words per speaker. For both clusters and singletons, the tautosyllabic 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. 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 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.

[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.

2.4. Articulatory segmentation

Articulatory segmentation consists in identifying the points in time where characteristic events such as onset of movement, achievement of target, and movement away from the target for a consonant or a vowel take place. For each consonant in the cluster of a cluster-initial or singleton consonant-initial word, the consonant(s), the subsequent vowel and the postvocalic consonant temporal landmarks were measured using the primary articulator(s) involved in their respective production. Thus, velar consonants (/k, g/) were measured using the most posterior TB sensor and labials (/p, b, m/) using the lip aperture (LA). LA is a derived signal computed as the Euclidean distance between upper and lower lip sensors. Coronals (/t, l, r, ɾ/) were measured using the TT sensor. Although the lateral and the rhotic/tap are often considered multigestural with a dorsal gesture in addition to the coronal gesture (Proctor, 2011), dorsal gestures are not consistently present during the production of the sonorants in our dataset. Thus, the only way to measure the data consistently was to use the tongue tip gesture which is consistently present throughout our data and reliably measured; moreover, as a consequence of our use of that gesture to measure the sonorants, our work is comparable to previous work on the same topic of the relation between syllabic organization and articulation (all such work uses the tongue tip; see for example Browman & Goldstein, 1988, Goldstein, Nam, Saltzman, & Chitoran, 2009, Marin & Pouplier, 2010, Marin & Pouplier, 2014, Hermes et al., 2013, Hermes et al., 2017). The tautosyllabic vowels /a, e, o/ 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 (developed by Mark Tiede at Haskins Laboratories). The algorithm first finds the peak velocities (to and fro the constriction) and the minimum velocity within a user-specified zoomed in temporal range. The achievement of target and the constriction release landmarks were then obtained by identifying the timestamps at which velocity falls below and rises above a 20% threshold of the local tangential velocity peaks. The gestural onset (onset) and gestural offset (offset) were obtained by identifying the timestamps at which velocity rises above and falls below a 20% threshold of the local tangential velocity peaks. Fig. 1 illustrates an example parse of the velar gesture in the initial consonant of the word /glato/ using the TB sensor. The top panel illustrates the TB movement trajectory in the vertical component and the bottom panel illustrates the tangential velocity of the TB sensor. The dashed vertical lines from left to right indicate the landmarks: gestural onset (onset), peak velocity to target (Peak Velocity to), achievement of target (Target), minimum velocity (maximum constriction), constriction release (Release), peak velocity away from the target (Peak Velocity fro), gestural offset (Offset).

Table 1
Stimuli. The orthographic "c" word-initially represents the phoneme /k/.

	Low vowel (/a/)		Mid vowel (/e/, /o/)	
Cluster	CCV	CV	CCV	CV
pl	plato	lato	plena	Lena
	'plate'	'to whip'	'full'	(proper name)
			plomo	lomo
			'lead'	'loin'
bl	blata	lato	bleque	leco
	(model of motorcycle)	'to whip'	'tar'	'nuts'
	blanda		bloque	loco
	'soft'		'block'	'crazy'
gl	Glato	lato	gleba	lema
	(province in Italy)	'to whip'	'mound of land'	'slogan/motto'
	Glana		globo	lomo
	(surname)		'balloon, globe'	'loin'
kl	clapas	lapa	clema	lema
	'stripping of unfertile land'	'Lap (from Lapland)'	'electrical connector'	'slogan/motto'
			clono	
			'clone'	lomo
				'loin'
pɾ	Prato	rato	presa	Rena
	(city in Italy)	'while, bit'	'pray'	(proper name)
			promo	romo
			'promotion'	'blunt'
kɾ	Crapa	rapa	crema	rema
	(name of Italian restaurant in Valencia)	'flower of olive tree'	'cream'	'to row'
			cromo	romo
			'chrome'	'blunt'
tɾ	trapo	rape	trecho	Recho
	'rag'	'monkfish'	'stretch'	(fishing company in Galicia)
			trono	roto
			'throne'	'broken'

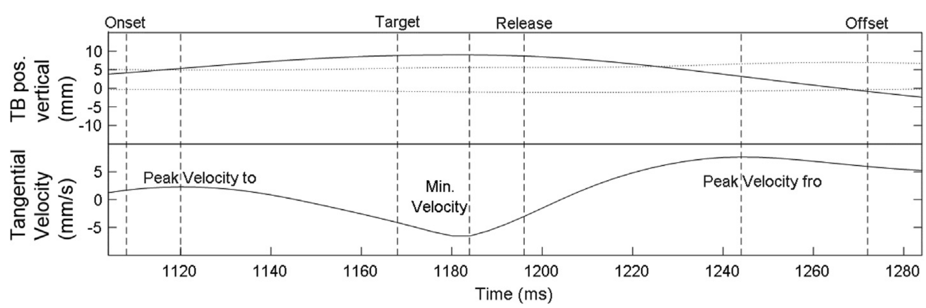

Fig. 1. Parsing of the velar gesture of the first consonant in the word /glato/. The top panel shows the tongue back (TB) movement trajectory in the superior-inferior (vertical) dimension and the bottom panel shows the tangential velocity of the TB sensor during the production of the velar consonant. The dashed vertical lines from left to right indicate the landmarks: gestural onset (Onset), peak velocity to target (Peak Velocity to), achievement of target (Target), minimum velocity (maximum constriction), constriction release (Release), peak velocity away from the target (Peak Velocity fro), gestural offset (Offset).

2.5. Data analysis

2.5.1. Inter-gestural timing and consonant duration

We quantify the timing between the two consonants C1 and C2 in a C1C2V using the articulatory interplateau interval (henceforth, IPI). IPI 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)). Consonant plateau duration was calculated as the interval between C release and C target. C plateau duration was normalized by dividing the raw plateau duration by the total duration of the cluster: C plateau normalized = C plateau duration/(C2 release – C1 target). Fig. 2 illustrates IPI and consonant plateau duration as defined above for the two consonants C1C2 of the /gl/ cluster in /glato/.

In the present work, we also examine the relation between IPI and the prevocalic consonant C2 in a C1C2 cluster and specifically how the duration of the prevocalic consonant responds to changes in duration of the IPI in a cluster. In a recent comparison between a subset of the data investigated in the present study, based on two speakers only and words beginning with /kla/ only, and Moroccan Arabic data of the same segmental sequence, Gafos, Roeser, Sotiropoulou, Hoole, & Zeroual (2020) find that a long lag in terms of IPI between the /k/ and the lateral must be compensated in Spanish by shortening of the lateral, but that no such relation is found in Moroccan Arabic. Gafos, Roeser, Sotiropoulou, Hoole, & Zeroual (2020) express this by saying that segments are appended linearly, one after another, in Arabic, but in Spanish adding a segment (/k/) to a sequence (/la/) or lengthening the /k/ in /kla/ has consequences for the spatio-temporal properties of its adjacent segments: increasing the lag between the constrictions of the

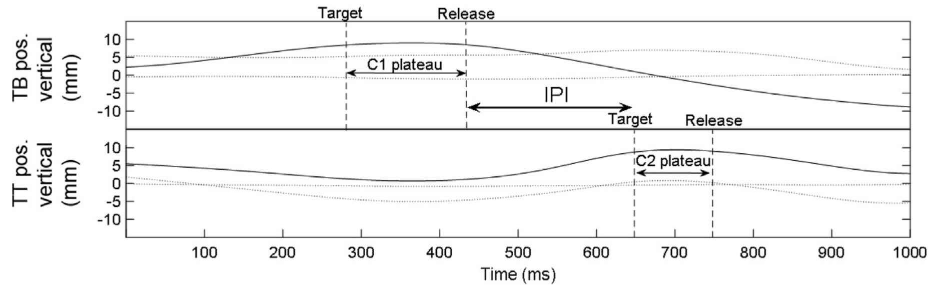

Fig. 2. Consonant plateaus for the first (C1) and second (C2) consonant and interplateau interval (IPI) of the C1C2 cluster in the word /glato/. The top panel shows the tongue back (TB) movement trajectory in the superior-inferior (vertical) dimension and the bottom panel shows the tongue tip (TT) movement trajectory in the superior-inferior (vertical) dimension. The intervals between the dashed vertical lines on the TB trajectory and on the TT trajectory demarcate the plateau duration of the initial consonant C1 and the second consonant C2. The interval between the release of the initial consonant C1 and the target of the second consonant C2 is the interplateau interval (IPI).

two prevocalic consonants (those of the stop and the lateral) implies shortening of the second consonant (the lateral) in Spanish. 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. We wish to assess whether the same relation seen in that earlier study in Gafos, Roeser, Sotiropoulou, Hoole, & Zeroual (2020) holds true when we examine data from more speakers and more sequences.

2.5.2. Interval-based indices

Past assessments of syllabic organization make use of the stability of certain intervals (described below) computed across CV, CCV stimuli pairs. For the stability analysis, two intervals were calculated for each stimulus observation. We first describe the right-delimiting landmarks (henceforth, anchors) used in defining these intervals. For thoroughness, in assessing the robustness of results, we used three different anchors: the target of the constriction of the postvocalic consonant (C^{tar}); that is, the final consonant in any CVC, CCVC stimulus over which these intervals were calculated, the maximum constriction of the postvocalic consonant (C^{max}), and the spatial extremum of the tautosyllabic 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 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 prevocalic consonants (hence, global) as opposed to ‘local’ for the second interval which is left-delimited by the constriction release of just the immediately prevocalic consonant. Fig. 3 illustrates the intervals global timing and local timing defined above.

A descriptive characterization of the data 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 lower 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.

2.5.3. Consonant-vowel timing

Consonant-vowel timing refers to the timing between the prevocalic consonant and the vowel. Relative timing is indexed by the lag between the release of the lateral and the maximum opening of the following vowel (CV lag) in the /lV/ and stop-/lV/ contexts (Goldstein et al., 2009; Marin, 2013). We use this measure to examine if and how the relative timing of the lateral and the vowel changes across different contexts, as in /lV/ versus stop-/lV/. Fig. 4 illustrates this CV lag measure as defined above in the /loko/ token.

2.5.4. Vowel initiation

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 slower and 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 (Browman & Goldstein, 1988, Honorof & Browman, 1995, Marin & Pouplier, 2010, Marin, 2013, among others).

The current work seeks to quantify vowel initiation so as to gain an additional window into possible effects of syllabic organization on articulation. Our measurements are based only on cases where the gestural onset of the vowel could be reliably

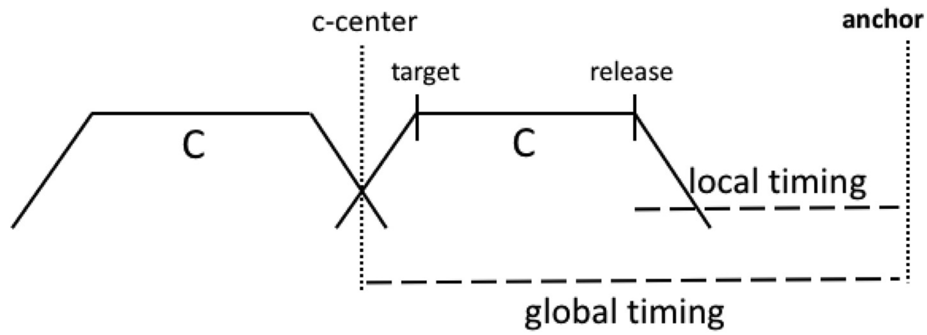

Fig. 3. Two intervals spanning over a sequence of a consonant cluster (CC) followed by a vowel. The global timing interval is delimited on the left by the c-center landmark of the CC cluster (defined as the mean of both consonants and hence 'global') and to the right by the anchor landmark which is chosen to be a time point close to the end of the vowel. The local timing interval is delimited on the left by the release of the immediate prevocalic consonant and to the right by the same anchor (as for the global timing interval).

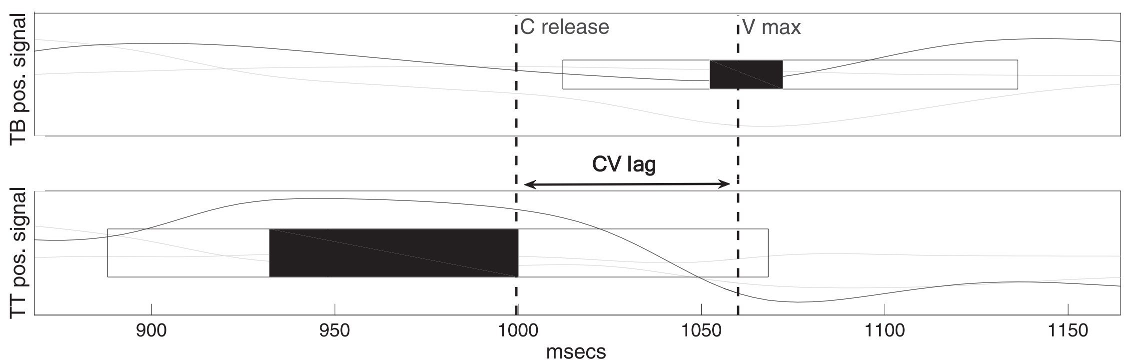

Fig. 4. Consonant-vowel (CV) lag: the top panel shows the tongue back (TB) movement trajectory in the superior-inferior (vertical) dimension and the bottom panel shows the tongue tip (TT) movement trajectory in the superior-inferior (vertical) dimension. The CV lag spans from the release of the consonant (C release) to the maximum opening of the vowel (V max).

parsed. Landmark identification for the vowel gesture 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, 7% for the stop-lateral cases and 10% for the stop-rhotics, 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 results we provide consist of vowel initiation patterns with respect to the prevocalic consonants in a CCV context where the initial consonant in the cluster is either a voiced or a voiceless stop. Fig. 5 illustrates an example of vowel initiation (V onset) with respect to the prevocalic C1C2 /kl/ cluster in /klapas/.

2.5.5. Statistical analysis

All statistical analyses were carried out using R Studio version 3.3.1 (RStudio Team, 2015). Linear mixed effects models were fitted to the data using the package *lmerTest*. The *lmerTest* estimates the degrees of freedom by using the Satterthwaite approximation providing relatively conservative *p*-values (Kuznetsova, Brockhoff, & Christensen, 2017). Since there were not enough levels for a random effects estimation for speaker (see Crawley, 2013), speaker was considered a fixed effect to account for data variance.[2] The linear mixed effects model was then fitted into an Omnibus Anova to test the significance of the model and specifically the significance of the interaction between factors. Post hoc pairwise comparisons were carried out using the *lsmeans* package (Lenth & Hervé, 2015). We accounted for multiple comparisons on the same dataset by adjusting the *p*-values using Tukey correction. Histograms were generated before modeling the data to test for the possible need to perform logit and/or square root transformations. The *symbox* function of the *car* package was used to determine which transformation should be carried out in order to better approach a normal distribution of the data. Separate models were fitted to the stop-lateral subset and the stop-rhotic subset of the data. For the stop-lateral subset of data and the respective lateral-vowel (CV) sequences, 15% of the data (220 out of 1425 items) were excluded due to technical problems with data recording and processing or due to cases where the parsing algorithm was failing to provide a reliable parse of some articulatory gestures. For the stop-rhotic subset

[2] Furthermore, an Anova comparison 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.

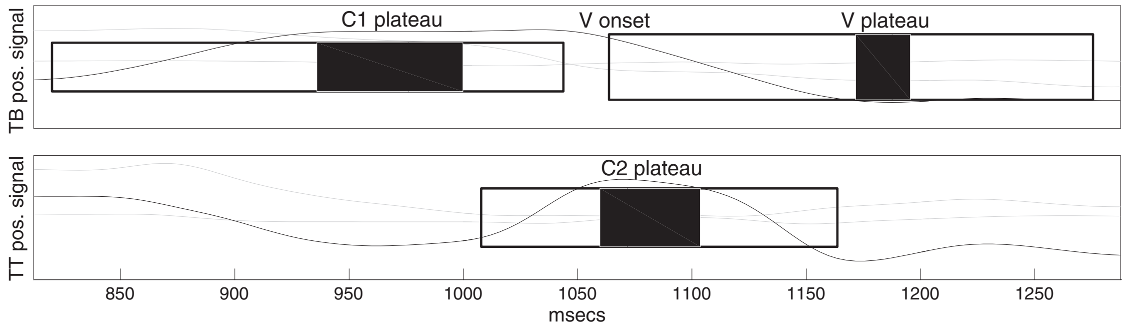

Fig. 5. Vowel initiation with respect to the prevocalic consonants in /klapas/. The top panel shows the tongue back (TB) movement trajectory in the superior-inferior (vertical) dimension and the bottom panel shows the tongue tip (TT) movement trajectory in the superior-inferior (vertical) dimension. The dark filled boxes indicate plateaus of the initial consonant (C1), the second consonant (C2) and the vowel (V). Vowel initiation (V onset) is indicated by the beginning of the white box which highlights the vowel parse on the TB movement trajectory.

of data and the respective rhotic-vowel (CV) sequences, 12% of the data (130 out of 1062 items) were excluded due to the same reasons reported above for stop-laterals.

3. Spatio-temporal properties in Spanish onsets

3.1. Interplateau interval in stop-lateral clusters

In this subsection, we examine how the interplateau interval (IPI) changes as a function of C1 voicing in stop-lateral clusters. Recall that the interplateau interval (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. A normalized IPI was also calculated 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 to stop-laterals (N = 592) with IPI (raw and normalized) as dependent variables, C1 voicing (voiced, voiceless) and speaker as fixed effects.[3] 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. Mean and standard deviation 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.

Raw IPI measures exhibit substantial variability as can be seen by their standard deviations in Table 2. Some of this variability is speaker-specific and cluster-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. Fig. 6 illustrates normalized IPI values as a function of C1 voicing. It is evident that when IPI is normalized by taking into account total cluster duration, regardless of C1 voicing, there is no difference in IPI.

[3] C1 place of articulation (labial, velar) has been used as a fixed effect in the statistical models in sections 3.1, 3.2, 3.3 but the current paper focuses on C1 voicing and we thus report results only on that. No significant interaction between C1 place and C1 voicing has been observed in any of the statistical models.

Table 2
Mean and standard deviation of IPI as a function of C1 voicing 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

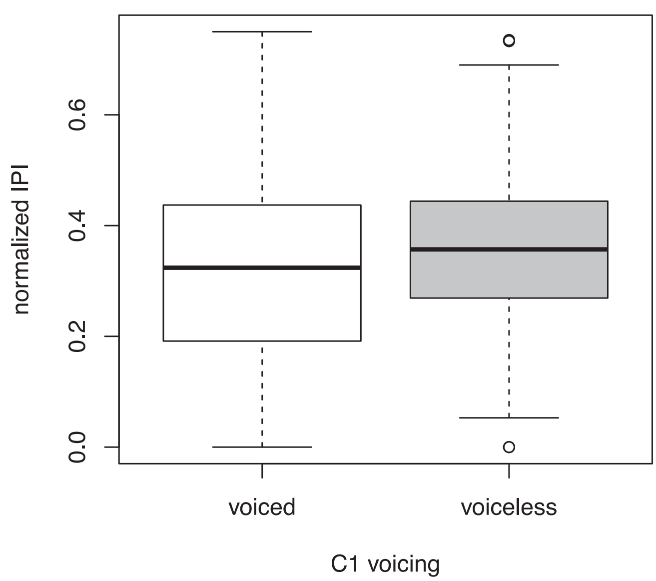

Fig. 6. Normalized interplateau interval (IPI) as a function of C1 voicing in stop-lateral clusters.

3.2. C1 duration in stop-lateral clusters

In this subsection, we examine how C1 plateau duration changes as a function of C1 voicing in stop-lateral clusters. The ways C1 plateau duration varies as a function of C1 voicing turns out to be useful when we examine how the lateral and the vowel (the two segments that follow the stop) react to changes in C1 plateau duration. For example, we will see that the duration of the lateral and the vowel initiation patterns are different depending on the C1 duration. Such relations among different parts of the string are examined in our study for the first time and as we argue throughout provide novel ways of diagnosing the nature of the organization throughout the CCV sequence even when other measures used in past work provide inconsistent or contradictory evidence.

Recall that 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. For the statistical analysis, linear mixed effects models were fitted to stop-laterals ($N = 592$) with C1 plateau duration (raw and normalized) as dependent variables, C1 voicing (voiced, voiceless) and speaker 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). With the normalized C1 plateau duration as dependent variable, the main effect of C1 voicing is maintained ($p = 0.01$). Fig. 7 illustrates normalized C1 plateau duration as a function of C1 voicing. The boxplots for C1 voiceless clusters show longer C1 plateau duration than the C1 voiced clusters.

Mean and standard deviation 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 3.

3.3. C2 lateral duration

In this subsection, we examine C2 lateral plateau duration as a function of C1 voicing. 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). For the statistical analysis, linear mixed effects models were fitted to stop-laterals ($N = 592$) with C2 plateau duration (raw and normalized) as dependent variables, C1 voicing (voiced, voiceless) and speaker 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. 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 ($\beta = -0.06$) than after a C1 voiced stop. Fig. 8 shows normalized C2 plateau duration as a function of C1 voicing.

Mean and standard deviation 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 4.

Segments in a cluster context have been found to be shorter than segments in the singleton context (Haggard, 1973). Table 5 presents mean and standard deviation 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 section 4.

3.4. IPI-C2 compensatory relation in stop-lateral clusters

Next, we examine the relation between two phonetic parameters, IPI and C2 lateral duration, in stop-lateral clusters. Specifically, we investigate how the duration of the lateral, the second consonant in the C1C2V sequence, responds to an increase in duration of the lag (IPI) between the two consonants. 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 region of the string. The presence of such effects would indicate that the different parts of CCV are not independently planned and produced and thus such effects, if present, offer evidence for global organization.

We find 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.[4] The compensatory relation between IPI and C2 duration can be seen in Fig. 9. The effect remains the same for both velars and labials.

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.

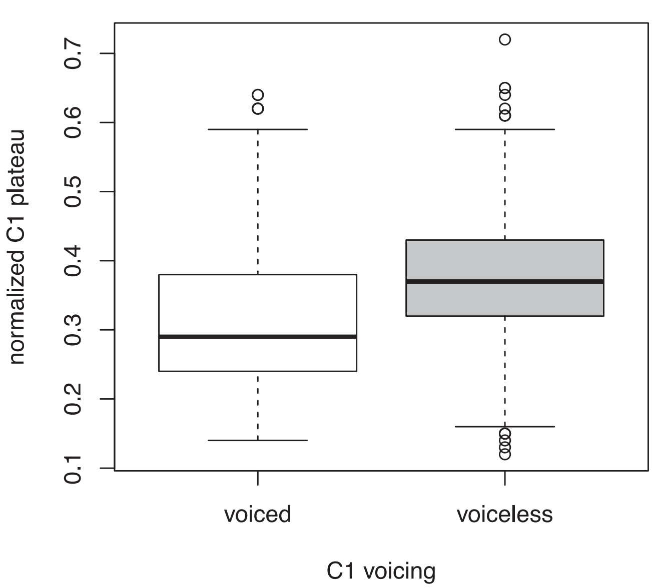

Fig. 7. Normalized C1 plateau duration as a function of C1 voicing. C1 plateau is greater for C1 voiceless stop clusters /pl, kl/ than for C1 voiced /bl, gl/.

[4] 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.

Table 3
Mean and standard deviation for C1 stop duration.

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

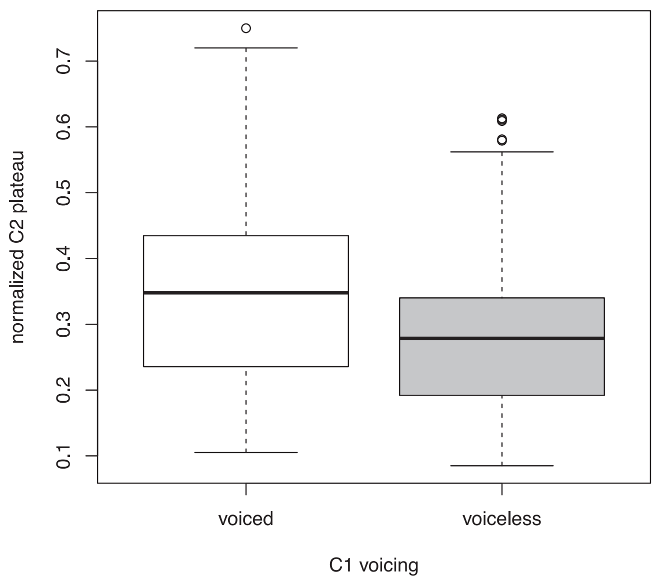

Fig. 8. Normalized C2 lateral plateau duration as a function of C1 voicing. The C2 lateral in voiceless stop-laterals is shorter than in voiced stop-laterals.

3.5. Stop-rhotic clusters

The stop-rhotic clusters in our dataset consist of the voiceless stop-tap clusters /pɾ, kɾ, tɾ/ (*N* = 466). 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. Table 6 provides mean and standard deviation 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).

Table 7 provides the duration of the 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 (Hualde, 2005; Lipski, 1990). These two segments have inherently different duration with the trill being longer than a tap. In the case of rhotics, the point is that the duration of the rhotic is substantially different across CV and CCV. Such a difference in duration between the segments is not met in the case of the lateral in CCV and CV. This large difference in the segments' duration in rhotics

Table 5
Mean and standard deviations for lateral duration in CV (/lV/) and CCV (stop-/lV/).

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

as opposed to laterals contributes to the emergence of a consistent global timing interval stability in the former as opposed to the latter (see Section 4). Our aim is to convey how stability-based heuristics behave in certain phonetic environments, that is, within the stop-rhotics and within the stop-laterals separately. Our aim is explicitly not to compare stop-rhotics and stop-laterals to one another (these contexts are not comparable because, as noted above, the rhotic is a different segment in the CV versus the CCV context, which is not the case for the lateral).

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 ($r(447) = 0.25$, $p < 0.0001$).

3.6. Interim summary

To summarize, our results confirm previous findings on the duration of consonants in complex onsets. Specifically, C1 plateaus for voiceless consonants were significantly longer than for voiced as per the well-known cross-linguistic generalization (Lisker, 1957). The results of the effects that C1 voicing has on the duration of the lateral are in line with Gibson, Fernández Planas, Gafos, and Remirez (2015) EPG study on Spanish 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 voicing. Specifically, we found that lateral duration and C1C2 interplateau interval (IPI) are 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.

Table 4
Mean and standard deviation for C2 lateral duration in CCV.

	Duration C2 (ms)				Normalized C2 duration			
	Labial	*sd*	Velar	*sd*	Labial	*sd*	Velar	*sd*
C1 voiced	40	19.9	43.3	20.5	0.36	0.12	0.35	0.17
C1 voiceless	38.3	18.3	36.6	13.9	0.27	0.09	0.27	0.11

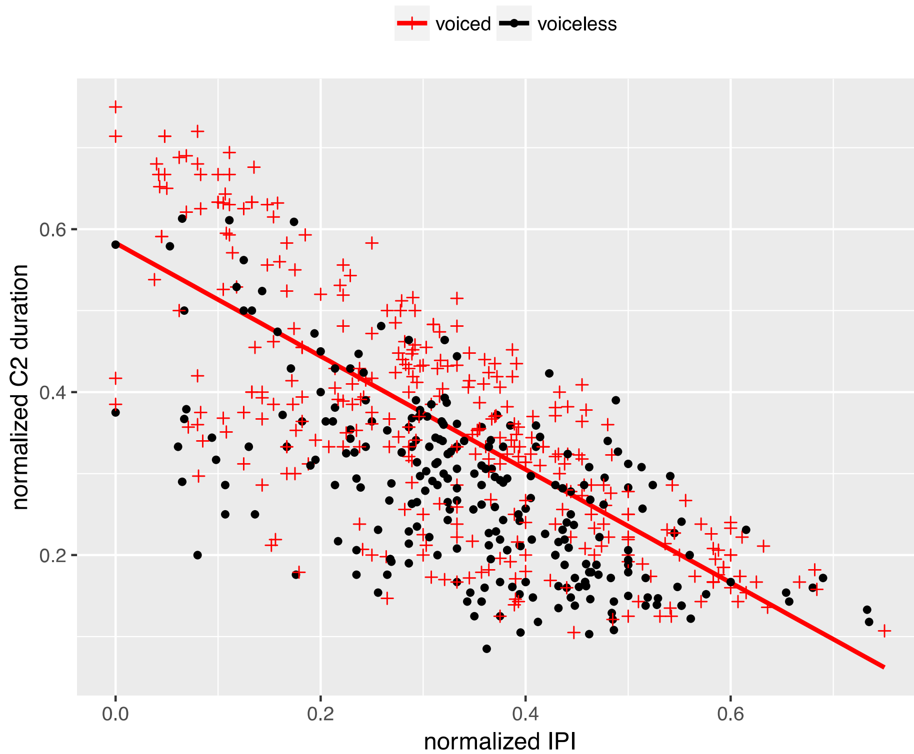

Fig. 9. Scatterplot showing the compensatory relation between normalized C2 lateral duration and normalized IPI for voiced and voiceless stop-lateral clusters. As IPI increases, the lateral's (C2) duration decreases ($r(587) = -0.70$, $p < 0.0001$).

Table 6
Mean and standard deviation for interplateau intervals (IPI).

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 7
Mean and standard deviation for prevocalic rhotic duration in CV and CCV.

Rhotic 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

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 that IPI in stop-lateral clusters is overall large regardless of the stop's voicing. Therefore, 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 by Bombien and Hoole (2013) before and we can verify with our own data, the IPI patterns are quite different from those in French and Spanish. Thus, voiced stop-lateral clusters in German show a very short, often negative IPI indicating temporal overlap between the plateaus of the consonants. In contrast, voiceless stop-lateral clusters in German show a larger IPI of approx. 30 ms. Fig. 10 illustrates the IPI difference between German and Spanish from our own data for a /gl/ cluster. In German, there is temporal overlap between the plateaus of the consonants, while in Spanish there is no overlap between the plateaus.

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 the language-particular implementation of the voicing contrast (such as the difference between oral-laryngeal coordination in German/English versus Spanish/French) conditions inter-gestural 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-

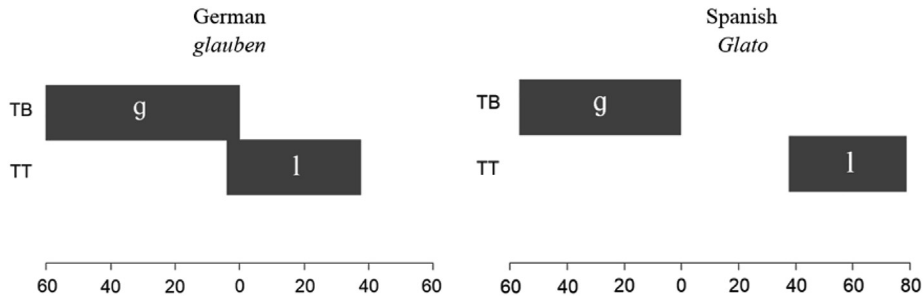

Fig. 10. Representative schemas of gestural overlap for the cluster /gl/ in German (lefthand side) and Spanish (righthand side). Spanish shows a substantial IPI or time lag between the plateaus (constriction phases) of the two consonants, as shown by the separation between the dark rectangles in the righthand schema, whereas in German the plateaus overlap in time.

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 of a vocoid like element, termed ‘Svarabhakti’/ ‘el elemento esvarabático’ (Lenz, 1892; Schmeiser, 2009), between the two consonants has been reported (see Quilis, 1970, Quilis, 1993, Levin, 1987, Harms, 1976, Hall, 2006, 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 Section 4).

3.7. 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 interval-based indices of syllabic organization, as have been employed in other work reviewed in Section 1, 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, 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.

3.7.1. Stop-lateral clusters

We begin with a first descriptive characterization of the data using the relative standard deviation (RSD) of the two intervals, global timing and local timing (as defined in subsection 2.5.2). Recall that the interval with the lowest RSD value 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 lower 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 the Appendix.

According to the interval-based indices of syllabic organization reviewed in Section 1 (Introduction), we expect Spanish clusters to exhibit global timing interval stability (RSD minima should be consistently found for global timing). However, the interval stabilities seem quite variable (see Appendix). By simply consulting the RSDs, 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 on stop-laterals ($N = 1205$). Furthermore, we investigate the effect of the stop's voicing on the interaction between cluster size and interval type. A linear mixed effects model was fitted with interval duration as a dependent variable. Interval duration was log transformed to approximate a normal distribution. Interval type, cluster size, C1 voicing, vowel and speaker were used as fixed effects. The variable “word”, which corresponds to each word in the corpus, was introduced as a random factor. 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}: $F(2) = 5.78$, $p = 0.003$). Vowel was not a significant factor in the model. The post hoc multiple comparisons provided the following results for each voicing condition. 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}: $p < 0.0001$, $\beta = -0.13$), while the local timing interval decreases (C^{tar}: $p < 0.0001$, $\beta = 0.13$). 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

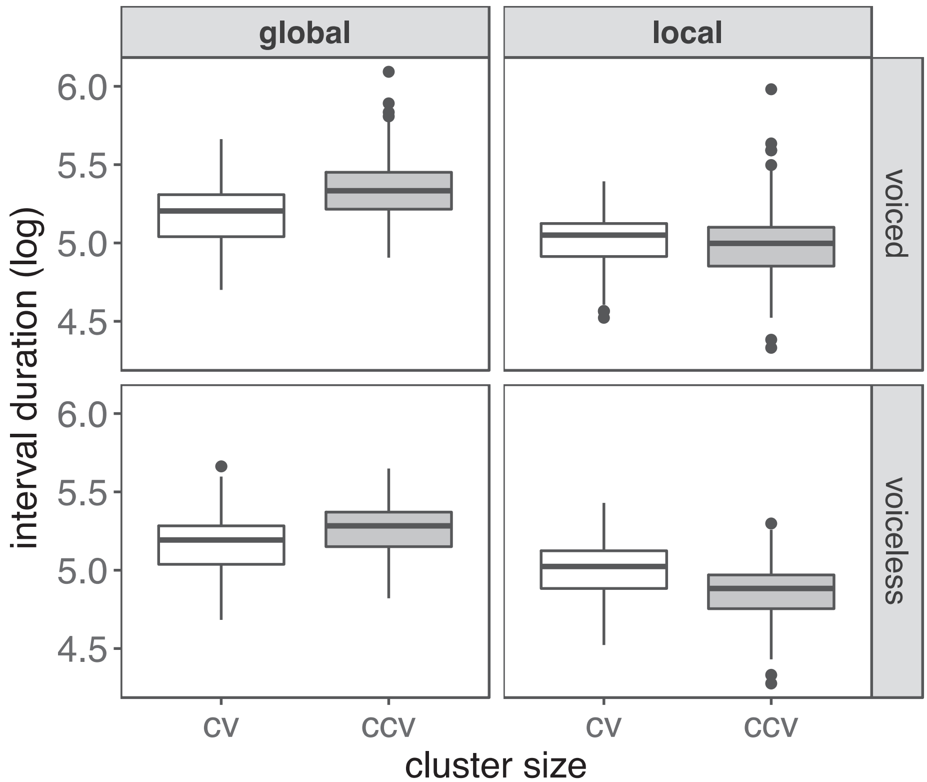

Fig. 11. Duration (in log scale) of the two intervals global timing and local timing for CV (white) and CCV (grey) words. For voiced stop-lateral CCV words (top), the difference between CV and CCV for the global timing interval is greater than the difference for the local timing interval. For voiceless 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.

(C^{max}: global timing $p = 0.0008$, $\beta = -0.10$, local timing $p < 0.0001$, $\beta = 0.13$; V^{max}: global timing $p = 0.0006$, $\beta = -0.77$, local timing $p < 0.0001$, $\beta = 1.31$). For voiced stop-laterals, the global timing was found to change significantly from CV to CCV (C^{tar}: $p < 0.0001$, $\beta = -0.16$, C^{max}: $p < 0.0001$, $\beta = -0.17$) while the local timing interval did not show any significant change (C^{tar}: $\beta = 0.04$, C^{max}: $\beta = 0.003$) using C^{tar} and C^{max} as anchor. When using the vowel as anchor, both the global and local timing intervals do not change significantly from CV to CCV in voiced stop-laterals (global: $\beta = -0.20$, local: $\beta = 0.24$).

Next, we examine within voiced and voiceless stop-laterals, whether there is an effect of cluster type on the way interval types change from CV to CCV. Therefore, we fitted separate linear mixed effects model within the voiced and voiceless stop-laterals with interval duration as a dependent variable and the interaction between interval type, cluster size and cluster type (levels: /pl, kl/ for the voiceless stop-lateral dataset, and levels: /bl, gl/ for the voiced stop-lateral dataset). 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/. The post hoc multiple comparisons provided the following results for each cluster within the voiced stop-laterals. For /bl/, the global timing interval increases significantly from CV to CCV (C^{tar}: global timing $p = 0.002$, $\beta = -0.17$; C^{max}: global timing $p = 0.003$, $\beta = -0.19$; V^{max}: global timing $p < 0.0001$, $\beta = -1.16$), 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}: $p < 0.0001$, $\beta = -0.16$; C^{max}: $p = 0.003$, $\beta = -0.13$; V^{max}: $p < 0.0001$, $\beta = -1.08$), and the local timing interval decreases (C^{tar}: $p = 0.001$, $\beta = 0.07$; C^{max}: $p = 0.03$, $\beta = 0.08$; V^{max}: $p < 0.0001$, $\beta = 0.88$). 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.[5]

[5] Although spirantization of voiced stops following vowels is a steadfast rule of Spanish phonology (Harris, 1983, Hualde, 2005), the way some subjects read the target word in the carrier sentence led to some of the consonants being produced as stops (because they paused before producing the target word) and others being produced as approximants (because there was no pause between the target word and the preceding vowel final word in the carrier phrase). 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 (spirantization does not take place in voiceless Cl clusters) and this may be a reason why local, not global, timing interval stability was found for the voiced C1 clusters (as the c-center landmark becomes more variable so does the global timing interval duration because that interval is left-delimited by the the c-center landmark). We, therefore, assessed in a further analysis whether excluding the spirantized cases would lead to the expected global timing interval stability. Typically, approximants are shorter in duration than stops. Based on the distribution of the duration of C1 voiced stops across speakers in our data and also on the fact that voiced stops in Spanish typically show duration of 70 ms (Martinez Celdrán, 1993, Gibson, Fernández Planas, Gafos, & Remirez, 2015), we decided to use the C1 plateau duration of 40 ms as a threshold. C1 voiced stops which are less than 40 ms were considered to be approximant realizations and thus excluded from the analysis for voiced stop-lateral clusters. The results of this analysis showed that even with a conservative elimination of these cases, the indices did not change; i.e., local timing interval stability was maintained.

Fig. 11 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).

To summarize, C1 voicing affects the way intervals change as the number of consonants increases from CV to CCV for stop-laterals. 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 (Ridouane, Hermes, & Hallé, 2014; Shaw et al., 2009, 2011).

3.7.2. Stop-rhotic clusters

To evaluate interval stability patterns for the stop-rhotic clusters (N = 962), 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/). The post hoc multiple comparisons provided the following results 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 more than the local timing interval (C^{tar}: global timing $\beta = -0.11$, local timing $\beta = 0.06$; C^{max}: global timing $\beta = -0.12$, local timing $\beta = 0.03$), while when using a vowel anchor the global timing changes less than the local timing interval (V^{max}: global timing $\beta = -0.14$, local timing $\beta = 0.26$). For the /kɾ/ cluster, although the two intervals do not change significantly, global timing changes slightly more than local timing from CV to CCV (C^{tar}: global timing $\beta = -0.13$, local timing $\beta = 0.10$; C^{max}: global timing $\beta = -0.12$, local timing $\beta = 0.08$). Using a vowel anchor, for /kɾ/, the local timing interval changes more than the global timing interval (local timing $\beta = 0.25$, global timing $\beta = -0.20$); the changes are, however, not significant. For /tɾ/, the local timing interval changes significantly (C^{tar}: local timing $p = 0.001$, $\beta = 0.26$; C^{max}: local timing $p = 0.001$, $\beta = 0.24$; V^{max}: local timing $p = 0.01$, $\beta = 0.57$), while the global timing interval does not change significantly (C^{tar}: $\beta = -0.12$; C^{max}: $\beta = -0.09$; V^{max}: $\beta = -0.14$). For none of the clusters was vowel identity (levels: low /a/, mid /e/, /o/) a significant factor. Fig. 12 illustrates the global timing and local timing intervals with C^{max} as anchor for each of the stop-rhotic clusters across speakers.

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 $p = 0.001$, $\beta = -0.12$, local timing $p = 0.0002$, $\beta = 0.14$; C^{max}: global timing $p = 0.001$, $\beta = -0.09$, local timing $p < 0.0001$, $\beta = 0.14$; V^{max}: global timing $p = 0.04$, $\beta = -0.16$, local timing $p = 0.0001$, $\beta = 0.36$). 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.

3.7.3. 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. Recall that 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. Speaker and the interaction between cluster size (CV, CCV) and C1 voicing were treated 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 ($p < 0.0001$, $\beta = 16.5$ ms). When the initial consonant is voiced, the /lV/ lag also decreases significantly from CV to CCV but to a lesser extent ($p = 0.01$, $\beta = 10$ 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.

3.7.4. Vowel initiation

The c-center organization pattern prescribes that the vowel starts somewhere around the c-center of the prevocalic onset cluster (Browman & Goldstein, 1988, 2000; Gafos, 2002; Honorof & Browman, 1995; Nam & Saltzman, 2003). 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 and Goldstein (1988:150) write 'let us make the following assumption: the (temporal) interval from the c-center to the final consonant anchor point is a measure of the activation interval of the vocalic gesture where: (a) the c-center corresponds to a fixed point early in the vocalic activation...' but also that 'We also assume that the actual movement for the vocalic gestures begins at the achievement of target of the first consonant in a possible initial cluster' (Browman & Goldstein, 1988:150). Honorof and Browman (1995: Fig. 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 and Saltzman

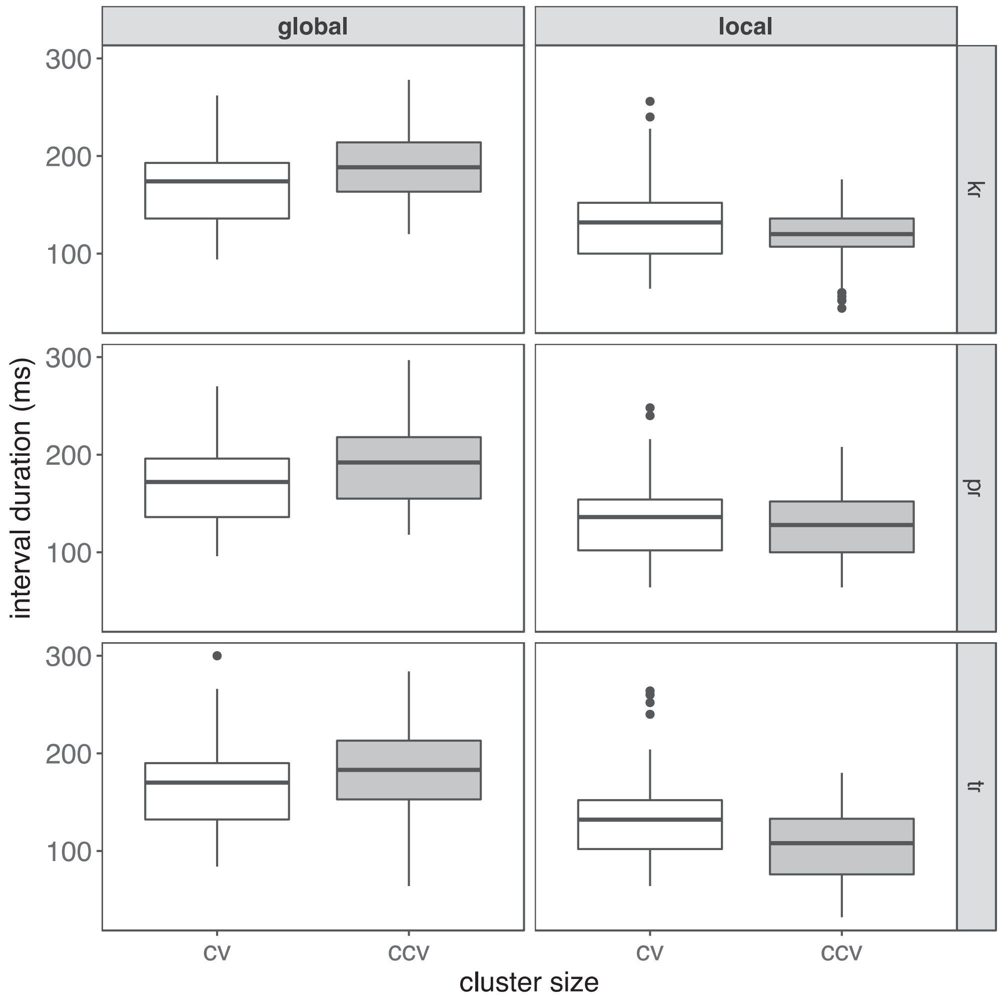

Fig. 12. Duration (in ms) of the intervals global timing and local timing for CV (white) and CCV (grey) words of the /kɾ/, /pɾ/, /tɾ/ clusters.

(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.

Fig. 13 shows vowel initiation in C1 voiced and C1 voiceless stop-lateral clusters. Fig. 14 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 sonorant respectively. The triangle indicates vowel initiation and the horizontal dotted line represents the c-center landmark. As can be seen in Fig. 13, 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.

For stop-rhotic clusters, as can be seen in Fig. 14, 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 these 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 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.

4. Multiplicity of phonetic indices for global organization

According to the interval-based indices of syllabic organization reviewed in Section 1 (Introduction), 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 partaking in the syllable. In what follows, we flesh out what is meant by global organization more

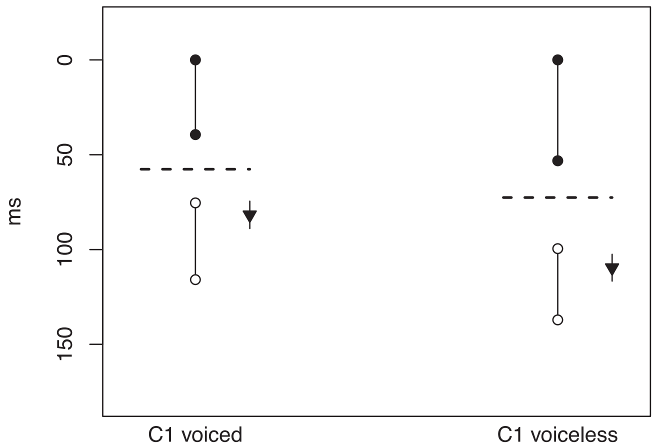

Fig. 13. Vowel onset in relation to prevocalic consonants in C1 voiced and C1 voiceless stop-lateral clusters (x-axis). The vertical lines denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. 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.

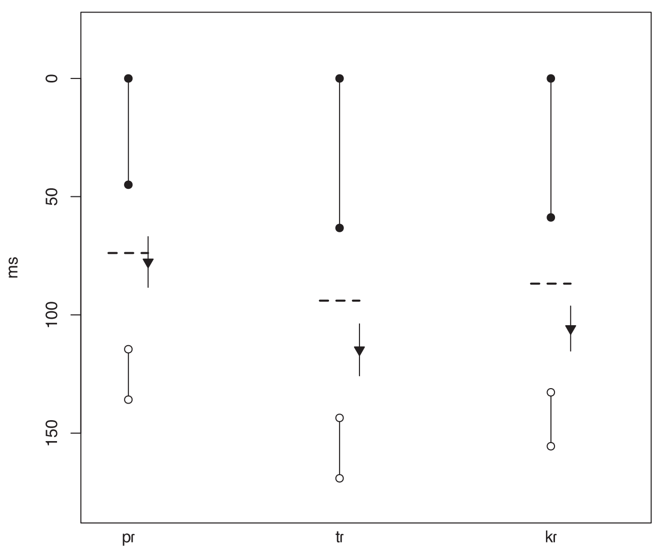

Fig. 14. Vowel onset in relation to prevocalic consonants in stop-rhotic clusters /pɾ, tɾ, kɾ/ (x-axis). The vertical lines denote intervals corresponding to gestural plateaus. Intervals delimited by black dots indicate the plateau onset and offset timestamps (y-axis) of the initial consonant. Intervals 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.

specifically for each type of cluster we have examined, namely, voiced stop-lateral, voiceless stop-lateral and voiceless stop-rhotic clusters.

4.1. Stop-lateral clusters

We begin by summarizing the evidence for the voiceless and voiced stop-lateral clusters. 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. The schema in Fig. 15 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 Fig. 15a versus 15c when it is clustered. The lateral in Fig. 15c is shorter and occurs later than the single lateral in 15a. Across the CV and CCV contexts (refer to Fig. 15a and c), the c-center (indicated as CC in the Figure) landmarks in /lV/ and voiceless /ClV/ are better aligned (as highlighted by the ellipsis hugging the CC landmarks) 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 (Fig. 15a and b). 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 Fig. 15a versus 15b when it is clustered with a voiced stop. The lateral in Fig. 15b shortens to some extent, but it does not shift later compared to the single lateral in 15a. Thus, shortening of the prevocalic lateral, as seen in Section 3.3, 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 ellipsis hugging the RE landmarks in Fig. 15a and b) than the CC landmarks between CV and CCV, thus resulting in stability of the local timing interval.

Therefore, it is 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 indicates 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

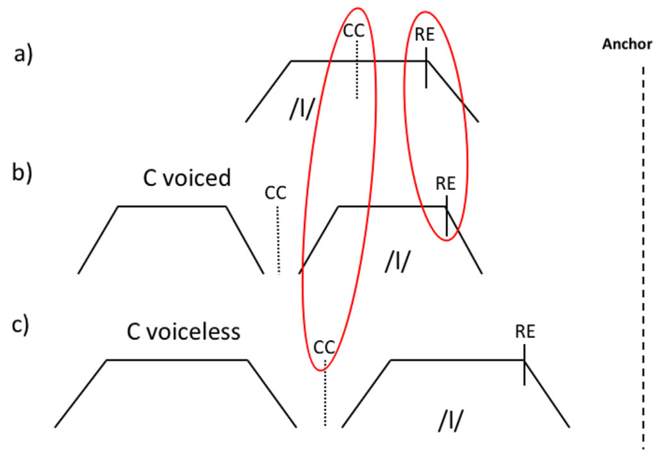

Fig. 15. Temporal configurations of (a) singleton lateral /l/, (b) voiced stop-/l/ cluster and (c) voiceless stop-/l/ cluster. The trapezoids represent consonant gestures. The landmark marked as CC corresponds to the c-center and RE to the right edge of the singleton consonant (C) or the consonant cluster (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).

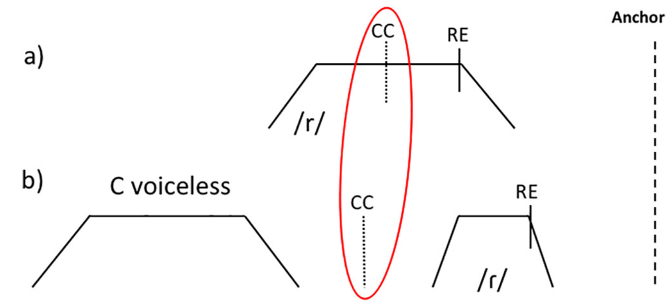

Fig. 16. Temporal configurations of a singleton trill /rV/ in (a) and a voiceless stop-/ɾV/ in (b). The trapezoids represent consonant gestures. The mnemonic CC corresponds to the c-center landmark and RE to the right edge landmark. The CC landmarks are better aligned than the RE landmarks across the /rV/, shown in (a), and stop-/ɾV/, shown in (b), contexts.

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.

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.

4.2. Stop-rhotic clusters

Consider now the evidence from 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 ms) 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. Fig. 16 illustrates the temporal readjustments that occur from /rV/ to /CɾV/. Compare the trill and the tap in Fig. 16a and b respectively. The tap is shorter and it occurs later with respect to the trill. The RE landmarks are thus misaligned and the CC landmarks are better aligned, as highlighted by the ellipsis, giving rise to the patterns we have observed in the stability analysis (Section 3.7.2) with the global timing interval remaining stable across CV and CCV.

For stop-rhotic clusters, we find no compensatory relation between IPI and prevocalic consonant duration in CCV. One reason for this may be that taps (the rhotics in second position of the relevant clusters) are so-called momentary articulations (Catford, 1977:130). The definitional characteristic of momentary articulations (as opposed to maintainable articulations as for fricatives) is that they have very short duration (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 preferably 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.[6] In particular, as we have seen in Fig. 14, 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 /

[6] Shortening of the prevocalic consonant from CV to CCV and change in the relative timing between rhotic and vowel from CV to CCV is not addressed (as opposed to the case of stop-lateral clusters) due to the rhotic being a trill in CV and a tap in CCV. That is, the tap is by its nature a much shorter segment than the trill, but these are two different segments (unlike for /l/ as in /lo/ - /glo/) and thus the shortening seen cannot be ascribed 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.

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 provides a rationale for why IPI in /pɾ/ is shorter than in /kɾ, tɾ/.

Let us elaborate on the reasons why for stop-rhotic clusters the indications for global organization are different and more specifically why they cannot be the same as for stop-lateral clusters. For stop-rhotics, any measure which assesses syllabic organization by combining both the CV and the CCV context, such as stability-based heuristics, CV relative timing across CV~CCV and 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 which in turn affects the calculation of intervals which include the duration of the prevocalic consonant. For instance, although the stability-based heuristics (in Section 3.7.2) 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 which is shorter because it is a tap in the stop-rhotic cluster context (and hence reduces the length of any interval that includes it). Apart from the vowel initiation measure, another measure which involves quantifying the cluster context only is the IPI-C2 relation. Even though the IPI-C2 relation can be assessed only in the cluster context, such a relation was not present in stop-rhotic clusters. However, that this relation was not detected in stop-rhotic clusters is not surprising. By its nature, the tap is a very short segment (around 25 ms). As such, it does not exhibit sufficient variability in duration. Such variability in a prerequisite for identifying a compensatory relation: given two units standing in a compensatory relation, changes in duration of one unit must be compensated by changes in duration of the other unit. Hence, when one of the units does not vary sufficiently, such a relation cannot be detected.

Due to all of the above points, we are left with the vowel initiation measure as an index for assessing syllabic organization in the stop-rhotics. As per this measure, as we have seen, 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. This vowel initiation measure is in fact a rather direct way of assessing syllabic organization. Recall the crucial assumption in the work of Browman and Goldstein (1988) about the timing of a vowel with respect to its (preceding) tautosyllabic cluster we reviewed in 3.7.4: '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' (p. 150). All subsequent work on the topic of articulatory indices of syllabic organization has kept with the status of this statement as an assumption. Our measures of vowel initiation herein provide the first systematic effort to assess it. The results are telling in the sense that, as we have seen for both the stop-lateral and stop-rhotic clusters, the vowel starts as close to the c-center of the cluster as permitted by the phonetic properties of the segments in the cluster.

Overall, then, Spanish provides robust evidence that complex onsets do not always show the expected global timing interval stability (for stop-laterals) and that when they do so (as in the case of stop-rhotics) the reason cannot be safely ascribed to the presence of a global organization. This is so because, as we have discussed, the prevocalic rhotic in CV and CCV contexts are different segments. 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 results in local timing interval stability as in voiced stop-lateral clusters while more lateral shortening results 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 clusters), 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.

5. Conclusion

This paper 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.

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. This reorganization is expressed not in terms of a single phonetic measure or index, but in terms of a set of properties (which we review below) and in terms of how these properties relate to one another. That is, we have shown explicitly on the basis of our results that one cannot seek a single or promote a privileged index of reorganization.

Recall that voiced stop-lateral clusters exhibit stability of the local timing interval while voiceless stop-lateral clusters exhibit stability of the global timing interval and yet both contexts share a set of readjustments which take place when a stop is added in front of the lateral (namely, lateral shortening, a change of relative timing between /l/ and the vowel, and a compensatory relation between IPI and lateral duration). It is the presence of this set of properties which serves as an indication of a global reorganization of the CCV sequence.

We review the specifics of the result outlined above, after highlighting those aspects of our study that seem crucial in demonstrating this result. The 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 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 Spanish data, 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. 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 decreases. For stop-rhotic clusters, which show quite large IPI (74 ms), we found vowel initiation close to the c-center of the cluster (10–20 ms).

Not only the existence but also the specificity of the observed changes (resulting from adding a consonant to the left of a CV to obtain a CCV) provides further evidence for global organization. Consider, in particular, the fact that the degree to which the changes enumerated above 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 point, let us contrast the observed effects in the voiced stop-lateral versus the 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/ as per the well known cross-linguisitic pattern of voiceless stops being longer than voiced stops) and the IPI is longer 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 observe, 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 stability-based heuristics which fail to show the expected result for the voiced stop-lateral clusters), there are other indices which do reflect a global organization.

In independent work, we have pursued comparable to the present analyses in Moroccan Arabic. As reviewed in the Introduction, Moroccan has consonant clusters with phonetic properties similar to Spanish but with distinct phonological organization. The phonetic profile of, for example, a /kla/ sequence is comparable across the two languages: there is an open transition between the stop and the lateral and the voiceless stops at the start of these clusters in both languages are unaspirated. The Moroccan, Spanish comparison thus offers a case where a set of segments have similar phonetic properties but distinct phonological organizations: /kl/ is a prototypical syllable onset in Spanish but not so in Moroccan and other dialects of Arabic where consonant clusters cannot form syllable onsets.

Despite the surface phonetic similarities to Spanish, when a /k/ joins /la/ to form /kla/, Moroccan does not show the set of readjustments which our present work documents in detail for Spanish (Gafos, Roeser, Sotiropoulou, Hoole, & Zeroual, 2020). For example, whereas Spanish shows substantial compression of the lateral when a /k/ joins /la/ to form /kla/, this is not the case in Moroccan. A more telling example is illustrated with the compensatory relation between IPI and lateral duration, where a long(er) lag between the /k/ and the lateral must be compensated by shortening of the lateral in Spanish. In Moroccan, we find no such relation between the two parameters (in fact, we find a weak positive relation when using raw duration). Thus, in contrast to Spanish, the different parts of the CC subsequence are independently produced (hence, local organization) in Moroccan: an expansion or contraction in some part (here, IPI) of that subsequence does not result in a contraction or expansion of the other part (here, lateral duration). Overall, adding /k/ to a sequence of segments /la/ produces effects that ripple through the inner sequence in Spanish but not so in Moroccan.

In parallel work, we are pursuing comparable analyses also in German, a language where global organization expressed by the stability of the global timing interval has not been found consistently in past work. Our analyses in German point to the same conclusion as in this study. Robust evidence for global organization emerges when one abandons the search for a (presumed) privileged phonetic index of that organization and considers instead relations among different phonetic parameters. These relations express effects where spatio-temporal modification in one local region of the sequence comes systematically with a change in another region of the sequence. Such compensatory effects indicate that the different parts of a CCV are not independently planned and produced (hence, globally organized).

Detection of these compensatory relations among different parameters requires, as we have emphasized earlier with our Spanish data, sufficient variability in the parameters so related. To obtain this in German, we included, along with word boundary (WB) condition in *Ich sah "Plage" an* (I see *"Plage"*), an utterance phrase boundary condition (UT) in *Zunächst sah ich Anna. "Plage" sagte sie* (First I saw Anna. She said *"Plage"*). Prosodic strengthening effects at the initial part of the /pl/ sequence, due to the presence of a stronger boundary at UT compared to WB, is the expected broad consequence. The crucial question is what happens to the rest of the sequence when expansion of the initial part of /pl/ changes the duration of the stop consonant and/or the IPI between the two consonants. Along with confirming the basic expectation of prosodic expansion, our results show that a host of other effects begin to emerge at the UT condition (but crucially not the WB condition): local timing stability (which had fueled concerns on whether evidence for global organization can be found in German; Brunner et al., 2014) worsens across /lV/~stop-/lV/, the vowel starts earlier with respect to the lateral's target, and as IPI and the stop's duration lengthen due to expansion of the stop-lateral subsequence via prosodic strengthening, the inner CV shortens to compensate. Hence, once sufficient variability is introduced, we find clear evidence that local perturbations affect the rest of the sequence (global organization) also in German.

Moreover, the presence of such effects contrasts sharply with what we find in (German) CCV sequences with a word boundary between the first two consonants of the relevant clusters /pl/ or /kl/ as in /knap#lagə/ or /pak#lagə/ (versus /plagə/ or /klagə/). Whereas we have proposed above that a Spanish, Moroccan comparison offers an appropriate opportunity for observing how the same or similar sequences of segments are phonetically expressed under different organizations, for German this comparison can be carried out without changing language or speaker. This is because the sequence /p-l-a/ is organized differently in /knap#lagə/ than in /plagə/, due to the absence of resyllabification across word boundaries in German.[7] Our analyses in the German C#CV condition indicate clearly that changes in a local region of the C#CV sequence do not produce effects which ripple through to other parts of the sequence. Thus, local timing stability is maintained across CV~C#CV, there is no change in vowel initiation relative to the lateral's target, and even though the lag between the consonants (in /p#l/) increases due to the boundary, the inner CV string in C#CV is not affected. In the C#CV condition, thus, local perturbations remain local (hence, local organization).

To conclude, conceptually, perhaps the most important implication of our results on Spanish is that there is no unique phonetic exponent of global organization. Uniformity in phonological organization (same organization presiding over sequences with varying segmental makeup as in /pl, bl, kl, gl, pɾ, kɾ, tɾ/ before the vowels /a, e, o/) does not imply uniqueness of phonetic exponents. Syllabic organization is simultaneously expressed by more than one phonetic exponent and these exponents enter into compensatory relations with one another in attestation of the global organization they instantiate.

Acknowledgments

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

Appendix

Table 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
vp01	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
	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
vp02	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

(*continued on next page*)

[7] Conversely, this within language (and speaker) comparison is not available in Spanish as it is in German because in Spanish (native) words cannot end in stops (except for the voiced coronal /d/ which however is commonly reduced or deleted; Hualde, 2005). As a consequence, the relevant contrasts in terms of segmental sequences are not constructible in Spanish.

(*continued*)

		C^{tar}		C^{max}		V^{max}	
		global	local	global	local	global	local
	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
vp03	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
	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
vp04	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
	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
vp05	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
	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
vp06	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

(*continued*)

	C^{tar}		C^{max}		V^{max}	
	global	local	global	local	global	local
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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wocn.2020.100995.

References

Barry, W. (1997). Another R-tickle. *Journal of the International Phonetic Association, 27*(1–2), 35–45.
Blecua Falgueras, B. (2001). *Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos (Doctoral dissertation)*. Spain: Universitat Autònoma de Barcelona.
Bombien, L. (2011). *Segmental and prosodic aspects in the production of consonant clusters (Doctoral dissertation)*. Germany: Ludwig-Maximilians-Universität Munich.
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.
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. (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.
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.
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, M. (2013). *The R book*. Chichester, West Sussex, United Kingdom: Wiley.
Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht, Netherlands: Kluwer.
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.
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., 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). Saabrucken, Germany.
Gafos, A., Roeser, J., Sotiropoulou, S., Hoole, P., & Zeroual, C. (2020). Structure in mind, structure in vocal tract. *Natural Language and Linguistic Theory, 38*, 43–75.
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 (2009)* (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. (1983). *Syllable structure and stress in Spanish*. Cambridge, MA: MIT Press.
Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: The case of 'impure s'. *Phonology, 30*, 1–25. https://doi.org/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 Tashlhiyt 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.
Hualde, J. I. (2005). *The sounds of Spanish*. New York City: Cambridge University Press.
Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.
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.
Lenth, R., Hervé, M. (2015). Package 'lsmeans'. cran.r-project.org/web/packages/lsmeans/ lsmeans.pdf.
Lenz, R. (1892). Chilenishe Studien. *Phonetische Studien, 5*, 272–293.
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.
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.
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.
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.
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.
Proctor, M. (2011). Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory phonology, 2*(2), 451–485.
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*. Brazil: Minas Gerais.
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.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. URL http://www.rstudio.com/.

Schmeiser, B. (2009). An acoustic analysis of intrusive vowels in Guatemalan Spanish /rC/ clusters. *University of Pennsylvania Working Papers 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), 2015. https://doi.org/10.1371/journal.pone.0124714. eCollection.

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