Open Library of Humanities

Articulatory overlap as a function of stiffness in German, English and Spanish word-initial stop-lateral clusters

Shihao Du, Department of Linguistics and Cognitive Sciences, University of Potsdam, Potsdam, Germany, shihao.du@uni-potsdam.de

Adamantios I. Gafos, Department of Linguistics and Cognitive Sciences, University of Potsdam, Potsdam, Germany; Haskins Laboratories, New Haven, Connecticut, USA, gafos@uni-potsdam.de

Recently, Roon, Hoole, Zeroual, Du, and Gafos (2021) have offered evidence that an empirical estimate of stiffness, a parameter in the dynamics hypothesized to control the kinematics of gestures, modulates overlap in consonant clusters of Moroccan Arabic. Here, we extend the Roon et al. (2021) study in two ways. First, empirically, we broaden the assessment of the role of stiffness in overlap by examining data from three languages, largely confirming the Roon et al. results with electromagnetic articulography recordings from stop-lateral clusters spoken by German, English, and Spanish participants. Second, we examine for the first time the extent to which C1 and C2 stiffness individually, as opposed to the combined measure of stiffness difference in earlier work, contributes to overlap with the main finding being that the significant effect of stiffness difference is mainly due to C2 stiffness; C1 stiffness is not as robust as C2 stiffness in modulating overlap. These results are placed in the context of other work on articulatory and perceptual influences on inter-consonantal timing as well as in the broader context of work on the coupling between perception and action in other domains of skilled action.

1. Introduction

1.1. Cluster overlap and rapidity

In the past three decades, a number of parameters have been identified to affect inter-segmental overlap in C1C2 clusters. These include cluster position within a word (Hardcastle, 1985; Byrd, 1996; Chitoran, Goldstein, & Byrd, 2002), place order of the involved segments (Chitoran, Goldstein, & Byrd, 2002; Gafos, Hoole, Roon & Zeroual, 2010), C1 voicing (Hoole, Bombien, Kühnert, & Mooshammer, 2009; Bombien & Hoole, 2013; Gibson, Sotiropoulou, Tobin, & Gafos, 2017, 2019; Pouplier, Pastätter, Hoole, Marin, Chitoran, Lentz, & Kochetov, 2022), C1 place (Yanagawa, 2006; Bombien & Hoole, 2013; Gibson et al., 2017, 2019; Pouplier et al., 2022), C2 manner (Gibson et al., 2017, 2019), vowel identity following the cluster (Lialiou, Sotiropoulou, & Gafos, 2021), speech rate (Byrd, 1994; Luo, 2017), age (Mücke, Hermes, & Tilsen, 2020), intonational or phrasal boundaries (Byrd & Saltzman, 1998; Hoole et al., 2009; Byrd & Choi, 2010; Cho, Lee, & Kim, 2014; Cho, 2016) and language-specific timing patterns (Yanagawa, 2006; Hoole et al., 2009; Shaw, Gafos, Hoole, & Zeroual, 2009; Gafos et al., 2010; Marin & Pouplier, 2010; Tilsen, Zec, Bjorndahl, Butler, L'Experance, Fisher, Heimisdottir, Renwick, & Sanker, 2012; Bombien & Hoole, 2013; Hermes, Mücke, & Grice, 2013; Marin, 2013; Pastätter & Pouplier, 2013; Brunner, Geng, Sotiropoulou, & Gafos, 2014; Hermes, Mücke, & Auris, 2017; Pastätter & Pouplier, 2017; Gibson et al., 2017, 2019; Sotiropoulou, Gibson, & Gafos, 2020 among others). However, little attention has been paid to the potential role of parameters in the underlying dynamics assumed to control the articulatory gestures (comprising these C1C2 clusters). Most notable among these parameters is stiffness, whose modulations directly affect durational aspects of the unfolding of gestures. Specifically, higher stiffness results in increased peak velocity to movement amplitude ratios and shorter movement durations. It seems intuitive, yet largely still unexplored, that overlap in C1C2 clusters would be directly affected by these 'rapidity' aspects of stiffness-controlled kinematics.

Browman and Goldstein (1990a) first pointed to the possibility that different degrees of overlap may be related to some notion of rapidity of the different oral articulators implicated in the consonants of the cluster. This suggestion was based on early work by Brown (1977) and Gimson (1962) who observed that the most common cases of place assimilation in English typically involve alveolar consonants assimilating to labials or velars, rather than the other way around. Browman and Goldstein (1990a) suggested that this propensity for assimilation might result from some gestures being more easily overlapped or "hidden" by other gestures due to differences in rapidity. Specifically, they postulated that a slower gesture, such as that implicated in a labial or a velar constriction, "might prove more difficult to hide" than a faster one as in a coronal, since the tongue tip enjoys a greater flexibility than the lips or more posterior parts of the tongue such as the tongue body implicated in velars (Browman & Goldstein, 1990a, p. 18).

The conjectured link between overlap and some notion of rapidity by Browman and Goldstein (1990a) was taken up by Jun (1995, 2004) in a proposal aimed at accounting for patterns of

place assimilation. In his typological study, Jun observed that coronals are the most likely to be targeted by place assimilation, but the least likely to be the trigger of place assimilation compared to non-coronals (Jun, 2004, p. 64). To explain such propensities in place assimilation, Jun proposes that in a C1C2 consonant cluster, when the rapidity of C2 is held constant, a more rapid C1 articulatory movement (**Figure 1A**) would result in more overlap between the gestures of C1 and C2 than a slower C1 (**Figure 1B**). This in turn would render the acoustic place cues of C1 more easily hidden by those of the overlapping C2 and thus perceptually less salient and more prone to assimilate in place to C2. Conversely, when the rapidity of C1 is held constant, a slower C2 (**Figure 1D**) would lead to more overlap between the gestures of C1 and C2 than a faster C2 (**Figure 1C**), as the C2 gesture would "intrude" into the unfolding of the C1 gesture. This would also have the effect of C2 obscuring C1 place cues acoustically, with C1 more likely to assimilate in place to C2 as a consequence. The propensity of (presumed more rapid) coronals to more frequently assimilate (than other consonants) to a following non-coronal corresponds to, according to the overlap schemes of Jun, the scenario described in **Figure 1A** (compared to **Figure 1B**). Likewise, the propensity of non-coronals to less frequently assimilate to a following coronal corresponds to the scenario laid out in **Figure 1C** (compared to **Figure 1D**). In short, Jun posits that the rapidities of both C1 and C2 should exert effects on the amount of overlap. Specifically, overlap should increase as the rapidity of C1 increases and that of C2 decreases; conversely, as the rapidity of C1 decreases and that of C2 increases, overlap should decrease.
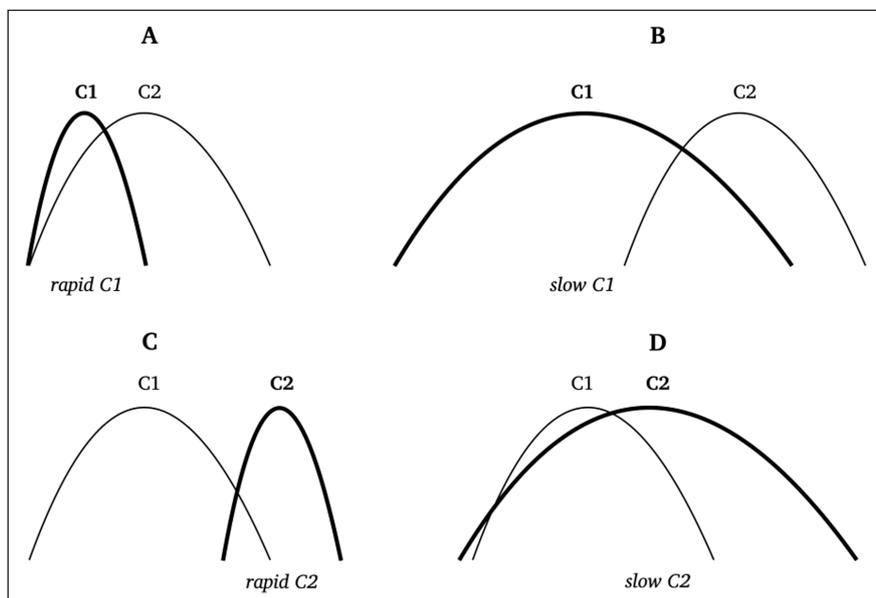
Figure 1: Schematic representation of Jun's proposal about the relationship between rapidity and overlap (recreated from Jun, 2004 (5) and (6); same as Figure 1 in Roon et al., 2021). In panels A and B, as the rapidity of C2 is held constant, a rapid C1 (A) results in more overlap than a slow C1 (B). Similarly, in panels C and D, as the rapidity of C1 is held constant, a slow C2 (D) results in more overlap than a rapid C2 (C).

Recently, Roon et al. (2021) tested Jun's proposal about the relationship between rapidity and overlap using electromagnetic articulography (EMA) data from Moroccan Arabic word-medial stop-stop clusters. Specifically, Roon et al. (2021) tested two predictions related to Jun's proposal. The first concerns an ordering of rapidities determined by the place of articulation as shown in (1). In Jun's proposal, asymmetries in assimilation reflect different degrees of saliency in place transition cues, which in turn are related to the rapidity of the gestures being targeted in assimilation. For Jun, propensities in place assimilation are to be derived from the saliency of place cues which is in turn, at least in part, determined by the articulatory rapidity of the gestures whose execution gives rise to these cues. Specifically, rapid gestures, such as those implicated in coronals (cor), are proposed to have brief and thus less salient place cues, whereas sluggish gestures, such as those implicated in non-coronals, are proposed to have more salient place cues.[1] Within the non-coronals, Jun reviews evidence pointing to velars (vel) having more salient cues than labials (lab) (Jun, 2004, pp. 63–65). Assuming that this notion of saliency has its basis in the rapidity of the gestures of these consonants (Roon et al., 2021),[2] this leads to the ordering of rapidity shown in (1).

(1)   **Prediction 1 (Roon et al., 2021, H-J1)**
       Ordering of rapidity:
       COR > LAB > VEL

Secondly, the rapidity ordering in (1) yields a further prediction in the form of a partial ordering in overlap between different cluster types shown in (2): COR-VEL > COR-LAB, LAB-VEL > LAB-COR, VEL-LAB > VEL-COR (Roon et al., 2021). The table in (2) illustrates how to derive this partial ordering from Jun's proposals in **Figure 1**. In the table in (2), for any given column, going from top to bottom, C1 rapidity increases (from VEL to LAB to COR) while C2 remains fixed, thus leading to increasingly more overlap among the cluster types in any given column of the table; this is because, as illustrated in **Figure 1**, panels A and B, in a C1C2 cluster, as the rapidity of C2 is held constant, a slow C1 (**Figure 1B**) results in more overlap than a fast C1 (**Figure 1A**). For example, looking at the leftmost column of the table in (2), the rapidity-based prediction is that

---

[1] "Typically, the underlying gesture with which coronals are realized is articulated more rapidly. That is, tongue tip gestures are rapid and thus have rapid transition cues; whereas tongue dorsum and lip gestures are more sluggish and thus give rise to long transitions." (Jun, 2004: p. 63).

[2] Jun's argument for the relative saliency between the non-coronals is more nuanced and does not call on rapidity *per se* (see discussion in Jun 2004: pp. 63–65). Roon et al. (2021), in their aim to assess Jun's predictions using articulatory data, assume that rapidity also underlies the relative saliency between labials and velars. However, we keep with the Roon et al. (2021) prediction in (1), due to continuity; this is the only prior research addressing some of the ideas in Jun (1995) using articulatory data (for a test of Jun's proposals using perceptual data see, especially, Winters, 2001, 2003). If we were to reformulate Prediction 1 in a way that does not specify an ordering among labials and velars, our results would still speak to that prediction (in particular, we find that this prediction as well as the version in (1) are not supported in our data).

the VEL-COR cluster /gl/ should be less overlapped than the LAB-COR cluster /bl/. Note that in (2) we leave cells that correspond to combinations of the same effector (e.g., VEL-VEL) empty, because C1, C2 individual rapidity measures cannot be reliable extracted in such combinations due to their homorganicity. Similarly, on the basis of the table in (2), we can project predictions by moving horizontally, from left to right. That is, looking now within each row of the table, C2 rapidity decreases (from COR to LAB to VEL), while C1 remains fixed, thus leading to increasingly more overlap among the cluster types in any given row of the table; this is because, as illustrated in Figure 1, panels C and D, in a C1C2 cluster, as the rapidity of C1 is held constant, a slow C2 (**Figure 1D**) results in more overlap than a faster C2 (**Figure 1C**). Overall, the closer a cluster type is located towards the upper left corner of the resulting table, the less overlapped it is, with VEL-COR predicted to be the least overlapped cluster type; the closer a cluster type is located towards the bottom right corner of the table, the more overlapped it is, with COR-VEL predicted to be the most overlapped cluster type. The predicted overlap patterns can thus be expressed in a partial ordering as in: COR-VEL > COR-LAB, LAB-VEL > LAB-COR, VEL-LAB > VEL-COR (where a comma separates cluster types for which no rapidity-based overlap prediction is made).

(2)   **Prediction 2 (adopted from Roon et al., 2021, H-J2, p. 3)**
         Partial ordering of overlap:
         COR-VEL > COR-LAB, LAB-VEL > LAB-COR, VEL-LAB > VEL-COR

C2 / C1		decreasing C2 rapidity, more overlap →		
		COR	**LAB**	**VEL**
increasing C1 rapidity, more overlap ↓	**VEL**	VEL-COR	VEL-LAB	
	LAB	LAB-COR		LAB-VEL
	COR		COR-LAB	COR-VEL

Roon et al. (2021) argued that Moroccan Arabic offers an ideal case for assessing these predictions because all place of articulation combinations ([bd, bt, bg, bk, dg, dk, tg, tk, db, gb, gd, gt, kb, kd, kt, tb]) are permitted in this language. Rapidity was quantified in Roon et al. (2021) in two ways. The first way operationalizes the notion of rapidity with the peak velocity of the closing phase movement for any given consonant. The second way employed instead the amplitude-normalized peak velocity, a measure that is assumed to index the parameter of "stiffness" in the underlying dynamics of gestures.

In assessing the effects of either of these two operationalizations of rapidity on overlap, Roon et al. (2021) quantified overlap in two different ways. The first measure of overlap was calculated as the ratio of the unfolding of C2 closing phase movement to the duration of C1

gestural plateau and termed as relative overlap. A second measure of overlap was quantified by the temporal interval between the initiation of C1 and C2 gestures, termed as onset lag (see §2.4.2 for reasons). In their results, Roon et al. (2021) found no systematic differences in rapidity based on articulator, regardless of whether rapidity was indexed by peak velocity or stiffness. Furthermore, the findings on overlap did not conform to the predicted partial ordering of overlap listed under (2), since LAB-VEL and LAB -COR were found to be significantly more overlapped than the other clusters, while the difference between COR-VEL and COR- LAB clusters was not reliable. Although neither of these two predictions was borne out in the empirical data, Roon et al. (2021) did find a strong correlation between overlap and the difference between C1 and C2 stiffness. In other words, they found that overlap increases as the stiffness of C2 decreases with respect to the stiffness of C1 independent of the identities of the involved oral articulators. Crucially, this lawful relationship between overlap and rapidity was only significant when rapidity was indexed by stiffness and not by peak velocity. Therefore, they concluded that stiffness, not peak velocity, modulates overlap in C1C2 clusters.

Even though Roon et al. (2021) identified stiffness as a systematic determinant of overlap in consonant clusters, their "results should be interpreted with some caution, since they come from one language" (Roon et al., 2021, p. 19). Hence, evidence from languages other than Moroccan Arabic, in particular those involving different coordination relations between consonantal gestures, is needed to solidify or, should contradictory evidence be found, rebut the conclusions drawn by Roon et al. (2021). Another potential drawback in the study of Roon et al. (2021) lies in their choice of the stiffness difference and the peak velocity difference of articulations as an alternative to Jun's original articulator-specific proposal. On the one hand, this choice, as the authors intended, indeed strips off the component of "inherent velocity" associated with different articulators in Jun's proposal, which did not find support in the empirical data, but maintained the spirit of Jun's idea that the rapidities of C1 and C2 both contribute to variances in overlap. On the other hand, this approach lumps C1 and C2 rapidities together into a single parameter quantified by either stiffness difference or peak velocity difference between C1 and C2. This prevents observation of any differential effects of the individual consonants on overlap (recall that Jun's original proposal predicts that both C1 and C2 rapidities modulate overlap, a hypothesis which is yet to be assessed with empirical data). Granted, modelling C1 and C2 rapidity as separate effects potentially risks ignoring within-token factors and variability across tokens, but this can be partially compensated by allowing by-cluster variance in the statistical models. Overall, it is worth exploring the rapidity effects of each consonant in a C1C2 cluster individually from the perspective of hypothesis testing. It is precisely the aforementioned two points which extend the work of Roon et al. (2021) that this present study aims to address.

Finally, whereas much of the literature on consonantal overlap tends to retain an articulatory perspective, there is evidence that the generalizations at play may have a perceptuomotor rather

than a purely articulatory basis (see, for instance, the proposed basis for the place order effect in Gafos et al., 2010). We similarly seek here, once the patterns in our findings are identified, to relate our results to other results in motor control that have a basis in the coupling between action and perception.

1.2. Hypotheses tested

Here, we present the two broad hypotheses this work intends to assess. The first broad hypothesis posits that the observations and conclusions drawn by Roon et al. (2021) from the Moroccan Arabic data can be extended to typologically different languages. In the present study, articulatory data from German, English, and Spanish stop-lateral clusters were analyzed to enable a cross-language comparison. In more specific terms, this first hypothesis has three components, corresponding to H-J1, H-J2, H-J3 (mistakenly written as H-3) in Roon et al. (2021) and stated in (3) below. The first and second components aim at the predictions of Jun (2004) on articulatory-specific inherent velocity and the projected partial ordering of overlap in clusters based on the presumed rapidity differences. Thus, the first component under (3) concerns the presence of consistent differences in rapidity among coronal, labial, and velar consonants. The second component under (3) aims at the predicted difference in overlap between velar-lateral (/kl/, /gl/) and labial-lateral clusters (/pl/, /bl/); as illustrated in (2), the former is predicted to be less overlapped than the latter, a prediction that was not verified in the Roon et al. (2021) results. The third component under (3) concerns the replicability of the relation between relative rapidity and overlap in Roon et al. (2021). Specifically, if the finding from Roon et al. (2021) is indeed valid across languages, then a robust (meaning significant across different overlap measures) correlation is expected between overlap and stiffness difference, but not between overlap and peak velocity difference.

(3)   **Hypothesis 1: Validity across datasets**
        Results of Roon et al. (2021) on Moroccan Arabic word-medial stop-stop clusters hold for German, English, and Spanish word-initial stop-lateral clusters. More specifically, three specific components are entailed here (all in agreement with Roon et al., 2021).
        • Component 1: There are no systematic differences in rapidity among COR, LAB, VEL consonants, regardless of whether rapidity is indexed by peak velocity or stiffness.
        • Component 2: There is no rapidity-based difference in overlap between labial-lateral clusters and velar-lateral.
        • Component 3: Overlap varies as a function of stiffness difference, rather than peak velocity difference, between C1 and C2.

Our second hypothesis aims to tease apart the separate contributions of C1 and C2 individual stiffness values on overlap. This hypothesis, stated in (4) below, was not assessed in Roon et al. (2021).

(4)   **Hypothesis 2: Individual consonant stiffness effects**
        Overlap in a cluster increases as the stiffness of C1 increases and that of C2 decreases.

If Jun's schemas about the relations between the rapidities of C1 and C2 (interpreted in the present work by using stiffness rather than peak velocity) and overlap (that overlap increases as C1 rapidity increases and C2 rapidity decreases; see Figure 1) comply with the data, then we should expect to see significant effects of both C1 and C2 stiffness on overlap. Crucially, the directions of the individual stiffness effects are expected to be opposite to each other: Higher rapidity, more overlap for C1 stiffness and higher rapidity, less overlap for C2 stiffness.

Note that the approach in Jun (1995, 2004), to which the above hypotheses relate, comprises a link between rapidity and overlap and another link between overlap and phonological patterns of assimilation. We hasten to point out that the hypotheses evaluated in the present work only concern the first link in Jun's proposal, the one regarding whether overlap between the gestures of the consonants in a C1C2 cluster is modulated by rapidity measures such as peak velocity and stiffness. Our work does not address the second link, that is, the issue of what drives assimilation. We return to this point in the discussion as well.

2. Method

2.1. Subjects and stimuli

The data analyzed in the present work are all subsets of articulatory data collected in previous experiments at the authors' institution. The German dataset consists in recordings from six adult German native speakers. The speakers were all between 20 and 33 years old. The data in the English dataset derive from six adult American English native speakers. The Spanish articulatory data were collected from five native speakers of Standard Peninsular Spanish, all of which were between 18- and 35-years-old. No participants in any of these studies reported any hearing or speech problems. The experimental procedures were approved by the Ethics Committee at the authors' institution. All the participants offered their informed consent to take part in the respective study and granted permission for publishing their anonymous data afterwards. Upon finishing the experiment, they were paid for their participation.

The German stimuli, shown in **Table 1** with their glosses, are all real disyllabic German words. Each word has one of the following stop-lateral clusters as its onset /bl/, /pl/, /gl/, /kl/, followed by the back low long vowel /a:/ (for /gl/ it is a diphthong /ai/). Each word is produced as part of three carrier phrases: a) *Ich sah ___ an* 'I see ___'; b) *Als ich Tom sah, ___ sagte er sofort* 'When I saw Tom, he said ___ immediately'; and c) *Zunächst sah ich Anna. ___ sagte sie* 'First I saw Anna. ___ said she', where "___" indicates the position of the stimulus word. Instances of these phrases with the required stimulus were projected on a computer screen with a random order among the stimuli words, and the participants were instructed to read the phrase out loud. The speakers were instructed to produce ten repetitions of each phrase for each stimulus word.

The English dataset, shown in **Table 1**, also contains predominantly real monosyllabic words in English as stimuli, except for one stimulus word which is disyllabic. The onset clusters are /bl/, /pl/ and /kl/. There is one stimulus with /bl/ as onset, four with /pl/ and three with /kl/. The tautosyllabic vowel following the stop-lateral cluster onset is one of the following: /ɑ/, /aɪ/, /eɪ/, /ɛ/, /ɪ/ and /iː/. Each of the stimuli was embedded in the carrier phrase *I saw ___ on the sign*. During the recording, the participants were instructed to produce eight repetitions for each phrase.

The Spanish stimuli, shown in **Table 1**, are real disyllabic words starting with the onset clusters /bl/, /pl/, /gl/ or /kl/. The tautosyllabic vowel following each of the four consonant clusters is either the back low vowel /a/, the mid front vowel /e/ or the mid back vowel /o/. Each stimulus word was embedded in a carrier phrase *Di ___ por favor* 'Say ___ please'. Each speaker produced seven to eleven repetitions of each stimulus.

The word-initial clusters in German, English, and Spanish stimuli are all in primary-stressed syllables except for the /bl/ cluster in the English stimulus "blockade," which has primary stress on the second syllable.

Cluster	/bl/	/pl/	/gl/	/kl/
German	*blasen* 'to blow'	*Plage* 'plague, nuisance'	*gleite* 'slide'	*Klage* 'lawsuit'
English	*blockade*	*plight* *played* *pledge* *plead*	/	*climb* *clip* *clean*
Spanish	*blata* (model of motorcycle) *blanda* 'soft' *bleque* 'tar' *bloque* 'block'	*plato* 'plate' *plena* 'full' *plomo* 'lead'	*Glato* (province in Italy) *Glana* (surname) *gleba* 'mound of land' *globo* 'balloon, globe'	*clapas* 'stripping of unfertile land' *clema* 'electrical connector' *clono* 'clone'

Table 1: Stimuli for all three languages. All clusters were word-initial.

2.2. Data acquisition

The German and Spanish articulatory data were registered using a Carstens AG501 three-dimensional Electromagnetic Articulograph (EMA) in the phonetics lab at the authors' institution. Three sensors were attached along the midsagittal line of the tongue on its upper surface: a tongue tip sensor (TT) located approximately 1 cm posterior to the actual apex of the tongue, a tongue mid sensor (TM) located 2 cm posterior to the tongue tip sensor, followed by a tongue

back sensor (TB) located 2 cm posterior to the tongue mid sensor. Three additional sensors were attached to the upper and lower lips and the jaw respectively. Reference sensors were placed on the upper incisor, behind the two ears on the left and right mastoid, and on the nose bridge to record the head movement. Three sensors were placed on a triangular bite place, which was held by the subjects between the upper and lower teeth to obtain a reference for the occlusal plane. Finally, the subjects were instructed to trace the shape of their hard palate using a sensor attached to the tip of their thumb. 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. During the recording, participants were instructed to read out phrases appearing on a computer screen located roughly one meter in front of them at a comfortable rate. The phrases were prompted by a separate computer outside the sound-proof booth, which also triggers the EMA system to start the recording. The kinematic movements of the sensors were recorded at a sampling rate of 250 Hz. Simultaneously, the acoustic data were also captured by a t.bone EM 9600 unidirectional microphone at a sampling rate of 48 kHz. After the recording was finished, data were corrected by subtracting the head movement captured by the reference sensors from the movement of all the other sensors. Last but not the least, the data were rotated for each subject based on their occlusal plane accordingly. The English corpus was acquired by the Northern Digital Inc. (NDI) Wave System. The same number of sensors was used for the English data collection as for the other datasets, and they were placed similarly to those used in the German and Spanish experiments. Articulatory data were recorded at a sampling rate of 400 Hz. Acoustic data were simultaneously acquired using the Schöps Colette modular system of microphones at a sampling rate of 25.6 kHz. For the reference sensors, the filter cut-off frequency was 5 Hz, while for all the other sensors it was 20 Hz. The raw data were processed by removing the contributions of head movement from the kinematic data of all the other sensors as well.

2.3. Articulatory segmentation

Articulatory segmentation denotes the process of identifying the timepoints, or temporal landmarks, at which the unfolding of an articulatory gesture for a consonant moves from one characteristic phase of movement to another. Hence, we need to first determine the correspondence between gestures and consonants. For each consonant in the word-initial clusters, its articulatory gesture was indexed by the movement of its primary oral articulator involved in its production. Thus, in all three languages the velar stops /g/ and /k/ were measured using the most posterior tongue back (TB) sensor; the labial stops /b/ and /p/ were measured using the lip aperture (LA) representing the distance between the upper and lower lips; and the alveolar lateral /l/ was measured using the tongue tip (TT) sensor.

The temporal landmarks of each gesture were identified automatically using the Matlab-based algorithm Mview developed at Haskins Laboratories by Mark Tiede. The user first selects a particular temporal range of interest in the entire recorded trajectory of a certain sensor

corresponding to the movement of the main articulator for the intended gesture, and then applies the gesture identification algorithm within that window. **Figure 2** gives an example of the tongue body and tongue tip gestures in the German cluster /kl/ with the landmarks indicated. The algorithm works by first finding the peak tangential velocity of the gestural movement towards (see 'Peak velocity to' in **Figure 2**) and away from (see 'Peak velocity fro' in **Figure 2**) the timestamp of maximal constriction (see definition below) as well as the minimal velocity during the constriction plateau (see 'Min. velocity' in **Figure 2**). The landmarks of gestural onset (see 'Onset' in **Figure 2**) and target (see 'Target' in **Figure 2**) are then defined as the timestamps at which the sensor first exceeds and then falls below the threshold of 20% of the maximum tangential velocity during the closing phase of the movement. On the other hand, the landmarks of constriction release (see 'Release' in **Figure 2**) and gestural offset (see 'offset' in **Figure 2**) are identified as the timestamps at which the sensor first exceeds and subsequently falls below the same 20% peak velocity threshold but of the opening phase movement. The period demarcated by the constriction target on the left and release on the right is the constriction plateau of the gesture (see the blue- and lila-filled boxes in **Figure 2**). The algorithm also calculated the so-called 'maximum constriction' of the movement, which was defined as the time point at which the tangential velocity of the articulator reached its minimum during closure (see 'Min. velocity' in **Figure 2**). The Euclidean distance between the position of the sensor at gestural onset and that at maximum constriction is the amplitude of the closing phase movement. This distance is crucial for the calculation of the amplitude-normalized peak velocity used in the present study.
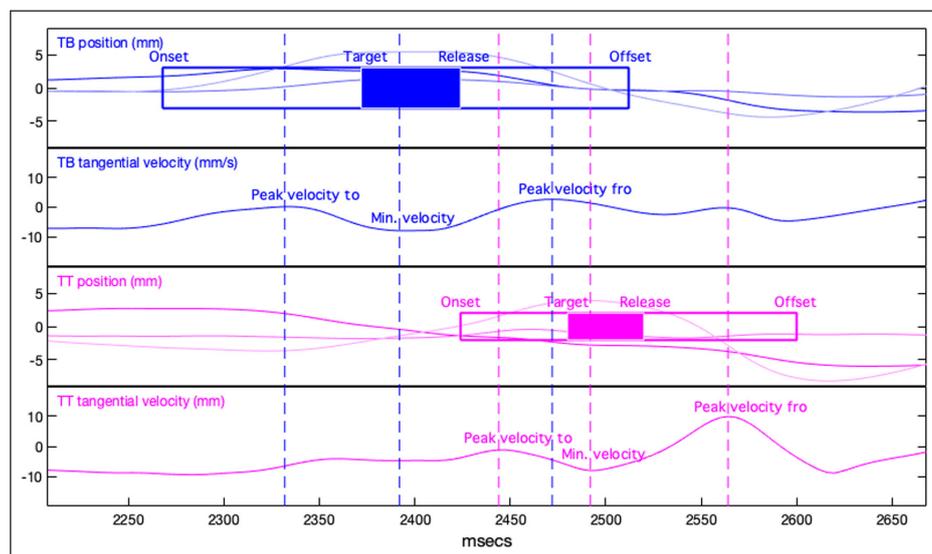
Figure 2: Temporal landmarks labeled by the Mview algorithm on the tongue body (TB) and tongue tip (TT) gestures of the word-initial /kl/ in the German word *Klage*. The first and third panels show the positional values of the vertical, horizontal and lateral dimensions from the TB and TT sensors over time respectively. The second and fourth panels show the tangential velocity for the two sensors.

2.4. Measurements

2.4.1. Rapidity of articulation

Following Roon et al. (2021), two measures of rapidity corresponding to Jun (2004)'s concept of inherent velocity are considered as candidates for contributing to overlap. The first is the maximum tangential velocity of the movement of the articulator towards forming a constriction. This is the measure that most closely corresponds to Jun's proposal for a rapidity parameter modulating overlap (Jun, 2004, p. 65). We will henceforth refer to this maximum velocity simply as peak velocity, denoted by $v\hat{}$ in the present study. This first rapidity measure was quantified for all primary oral articulators involved in the production of the two consonants in stop-lateral clusters.

However, the peak velocity of a movement is well-known to covary with the spatial excursion or amplitude of the movement (Kozhevnikov & Chistovich, 1965; Ohala, 1970; Kent & Moll, 1972; Kuehn & Moll, 1976; Ostry, Keller, & Parush, 1983; Ostry & Munhall, 1985; Kelso, Vatikiotis-Bateson, Saltzman, & Kay, 1985; Guenther, 1995), with higher amplitudes having higher peak velocities (with some funning out of the linearity at the extreme values of amplitude; see Kuberski & Gafos, 2019; Sorensen & Gafos, 2016). Therefore, a second measure of rapidity, which takes movement amplitude into account, was considered. This second measure is the amplitude-normalized peak velocity (Roon et al., 2021), computed by dividing the peak velocity $v\hat{}$ by the amplitude A of the movement during the closing phase of the consonant as shown in (5) below.

(5)   Amplitude-normalized peak velocity: $k' = v\hat{} / A$

This ratio of peak velocity to amplitude, which we denote as k', has been used extensively in the literature to characterize a wide number of linguistic and non-linguistic contrasts,[3] and tends to be associated in that literature with the parameter k in the well-known linear mass-spring dynamic model of gestures given by the equation $m\ddot{x} + b\dot{x} + k(x - x_0) = 0$ (Fowler, Rubin, Remez, Turvey, 1980; Browman & Goldstein, 1984, 1986, 1990b; Saltzman & Munhall, 1989; Pouplier, 2020). Our notation indicates that we wish to maintain a distinction between the (empirical) measure k' and the (theoretical) model parameter k. This is because the dynamics expressed by the model's equation $m\ddot{x} + b\dot{x} + k(x - x_0) = 0$ assumes Hooke's Law and in this law stiffness is identified with the ratio between peak acceleration (not peak velocity) and amplitude. Nevertheless, we will keep with the long tradition of referring to the amplitude-normalized peak

---

[3] Thus, for instance, stiffness as measured by the peak velocity to amplitude ratio has been implicated in spatio-temporal modulations at prosodic boundaries (Beckman & Edwards, 1992; Edwards, Beckman, & Fletcher, 1991; Byrd & Saltzman, 1998), in speech rate differences (Kelso et al., 1985; Kühnert & Hoole, 2004), in distinctions between normal speech and ataxic dysarthria speech disorders (Ackermann, Hetrich, & Scharf, 1995), and in differences among clear versus normal speech (Perkell et al., 2002).

velocity as stiffness. We note furthermore that the usefulness of k' in characterizing various phenomena, as evidenced by its wide use in the literature, should not be tied to the whether the linear mass-spring model is 'the right' model of speech gestures. In other words, the validity of k' as a predictor of overlap in our study or as a parameter characterizing distinctions in any of the other phenomena it has been implicated is independent of the validity of that model. In nonlinear models, the notion of stiffness may no longer correspond to a single parameter, but to mixtures of parameters associated with higher nonlinear orders, dependent on the precise structure of the model (Sorensen & Gafos, 2016; Sorensen & Gafos, 2022). Yet, since nonlinearities are typically assumed to be small, contributions of these higher order components are small too, resulting in k (and thus k') being a good approximation to the true stiffness.

Using the two measures of rapidity above, peak velocity and amplitude-normalized peak velocity, Roon et al. (2021) did not find any reliable evidence for the notion of inherent velocity, that is, for the idea that articulators (lips, tongue tip, tongue body) can be differentiated with respect to their articulator-specific rapidities. Roon et al. (2021) then proceeded to study the extent to which the rapidity difference between C1 and C2 modulates overlap in Moroccan Arabic, using two measures: the difference between the peak velocity of C1 and that of C2, as shown in (6), and the difference between the stiffness of C1 and that of C2, as shown in (7). Significant effects were found for the stiffness difference on overlap, but not for the peak velocity difference. In order to test the generalizability of these findings from Roon et al. (2021) to the three languages studied in the present work, we adopt these two measures of relative articulator rapidity in the present study. In each of these difference measures, positive (negative) values indicate that the stop consonant in a stop-lateral cluster (C1) has a greater (lower) peak velocity $\hat{v}$ or stiffness k' value than the lateral (C2).

(6)   Peak velocity difference $= \hat{v}_{C1} - \hat{v}_{C2}$

(7)   Stiffness difference $= k'_{C1} - k'_{C2}$

Going beyond the assessment in Roon et al. (2021), we will also examine the effects of C1 and C2 individual stiffness on overlap for reasons explained in §1.1 and §1.2, in order to test Hypothesis 2.

2.4.2 Articulatory overlap

To assess the robustness of the different presumed predictors of overlap, three different indices of articulatory overlap were used: relative overlap, onset lag, and absolute overlap. The former two were inherited from Roon et al. (2021) to ensure cross-study comparability. Relative overlap is quantified by the proportion of C2's closing phase movement that is coextensive with the movement of C1 before its constriction release. Since this is not a simple temporal interval,

but rather a ratio of one interval to another, this overlap measure is denoted as "relative". It is defined by Gafos et al. (2010) and Roon et al. (2021) as in (8):

(8)   Relative overlap $= 1 - (\text{Onset}_{C2} - \text{Target}_{C1}) / (\text{Release}_{C1} - \text{Target}_{C1})$

which can be further simplified into (9):

(9)   Relative overlap $= \text{Release}_{C1} - \text{Onset}_{C2} / (\text{Release}_{C1} - \text{Target}_{C1})$

By subtracting the timestamp of C2 onset from that of C1 release and then dividing by the duration of C1 plateau, relative overlap provides a way to quantify how early C2 movement is initiated within the constriction plateau of C1, while at the same time normalizing the lag between the release of C1 constriction and the onset of C2 by the plateau duration of C1 to account for speech rate variance. Therefore, higher values of relative overlap indicate that C2 starts early during C1 plateau, thus showing more overlap between the two consonants, whereas lower values indicate the opposite. **Table 2** summarizes what different values of relative overlap indicate about the relative timing between the landmarks of the two consonants.

C2 Onset position[4]	Relative overlap value	Indication
A	> 1	C2 onset before C1 plateau
B	= 1	C2 onset coincides with C1 onset
C	(0, 1)	C2 onset during C1 plateau
D	= 0	C2 onset coincides with C1 release
E	< 0	C2 onset after C1 release

Table 2: Indication of different values of relative overlap

As pointed out by Roon et al. (2021), differences in relative overlap may be modulated by both the rapidities of C1 and C2 (as in Jun's proposal) and differences in C1 plateau duration (not accounted for in Jun's proposal), since this measure of overlap uses C1 plateau duration as the denominator. In other words, it is conceivable that a higher value of relative overlap does not result exclusively from any effect of the rapidity of C1 and C2, but simply from a shorter C1 plateau duration, that is, the denominator in the equation (9), with the effect of shorter plateau duration resulting in higher overlap. Consider the schema in **Figure 3**: C2 is coextensive with C1 plateau in **Figure 3A** as much as it is **Figure 3B** (represented by the grey boxes). Yet because the gesture of C1 (in red) in **Figure 3B** has a shorter plateau duration than that of C1 in **Figure 3A**,

---

[4] The letters 'A, B, C, D, E' refer to different C2 onset timestamps relative to the unfolding of C1, as further specified under the 'Indication' column, which in turn determines the quantitative range the value of relative overlap falls in. The same conventions apply to Table 3 and Table 4.

the lower scenario would show a larger relative overlap value compared to the upper scenario. Granted, this larger relative overlap could also result from the more rapid articulation of C1 in **Figure 3B**, but the effect of C1 rapidity on relative overlap would not be able to be teased out from that of the shorter plateau duration.
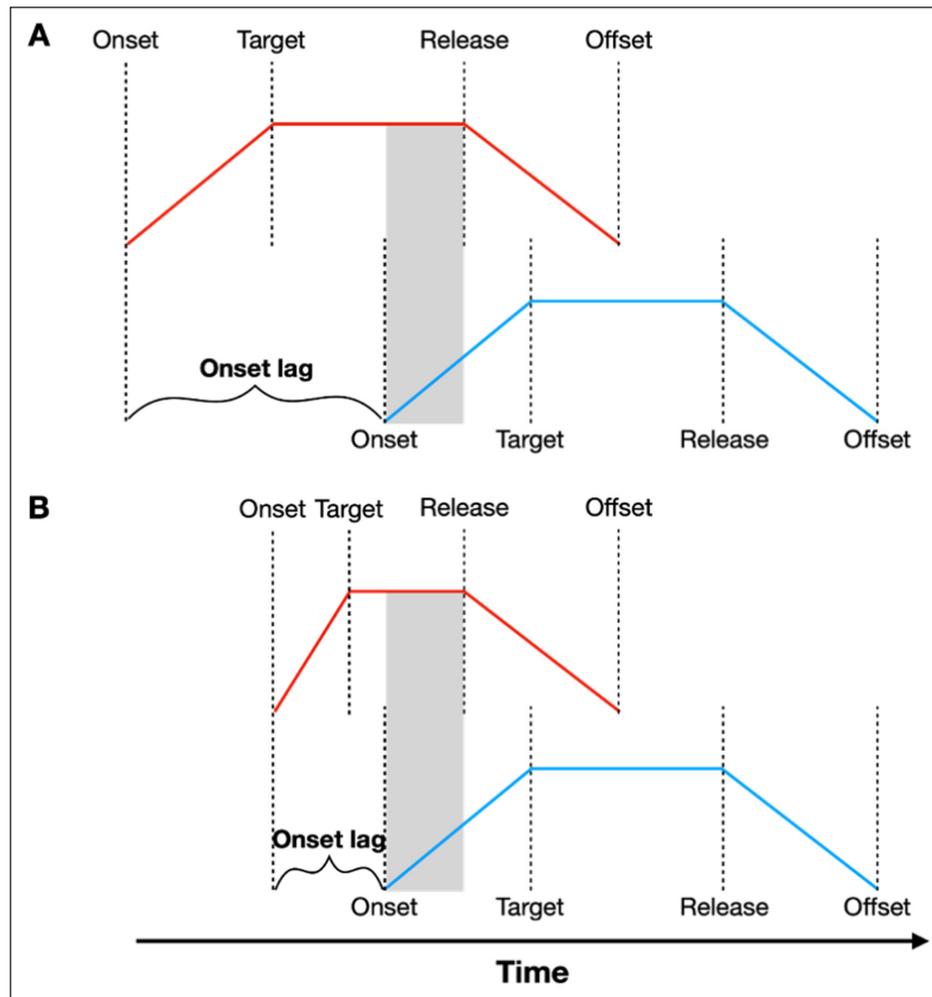
Figure 3: Relative overlap reflects both the effect of rapidity and that of C1 plateau duration. Therefore, onset lag is used as a compensatory measure of overlap.

To compensate for this potential drawback of the measure of relative overlap and assess the robustness of different notions of rapidity as predictors of overlap, we followed Roon et al. (2021) and also employed a simpler lag time measure of overlap defined as the temporal distance between the onsets of the two gestures and referred to as onset lag (as shown in **Figure 3**). Onset lag was computed by subtracting the timestamp of C2 onset from that of C1 onset, as in Roon

et al. (2021), in order to maintain the convention that larger values indicate more and smaller values less overlap.

(10)   Onset lag $=$ Onset$_{C1}$ – Onset$_{C2}$

C2 Onset position	Onset Lag value	Indication
A	> 0	C2 onset before C1 onset
B	= 0	C2 onset coincides with C1 onset
C	< 0	C2 onset after C1 onset

Table 3. Indication of different values of onset lag.

In addition to the two measures of overlap employed in Roon et al. (2021), we also quantified overlap by a simpler version of the relativized overlap measure without normalization. This is simply the time lag between C2 onset and C1 release, as in (11). This measure, which is denoted as absolute overlap in the present work, also teases out the potential effect of C1 plateau on overlap and allows us to assess the robustness of any effects over the specifics of the different overlap measures used.

(11)   Absolute overlap $=$ Release$_{C1}$ – Onset$_{C2}$

C2 Onset position	Onset Lag value	Indication
A	> 0	C2 onset before C1 release
B	= 0	C2 onset coincides with C1 release
C	< 0	C2 onset after C1 release

Table 4: Indication of different values of absolute overlap.

3. Results

In total, 658 German, 318 English, and 511 Spanish tokens were selected for the current study.[5] For each of the measurements described in §2.4, data points located more than three

---

[5] The total number of tokens measured was 1541 (707 German, 318 English, 516 English). A first-pass automatic segmentation indicated that 54 from these tokens, deriving from the German datasets, had a C1 onset starting at least 250 ms earlier than C2 onset. Most of these tokens were produced in the German carrier phrases where an utterance boundary preceded C1, *Als ich Tom sah, ___ sagte er sofort* 'When I saw Tom, he said ___ immediately' and *Zunächst sah ich Anna. ___ sagte sie* 'First I saw Anna. ___ said she'. In these contexts, the closing gesture of C1 attained its constriction target and stayed put there for a prolonged period, in an anticipatory fashion before enunciating the rest of the required utterance, thus dramatically extending C1's plateau (beyond its expected range in running speech) and rendering some of the overlap measures we employ in our study inapplicable. A similar issue of overextended

standard deviations away from the mean of the measurement were excluded. This led to an exclusion of 95 tokens (47 German, 17 English, 31 Spanish). The rest of the 1392 tokens went into the subsequent analyses. **Table 4** shows the number of tokens for each cluster in each language.

Cluster	/bl/	/pl/	/gl/	/kl/
German	155	157	142	157
English	38	156	/	124
Spanish	163	122	104	91

**Table 4:** Number of tokens for the analyses by language and cluster.

3.1. Hypothesis 1: Validity across datasets

Our aim here is to assess whether the findings of Roon et al. (2021) can be extended to the three languages in the present study. If this is indeed the case, then we should find no evidence for a consistent ordering of articulator-specific inherent velocities and no effect of peak velocity on overlap. Instead, an effect of relative rapidity indexed specifically by stiffness difference is expected on overlap.

First, we looked at whether the notion of inherent velocity finds any basis in our data and specifically whether inherent velocities follow the order in (1), COR > LAB > VEL, from most rapid to less rapid. **Figure 4A** displays the peak velocity for velars (VEL), labials (LAB), and coronals (COR). The results for COR are shown separately for each context as defined by C1 place, with separate violin plots for (VEL-)COR and (LAB-)COR. **Figure 4B** does the same for stiffness across German, English, and Spanish. In our datasets, LAB seems to have both higher peak velocity and higher stiffness than COR, with the difference between the two being the largest in Spanish when rapidity is indexed by peak velocity (rightmost panel in **Figure 4A**). COR generally has lower peak velocity than VEL and LAB, but the differences are mitigated when rapidity is measured by stiffness.

C1 plateau durations was met also in 5 Spanish tokens and was caused, in this case, by the high front vowel /i/ preceding a C1 /g/ or /k/. Since the vocalic gesture of the preceding /i/ involves a movement of the same effector (the tongue body) as for /g, k/, the TB sensor is elevated to a region very close to its consonantal target before the consonant. This results in no robust kinematic signatures for separating the plateau of the consonant from that of the earlier vocalic movement, which in turn leads to an unusually long plateau for the C closure on the trajectory of the TB sensor using the standard algorithmic identification procedures. These Spanish tokens and the aforementioned German tokens were removed from further analysis.
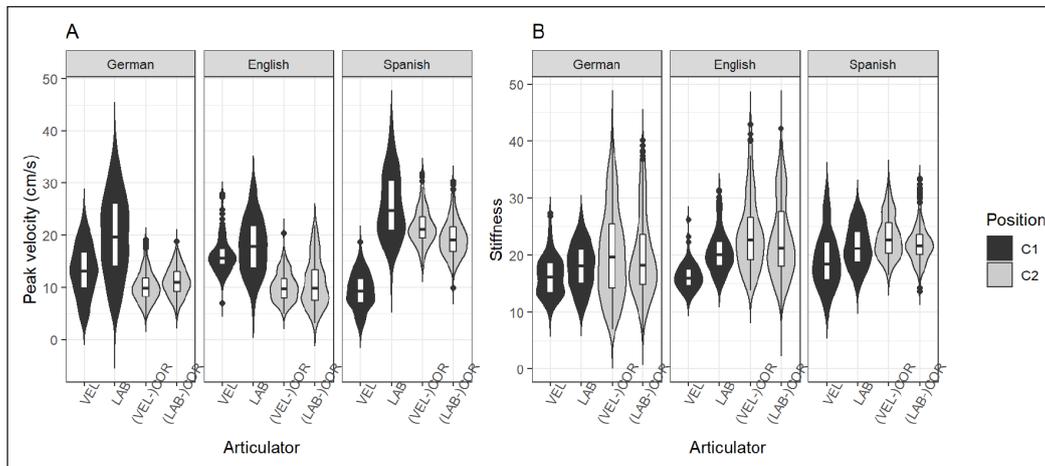
Figure 4: Peak velocity (A) and stiffness (B) values, by primary articulator and cluster position in our German, English and Spanish datasets.[6]

To determine which differences shown in the figures were significant, two linear mixed-effects models (Baayen, Davidson, & Bates, 2008; Gelman & Hill, 2007) were fit to each individual language using the LME4 package (Bates, 2005; Bates et al., 2020) for R (R Development Core Team, 2018). One model had peak velocity as the dependent variable and the other had stiffness as the dependent value. Articulator was modeled as the fixed effect. SPEAKER was modeled as random effect allowing varying intercepts and varying slopes for the effect of articulator by speaker. Omnibus test statistics for the fixed effects were determined using a type III analysis of variance with Satterthwaite's method from the ANOVA function of the stats package (R Development Core Team, 2018) in R. The results of the statistical models are listed in **Table 5**.

---

[6] One of the reviewers pointed out that Spanish COR peak velocity is almost double that of German and Spanish, whereas the difference disappears in stiffness, and thus suspected that there is a difference in the amplitude of the lateral movement. We examined the data accordingly and did find that the COR amplitude in Spanish (around 9 mm) is indeed nearly double that for German and English (around 5 mm). We think there are two reasons for this difference in amplitude across languages (Spanish vs. German and English). First, the preceding high front vowel in the Spanish recordings increases the distance between the position of the TT sensor at the timestamp of the preceding vowel target and its position at the following lateral target. When /i/ is articulated, the tongue tip is tucked further down behind the lower incisors due to the raised tongue body, compared to the tongue tip position of the preceding low vowel in German and English. Second, while in German and English the lateral is laminal, in Spanish it is apical (tongue tip is raised up, Ladefoged & Maddieson, 1996, p. 189). In both cases, an EMA sensor, such as TT in our case, attached about 2 cm behind the tongue apex traces a clear consonantal gesture but its trajectory differs due to the mid-sagittal shape differences between apical and laminal constrictions (see Gafos, 1996, Ch. 4). Since in Spanish it is the tongue tip that raises to make contact whereas in German and English it is the tongue blade, COR /l/ in Spanish has to travel an extra distance to attain the constriction target compared to COR /l/ in German and English. In short, when /l/ is articulated in Spanish, it starts from a lower position due to the preceding high vowel (reason 1) and ends at a higher position due to the Spanish lateral being apical (reason 2), hence the larger amplitude values for Spanish laterals than for German and English laterals.

Effect	df	Peak velocity (Figure 3A)					Stiffness (Figure 3B)				
		Sum Sq	Mean Sq	Den. df	*F*	*p*	Sum Sq	Mean Sq	Den. df	*F*	*p*
German											
Art.	3	**1391**	**464**	**5.07**	**26.95**	**0.002**	338	113	4.95	4.15	0.081
English											
Art.	3	**425**	**142**	**4.92**	**16.01**	**0.006**	**3931**	**1310**	**6.61**	**51.75**	**0.000**
Spanish											
Art.	3	**10315**	**3438**	**3.94**	**278.01**	**0.000**	**410**	**137**	**4.04**	**13.27**	**0.015**

Table 5: Results of the linear mixed-effect models evaluating differences in peak velocity (left) and stiffness (right) by articulator (Art.) for each language. Significant effects are in bold.

There were significant effects of articulator across languages for peak velocity and significant effects of articulator in English and Spanish for stiffness. Post-hoc differences in articulator pairs within language were assessed using estimated marginal means (Searle, Speed, & Milliken, 1980) provided by the EMMEANS package (Lenth, Singmann, Love, Buerkner, & Herve, 2019) for R. **Table 6** shows the results, where pairs with a Tukey-adjusted *p*-value < 0.05 were considered reliable and are shown in bold.

In German, LAB had a significantly higher peak velocity than both VEL and COR (independent of C1 place). The differences between the same two sets of articulators when stiffness is used as the index of rapidity were however not significant. The same goes for English: although VEL had significantly lower stiffness than both LAB and COR (independent of C1 place), and LAB had significantly higher peak velocity than COR (independent of C1 place), their differences in the other rapidity measure were not significant (except for the difference between VEL and (VEL-)COR, which was significant across both rapidity measures). Similarly, Spanish VEL had significantly lower peak velocity than both LAB and COR (independent of C1 place). However, the differences were not significant in stiffness (except for the difference between VEL and (VEL-)COR, which was significant across both rapidity measures). Finally, C1 place had no effect on the rapidity of COR at all regardless of which rapidity measures were used as the index. In general, even though there are some significant differences in rapidity measures between certain articulators for each language, the majority of those differences was not consistent across both rapidity measures and also not consistent across languages. More importantly, they did not conform to the rapidity ordering (COR > LAB > VEL) derived from Jun's proposal, regardless of rapidity index or language. Hence, we conclude that, similar to the findings in Roon et al. (2021), there is no reliable evidence for the concept of inherent velocity and its ordering as in Prediction 1 in the current data (Component 1, Hypothesis 1).

Contrast	Peak Velocity (Figure 3A)					Stiffness (Figure 3B)				
	est.	SE	df	t ratio	p	est.	SE	df	t ratio	p
German										
VEL vs. LAB	**−6.43**	**1.40**	**5.00**	**−4.59**	**0.022**	−1.73	1.14	5.00	−1.51	0.494
VEL vs. (VEL-)COR	3.06	1.47	5.00	2.08	0.276	−4.38	1.72	5.00	−2.54	0.167
VEL vs. (LAB-)COR	2.33	1.04	5.00	2.25	0.230	−3.49	2.01	5.00	−1.73	0.398
LAB vs. (VEL-)COR	**9.49**	**2.03**	**5.00**	**4.68**	**0.020**	−2.65	2.07	5.00	−1.28	0.609
LAB vs. (LAB-)COR	**8.76**	**1.39**	**5.00**	**6.33**	**0.006**	−1.76	1.62	5.00	−1.08	0.715
(VEL-COR) vs. (LAB-COR)	−0.73	0.75	4.99	−0.96	0.776	0.90	1.54	5.00	0.58	0.934
English										
VEL vs. LAB	−1.71	2.32	5.00	−0.74	0.880	**−4.48**	**0.82**	**4.87**	**−5.50**	**0.011**
VEL vs. (VEL-)COR	**5.96**	**1.53**	**5.00**	**3.89**	**0.041**	**−7.45**	**0.80**	**4.70**	**−9.27**	**0.001**
VEL vs. (LAB-)COR	5.09	2.29	5.00	2.22	0.238	**−6.64**	**0.94**	**4.92**	**−7.03**	**0.004**
LAB vs. (VEL-)COR	**7.67**	**1.60**	**5.00**	**4.78**	**0.018**	−2.96	0.97	4.93	−3.06	0.097
LAB vs. (LAB-)COR	**6.80**	**1.58**	**5.00**	**4.30**	**0.028**	−2.16	1.20	5.00	−1.80	0.372
(VEL-COR) vs. (LAB-COR)	−0.87	1.65	5.00	−0.53	0.948	0.81	1.22	4.97	0.66	0.907
Spanish										
VEL vs. LAB	**−17.40**	**1.50**	**3.96**	**−11.57**	**0.001**	−2.48	2.17	3.99	−1.15	0.685
VEL vs. (VEL-)COR	**−13.11**	**1.29**	**3.91**	**−10.16**	**0.002**	**−4.09**	**0.93**	**3.74**	**−4.42**	**0.043**
VEL vs. (LAB-)COR	**−11.07**	**1.91**	**3.98**	**−5.81**	**0.015**	−3.14	1.08	3.92	−2.91	0.138
LAB vs. (VEL-)COR	4.30	2.28	3.99	1.88	0.361	−1.61	1.79	3.98	−0.90	0.807
LAB vs. (LAB-)COR	6.33	2.56	4.00	2.47	0.205	−0.66	1.37	4.00	−0.48	0.960
(VEL-COR) vs. (LAB-COR)	2.03	0.94	3.87	2.16	0.280	0.95	0.98	3.88	0.97	0.772

Table 6: Pairwise comparisons of estimated marginal means for articulators within language with movement rapidity indexed by peak velocity (left) and stiffness (right). Significant effects are in bold.

Next, we inspected whether there were robust differences in overlap between LAB-COR and VEL-COR clusters for the three languages. Specifically, LAB-COR was expected to show more overlap than VEL-COR as per (2). Given that support for an ordering of articulator-specific inherent velocity was not obtained for the current data, there was no rapidity-difference basis on which to expect LAB-COR and VEL-COR clusters to differ significantly in overlap. Thus, we predicted that either no reliable differences or differences in the non-expected direction exist in overlap between LAB-COR and VEL-COR across languages (that is, VEL-COR has more overlap than LAB-COR). **Figure 5A** shows relative overlap by cluster for each language, **Figure 5B** does the same for onset lag, and **Figure 5C** for absolute overlap. Across languages, clusters with voiced C1 (/bl, gl/) show more overlap than their voiceless counterparts, especially for English and Spanish. Additionally, English and Spanish clusters show a potential effect of C1 place, whereby LAB-COR clusters (/bl/, /pl/) appear to be more overlapped than VEL-COR clusters (/gl/, /kl/).
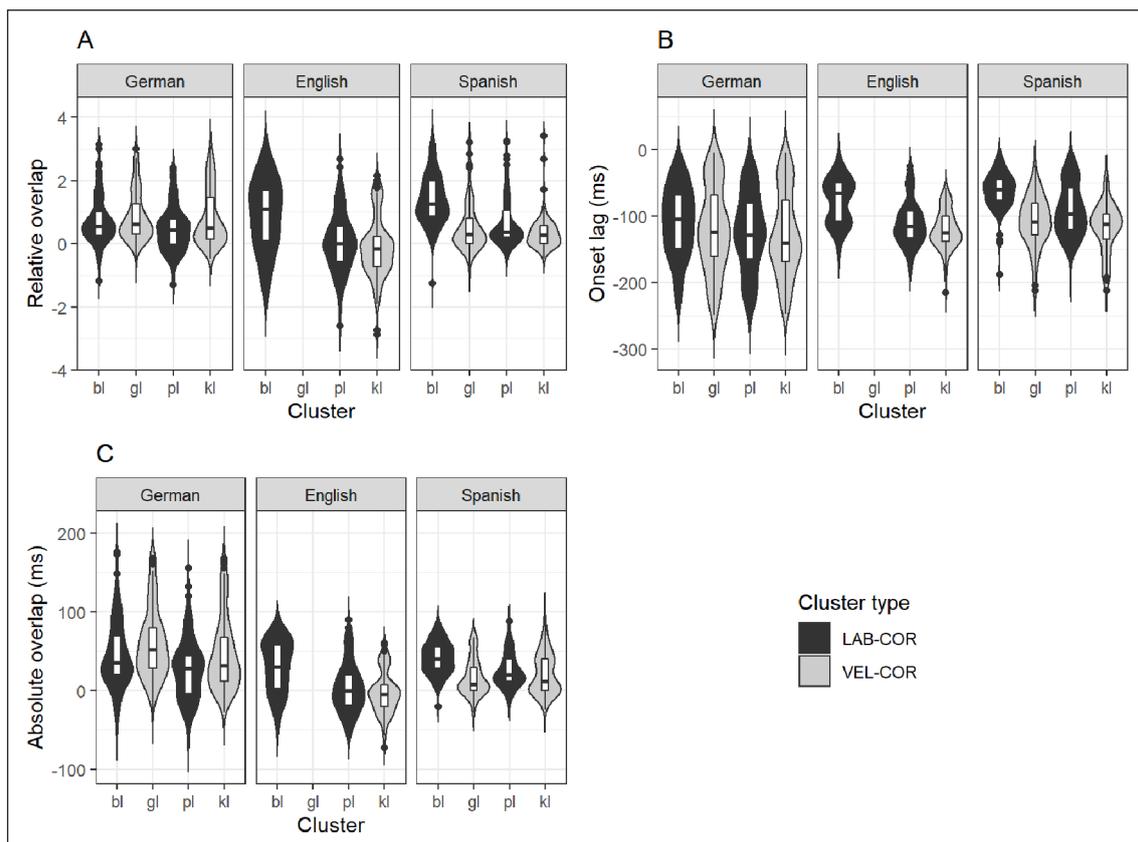
Figure 5: Relative overlap (A), onset lag (B) and absolute overlap (C) by cluster (i.e., /bl/, /gl/, /pl/, /kl/) in our German, English and Spanish datasets.

Contrast	Relative Overlap (Figure 5A)					Onset Lag (Figure 5B)					Absolute Overlap (Figure 5C)				
	est.	SE	df	t	p	est.	SE	df	t	p	est.	SE	df	t	p
German															
(Intercept)	0.77	0.15	5.51	4.16	0.003	−108.23	10.54	5.59	−10.27	0.000	47.08	8.77	5.24	5.37	0.024
VEL-COR	0.09	0.15	7.30	0.59	0.575	−12.74	7.24	21.31	−1.76	0.093	12.25	9.39	6.53	1.31	0.236
C1 voicing	**−0.31**	**0.09**	**17.80**	**−3.34**	**0.004**	**−16.72**	**6.19**	**267**	**−2.70**	**0.007**	**−20.34**	**5.75**	**12.34**	**−3.54**	**0.004**
VEL-COR:C1 voicing	**0.26**	**0.11**	**599**	**2.30**	**0.022**	10.69	8.81	602	1.21	0.226	6.76	6.48	598	1.04	0.297
English															
(Intercept)	0.87	0.28	5.31	3.13	0.024	−79.13	10.95	5.10	−7.23	0.001	30.97	9.78	5.15	3.17	0.024
VEL-COR	−0.21	0.18	3.80	−1.20	0.301	−10.25	8.47	4.95	−1.21	0.281	−7.76	6.88	4.41	−1.13	0.317
C1 voicing	**−0.87**	**0.26**	**5.32**	**−3.28**	**0.020**	**−31.88**	**8.70**	**5.30**	**−3.66**	**0.013**	**−28.46**	**8.37**	**5.14**	**−3.40**	**0.018**
Spanish															
(Intercept)	1.44	0.27	4.04	5.34	0.006	−59.99	10.74	4.04	−5.59	0.005	42.50	5.96	4.11	7.13	0.002
VEL-COR	**−0.99**	**0.17**	**5.19**	**−5.77**	**0.002**	**−54.32**	**4.41**	**4.76**	**−12.31**	**0.000**	**−25.28**	**4.30**	**6.26**	**−5.88**	**0.001**
C1 voicing	**−0.76**	**0.07**	**66.03**	**−11.32**	**0.000**	**−29.55**	**4.82**	**5.58**	**−6.13**	**0.001**	**−17.57**	**4.18**	**4.83**	**−4.20**	**0.009**
VEL-COR:C1 voicing	**0.54**	**0.10**	**436**	**5.27**	**0.000**	**17.44**	**4.88**	**347**	**3.58**	**0.000**	**19.02**	**3.56**	**447**	**5.34**	**0.000**

Table 7: Results of the linear mixed-effects models with cluster type as a predictor of relative overlap (upper left), onset lag (upper right) and absolute overlap (lower left) within language. Significant effects compared to LAB-COR clusters are bolded.

Du and Gafos: Articulatory overlap as a function of stiffness in German, English and Spanish word-initial stop-lateral clusters

23

Three linear mixed-effects models were fit to the dataset of each language individually with either relative overlap or onset lag as the dependent variable to determine the reliability of the differences. Cluster type and C1 voicing were modeled as fixed effects for all models with LAB-COR cluster and voiced C1 as the reference level. Random intercepts and slopes for both fixed effects were included for. For German and Spanish, an additional interaction between the fixed effects was also included in each model (English data do not allow this since /gl/ is lacking). **Table 7** shows that there was no reliable difference between the two cluster types in German and English regardless of whether overlap was indexed by relative overlap, onset lag or absolute overlap, but in Spanish the same difference was significant across all three overlap measures. C1 voicing revealed a consistently significant effect on all overlap measures across languages, such that clusters with voiced C1 show more overlap than their voiceless counterparts. In addition, Spanish clusters showed a significant interaction between cluster type and C1 voicing across overlap indices. Ad-hoc pairwise comparisons confirmed that the significant interaction in Spanish was due to Spanish /bl/ being significant more overlapped than all the other three clusters across overlap indices.

Lastly but most importantly, we turned to the effects of relative rapidity, where Roon et al. (2021) found the stiffness difference, rather than the peak velocity difference, between C1 and C2 to be a reliable predictor of overlap. **Figure 6** shows the relationship between relative rapidity and overlap by cluster type (horizontally) and language (vertically). **Figure 6A** plots peak velocity difference in the two cluster types (left two columns) and stiffness difference in the two cluster types (right two columns) against relative overlap, **Figure 6B** does the same for onset lag, and **Figure 6C** for absolute overlap. A colored regression line was overlayed for every speaker on each subplot.

Across languages, cluster types, and speakers, stiffness difference mostly showed a positive correlation with the three overlap indices (right two columns in **Figure 6A, 6B, 6C**) with the exception of Spanish VEL-COR clusters (the subplot in the bottom right corner of **Figure 6A, 6B, 6C**) and the VEL-COR clusters of one English speaker (the last subplots in the second rows of **Figure 6A, 6C**). The correlation between peak velocity difference and overlap was much less uniform compared to the former relation and showed substantial individual variability across speakers within each language (left two columns in **Figure 6A, 6B, 6C**). Six linear mixed models (three overlap measures by two relative rapidity measures) were fit to the data for each language (for a total of 18 linear mixed models). Peak velocity difference or stiffness difference (both continuous), cluster type, and their interaction were modelled as fixed effects. LAB-COR was designated as the base level of cluster type. Random intercepts were allowed for both speaker and stimulus. Speaker-specific random slopes for cluster type were also included. The results of all models are listed in **Table 8**.
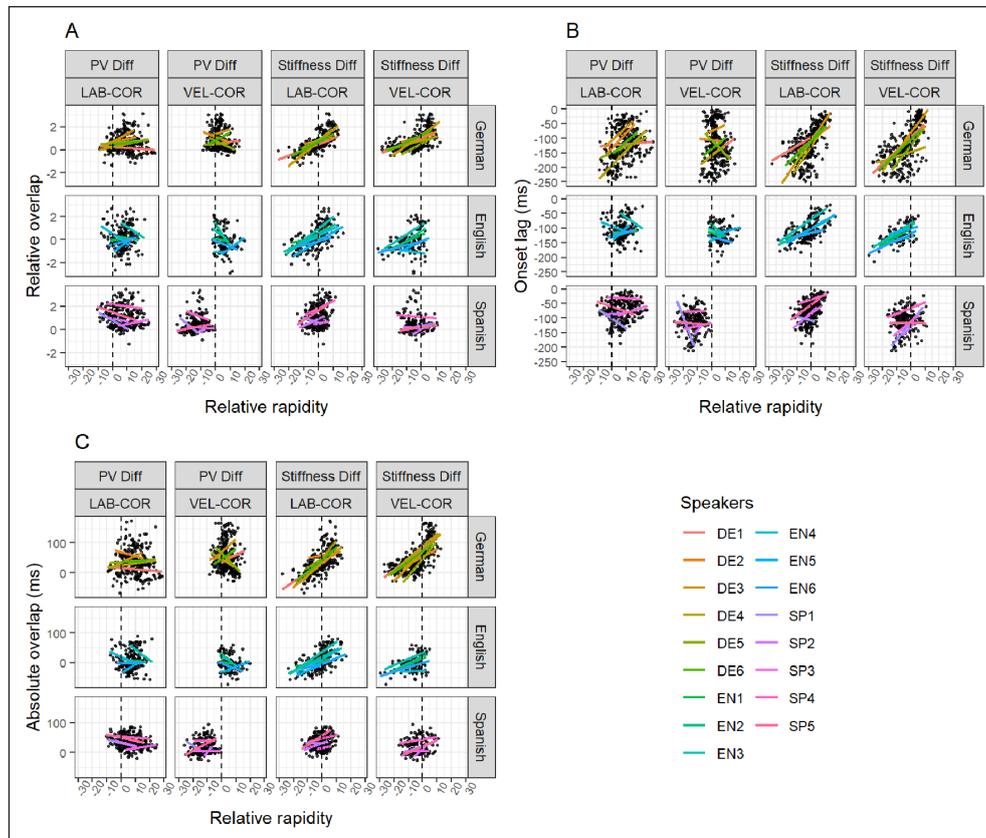
Figure 6: Relationship between within-token C1-C2 relative rapidity and relative overlap (A), onset lag (B) and absolute overlap (C) across our German, English and Spanish datasets. In the left two columns of each plot (LAB-COR and VEL-COR respectively), relative rapidity is indexed by C1-C2 peak velocity difference, whereas in the right two columns (LAB-COR and VEL-COR respectively) it is indexed by C1-C2 stiffness difference. Colored regression lines correspond to speakers.[7]

The statistical results confirmed that effects of stiffness difference on the three overlap indices were all significant across all three languages and both cluster types (LAB-COR, VEL-COR) assessed here.[8] This was not true for peak velocity difference, which had significant effects only

---

[7] One reviewer noted that Spanish LAB-COR and VEL-COR clusters form two clouds in the plots for peak velocity difference in Figure 6A, B and C. This is mainly due to Spanish VEL having significantly lower peak velocity than Spanish LAB and COR, so that VEL minus COR (VEL-COR peak velocity difference) has much lower values (almost always below zero) than LAB minus COR (LAB-COR peak velocity difference, mostly above zero). This can be seen in Figure 4 and Table 6.

[8] Note that despite the significant interactions between stiffness difference and VEL-COR clusters in German (relative overlap) and Spanish (relative overlap and absolute overlap), which indicated that the effect size of stiffness difference in VEL-COR was significantly smaller than that in LAB-COR, post-hoc analyses on VEL-COR clusters alone confirmed that the effect of stiffness difference still reached significance (p-value < 0.05) in German and Spanish VEL-COR clusters.

in German LAB-COR clusters when using relative overlap and onset lag, but not absolute overlap, in English LAB-COR clusters when using absolute overlap, but not relative overlap and onset lag, and in Spanish VEL-COR clusters when using absolute overlap, but not relative overlap and onset lag, based on post-hoc analyses. Therefore, the effect of peak velocity difference on overlap was not robust across languages, clusters and overlap measures.

Overall, stiffness difference does appear as a more robust predictor of overlap than peak velocity difference across languages, consistent with the findings of Roon et al. (2021) in Moroccan Arabic (Component 3, Hypothesis 1).

3.2. Hypothesis 2: Individual consonant stiffness effects

While Roon et al. (2021) first established a robust effect of stiffness difference on overlap, a result we replicated and extended here to new languages, their measure (the difference between the two stiffness values of the two consonants) yokes the stiffness values of the two consonants in one independent variable. As such, this measure cannot address whether both C1 and C2 stiffness, individually, contribute to overlap in a way predicted by Jun (2004). It is this prediction we take up next. Specifically, the proposal in Jun (2004) predicts that an increase in overlap should be accompanied by increasing C1 stiffness and/or decreasing C2 stiffness. To test this hypothesis, we modelled both C1 and C2 stiffnesses and their interactions with cluster type as fixed effects. Relative overlap, onset lag and absolute overlap remained as dependent variables. Speaker and stimulus were included as random effects enabling varying intercepts. Random slopes of cluster type by speaker were allowed as well. **Figure 7** displays the relations between C1 or C2 stiffness with relative overlap (A), onset lag (B) and absolute overlap (C) respectively divided by cluster types (horizontally) and languages (vertically). Colored regressions lines were added for speakers. As seen in **Figure 7**, while C2 stiffness shows a consistent negative correlation with all overlap measures across languages and speakers except for some in Spanish VEL-COR (right two columns in **Figure 7A, 7B, 7C**), C1 stiffness only shows a relatively clear relation with onset lag (left two columns in **Figure 7B**) and even here not without inconsistency among speakers in German and Spanish VEL-COR. Results of the statistical analyses are shown in **Table 9**.

Statistical results provide evidence for C2 stiffness as a robust predictor of overlap, as its effect on overlap was significant across languages, cluster types, and overlap indices. The only exception was Spanish VEL-COR clusters when overlap was indexed by one of the three overlap measures, that of onset lag. Post-hoc analysis done on Spanish VEL-COR clusters alone revealed that the effect of C2 stiffness on onset lag approached statistical significance (p-value $= 0.064$). Crucially, in all cases the correlation between C2 stiffness and the three overlap measures was negative for both cluster types across all three languages, confirming the prediction of Jun's proposal regarding the effect of C2 rapidity on overlap (Hypothesis 2). On the other hand, C1

PV Difference

effect	Relative overlap (Figure 6A)					Onset lag (Figure 6B)					Absolute overlap (Figure 6C)				
	est.	SE	df	t	p	est.	SE	df	t	p	est.	SE	df	t	p
German															
(Intercept)	0.45	0.20	7.21	2.27	0.057	−143.97	13.97	7.18	−10.30	0.000	33.29	11.42	5.36	2.92	0.031
PV Diff.	**0.02**	**0.01**	**585.57**	**3.16**	**0.002**	**3.10**	**0.42**	**563.67**	**7.30**	**0.000**	0.40	0.34	589.40	1.17	0.245
VEL-COR	0.38	0.21	5.27	1.78	0.132	16.11	10.37	3.06	1.55	0.216	18.32	14.99	4.47	1.22	0.282
PV Diff.:VEL-COR	−0.02	0.01	364.36	−1.52	0.129	**−1.83**	**0.821**	**448.17**	**−2.23**	**0.027**	−0.17	0.73	334.73	−0.23	0.815
English															
(Intercept)	0.36	0.28	10.46	1.29	0.223	**−103.29**	**10.33**	**10.32**	**−1.00**	**0.000**	14.73	0.84	9.50	1.50	0.167
PV Diff.	−0.03	0.02	252.75	−1.72	0.087	−0.17	0.48	276.71	−0.35	0.727	**−0.93**	**0.44**	**278.02**	**−2.12**	**0.035**
VEL-COR	−0.35	0.38	11.89	−0.93	0.371	−12.35	14.46	11.62	−0.85	0.410	−13.74	12.54	11.63	−1.10	0.295
PV Diff.:VEL-COR	−0.01	0.03	192.67	−0.39	0.700	−0.75	0.87	182.10	−0.86	0.390	−0.09	0.81	180.51	−0.11	0.909
Spanish															
(Intercept)	1.09	0.29	5.95	3.74	0.010	−76.27	12.42	5.85	−6.14	0.001	35.09	6.98	6.01	5.03	0.002
PV Diff.	0.00	0.01	384.15	0.40	0.687	0.56	0.31	61.28	1.81	0.075	−0.04	0.23	156.22	−0.15	0.880
VEL-COR	−0.55	0.28	22.63	−1.98	0.060	**−37.31**	**10.39**	**34.71**	**−3.59**	**0.001**	−1.65	7.07	32.46	−0.23	0.817
PV Diff.:VEL-COR	0.01	0.01	245.68	0.82	0.413	−0.11	0.51	31.78	−0.21	0.838	**1.30**	**0.39**	**88.57**	**3.32**	**0.001**

(Contd.)

Stiffness Difference

effect	est.	SE	df	t	p	est.	SE	df	t	p	est.	SE	df	t	p
German															
(Intercept)	0.74	0.12	2.16	6.12	0.021	−107.95	9.74	5.36	−11.08	0.000	42.78	8.83	2.42	4.85	0.027
Stiff. Diff.	**0.07**	**0.00**	**573.29**	**17.50**	**0.000**	**5.01**	**0.37**	**502.69**	**13.56**	**0.000**	**3.65**	**0.24**	**215.69**	**15.34**	**0.000**
VEL-COR	0.32	0.22	4.85	1.44	0.212	2.67	11.65	3.71	0.23	0.831	26.25	14.98	4.38	1.75	0.149
Stiff. Diff.:VEL-COR	**−0.02**	**0.01**	**517.53**	**−3.34**	**0.001**	−0.72	0.48	450.02	−1.48	0.139	0.22	0.33	490.46	0.67	0.504
English															
(Intercept)	0.34	0.22	8.88	1.58	0.150	−99.07	8.27	9.14	−11.98	0.000	13.41	7.68	8.82	1.75	0.115
Stiff. Diff.	**0.07**	**0.01**	**286.69**	**9.47**	**0.000**	**2.55**	**0.22**	**286.89**	**11.85**	**0.000**	**2.39**	**0.21**	**287.58**	**11.33**	**0.000**
VEL-COR	−0.13	0.29	9.45	−0.44	0.671	−3.75	9.92	9.47	−0.38	0.714	−5.58	9.10	9.37	−0.61	0.554
Stiff. Diff.:VEL-COR	−0.01	0.01	286.99	−1.06	0.292	−0.05	0.35	285.79	−0.14	0.886	−0.60	0.35	285.67	−1.73	0.084
Spanish															
(Intercept)	1.16	0.30	6.28	3.92	0.007	−70.66	11.64	7.55	−6.07	0.000	36.34	8.28	5.61	4.39	0.005
Stiff. Diff.	**0.08**	**0.01**	**462.90**	**9.53**	**0.000**	**3.29**	**0.35**	**446.85**	**9.29**	**0.000**	**2.31**	**0.27**	**464.35**	**8.60**	**0.000**
VEL-COR	**−0.71**	**0.31**	**10.00**	**−2.30**	**0.044**	**−34.77**	**12.04**	**13.51**	**−2.89**	**0.012**	−13.67	9.23	7.61	−1.50	0.174
Stiff. Diff.:VEL-COR	**−0.06**	**0.01**	**468.26**	**−5.20**	**0.000**	−0.17	0.48	467.16	−0.34	0.731	**−0.84**	**0.36**	**471.74**	**−2.34**	**0.020**

Table 8: Results of the 18 (6 models by 3 languages) linear mixed-effects models assessing the relation between overlap and relative rapidity. For each language, 6 models were fitted, each model assessing the relation between one of the three overlap measures and one of the two relative rapidity measures. Significant effects are bolded.

stiffness was positively associated with overlap only for onset lag but not for relative overlap and absolute overlap across languages. Apart from onset lag, the only cases where C1 stiffness also showed a significant effect were German LAB-COR clusters (relative overlap) and English LAB-COR clusters (absolute overlap). For the last two cases, the effect of C1 stiffness in the corresponding VEL-COR clusters not only did not reach significance, but also had opposite signs compared to its effect in LAB-COR clusters. Specifically, in German and English LAB-COR clusters, C1 stiffness was positively associated with relative overlap and absolute overlap respectively, whereas in VEL-COR clusters of the two languages it was negatively associated with overlap.
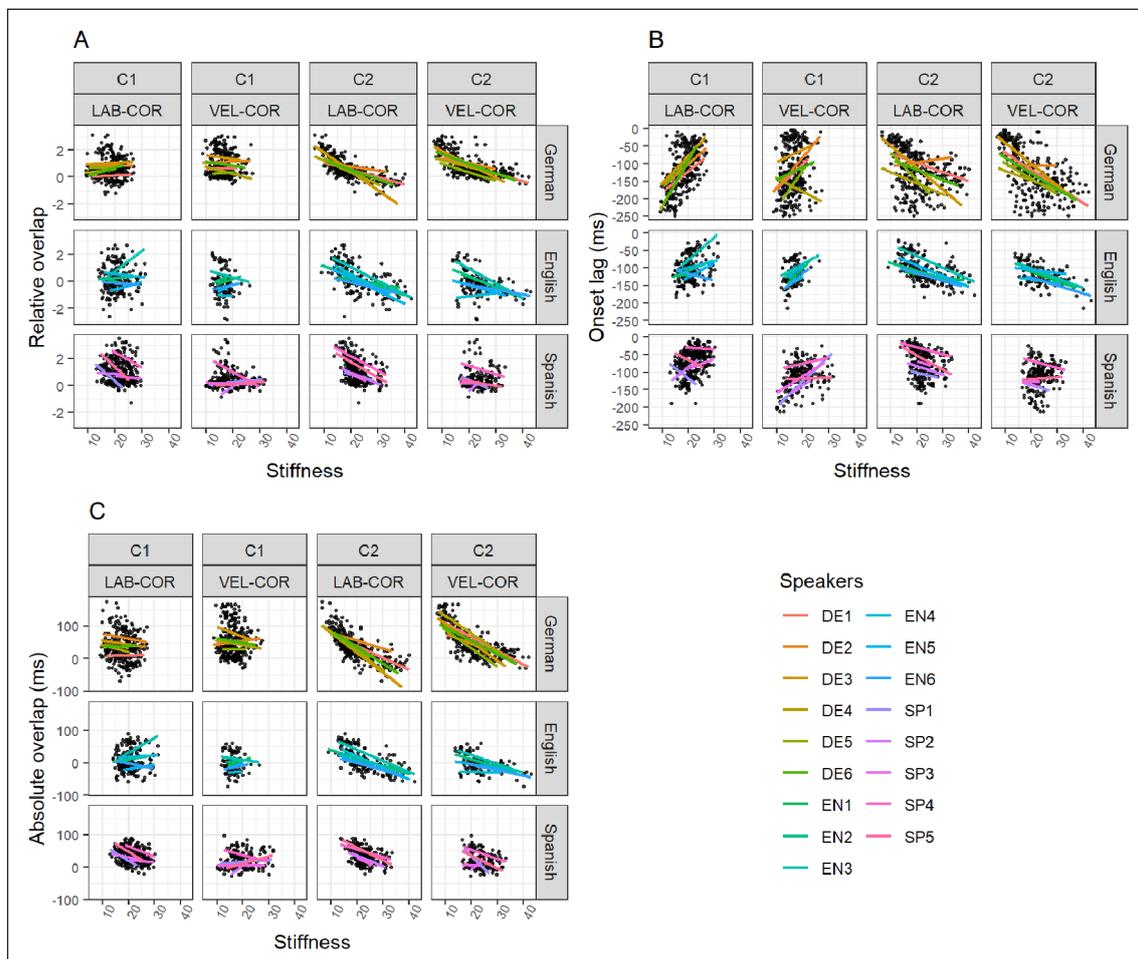
Figure 7: Relationship between C1 or C2 stiffness and relative overlap (A), onset lag (B) and absolute overlap (C) across our German, English and Spanish datasets. C1 stiffness is shown in the left two columns of each plot (LAB-COR and VEL-COR respectively) and C2 stiffness is shown in the right two columns (LAB-COR and VEL-COR respectively). Colored regression lines correspond to speakers.

effect	Relative overlap (Figure 7A)					Onset lag (Figure 7B)					Absolute overlap (Figure 7C)				
	est.	SE	df	t	p	est.	SE	df	t	p	est.	SE	df	t	p
German															
(Intercept)	1.48	0.19	12.23	7.77	0.000	−211.37	15.82	29.83	−13.36	0.000	126.37	13.02	9.81	9.71	0.000
C1 Stiff.	0.04	0.01	69.39	4.96	0.000	9.52	0.68	244.28	14.01	0.000	0.11	0.43	536.32	0.26	0.799
C2 Stiff.	−0.08	0.00	55.84	−18.37	0.000	−3.86	0.38	238.98	−10.23	0.000	−4.65	0.24	541.43	−19.52	0.000
VEL-COR	0.89	0.30	19.48	2.99	0.007	102.61	23.21	36.09	4.42	0.000	30.19	18.77	10.35	1.61	0.138
C1 Stiff.:VEL-COR	−0.05	0.01	272.44	−4.18	0.000	−5.07	1.01	481.24	−5.04	0.000	−0.55	0.62	593.01	−0.89	0.376
C2 Stiff.:VEL-COR	0.01	0.01	168.14	2.30	0.023	−0.38	0.50	387.40	−0.762	0.447	−0.06	0.31	586.06	−0.20	0.844
English															
(Intercept)	1.38	0.44	115.03	3.13	0.000	−105.23	14.23	65.76	−7.40	0.000	45.31	13.36	79.40	3.39	0.001
C1 Stiff.	0.03	0.02	265.89	1.81	0.071	2.79	0.49	283.84	5.68	0.000	1.16	0.48	283.76	2.43	0.016
C2 Stiff.	−0.08	0.01	286.05	−9.79	0.000	−2.49	0.24	284.96	−10.53	0.000	−2.67	0.23	285.33	−11.56	0.000
VEL-COR	0.36	0.75	174.13	0.47	0.636	−41.38	22.62	139.35	−1.83	0.069	−0.12	21.83	159.38	−0.01	0.996
C1 Stiff.:VEL-COR	−0.05	0.04	273.27	−1.44	0.151	1.94	1.03	270.38	1.90	0.059	−1.29	1.00	276.93	−1.28	0.202
C2 Stiff.:VEL-COR	0.01	0.01	284.09	0.90	0.369	0.30	0.38	282.78	0.79	0.430	0.63	0.37	282.65	1.69	0.093

Spanish

	Model 1					Model 2					Model 3				
(Intercept)	3.30	0.44	25.85	7.49	0.000	−96.40	18.86	49.36	−5.11	0.000	118.36	12.28	40.41	9.64	0.000
C1 Stiff.	−0.00	0.01	289.90	−0.17	0.869	**4.27**	**0.65**	**356.51**	**6.57**	**0.000**	−0.76	0.44	153.39	−1.75	0.083
C2 Stiff.	**−0.10**	**0.01**	**457.66**	**−11.60**	**0.000**	**−3.08**	**0.37**	**456.51**	**−8.26**	**0.000**	**−3.03**	**0.26**	**459.89**	**−11.50**	**0.000**
VEL-COR	**−1.49**	**0.51**	**182.96**	**−2.93**	**0.004**	**−77.367**	**24.92**	**93.77**	**−3.11**	**0.003**	−29.81	15.52	188.90	−1.92	0.056
C1 Stiff.:VEL-COR	−0.00	0.02	380.92	−0.16	0.875	0.12	0.77	416.66	0.16	0.876	1.00	0.52	260.56	1.93	0.055
C2 Stiff.:VEL-COR	**0.04**	**0.01**	**460.10**	**2.66**	**0.008**	**1.89**	**0.63**	**460.97**	**2.98**	**0.003**	−0.31	0.45	463.89	−0.70	0.483

Table 9: Results of the 9 linear mixed-effects models assessing the relation between overlap and individual stiffness effects on overlap. For each language, 3 models were fitted, each model assessing the relation between one of the three overlap measures and C1/C2 stiffness. Significant effects are bolded.

In short, C2 stiffness is a more robust predictor of overlap compared to C1 stiffness for two reasons. Firstly, it has a significant effect on overlap across the three overlap measures and the three languages (only one exception out of 18 cluster type-and-overlap combinations). Secondly, the prediction of Hypothesis 2 on the effect direction of C2 stiffness (negative) was validated by the empirical results across languages and overlap indices, whereas that of C1 stiffness was not, given that significant effects of C1 stiffness had different effect directions for the LAB-COR clusters in German and English when relative overlap and absolute overlap were used as the overlap index respectively.

4. Discussion

The findings of Roon et al. (2021) regarding rapidity, overlap, and the effect of relative rapidity for Moroccan Arabic stop-stop clusters were successfully replicated for German, English, and Spanish stop-lateral clusters in the current study. We enumerate here these findings. First, stiffness did not seem to be an inherent property of the individual oral articulators. Second, stiffness provided a better quantification of relative rapidity when predicting overlap than peak velocity did. In addition to these successful replications of the previous results, the current work inspected more closely the individual effects of C1 and C2 stiffness on overlap and found that C2 stiffness was more robust in predicting overlap than C1 stiffness was. In what follows, we discuss each of the findings in more detail.

A first finding is that the rapidities of coronals, labials, and velars did not differ systematically and robustly from each other in stop-lateral clusters of German, English, and Spanish. This goes against the notion of an articulator-specific inherent velocity proposed by Browman and Goldstein (1989, p. 87) and Jun (2004, p. 65), who claimed that different oral articulators have inherently different rapidities. The absence of significant differences in peak velocity and stiffness among the three articulators finds an early precursor in the work Kuehn and Moll (1976). Kuehn and Moll reported that when differences in displacement (or movement amplitude) were taken into account, coronal, labial, and velar gestures all have similar velocities. As in Roon et al. (2021) and in the present study, thus, Kuehn and Moll concluded that rapidity does not depend on the identity of the oral articulators (Kuehn & Moll, 1976, p. 316).

Since significant differences in rapidity measures were not obtained for the three articulators across languages, there was no rapidity-difference basis on which to expect LAB-COR and VEL-COR clusters to differ significantly in overlap. Our results are largely consistent with this: LAB-COR and VEL-COR clusters in German and English did not show a significant difference in overlap contrary to Prediction 2 (see §1.1). Spanish, on the other hand, did reveal a significant difference between the two cluster types across overlap indices, which was mainly contributed by /bl/ being more overlapped than all the other three clusters. However, since no robust difference in rapidity measures was found for Spanish LAB and VEL, we do not interpret the difference in

overlap as having a rapidity-difference basis. Given that the connection between the typological patterns in place assimilation and rapidity, as suggested by Browman and Goldstein (1990a) and Jun (2004), we did not find any basis in the data, and it seems reasonable to conclude that articulator-specific rapidity alone cannot be the source of asymmetries in assimilation patterns. Our current study does not address the potential link between typological patterns in assimilation and overlap, as it is concerned only with evaluating the relation between overlap and rapidity. Nevertheless, we point out that other work has cast doubt on the assumption that asymmetrical patterns in place assimilation result solely from differences in intersegmental overlap. Chen (2003) studied the effects of gestural overlap and segmental reduction on the recoverability of C1 in coronal-labial and labial-coronal clusters using computational modeling. The stimuli in this study were generated using GEST, a computational model of gestural structure (Browman & Goldstein, 1990a; Gafos, 2002), with corresponding acoustics derived from the Haskins articulatory synthesizer (Rubin, Baer, & Mermelstein, 1981). The "listener" model, on the other hand, was an algorithm that determines the vocal tract area functions based on the first three formant frequencies of a synthesized speech signal (Chen, 2003, p. 2822). Chen found that increasing overlap in coronal-labial clusters decreases recoverability rates for C1, whereas increasing overlap in labial-coronal clusters has little effect on C1 recovery (Chen, 2003, p. 2823). Thus, it seems that increasing overlap is not uniformly associated with weaker perception of the overlapped C1 across all cluster types. This implies that place assimilation patterns cannot be solely based on differences in overlap, but must take other factors into consideration, such as differences in the acoustic cues of different places of articulation (Winitz, Scheib, & Reeds 1972; Kuehn & Moll, 1976; Ohala, 1990; Surprenant & Goldstein, 1998; Kochetov & So, 2007), segment frequency, and potentially language-specific phonotactic knowledge (Winters, 2001, 2003).

The second finding of the current study, which is in line with Roon et al. (2021), is that stiffness, or amplitude-normalized peak velocity, is a better candidate for predicting overlap compared to the peak velocity of the closing gesture. In this regard, let us return to the relation between overlap and velocity, as depicted in the schemas of Jun (2004) shown in **Figure 1**. The fact that peak velocity was not by itself a predictor of overlap, unlike stiffness, is somewhat puzzling. Note that in these schemas all gestures are depicted as having the same movement amplitude. That is, in all of the Jun (2004) schemas, the distance the oral articulator travels to reach its target is depicted as being the same, regardless of any differences in peak velocity. The important consideration missed in these schemas, as already discussed in §2.4, is that the peak velocity of an articulatory movement covaries with its amplitude (Ostry et al., 1983; Parush, Ostry, & Munhall 1983; Kelso et al., 1985; Bootsma & van Wieringen, 1990). To be more specific, in Jun's schemas, the C2 articulation in **Figure 1**, schema C, has approximately the same amplitude as the C2 articulation in **Figure 1**, schema D, albeit their peak velocities being drastically different. This depiction does not agree with the phonetic reality, in that higher peak

velocity has been observed to be accompanied by higher amplitude and vice versa. Hence, a more realistic schematic representation of the **Figure 1** schemas C and D should be as shown in **Figure 8** below.
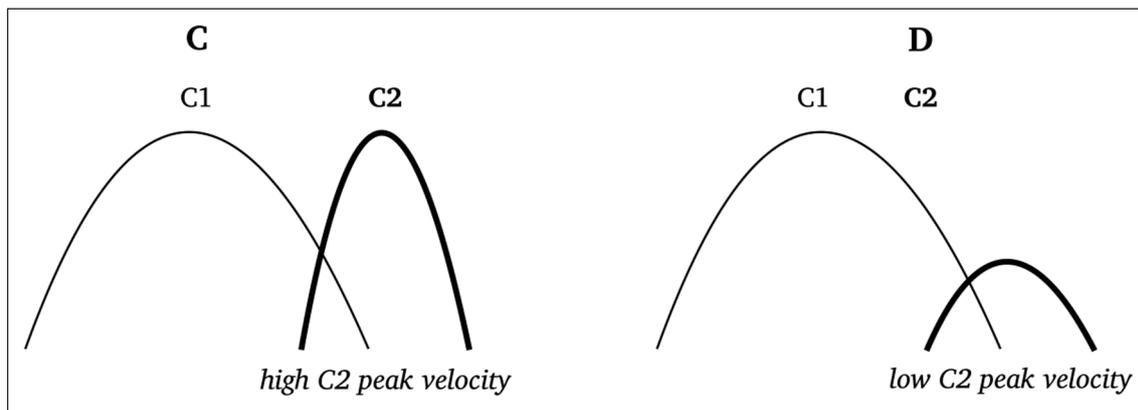
Figure 8: A more realistic revision of the schemas in Figure 1C and 1D.

Figure 8 respects the velocity-amplitude relation by depicting a low C2 peak velocity along with a low amplitude, as in schema C, and a high C2 peak velocity has a high amplitude, as in D. Note that even though the C2 articulation in schema C is faster than the C2 articulation in schema D, according to the results of the current study the two scenarios should show roughly the same overlap between the two segments because C2 stiffness is essentially the same across the two schemas (recall that stiffness is given by the ratio of peak velocity to amplitude); the same holds for C1 stiffness across the two schemas (to the extent that C1 stiffness also contributes to overlap). As the amounts of overlap in schemas C and D do not differ, we see that stiffness is evidently a better indicator of overlap than peak velocity.

The comparison intended by Jun (2004) between a scenario with a high C2 rapidity and one with a low C2 rapidity should then be captured by the schemas in **Figure 9**, where rapidity is indexed by stiffness in our sense. In **Figure 9**, stiffness should not be seen as simply equal to the slope of C2 closing phase movement, but as the ratio of that slope to the amplitude of the movement, a measure that is not easy to convey through such static schemas. Viewed in this way, the basic idea of Jun (2004) behind these schemas, namely, that a certain measure related to the rapidity of the consonants in a C1C2 cluster contributes to the amount of overlap in the cluster, is still fully consistent with the results of the current study. As per the prediction of this idea, the cluster in schema D should have more overlap (since overlap increases as C2 stiffness lowers) than in schema C, and this is indeed what the revised depictions in **Figure 9** convey.
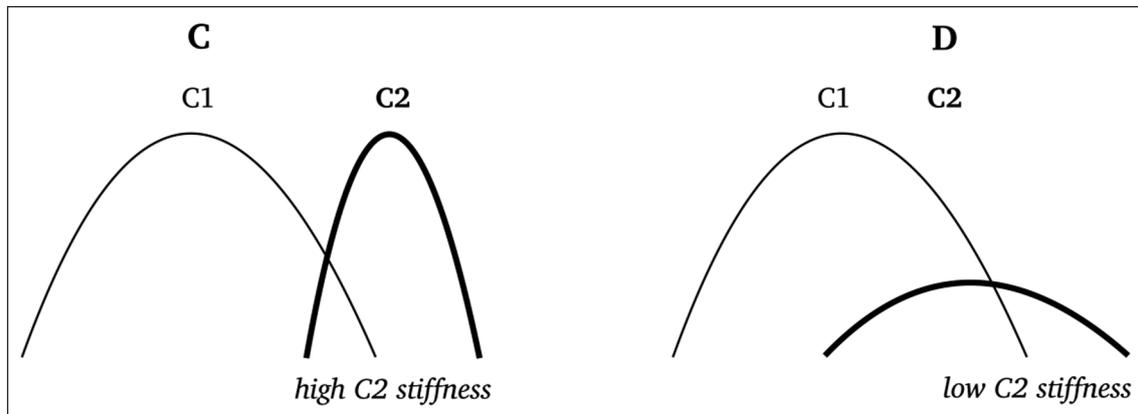
Figure 9: Comparison between a C1C2 sequence with high C2 stiffness and another with low C2 stiffness when C1 is held constant.

Extending the investigation of Roon et al. (2021), the present study also assessed any effects of, separately, C1 and C2 rapidity on overlap. Recall that C2 stiffness was a significant predictor of overlap regardless of whether overlap was indexed by relative overlap onset lag or absolute overlap, whereas C1 stiffness was only significant across languages when onset lag was used to quantify overlap. In fact, the significant effects of C1 stiffness on onset lag were hardly surprising, given that onset lag by definition contains a part of, if not the whole (depending on the timestamp of C2 onset), C1 movement duration, and movement duration is known to vary inversely with the ratio of peak velocity to amplitude (or stiffness) by previous studies (Munhall, Ostry, & Parush, 1985; Ostry & Munhall, 1985; Ostry, Cooke, & Munhall, 1987). To be more explicit, given that stiffness is negatively correlated to movement duration, a greater C1 stiffness should necessarily result in a shorter C1 movement duration. Assuming C2 onset is held constant, a shorter C1 movement duration would lead to a correspondingly shorter lag between C1 and C2 onsets, reflected in numerically larger onset lag values (less negative, recall that onset lag is computed by subtracting the timestamp of C2 onset from that of C1). Therefore, C1 stiffness can only have a positive correlation with onset lag, which was confirmed by our results. However, C2 stiffness does not have such a necessary relation with onset lag. Thus, its significant effects on onset lag across languages must reflect that, unlike C1 stiffness, it genuinely explained variances in overlap. Moreover, in terms of effect direction and consistency thereof, we see a drastic contrast between the effects of C1 versus C2 stiffness: while the effect of C2 stiffness, in the vast majority of cases, was significant and negative across all three languages regardless of the indexation of overlap confirming Hypothesis 2 (overlap should increase as C2 rapidity decreases), that of C1 stiffness, in the sporadic cases where C1 stiffness did contribute to overlap, had inconsistent effect directions across cluster types. That is, while in those few cases where C1 stiffness was positively associated with overlap for one cluster type, it was negatively correlated for the other

cluster type, contrary to Hypothesis 2 (overlap should increase as C1 rapidity increases). These findings lead to the conclusion that, in a C1C2 cluster, the amount of overlap between the two consonantal gestures is more consistently dependent on the rapidity of C2 rather than that of C1. In other words, C2 stiffness is a more robust predictor of overlap than C1 stiffness.

This finding goes against the prediction implicit in Jun's schemas given in **Figure 1** earlier that both C1 and C2 rapidities should affect overlap (specifically, that overlap should increase as C1 rapidity decreases and as C2 rapidity increases). In fact, this is not an entirely new observation. Gafos et al. (2020), on the basis of data from the Moroccan Arabic consonant clusters [bd, db, dg, gd, br, rb, kr, rk, kl, lk, lb, nk] in word initial, medial, and final positions, reported a similar negative correlation between amplitude-normalized peak velocity of C2 (stiffness in the current study) and overlap (absolute overlap in the current study). **Figure 10** demonstrates that the current data from German, English, and Spanish reveal essentially the same negative correlation between C2 stiffness and C2 onset time relative to C1 release with some quantitative differences across languages.[9] Gafos et al. (2020) describe this relation from the perspective of a perception-production interplay, as "the earlier C2 starts while C1 is active, the more C2 slows down, apparently to ensure C2 attains its target after the release of C1," thus securing the audibility of C1. Crucially, this reasoning not only explains the relation between C2 rapidity and its gestural initiation with respect to C1 constriction release, or in other words overlap between C1 and C2, but also offers an explanation of why C1 rapidity does not affect overlap in a similar way to that seen for C2: if the stiffness of an articulator is adjusted as a function of the temporal lag between its gestural onset and the constriction release of a preceding gesture, then it follows that this relation should always be progressive, meaning that it should only surface for the second segment in a two-segment sequence (C2 in a C1C2 sequence), but never for the first segment (C1 in a C1C2 sequence). Since word-initial C1 is the first segment in a sequence of gestures, no constriction has been achieved in the vocal tract before C1 is articulated with which C1 articulation has to temporally coordinate. Hence C1 can unfold under whatever constraints (speech rate, prosodic specifications, etc.) determine its articulatory realization, but once C1 has been produced, then C2 has to respond to the particulars of the realization of C1. It is this sequential ordering, namely the fact that C1 precedes C2, that leads to their distinct relations with overlap in the C1C2 cluster (C2 but not C1 stiffness is a consistent predictor of overlap).

---

[9] Whence the apparent linearity of the relation in the Spanish subset? In Spanish C1C2 clusters, C2 onset occurs mostly earlier (datapoints to the right of the 0 on the x-axis) than C1 release, while in German and English clusters, C2 onset is distributed more freely to occur both earlier and later (datapoints to the left of the 0 on the x-axis) than C1 release. This more restricted range of datapoints results in squeezing the depicted relation to a flatter slope in the Spanish subset.
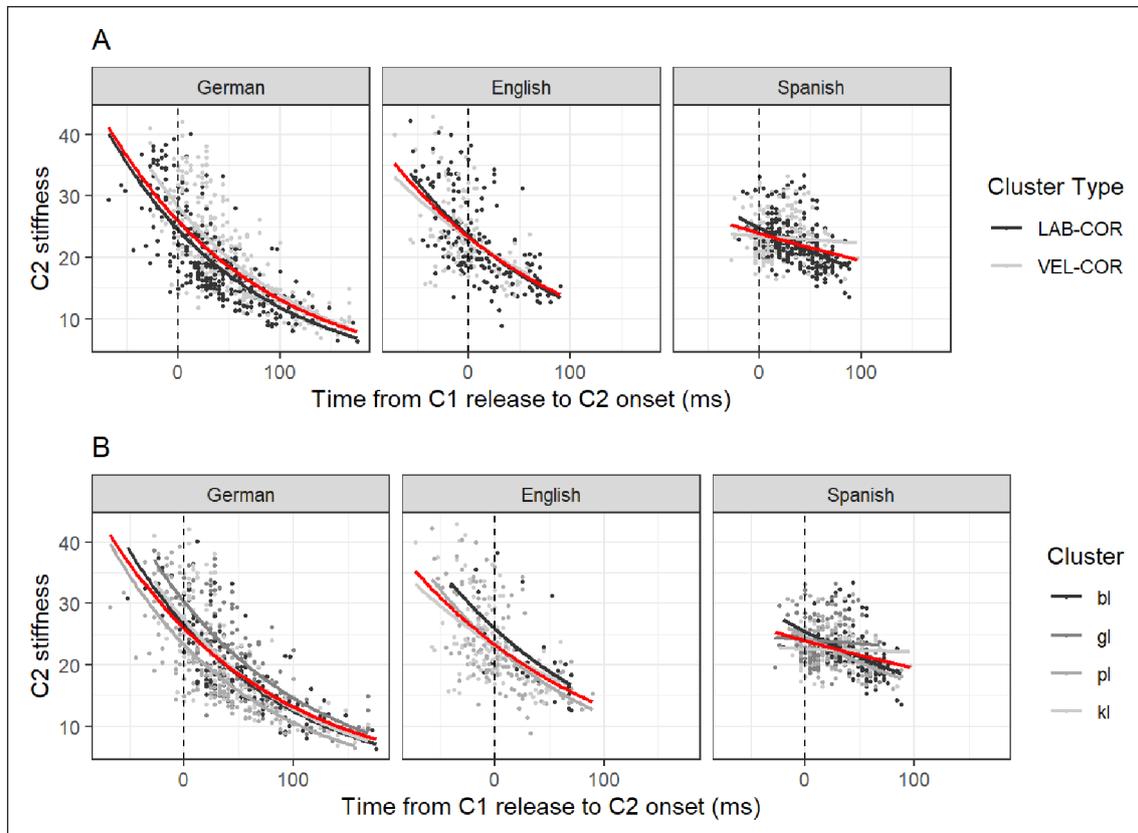
Figure 10: Relationship between C2 stiffness and the temporal interval from C2 onset to C1 release in German, English and Spanish word-initial /Cl/ clusters grouped up by cluster types (A) and clusters (B). Data points to the left versus right of the 0 on the x-axis represent instances of "late" versus "early" C2s, that is, instances where C2 onset occurs after versus before C1 release.

Placing our results in a broader context, let us note that the compensatory nature of the relation between onset of movement of C2 and its stiffness (the greater the overlap, or in other words, the earlier the initiation of C2 with respect to C1, the lower its stiffness; and vice versa) as reported in Gafos et al. (2020) and replicated in our data finds parallels in other human limb movement studies. For example, Bootsma and van Wieringen (1990) observed a similar compensatory relation in an attacking forehand stroke in table tennis. Specifically, they found that "the closer in time the ball was to the player when the movement was initiated, the more force was applied during the stroke." In such a scenario, the distance between the ball and the player is analogous to how early the onset of C2 movement is relative to C1 (earlier C2 onset translates to larger ball-player distance) while the force of the attacking stroke is analogous to stiffness (lower stiffness translates to weaker force). Additional examples can be also drawn from pistol shooting (Aratyunyan, Gurfinkel, & Mirsky, 1969), juggling (Beek, 1989), and timing in other skilled actions. In all these examples, the recurrent common theme is the tight coupling

between action and perception and the inter-dependence of control parameters across individual events in the performance of a skilled spatio-temporal coordination. Let us make this parallel more explicit. In the case of a C1C2 cluster in speech, if C2 starts early within the lifetime of C1, then the stiffness of C2 is lowered so that some lag in time is achieved between the plateaus of the two consonants. **Figure 11** offers a schematic illustration. When C2 starts early within C1 (corresponding to the schema in **Figure 11** with Onset B as the beginning of C2), if C2's stiffness were not lowered, with C2 rising swiftly towards its plateau, this would result in no gap between the plateaus of C1 and C2, thus resulting in a failure to perceptually recover C1's identity. Similarly, in the table tennis context, initiating an attacking stroke with less force when the ball is close to the player would result in failure to achieve ball-racket contact. The ubiquity of this perception-action interaction in these and other examples from skilled action suggests that our finding is not unique to speech production but is a common principle of skilled action in general.
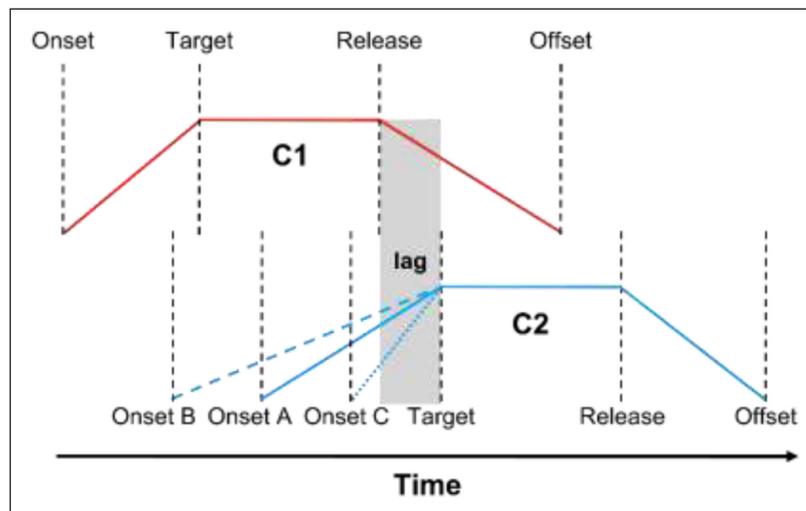
Figure 11: To maintain some lag in time between C1 and C2 so as to ensure C1 perceptibility, the earlier C2 starts within C1 (compare Onset B to Onset A), the lower its stiffness (slower C2 rise to its target). Conversely, the later C2 starts within C1 (compare Onset C to Onset A), the higher its stiffness (faster C2 rise to its target).

5. Conclusion

The present study both confirms and extends earlier findings (Roon et al., 2021) on the factors contributing to overlap in consonant clusters. Specifically, the present study confirms the findings of Roon et al. (2021) in that there seems to be no reliable evidence in support of the notion of articulator-specific inherent velocity and that there is no reliable effect of velocity on overlap in clusters. Instead, a different notion of rapidity as indexed by the stiffness difference between C1 and C2 was found to be a more robust predictor of overlap than the peak velocity difference. The

present study also extends these findings in two ways. First, it employs data from three languages as opposed to data just from Moroccan Arabic (Roon et al., 2021). Second, it examines here for the first time the extent to which C1 and C2 stiffness (individually) contribute to overlap with the main finding being that the significant effect of stiffness difference observed on overlap is mainly contributed by C2 stiffness (C1 stiffness is not as robust as C2 stiffness in predicting overlap). Given the scarcity of cross-language comparisons in overlap with Electromagnetic Articulography data and specifically the dearth of knowledge on what parameters in the dynamics of gestures may modulate that overlap, the present study serves as a first step in evaluating theoretical hypotheses (here, Jun, 2004) across languages and points to a new more promising parameter that can serve as a predictor of overlap (namely, stiffness as opposed to peak velocity, and more specifically, C2 stiffness as opposed to C1 stiffness).

Acknowledgements

Funding Information

Competing Interests

The authors have no competing interests to declare.

Author Contributions

The authors contributed equally to this work.

References

Ackermann, H., Hertrich, I., & Scharf, G. (1995). Kinematic analysis of lower lip movements in ataxic dysarthria. *Journal of Speech, Language, and Hearing Research, 38*(6), 1252–1259. DOI: https://doi.org/10.1044/jshr.3806.1252

Aratyunyan, G. H., Gurfinkel, V. S., & Mirsky, M. L. (1969). Investigation of aiming at a target. *Biophysics, 13*, 536–538.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. DOI: https://doi.org/10.1016/j.jml.2007.12.005

Bates, D. M. (2005). Fitting linear mixed models in R. *R News, 5*, 27–30. DOI: https://doi.org/10.18637/jss.v067.i01

Bates, D. M., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, I., Fox, J. (2020). *lme4: Linear mixed-effects models using 'Eigen' and S4* (Version 1.1-23). Retrieved from https://cran.r-project.org/web/packages/lme4/index.html

Beckman, M. E., & Edwards, J. (1992). Intonational categories and the articulatory control of duration. In Tohkura, Y., Vatikiotis-Bateson, E., Sagisaka, Y. (Eds.), *Speech perception, production, and linguistic structure*. Ohmsha, Tokyo, pp. 359–375.

Beek, P. J. (1989). Timing and phase locking in cascade juggling. *Ecological Psychology, 1*(1), 55–96. DOI: https://doi.org/10.1207/s15326969eco0101_4

Bombien, L., & Hoole, P. (2013). Articulatory overlap as a function of voicing in French and German consonant clusters. *The Journal of the Acoustical Society of America, 134*(1), 539–550. DOI: https://doi.org/10.1121/1.4807510

Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 21–29. DOI: https://doi.org/10.1037/0096-1523.16.1.21

Browman, C. P., & Goldstein, L. (1984). Dynamical modeling of phonetic structure. *Haskins Laboratories Status Report in Speech Research SR-79/80*, 1–18.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219–252. JSTOR. DOI: https://doi.org/10.1017/S0952675700000658

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*(2), 201–251. DOI: https://doi.org/10.1017/S0952675700001019

Browman, C. P., & Goldstein, L. (1990a). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology* (1st ed., pp. 341–376). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511627736.019

Browman, C. P., & Goldstein, L. (1990b). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics, 18*, 299–320. DOI: https://doi.org/10.1016/S0095-4470(19)30376-6

Brown, G. (1977). *Listening to spoken English* (2 ed.). Routledge. DOI: https://doi.org/10.4324/9781315538518

Brunner, J., Geng, C., Sotiropoulou, S., & Gafos, A. (2014). Timing of German onset and word boundary clusters. *Laboratory Phonology, 5*(4). DOI: https://doi.org/10.1515/lp-2014-0014

Byrd, D. (1994). *Articulatory timing in English consonant sequences* [PhD dissertation]. UCLA.

Byrd, D. (1996) Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24*, 209–244. DOI: https://doi.org/10.1006/jpho.1996.0012

Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics – the interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Phonology and Phonetics [PP]*. DE GRUYTER MOUTON. DOI: https://doi.org/10.1515/9783110224917.1.31

Byrd, D., & Saltzman, E. L. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics, 26*, 173–199. DOI: https://doi.org/10.1006/jpho.1998.0071

Chen, L. H. (2003). Evidence for the role of gestural overlap in consonant place assimilation. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *[ICPhS-15] 15th International Congress of Phonetic Sciences, Barcelona, Spain, August 3–9, 2003*. http://www.internationalphoneticassociation.org/icphs/icphs2003

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven, T. Rietveld, & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 419–447). Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197105.419

Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface: Prosodic boundary strengthening. *Language and Linguistics Compass*, *10*(3), 120–141. DOI: https://doi.org/10.1111/lnc3.12178

Cho, T., Lee, Y., & Kim, S. (2014). Prosodic strengthening on the /s/-stop cluster and the phonetic implementation of an allophonic rule in English. *Journal of Phonetics*, *46*, 128–146. DOI: https://doi.org/10.1016/j.wocn.2014.06.003

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *The Journal of the Acoustical Society of America, 89*, 369–382. DOI: https://doi.org/10.1121/1.400674

Fowler, C. A., Rubin, P., Remez, R. E., Turvey, M. T. (1980). Implications for speech production of a general theory of action, in: Butterworth, B. (Ed.), *Language production.* volume 1, pp. 373–420.

Gafos, Adamantios. (1996). *The articulatory basis of locality in phonology* [PhD dissertation]. Johns Hopkins University. [Published 1999, Outstanding Dissertations in Linguistics, Routledge Publishers.]

Gafos, A. I. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory*, *20*, 269–337. DOI: https://doi.org/10.1023/A:1014942312445

Gafos, A. I., Hoole, P., Roon, K., & Zeroual, C. (2010). Variation in overlap and phonological grammar in Moroccan Arabic clusters. In C. Fougeron, B. Kühnert, M. D'Imperio, & Nathalie Vallée (Eds.), *Laboratory Phonology* (Vol. 10, pp. 657–698). DE GRUYTER MOUTON. DOI: https://doi.org/10.1515/9783110224917.5.657

Gafos, A. I., Roeser, J., Sotiropoulou, S., Hoole, P., & Zeroual, C. (2020). Structure in mind, structure in vocal tract. *Natural Language & Linguistic Theory*, *38*(1), 43–75. DOI: https://doi.org/10.1007/s11049-019-09445-y

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge/New York: Cambridge University Press. DOI: https://doi.org/10.1007/CBO9780511790942

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).

Gibson, M., Sotiropoulou, S., Tobin, S., & Gafos, A. I. (2019). Temporal aspects of word initial single consonants and consonants in clusters in Spanish. *Phonetica*, *76*(6), 448–478. DOI: https://doi.org/10.1159/000501508

Gimson, A. C. (1962). *An introduction to the pronunciation of English.* London: Edward Arnold Publishers, Ltd.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*(3), 594–621. DOI: https://doi.org/10.1037/0033-295X.102.3.594

Hardcastle, W. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication, 4*, 247–263. DOI: https://doi.org/10.1016/0167-6393(85)90051-2

Hermes, A., Mücke, D., & Auris, B. (2017). The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics*, *64*, 127–144. DOI: https://doi.org/10.1016/j.wocn.2017.05.004

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

Hoole, P., Bombien, L., Kühnert, B., & Mooshammer, C. (2009). Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In G. Fant, H. Fujisaki, & J. Shen (Eds.), *Frontiers in phonetics and speech science* (pp. 275–286). The Commercial Press.

Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach* [PhD dissertation]. University of California, Los Angeles.

Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically Based Phonology* (pp. 58–86). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486401.003

Kelso, J. A. S., Vatikiotis-Bateson, E., Saltzman, E. L., & Kay, B. (1985). A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling. *The Journal of the Acoustical Society of America*, *77*(1), 266–280. DOI: https://doi.org/10.1121/1.392268

Kent, R. D., & Moll, K. L. (1972). Cinefluorographic analyses of selected lingual consonants. *Journal of Speech and Hearing Research, 15*(3), 453–473. DOI: https://doi.org/10.1044/jshr.1503.453

Kochetov, A., & So, C. K. (2007). Place assimilation and phonetic grounding: A cross-linguistic perceptual study. *Phonology, 24*(3), 397–432. DOI: https://doi.org/10.1017/S0952675707001273

Kozhevnikov, V. A., & Chistovich, L. A. (1965). *Speech: Articulation and perception*. Nauka.

Kuberski, S. R., & Gafos, A. I. (2019). The speed-curvature power law in tongue movements of repetitive speech. *PLOS ONE, 14*(3), e0213851. DOI: https://doi.org/10.1371/journal.pone.0213851

Kuehn, D. P., & Moll, K. L. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics, 4*(4), 303–320. DOI: https://doi.org/10.1016/S0095-4470(19)31257-4

Kühnert, B., & Hoole, P. (2004). Speaker-specific kinematic properties of alveolar reductions in English and German. *Clinical Linguistics & Phonetics*, *18*(6–8), 559–575. DOI: https://doi.org/10.1080/02699200420002268853

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, OX, UK: Blackwell Publishers.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *R package 'emmeans': Estimated marginal means, aka least-squares means*. Retrieved from https://github.com/rvlenth/emmeans

Lialiou, M., Sotiropoulou, S., & Gafos, A. I. (2021). *Spatiotemporal coordination in word-medial stop-lateral and s-stop clusters of American English*. Phonetica, *78*(5–6), 385–433. DOI: https://doi.org/10.1515/phon-2021-2010

Luo, S. (2017). Gestural overlap across word boundaries: Evidence from English and Mandarin speakers. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique, 62*(1), 56–83. DOI: https://doi.org/10.1017/cnj.2016.37

Marin, S. (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics, 41*(3–4), 211–227. DOI: https://doi.org/10.1016/j.wocn.2013.02.001

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control, 14*(3), 380–407. DOI: https://doi.org/10.1123/mcj.14.3.380

Mücke, D., Hermes, A., & Tilsen, S. (2020). Incongruencies between phonological theory and phonetic measurement. *Phonology, 37*(1), 133–170. DOI: https://doi.org/10.1017/S0952675720000068

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance, 11*(4), 457–474. DOI: https://doi.org/10.1037/0096-1523.11.4.457

Ohala, J. (1970). WPP, No. 15: Aspects of the Control and Production of Speech. *UCLA: Department of Linguistics*. https://escholarship.org/uc/item/1859f9tk

Ostry, D. J., Cooke, J. D., & Munhall, K. G. (1987). Kinematic form of limb and speech movements. In G. N. Gantchev, B. Dimitrov, & P. Gatev (Eds.), *Motor Control* (pp. 209–213). Springer US. DOI: https://doi.org/10.1007/978-1-4615-7508-5_35

Ostry, D. J., Keller, E., & Parush, A. (1983). Similarities in the control of the speech articulators and the limbs: Kinematics of tongue dorsum movement in speech. *Journal of Experimental Psychology: Human Perception and Performance, 9*(4), 622–636. DOI: https://doi.org/10.1037/0096-1523.9.4.622

Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America, 77*(2), 640–648. DOI: https://doi.org/10.1121/1.391882

Parush, A., Ostry, D. J., & Munhall, K. G. (1983). A kinematic study of lingual coarticulation in VCV sequences. *The Journal of the Acoustical Society of America, 74*(4), 1115–1125. DOI: https://doi.org/10.1121/1.390035

Pastätter, M., & Pouplier, M. (2013). Temporal coordination of sibilants in Polish onset clusters. *The Journal of the Acoustical Society of America, 134*(5), 4201–4201. DOI: https://doi.org/10.1121/1.4831417

Pastätter, M., & Pouplier, M. (2017). Articulatory mechanisms underlying onset-vowel organization. *Journal of Phonetics, 65*, 1–14. DOI: https://doi.org/10.1016/j.wocn.2017.03.005

Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America, 112*(4), 1627–1641. DOI: https://doi.org/10.1121/1.1506369

Pouplier, M. (2020). Articulatory phonology. In M. Pouplier, *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.745

Pouplier, M., Pastätter, M., Hoole, P., Marin, S., Chitoran, I., Lentz, T. O., & Kochetov, A. (2022). Language and cluster-specific effects in the timing of onset consonant sequences in seven languages. *Journal of Phonetics, 93*, 101153. DOI: https://doi.org/10.1016/j.wocn.2022.101153

R Development Core Team. (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Roon, K. D., Hoole, P., Zeroual, C., Du, S., & Gafos, A. I. (2021). Stiffness and articulatory overlap in Moroccan Arabic consonant clusters. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 12*(1), 8. DOI: https://doi.org/10.5334/labphon.272

Rubin, P., Baer, T., & Mermelstein, P. (1981). An articulatory synthesizer for perceptual research. *The Journal of the Acoustical Society of America, 70*(2), 321–328. DOI: https://doi.org/10.1121/1.386780

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382. DOI: https://doi.org/10.1207/s15326969eco0104_2

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician, 34*(4), 216–221. DOI: https://doi.org/10.1080/00031305.1980.10483031

Shaw, J., Gafos, A. I., Hoole, P., & Zeroual, C. (2009). Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability in articulation. *Phonology*, *26*(1), 187–215. DOI: https://doi.org/10.1017/S0952675709001754

Sorensen, T., & Gafos, A. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology, 28*(4), 188–215. DOI: https://doi.org/10.1080/10407413.2016.1230368

Sorensen, T., & Gafos, A. (2022). The relation between gestures and kinematics. In F. Breit, B. Botma, M. van't Veer, & M. van Oostendorp (Eds.), *Primitives of Phonological Structure* (Ch. 10). Oxford University Press.

Sotiropoulou, S., Gibson, M., & Gafos, A. (2020). Global organization in Spanish onsets. *Journal of Phonetics*, *82*, 100995. DOI: https://doi.org/10.1016/j.wocn.2020.100995

Surprenant, A. M., & Goldstein, L. (1998). The perception of speech gestures. *The Journal of the Acoustical Society of America*, *104*(1), 518–529. DOI: https://doi.org/10.1121/1.423253

Tilsen, S., Zec, D., Bjorndahl, C., Butler, B., L'Experance, M.-J., Fisher, A., Heimisdottir, L., Renwick, M., & Sanker, C. (2012). *A cross-linguistic investigation of articulatory coordination in word-initial consonant clusters*. DOI: https://doi.org/10.5281/ZENODO.3726937

Winitz, H., Scheib, M. E., & Reeds, J. A. (1972). Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *The Journal of the Acoustical Society of America, 51*(4B), 1309–1317. DOI: https://doi.org/10.1121/1.1912976

Winters, S. J. (2001). VCCV perception: Putting place in its place. *Ohio State Working Papers in Linguistics, 55*, 70–87.

Winters, S. J. (2003). *Empirical investigations into the perceptual and articulatory origins of cross-linguistic asymmetries in place assimilation* [PhD dissertation]. The Ohio State University. https://linguistics.osu.edu/sites/linguistics.osu.edu/files/dissertations/winters2003.pdf

Yanagawa, M. (2006). *Articulatory Timing in first and second language: A cross-linguistic study* [PhD Dissertation]. Yale University.