

JOURNAL ARTICLE

# Stiffness and articulatory overlap in Moroccan Arabic consonant clusters

Kevin D. Roon<sup>1,2</sup>, Philip Hoole<sup>3</sup>, Chakir Zeroual<sup>4</sup>, Shihao Du<sup>5</sup> and Adamantios I. Gafos<sup>2,5</sup>

<sup>1</sup> Program in Speech-Language-Hearing Sciences, CUNY Graduate Center, New York, US

<sup>2</sup> Haskins Laboratories, New Haven, Connecticut, US

<sup>3</sup> Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München, München, DE

<sup>4</sup> Université Mohammed Premier, Oujda, MA

<sup>5</sup> Department Linguistik und Kognitionswissenschaften, Universität Potsdam, Potsdam, DE

Corresponding author: Kevin D. Roon ([kroon@gc.cuny.edu](mailto:kroon@gc.cuny.edu))

It has been claimed that patterns of regressive place assimilation in consonant clusters are attributable to the ‘inherent velocities’ of the primary oral articulators involved. The present study used articulatory data from Moroccan Arabic to evaluate whether there were reliable differences in peak velocity or measured stiffness based on primary oral articulator. Evidence for such differences was limited. We also evaluated whether the degree of overlap in clusters could be predicted by the combination of articulators in a C1C2 cluster. Some significant differences were found, but they did not align with the assumed differences in inherent rapidity of the articulators. Lastly, we evaluated whether the amount of overlap was influenced by the relative peak velocities of the closing gestures of two-consonant clusters. It was not. However, when the dynamical control parameter of stiffness of the same closing gestures was used instead of peak velocity, a robust effect on overlap was found. This result suggests that the relative settings of the control parameter of stiffness influence degree of overlap in a cluster, not peak velocity, which is too subject to contextual effects such as movement amplitude.

**Keywords:** Stiffness; cluster overlap; place assimilation; Moroccan Arabic; speech production

## 1. Introduction

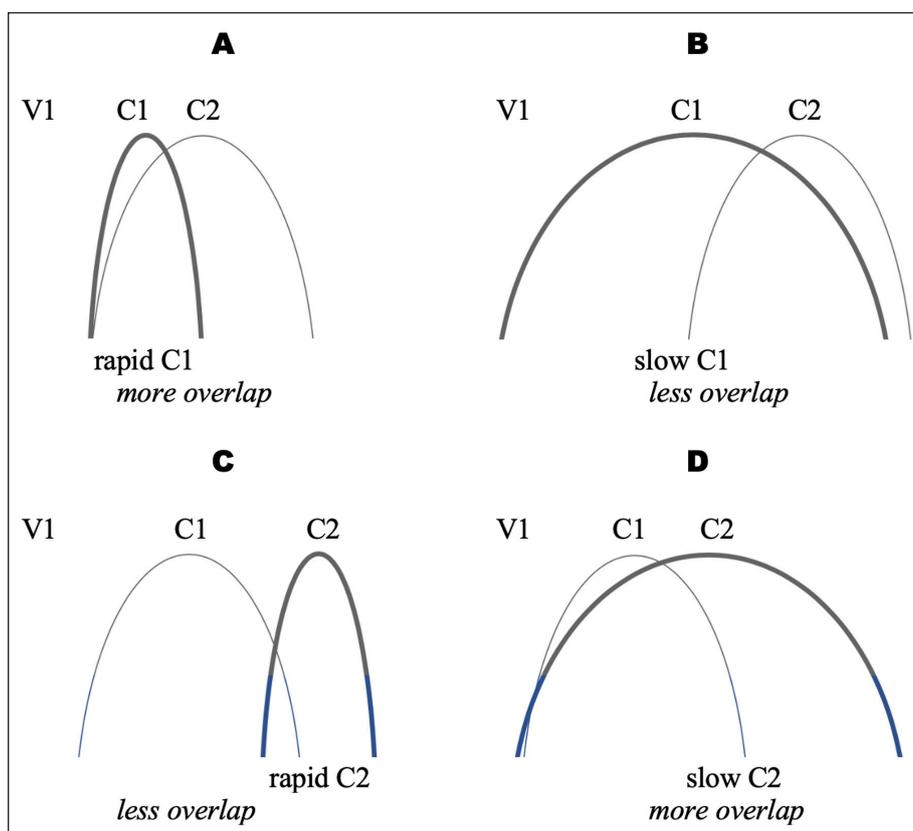
In many languages, two consonants in a cluster will agree in place, with the first consonant changing to match the place of the second (see, e.g., Jun, 2004, and references therein). It has been claimed (Jun, 1995, 2004; Monahan, 1993) that certain places are more likely to be triggers of place assimilation than others. Jun (2004) discusses the fact that coronals are the least likely to be triggers of place assimilation and velars are the most likely, providing an example from Korean:

- (1) Korean place assimilation (data originally from Jun, 1996)
  - a. /ip + ko/ → /[ikko] ‘wear and’
  - b. /ip + tolok/ → [iptolok], \*[ittolok] ‘wear + causative marker’

Jun (2004) proposes that, given a cluster containing two consonants VC1C2, C1 is more likely to assimilate to the place of C2 when there is more overlap between the articulations of the consonants than when there is less. The proposal by Jun (2004) is a phonetically-based approach to phonological assimilation, motivated by the effects of overlap on

perception. The proposal has two components. The first component is that overlap is important in explaining patterns of assimilation because contemporaneous articulator movements of C2 and C1 potentially negatively influence the perceptibility of acoustic information about the place of C1 (specifically, formant transitions) during the transition from the vowel to C1. This component of the proposal is illustrated in **Figure 1** with two examples in post-vocalic VC1C2 clusters. First, given two clusters with C2 having the same rapidity, the cluster with a more rapid C1 gesture (**Figure 1A**) will show more overlap of C1 by C2 than the cluster with a slower C1 (**Figure 1B**: Jun, 2004, p. 63). Second, given two clusters with C1 having the same rapidity, the cluster with a slower C2 gesture (**Figure 1D**) will show more overlap of C1 by C2 than the cluster with a more rapid C2 (**Figure 1C**: Jun, 2004, p. 65). These four cases can be summarized by saying that when C1 is more rapid compared to C2, there is more overlap, and when C2 is more rapid compared to C1, there is less overlap.

The second component of the proposal in Jun (2004) states that the amount of overlap between two consonants in a cluster depends on the ‘inherent velocities’ of the particular oral articulators involved (Jun, 2004, p. 65). Specifically, the proposal assumes that the tongue tip (TT) has the fastest inherent velocity, followed by the lower lip (LL), and then the tongue body (TB). According to Jun (2004), coronal C1s are typologically more likely to be targets of assimilation because the context in which they are found corresponds to the case in **Figure 1A** (as opposed to **Figure 1B**). Conversely, coronal C2s are typologically less likely to be triggers of assimilation because the context in which they are found corresponds to the case in **Figure 1C** (as opposed to **Figure 1D**).



**Figure 1:** The predictions of Jun (2004, numbered examples 5 and 6) for effects of relative rapidity on consonantal overlap in VC1C2 clusters. Top panel: When rapidity of C2 is held constant, a rapid C1 results in more overlap (A) than a slow C1 (B). Bottom panel: When rapidity of C1 is held constant, a slow C2 results in more overlap (D) than a rapid C2 (C).

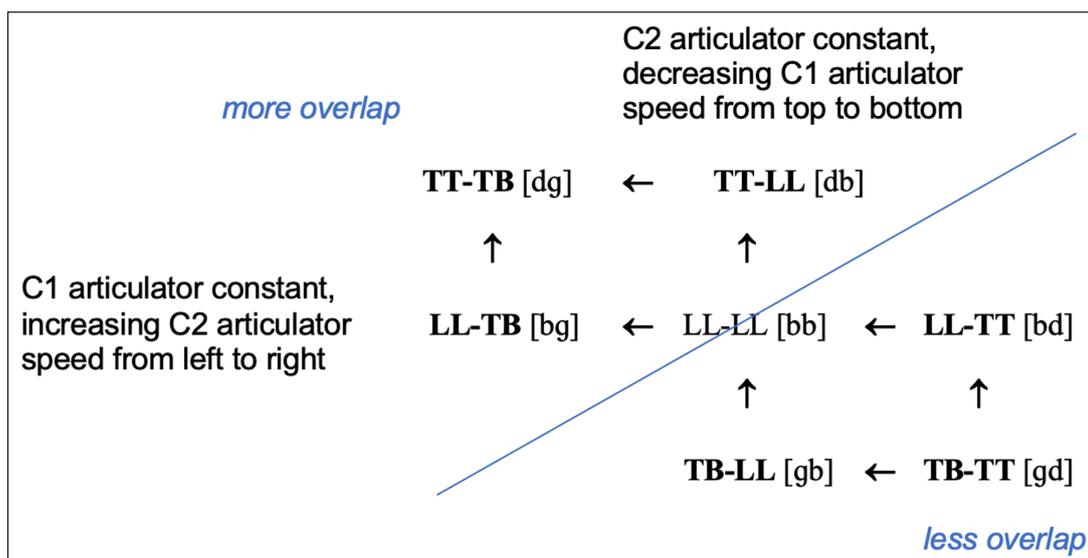
These two components of the proposal made by Jun (2004) have not been tested experimentally. The first component presents a clear testable hypothesis:

*H-J1:* Articulations of tongue tip consonants should be faster than articulations of lower lip and tongue body consonants, and articulations of lower lip consonants should be faster than those of tongue body consonants, all things being equal.

Assuming H-J1 is supported, then the amount of overlap in a heterorganic cluster should be determinable by the combination of articulators in a cluster (cluster type). The expected overlap by cluster type is illustrated in **Figure 2**.

The middle column of **Figure 2** shows the expected overlap when C2 is always a lower lip consonant (corresponding to **Figure 1A–B**). More overlap is expected when C1 is a tongue tip consonant than when C1 is also a lower lip consonant, since the tongue tip is putatively more rapid than the lower lip. More overlap is expected when C1 is a lower lip consonant than when C1 is a tongue body consonant, since the lower lip is putatively more rapid than the tongue body. The middle row of **Figure 2** shows the expected overlap when C1 is always a lower lip consonant (corresponding to **Figure 1C–D**). More overlap is expected when C2 is a tongue body consonant than when C2 is also a lower lip consonant, since the lower lip is putatively more rapid than the tongue tip. More overlap is expected when C2 is a lower lip consonant than when C1 is a tongue tip consonant, since the lower lip is putatively more rapid than the tongue tip. By the same reasoning, tongue tip-tongue body clusters should have more overlap than lower lip-tongue body clusters and tongue tip-lower lip clusters. Tongue body-tongue tip clusters should have less overlap than tongue body-lower lip clusters and lower lip-tongue tip clusters. No prediction is made about the overlap between lower lip-tongue body clusters and tongue tip-lower lip clusters, or between tongue body-lower lip clusters and lower lip-tongue tip clusters. Jun's proposal therefore is tantamount to the following hypothesis.

*H-J2:* The amount of overlap in a heterorganic cluster is predicted to decrease by cluster type in the following way: TT-TB > {LL-TB, TT-LL} > {TB-LL, LL-TT} > TB-TT, where no prediction is made between pairs within curly brackets.



**Figure 2:** Expected overlap by cluster type based on the inherent velocities of the constituent articulators, per Jun (2004).

It is possible that H-J1 is supported empirically while H-J2 is not. This would undermine Jun's proposal that the typological patterns of place assimilation are attributable to inherent properties of the articulators involved in producing a cluster. If empirical evidence does not support H-J1, then H-J2 is likely moot, since H-J2 is derived based on the assumption that H-J1 is true. However, in this case, that would mean rejecting the insight from Jun that the relative rapidity of the articulations involved in producing a cluster has a significant influence on overlap because of the assumption that the rapidity of the articulations is solely a function of the articulators involved. It is possible to propose an alternative to H-J2 that maintains that the amount of overlap in a cluster will be a function of the relative rapidity of the articulations involved, without any reference to articulator.

*H-3:* The amount of overlap in a cluster increases as the rapidity of C1 increases compared to C2.

The goal of the present study was to test these hypotheses, since no work has been done to test experimentally the proposals by Jun (2004). In order to test these hypotheses, however, other known influences on cluster overlap needed to be taken into consideration, several issues of quantification needed to be addressed, and a suitable language for testing these predictions had to be identified. We address each of these topics below.

### 1.1. Test language: Moroccan Arabic

The present study used articulatory data from speakers of Moroccan Arabic (MA) gathered using 3D electromagnetic articulography ('EMA'; Hoole, Zierdt, & Geng, 2003).

MA is a highly appropriate language for studying timing in consonant clusters since it permits a complex set of consonant combinations. Looking only at two-consonant stop-stop sequences; labial, coronal, and dorsal places of articulation are attested in all six combinations word-initially, word-medially, and word-finally except for word-final coronal-labial combinations, as shown in **Table 1**. Several manners of coronal consonants—liquids, a rhotic tap, stops, and fricatives—are also widely found as members of two-consonant clusters. The freedom of clustering permitted in MA therefore offers a particularly apt case study for testing hypotheses regarding overlap, since it is possible within-speaker to independently control for and test the effects of several different factors affecting overlap.

### 1.2. Quantification

There are two aspects of articulation that are crucial to the proposal made by Jun (2004): the 'rapidity' of an articulation (which is a function of the 'inherent velocity' of the primary articulator), and overlap. However, no quantification of either of these terms

**Table 1:** Six place combinations in representative stop-stop clusters in Moroccan Arabic.

	Word initial	Word medial	Word final
labial + coronal	btasim	nabta	ssibt
coronal + labial	tbiʃ	ratba	kɪdb
labial + dorsal	bkat	sabga	–
dorsal + labial	kbaʃ	rakba	–
coronal + dorsal	dgig	hadga	fūdg
dorsal + coronal	gdat	ʃagda	ʃūgd

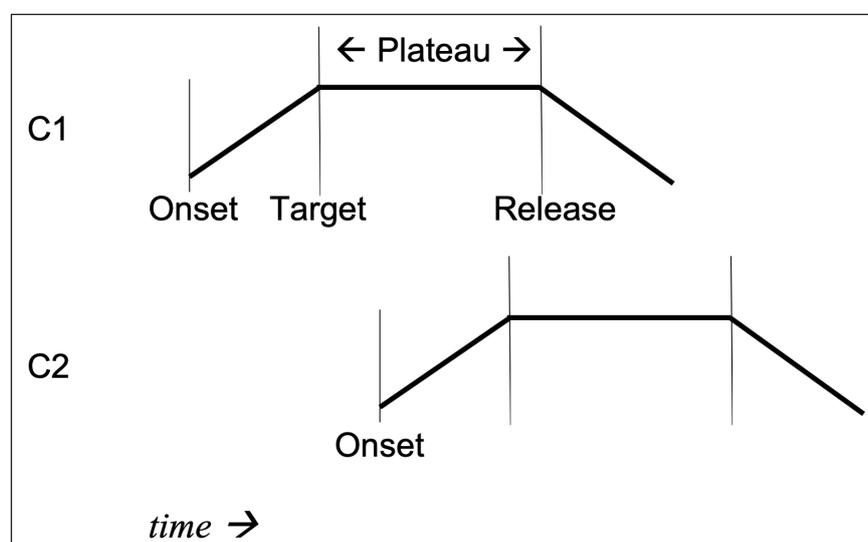
is made explicit by Jun (2004), and there are multiple ways to quantify both terms. We motivate two different quantifications for each term that we used to refine and test the present hypotheses.

In order to discuss these quantifications, we first discuss how articulator movements themselves were identified and quantified. The articulator movement measured for each consonant was indexed by the EMA receiver corresponding to the consonant's primary oral articulator—lower lip: [b]; tongue tip: [d, t]; tongue body: [g, k]. Following Gafos (2002, p. 296), we assumed that the oral gesture of a consonant drives the timing relations among consonants.

Each gesture can be characterized by certain temporal landmarks, shown in **Figure 3**. The Onset of movement is the point in time where the articulator starts moving toward its required constriction, Target is the point in time when that constriction is achieved, and Release is the point in time when the articulator moves away from the constriction. The Plateau of the articulation is the interval between Target and Release when the articulator is maintaining the required constriction. These landmarks were calculated using MATLAB-based code for analyzing EMA data ('MVIEW'), developed at Haskins Laboratories by Mark Tiede. Onsets were identified algorithmically as the point at which the EMA sensor exceeded 20% of the maximum tangential velocity of the articulator during the closing phase, Targets were identified as the point at which the EMA sensor subsequently went below 20% of the same maximum tangential velocity, and Releases were identified as the point at which the EMA sensor exceeded 20% of the following maximum tangential velocity (comparable to the methods used by, e.g., Chitoran, Goldstein, & Byrd, 2002; Gafos, Hoole, Roon, & Zeroual, 2010; Oliveira & Teixeira, 2007, and many others). Further technical details about the acquisition and processing of the EMA data are included in the Methods section below.

### 1.2.1. Rapidity

Any targeted movement can be characterized by three basic kinematic properties: movement duration  $T$ , movement amplitude  $A$ , and peak velocity  $\hat{v}$  (Nelson, 1983). One way to calculate the rapidity of a given articulatory movement is to use the peak velocity  $\hat{v}$  of a fleshpoint on the relevant articulator as it moves to effect the constriction associated with the consonant. This measure, when applied to appropriate datasets that bring out the comparisons shown in **Figure 1**, would seem to offer a most direct test of



**Figure 3:** Articulatory landmarks used for calculating stiffness, peak velocity, and overlap.

the claims on overlap of Jun (2004). We therefore evaluated H-J1 using the peak velocity of the closing movement associated with each consonant as the index of the articulation's rapidity ( $H-J1_{pv}$ ).

However, the use of peak velocity  $\hat{v}$  as an index of articulator rapidity is potentially problematic given two empirically well-documented relations between  $\hat{v}$  and the other two kinematic properties of movement duration  $T$  and movement amplitude  $A$ . The first relation is that a movement's peak velocity  $\hat{v}$  covaries with its amplitude  $A$  (Kent & Moll, 1976; Kuehn & Moll, 1976). This relation has been reported for a large variety of consonant-vowel and vowel-consonant sequences involving movements of tongue body, tongue tip, lips, and jaw (Ostry, Keller, & Parush, 1983) and has been described as an overall linear correlation (Kuberski & Gafos, 2019; Ostry & Munhall, 1985), with  $A$ - $\hat{v}$  slopes steeper for faster than for slower speech rates and with decreasing strength of covariation as  $A$  increases (Vatikiotis-Bateson & Kelso, 1990, 1993). This relationship between  $A$  and  $\hat{v}$  is also predicted by the DIVA model of Guenther (1995). The other relation ties together all three kinematic properties: The ratio of peak velocity to amplitude,  $\hat{v}/A$ , varies inversely with movement duration  $T$  (Kuberski & Gafos, 2019; Munhall, Ostry, & Parush, 1985; see also Sorensen & Gafos, 2016, for modeling) across manipulations of stress, vowel, and consonant identity (Fuchs, Perrier, & Hartinger, 2011; Ostry & Munhall, 1985). This ratio, which we denote as  $k'$  (see below for further discussion), has been used as an index of articulatory stiffness (Kelso, Vatikiotis-Bateson, Saltzman, & Kay, 1985; Munhall et al., 1985; and others), a control parameter proposed in the motor-control literature that modulates the time-space behavior of an articulator in various ways (Cooke, 1980).

Stiffness is incorporated formally in both Articulatory Phonology (AP, Browman & Goldstein, 1986, et seq.) and the Task Dynamic Model of speech production (TDM, Browman & Goldstein, 1990; Munhall et al., 1985; Saltzman & Munhall, 1989), and is therefore relevant to theories of phonological representation and its phonetic implementation. Various researchers have investigated how stiffness may express different prosodic phenomena, including final lengthening (Edwards, Beckman, & Fletcher, 1991), timing across prosodic boundaries (Byrd & Saltzman, 1998), and intonation (Beckman & Edwards, 1992). An explanation of the task-dynamical system used in AP and the specification of gestures within the system is provided in Saltzman and Munhall (1989) and Browman and Goldstein (1990). Perhaps most relevant to the issues addressed in the present study is that Browman and Goldstein (1989, p. 229) raised the possibility that different articulators may have different inherent stiffness.

The TDM provides an invariant means of describing a dynamical system, parameterized by the variables of rest position ( $x_0$ ), stiffness ( $k$ ), and damping ( $b$ ). In this model, velocity is instantaneous velocity ( $\dot{x}$ , which, along with acceleration and position, are obtained by the TDM from sensory-motor feedback in order to compute what articulator movements are necessary to achieve the task) that the TDM factors into calculating what articulations are required, but it is not manipulated directly in the system. Damping for non-laryngeal gestures in the TDM is set to a constant critical value. The variable for stiffness  $k$ , on the other hand, can be manipulated in the model. Manipulating stiffness with everything else held constant has the effect that gestures with higher stiffness values result in "shorter duration movements and greater peak velocity/movement amplitude ratios" (Munhall et al., 1985). Measured stiffness was therefore calculated for the closing movement of each of the articulations in each consonant as in (2):

$$(2) \quad \textit{Measured stiffness: } k' = \hat{v}/A$$

We also evaluated H-J1 using the measured stiffness of the closing movement associated with each consonant as another index of the articulation's rapidity. Given these two quantifications of rapidity, H-3 can be refined in two different formulations:

H-3<sub>pv</sub>: The amount of overlap in a cluster increases as the peak velocity of C1 increases compared to the peak velocity of C2.

and

H-3<sub>stiffness</sub>: The amount of overlap in a cluster increases as the measured stiffness of C1 increases compared to the measured stiffness of C2.

### 1.2.2. Overlap

Two indices of overlap were calculated, based on the landmarks in articulator trajectories shown in **Figure 3**. One measure of overlap used in the present study ('Relative Overlap') was calculated as the difference between the time points of C2 Onset and C1 Target, normalized by the duration of the C1 Plateau, as shown in **Figure 3** and defined in (3).

$$(3) \quad \text{Relative overlap} = 1 - ((\text{time}_{\text{OnsetC2}} - \text{time}_{\text{TargetC1}}) / (\text{time}_{\text{ReleaseC1}} - \text{time}_{\text{TargetC1}}))$$

This index of overlap was chosen for the same reasons outlined in Jun (2004) and Chitoran et al. (2002): It provides a reasonable means of measuring the potential effect of C2 movement on the acoustic outcomes of C1. This index is the same as the relative measure used by Gafos et al. (2010) and comparable to the one used by Chitoran et al. (2002) in that it quantifies overlap as the proportion of the C1 Plateau that is coextensive with the C2 gesture, as a proportion of the C1 Plateau.

It is possible that the velocity and/or stiffness of an articulator movement may affect the length of its Plateau. Since the Relative Overlap measure described here has Plateau of C1 as the denominator, any effects of velocity and/or stiffness on overlap using this measure may reflect differences in C1 Plateau duration rather than differences in overlap (all other things being equal, the faster C1 gestures will yield higher overlap measures). To address this, the simple lag time between the onsets of both gestures in milliseconds ('Onset Lag') was also used as a second measure of overlap, with the onset time of C2 subtracted from the onset time of C1 so that greater lag values indicate more overlap (or, putting this another way, given that values are typically negative, the more negative the values, the less overlap), calculated as in (4).

$$(4) \quad \text{Onset Lag} = \text{time}_{\text{OnsetC1}} - \text{time}_{\text{OnsetC2}}$$

We evaluated hypotheses H-J2, H-3<sub>pv</sub>, and H-3<sub>stiffness</sub> with overlap first indexed as Relative Overlap (3) and then indexed as Onset Lag (4), since there is no agreed-upon norm for indexing overlap, and Jun (2004) does not make this explicit.

### 1.3. Influences on overlap

Since the primary goal of the present study was to investigate the potential influence of relative rapidity of articulation on overlap, other known influences on overlap in clusters had to be taken into account. We summarize below some known influences on overlap in general, as well as on some specific findings from Gafos et al. (2010), who reported on

cluster overlap from a different subset of the data from two of the speakers analyzed in the present study.

Word-initial clusters are typically less overlapped than word-medial, intervocalic clusters. This effect has been observed for English (Byrd, 1996; Hardcastle, 1985), Tsou (Wright, 1996), and Georgian (Chitoran et al., 2002). There are two perception-based motivations for this effect given by Chitoran et al. (2002). First, word-initial clusters are less overlapped because they may also be utterance-initial and therefore lack a preceding vowel to provide cues to help in the identification of the first consonant in the cluster (see Redford & Diehl, 1999, for more discussion). Second, lexical access relies heavily on word-initial phonetic detail (Marslen-Wilson, 1987). Gafos et al. (2010) compared overlap in stop-stop clusters for two MA speakers. Word position effects were found for Speaker 1 in that word-initial clusters were significantly less overlapped than word-medial and word-final clusters. The overlap in word-medial clusters was not significantly different from word-final clusters. Speaker 2 showed no word position effects.

Previous studies (Byrd, 1992, 1996; Chitoran et al., 2002; Hardcastle & Roach, 1979; Son, 2008; Surprenant & Goldstein, 1998; Zsiga, 1994) have shown that front-to-back clusters (that is, clusters where the place of articulation of C1 is anterior to that of C2, e.g., [tk], [pt]) tend to show more overlap on average than back-to-front clusters (e.g., [kt], [tp]) than in front-to-back clusters. The proposed explanation for this effect is perceptually based: Starting the C2 closure of a back-to-front cluster too soon will hide the acoustic information needed for the listener to identify C1, whereas this is not true for front-to-back clusters. Results from other studies (Chitoran & Goldstein, 2006; Kühnert, Hoole, & Mooshammer, 2006; Zeroual, Hoole, Gafos, & Esling, 2014) have called into question the generality and/or the perceptual motivation of the place order effect. To illustrate, Gafos et al. (2010) analyzed only stop-stop clusters within each word position and speaker, and found that Speaker 2 showed significant place order effects consistent with findings in other studies: Front-to-back stop-stop clusters were significantly more overlapped than back-to-front clusters word-initially and word-medially. There was no significant difference word-finally. The overlap patterns for Speaker 1 were significantly different both from what had been reported in other studies and from Speaker 2. In particular, contrary to the expected place order effect, Speaker 1 had significantly more overlap for back-to-front clusters, both as a main effect (across all word positions) and word-initially. Gafos et al. (2010) concluded that this was related to the fact that Speaker 1 had less cluster overlap than Speaker 2. Such results along with questions about the validity of an unqualified or universal place order effect in previous studies across languages and speakers led them to propose that place order effects are in fact relativized: If there are two environments (say, two different speakers or two different phonetic contexts) in which place order effects may arise and they arise in only one, the place order effect will be found in the environment where there is more overlap. The reasoning behind the hypothesis is that it is only in the more overlapped environment where recoverability is at stake and thus the effect should emerge; in a less overlapped environment, recoverability of acoustic information is less threatened and therefore the place order effects on overlap need not come into play.

Lastly, the amount of overlap in clusters may also be to some degree idiosyncratic to individual speakers. Gafos et al. (2010) showed that there was a significant difference between speakers in the amount of overlap in stop-stop clusters, with Speaker 1 having less overlapped stop-stop clusters than Speaker 2. This held overall and within each word position.

The above influences on overlap were therefore taken into consideration in evaluating the hypotheses tested in the present study, both in the selection of stimuli and in the design of the statistical models (see the Methods section for further details).

Verifying the validity of H-J2, H-3<sub>pv</sub>, or H-3<sub>stiffness</sub> would be useful in at least two ways. First, these hypotheses make predictions about overlap in cases where other hypotheses in the literature about overlap make no predictions. Comparing, for example, word-medial [gd] and [gb] clusters: Neither the word position nor the place order hypotheses make any predictions about the expected amount of articulatory overlap since the clusters are in the same word position and are both back-to-front. However, H-J2 predicts more overlap for [gb] than for [gd]. The two formulations of H-3 would also make predictions about these overlap in these clusters, not based on the articulator, but rather on the relative peak velocity or relative stiffness of the closing movements of C1 and C2.

Second, overlap is a continuum regardless of whether measure (3) or (4) is used, and there are multiple influences on the amount of overlap that were observed in any given cluster. The hypotheses tested here may possibly account for cases where observed articulatory overlap runs contra these other (word position and place order) hypotheses. For example, if [bd] clusters are observed to have less overlap than [db] clusters, this is not expected based on the place order hypothesis, but it could be accounted for by one of the present hypotheses. This could potentially explain the nature of the distributions of overlap more adequately, rather than unpredicted cases being viewed as exceptions or arising from noise.

## 2. Materials and methods

### 2.1. Speakers

Six native speakers of the Oujda dialect of Moroccan Arabic (spoken in Northeast Morocco, near the Algerian-Moroccan border) were recorded. The speakers (one female, five male) ranged in age from 25 to 38. All speakers provided written informed consent to take part in the experiment, and were paid for their participation in the experiment. The experimental procedures were approved by the Ethics Committee of Ludwig-Maximilians-Universität.

### 2.2. 3D Electromagnetic Articulography

The movement of speech articulators was tracked with 3D EMA at 200 Hz sampling rate. Concurrent audio recordings sampled at 24 kHz were made. Recordings were made at the EMA lab in the Institut für Phonetik und Sprachliche Kommunikation (now the Institut für Phonetik und Sprachverarbeitung), Ludwig-Maximilians-Universität München, Germany. EMA receivers included for analysis in the current study were attached to the speaker's lower lip, tongue tip, and tongue body. EMA receivers were also affixed at the nasion, right and left mastoid, and upper incisor to allow for head correction of the movement of the sensors used in analyses. The data were processed following Hoole and Zierdt (2010).

### 2.3. Stimuli

There are two types of consonant sequences types in MA: clusters with no inter-consonantal vowel (CC, e.g., [kbaʃ]) and sequences of two consonants where an optional schwa-like vocoid can appear between the two consonants (C<sup>^</sup>C, e.g., [k<sup>^</sup>bda]). The precise nature of this vocoid and the phonology that accounts for it is a matter of some debate (cf. Dell & Elmedlaoui, 2002; Gafos et al., 2010; Heath, 1987). C<sup>^</sup>C sequence types were therefore not included in the present study: Only unambiguous CC clusters were analyzed. Another consideration in selecting stimuli for the present experiment is that Gafos et al. (2010) found that word-medial clusters were significantly more overlapped than word-initial and word-final clusters. The same study also found that the place-order effect on overlap was more likely to be found in contexts where clusters are otherwise more overlapped than in those where they are less so. If the same is true for the effects of differences in peak

velocity and/or measured stiffness of the consonants in a cluster, these effects should most likely be observed in word-medial clusters. Therefore, all stimuli were real words containing two-consonant, word-medial clusters, where both the preceding and following vowel were [a]. Restricting the stimuli to word-medial clusters allowed us to control for word position rather than increase the complexity of our analyses by including another predictor.

The data analyzed in the present study are a subset of articulatory recordings from three corpora that were collected for other experiments. The stimuli are shown in **Table 2**, broken down by corpus. The stimuli in corpus 1 were produced by Speaker 1. Those in corpus 2 were produced by Speakers 2 and 3. Those in corpus 3 were produced by Speakers 4, 5, and 6. Stimuli were presented on a computer screen in Arabic script, including diacritics indicating full vowels, within a carrier phrase. Articulatory and acoustic recordings were made as the speakers read these stimuli from the computer screen. The presentation order of all stimuli was pseudo-randomized for each speaker. All speakers produced each of their stimuli five times, such that there were always other stimuli intervening between any two productions of a given stimulus.

### 3. Results

A total of 566 productions were recorded.  $v^{\wedge}$  and  $k'$  were calculated for both consonants in each token. Within each token, the difference of the  $v^{\wedge}$  and  $k'$  values of C1 and C2 were calculated to determine the PV Difference and Stiffness Difference in each token, as in (5) and (6), meaning that positive values indicate that C1 had a greater  $v^{\wedge}$  or  $k'$  value than that of C2, and negative values indicate that C1 had a smaller  $v^{\wedge}$  or  $k'$  value than that of C2. The values for each token as output from MVIEW and calculated therefrom are available in the file indicated in Supplementary Material 1.

**Table 2:** Stimuli for all speakers. All clusters were word-medial. See text for details of the different corpora.

Place order	Cluster	Stimuli		
		Corpus 1	Corpus 2	Corpus 3
Front-to-back	[bd]	ʒabda	ʒabda	ʒabda
	[bt]	nabta		nabta
	[bg]	sabga	sabga	sabga
	[bk]			sabka
	[dg]	ħadga	ħadga	fadga
	[dk]			ʃadka, fatka
	[tg]			ʃatga, ratga
	[tk]	ʃatka	ʃatka	ʃatka
Back-to-front	[db]	kadba	kadba	nadba
	[gb]		ragba	ragba
	[gd]	ʕagba	ragba	ragda
	[gt]			ragta, sagta
	[kb]	rakba		rakba
	[kd]			ħakda
	[kt]	sakta	sakta	sakta
	[tb]	ratba		ratba

$$(5) \quad \text{PV Difference: } \hat{v}_{C1} - \hat{v}_{C2}$$

$$(6) \quad \text{Stiffness Difference: } k'_{C1} - k'_{C2}$$

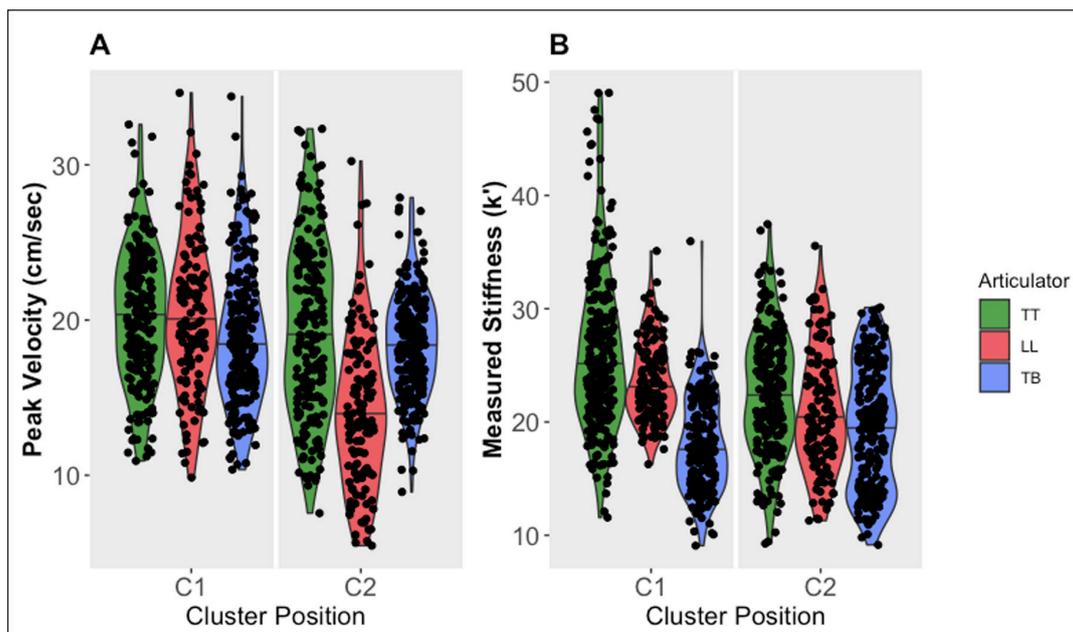
Twenty productions (3.4% of the data) were removed due to having an outlier value ( $\geq 3$  standard deviations away from the mean) of any of the following values:  $\hat{v}$  of C1,  $\hat{v}$  of C2, PV Difference,  $k'$  of C1,  $k'$  of C2, Stiffness Difference, Relative Overlap, or Onset Lag. These tokens were discarded on the assumption that these differences were due to measurement error, leaving 546 tokens for analyses. Each token was also classified by place order. **Table 3** shows the number of tokens included for each speaker and place order.

**Table 3:** Number of tokens (546 total) included in the analyses, by Place Order within Speaker.

Speaker	1	2	3	4	5	6
<b>Total tokens</b>	<b>50</b>	<b>31</b>	<b>46</b>	<b>124</b>	<b>149</b>	<b>146</b>
<i>Front-to-back</i>	24	12	25	69	77	80
<i>Back-to-front</i>	26	19	21	55	72	66

### 3.1. Articulator-specific rapidity

First we tested H-J1, which posits that there should be inherent differences in the rapidity of articulations based on the primary oral articulator, with the tongue tip being fastest, lower lip being intermediary, and tongue body being slowest. **Figure 4A** shows the data relevant for H-J1 using peak velocity, and **Figure 4B** shows the data relevant for H-J1 using measured stiffness. The distributions are broken out by cluster position (C1 or C2) to take into consideration the potential effect of position on rapidity, and because the patterns of assimilation discussed by Jun (2004) pertain to cluster positions separately, as shown in **Figure 1** (A-B versus C-D).



**Figure 4:** Peak velocity (A) and measured stiffness (B) values, by primary articulator and cluster position.

The peak velocity of the closing movements showed very little difference by articulator when in C1 position, though numerically they did pattern as predicted by H-J1. Peak velocities were slower in general for C2 consonants, especially for lower lip movements, which were slower not just compared to lower lip peak velocities in C1, but compared to the other two articulators in C2. This pattern is inconsistent with the prediction of H-J1. The numerical pattern for measured stiffness was consistent with the prediction of H-J1, and as with the peak velocities, stiffness values were lower in C2 position than in C1.

Two linear mixed-effects models (Baayen, Davidson, & Bates, 2008; Gelman & Hill, 2007) were fit to test whether the differences shown in **Figure 4** were significant, using the LME4 package (Bates, 2005; Bates et al., 2020) for R (R Development Core Team, 2018). Speaker and stimulus were modeled as random effects, with one model having peak velocity as its predicted value and the second for measured stiffness. The fixed effects in the models were articulator and cluster position (C1 or C2), as well as the interaction between the two. Each model included a random slope for articulator by speaker. Omnibus test statistics for the fixed effects (articulator, cluster position, and their interaction) 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. Post-hoc differences in articulator-cluster position combinations were assessed using estimated marginal means (Searle, Speed, & Milliken, 1980) using the EMMEANS package (Lenth, Singmann, Love, Buerkner, & Herve, 2019) for R. Degrees of freedom were calculated using the Kenward-Roger method. The R script used to fit these and all of the mixed-effects models, as well as to generate all of the figures, in the present study are available in the file indicated in Supplementary Material 2.

Results of the ANOVA are shown in **Table 4**. There were no significant differences in peak velocity based on articulator, though cluster position and the interaction were significant. For measured stiffness, all three fixed effects were significant.

**Table 5** shows the results of the comparisons of all articulator-cluster position combinations, providing more insight into the results of the ANOVA shown in **Table 4**. Pairs with a Tukey-adjusted  $p$  value  $< 0.05$  were deemed reliable and are shown in bold. For peak velocity, there were no reliable differences based on articulator in C1 position. In C2 position, the only reliable difference between articulators was that the lower lip had lower peak velocity than the tongue body. The peak velocity of the lower lip was significantly higher in C1 position than in C2, but there was no difference based on cluster position for the other two articulators. The only other significant difference was that the lower lip in C2 position had lower peak velocity than the tongue body in C1. There was no support for H-J1 based on peak velocity.

**Table 4:** Results of the Type III analysis of variance of the mixed-effect model evaluating differences in peak velocity (left) and measured stiffness (right) by articulator (Art.), cluster position (C. Pos.), and their interaction. Significant effects are bolded.

Effect	df	Peak Velocity (Figure 4A)					Measured Stiffness (Figure 4B)				
		Sum Sq	Mean Sq	Den. df	F	p	Sum Sq	Mean Sq	Den. df	F	p
Art.	2	45.4	22.69	8.7	1.5	0.267	<b>507.9</b>	<b>254.0</b>	<b>4.8</b>	<b>17.5</b>	<b>0.006</b>
C. Pos.	1	<b>1660.6</b>	<b>1660.6</b>	<b>1050.2</b>	<b>112.7</b>	<b>0.000</b>	<b>1092.5</b>	<b>1092.5</b>	<b>1069.7</b>	<b>75.1</b>	<b>0.000</b>
Art. :C. Pos.	2	<b>1455.3</b>	<b>727.6</b>	<b>1050.2</b>	<b>49.4</b>	<b>0.000</b>	<b>715.6</b>	<b>357.8</b>	<b>323.8</b>	<b>24.6</b>	<b>0.000</b>

**Table 5:** Comparison of estimated marginal means for articulator-cluster position pairs, based on the results of the linear mixed-effects models analyzing the relationship between articulator and cluster position with movement rapidity, as indexed by peak velocity (left) and measured stiffness (right). The rows under the headings ‘Within C1’ and ‘Within C2’ show the results for comparisons across articulator within a cluster position. The rows under the heading ‘C1 vs. C2 (same articulator)’ show the results of comparing the effect of cluster position within each articulator. The rows under the heading ‘C1 vs. C2 (different articulators)’ show the results for comparisons across both cluster positions and articulators. Significant differences are bolded.

Contrast	Peak Velocity (Figure 4A)					Measured Stiffness (Figure 4B)				
	est.	SE	df	t ratio	p	est.	SE	df	t ratio	p
<b>Within C1</b>										
TT vs. LL	-0.90	1.43	9.31	-0.63	0.986	-2.61	1.53	6.61	-1.70	0.570
TT vs. TB	0.30	1.29	8.71	0.24	1.000	<b>6.19</b>	<b>1.20</b>	<b>8.53</b>	<b>5.18</b>	<b>0.006</b>
LL vs. TB	1.19	1.03	14.92	1.15	0.852	<b>8.80</b>	<b>1.30</b>	<b>7.46</b>	<b>6.76</b>	<b>0.002</b>
<b>Within C2</b>										
TT vs. LL	4.42	1.43	9.31	3.09	0.094	1.84	1.53	6.65	1.20	0.824
TT vs. TB	-0.08	1.24	8.71	-0.06	1.000	3.50	1.20	8.54	2.93	0.126
LL vs. TB	<b>-4.50</b>	<b>1.03</b>	<b>14.92</b>	<b>-4.36</b>	<b>0.006</b>	1.66	1.30	7.42	1.27	0.791
<b>C1 vs. C2 (same articulator)</b>										
TT	0.90	0.38	1051.26	2.36	0.171	1.57	0.66	291.9	2.40	0.161
LL	<b>6.22</b>	<b>0.49</b>	<b>1051.26</b>	<b>12.70</b>	<b>0.000</b>	<b>6.02</b>	<b>0.72</b>	<b>832.5</b>	<b>8.32</b>	<b>0.000</b>
TB	0.53	0.37	1051.26	1.44	0.703	-1.12	0.66	304.0	-1.70	0.531
<b>C1 vs. C2 (different articulators)</b>										
C1 TT vs. C2 LL	5.32	1.43	9.31	3.72	0.038	3.41	1.50	6.14	2.27	0.326
C1 TT vs. C2 TB	0.82	1.24	8.71	0.66	0.982	<b>5.07</b>	<b>1.09</b>	<b>5.95</b>	<b>4.65</b>	<b>0.026</b>
C1 LL vs. C2 TB	1.72	1.03	14.92	1.66	0.575	<b>7.68</b>	<b>1.27</b>	<b>6.78</b>	<b>6.04</b>	<b>0.005</b>
C2 TT vs. C1 LL	-1.80	1.43	9.31	-1.25	0.802	-4.18	1.50	6.11	-2.79	0.184
C2 TT vs. C1 TB	-0.61	1.24	8.71	-0.49	0.995	<b>4.62</b>	<b>1.09</b>	<b>6.08</b>	<b>4.23</b>	<b>0.038</b>
C2 LL vs. C1 TB	<b>-5.03</b>	<b>1.03</b>	<b>14.92</b>	<b>-4.87</b>	<b>0.002</b>	2.78	1.27	6.76	2.19	0.349

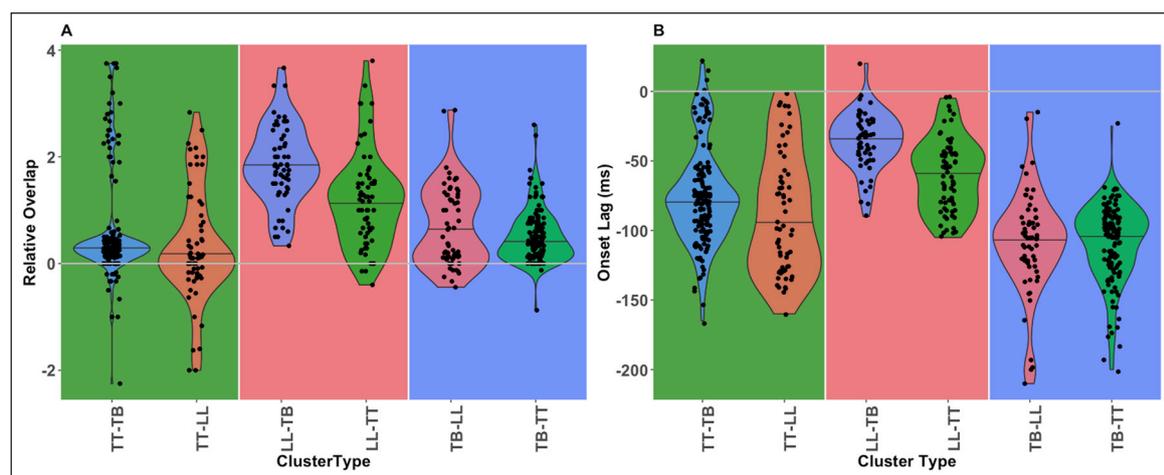
In C1 position, the tongue body had significantly lower measured stiffness values than either the tongue tip or lower lip. There were no significant differences between any articulators in C2 position. Similar to peak velocity, the lower lip had significantly lower stiffness values in C2 position than in C1. The tongue body also had lower measured stiffness than the other two articulators when it was in C2 position and the other articulators were in C1. The tongue body in C1 position also had significantly lower stiffness than the tongue tip in C2 position. There was therefore limited support for H-J1 based on the measured stiffness of the articulators: The tongue body had lower stiffness than the other two articulators, but only within C1 position or sometimes across cluster positions. Interpreting the cross-position differences is complicated by the main effect of Cluster Position (with C2 having lower stiffness values). The significant effect of articulator in the ANOVA was therefore due to these differences between the tongue body and the other articulators, given that the predicted difference in measured stiffness between the tongue tip and lower lip was found neither within nor across cluster positions.

### 3.2. Overlap based on cluster type

In this section we test hypothesis H-J2. We have noted already that H-J2 is predicated on H-J1, and we have just shown that there is limited support for H-J1 in these data. Nevertheless, there is still a possibility that articulator-specific differences may be evident only when consonants are analyzed using measures that are within-token. Looking at within-token overlap within cluster types is the most direct test of hypothesis H-J2. **Figure 5A** and **B** show the Relative Overlap and Onset Lag, respectively, with the cluster types organized such that the amount of overlap should decrease from left to right, according to hypothesis H-J2.

The patterns of overlap do not correspond to those expected based on H-2J, regardless of the index of overlap that was used. Most notably, the two tongue tip-initial cluster types have much less overlap than the two lower lip-initial cluster types, comparable to the amount of overlap with the two tongue body-initial cluster types, which were expected to have the least overlap. Apart from the tongue tip-initial clusters, overlap by cluster type patterns more or less as expected.

The reliability of the differences shown in **Figure 5** was evaluated with two linear mixed effects models in which cluster type was the fixed predictor (with six levels), one of Relative Overlap and another in which cluster type was the fixed predictor of Onset Lag. Random intercepts were included for speaker and stimulus. Full pairwise comparisons of all cluster types would require statistical power beyond what is available in our data. Therefore, we set the cluster type tongue tip-tongue body as the reference level, since this cluster type was predicted to have the highest overlap according to H-J2 yet had numerically lower overlap than lower lip-initial clusters. Structuring the model this way allowed us to determine whether tongue tip-tongue body clusters had reliably more overlap than tongue tip-lower lip clusters (per H-J2), whether the greater overlap for the two lower lip-initial clusters was reliable (contra H-J2), and whether the tongue body-initial clusters had less overlap than the tongue tip-tongue body clusters (per H-J2). Place Order could not be included as a predictor because it is too correlated with cluster type.



**Figure 5:** Relative Overlap (A) and Onset Lag (B), by cluster type (i.e., combination of primary oral articulators, where TT = tongue tip, LL = lower lip, and TB = tongue body). Distributions of overlap values by cluster type are arranged such that the amount of overlap predicted by H-J2 (see Figure 2) should be greatest for the leftmost type (tongue tip-tongue body) and decrease by type going rightwards, with tongue body-tongue tip clusters having the least overlap. Colors correspond to the primary oral articulator, with the background color indicating the C1 articulator and the color of the violin plot indicating the C2 articulator.

**Table 6:** Results of the linear mixed-effects models with cluster type as a predictor of Relative Overlap (left) and Onset Lag (right). Significant effects compared to TT – TB clusters are bolded.

Effect	Relative Overlap (Figure 5A)					Onset Lag (ms, Figure 5B)				
	est.	SE	df	t	p	est.	SE	df	t	p
(Intercept)	<b>0.67</b>	<b>0.16</b>	<b>10.51</b>	<b>4.22</b>	<b>0.002</b>	<b>-76.20</b>	<b>7.63</b>	<b>7.95</b>	<b>-9.99</b>	<b>0.000</b>
TT – LL	-0.33	0.18	19.55	-1.82	0.085	-10.26	7.18	17.36	-1.43	0.171
LL – TB	<b>1.17</b>	<b>0.20</b>	<b>15.56</b>	<b>5.92</b>	<b>0.000</b>	<b>38.49</b>	<b>8.03</b>	<b>14.36</b>	<b>4.79</b>	<b>0.000</b>
LL – TT	<b>0.59</b>	<b>0.19</b>	<b>14.80</b>	<b>3.04</b>	<b>0.008</b>	<b>18.08</b>	<b>7.93</b>	<b>13.71</b>	<b>2.28</b>	<b>0.039</b>
TB – LL	0.06	0.20	15.34	0.31	0.765	<b>-33.85</b>	<b>7.99</b>	<b>14.12</b>	<b>-4.24</b>	<b>0.001</b>
TB – TT	-0.15	0.14	18.77	-1.05	0.306	<b>-37.50</b>	<b>5.77</b>	<b>17.02</b>	<b>-6.51</b>	<b>0.000</b>

**Table 6** shows that there was no reliable difference in overlap between tongue tip-tongue body and tongue tip-lower lip clusters (not supporting H-J2), and that the overlap for the two lower lip-initial clusters was reliably greater than for the tongue tip-tongue body clusters (contra H-J2), regardless of whether overlap was indexed by Relative Overlap or Onset Lag. Both tongue body-initial clusters had less overlap than the tongue tip-tongue body clusters (per H-J2), but only when overlap was indexed by Onset Lag. There was no reliable difference when indexed by Relative Overlap.

### 3.3. Effects of Peak Velocity Difference and Stiffness Difference

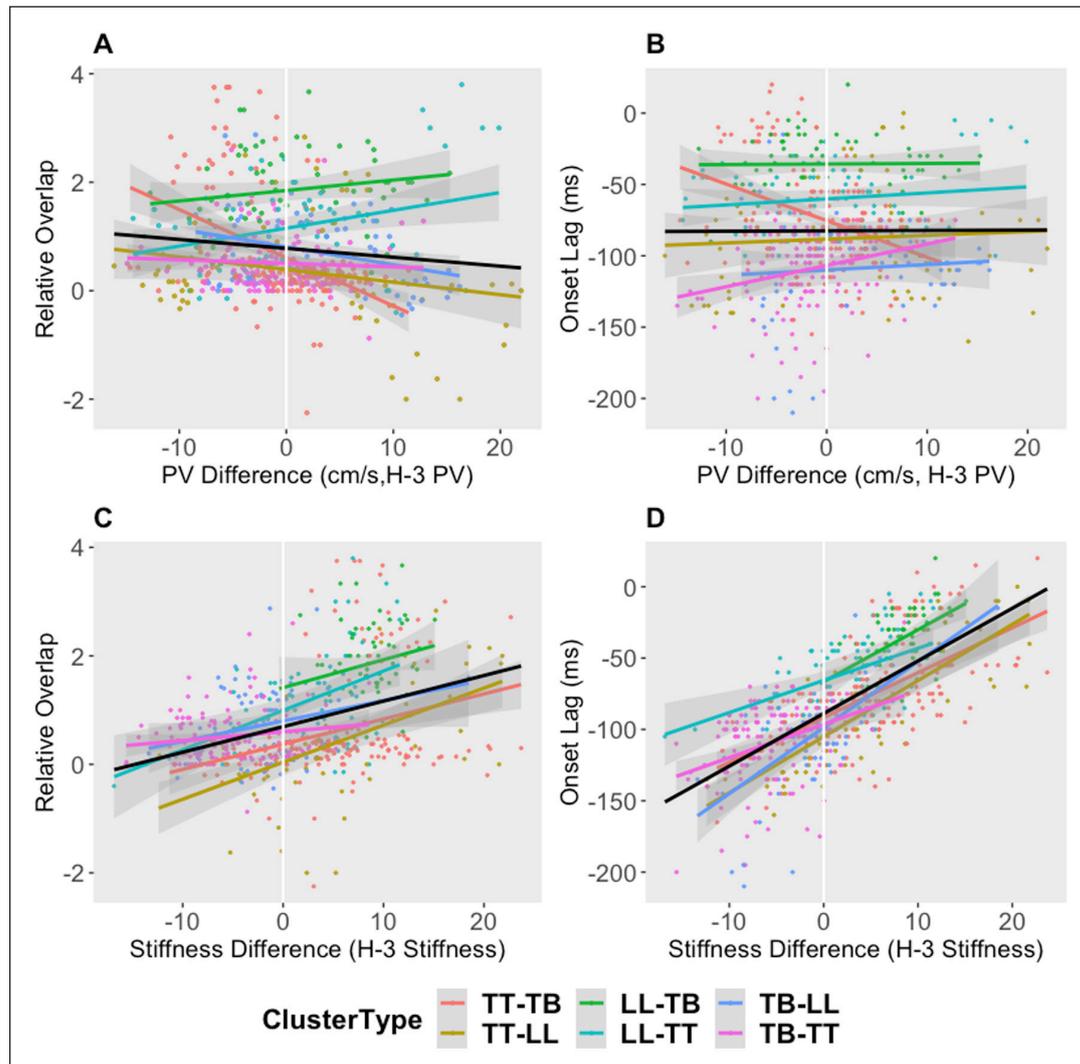
Lastly, we tested  $H-3_{pv}$  and  $H-3_{stiffness}$ , using both Relative Overlap and Onset Lag as indexes of overlap. The relationship between PV Difference and Stiffness Difference, and Relative Overlap and Onset Lag is shown in **Figure 6**. The black lines are linear regression lines across all cluster types. The colored lines are linear regression lines within cluster type.

Four linear effects models were fit to the data, one for each relationship shown in **Figure 6**. Speaker and stimulus were modeled as random effects. The fixed effects were Peak Velocity Difference or Stiffness Difference (both continuous), and Place Order (categorical). Cluster type was not included as a predictor, as the models would not converge if it was added. Since Gafos et al. (2010) found Place Order effects were speaker-dependent for a subset of the speakers included in the present study, speaker-specific random slopes for Place Order were also included in the model. The results of all four models are shown in **Table 7**.

Results of the statistical models confirm that Stiffness Difference was a significant predictor of overlap, regardless of whether Relative Overlap or Onset Lag was used to index overlap, with overlap increasing as Stiffness Difference increased. These results provide strong support for hypothesis  $H-3_{stiffness}$ . On the other hand, no support was found for hypothesis  $H-3_{pv}$ : Peak Velocity Difference was not a significant predictor of overlap, regardless of whether Relative Overlap or Onset Lag is used to index overlap. Place Order was not a significant predictor of overlap other than in the model that included Peak Velocity Difference as a predictor of Onset Lag (**Figure 6B/Table 5-B**), where it predicted greater overlap in front-to-back clusters, as expected.

## 4. Discussion

Jun (2004) proposed that typological patterns of regressive place assimilation can be explained by differences in overlap in different types of clusters. According to this proposal, when clusters are more overlapped, acoustic information as to the place of the first consonant is diminished and place assimilation is more likely. Differences in overlap arise, according to this proposal, based on the ‘inherent velocities’ of the primary oral



**Figure 6:** Relationship between within-token C1–C2 Peak Velocity Difference and Relative Overlap (A) and Onset Lag (B). Relationship between within-token C1–C2 Stiffness Difference and Relative Overlap (C) Onset Lag (D). Colored regression lines correspond to cluster types. The black regression lines are fit across all cluster types.

**Table 7:** Results of the four linear mixed-effects models of influences on overlap (letters A–D correspond to the data shown in Figure 6). Significant effects are bolded.

	Relative Overlap					Onset Lag (ms)				
	A					B				
<b><i>H-3<sub>PV</sub></i></b>										
<b>effect</b>	<b>est.</b>	<b>SE</b>	<b>df</b>	<b>t</b>	<b>p</b>	<b>est.</b>	<b>SE</b>	<b>df</b>	<b>t</b>	<b>p</b>
(Intercept)	<b>0.89</b>	<b>0.19</b>	<b>10.76</b>	<b>4.52</b>	<b>0.001</b>	<b>-68.54</b>	<b>7.64</b>	<b>10.12</b>	<b>-8.98</b>	<b>0.000</b>
PV Diff.	-0.01	0.01	487.49	-1.48	0.139	-0.07	0.24	424.64	-0.29	0.771
PO: Back-to-front	-0.33	0.27	11.06	-1.22	0.247	<b>-35.65</b>	<b>9.27</b>	<b>13.09</b>	<b>-3.85</b>	<b>0.002</b>
<b><i>H-3<sub>stiffness</sub></i></b>										
<b>effect</b>	<b>est.</b>	<b>SE</b>	<b>df</b>	<b>t</b>	<b>p</b>	<b>est.</b>	<b>SE</b>	<b>df</b>	<b>t</b>	<b>p</b>
(Intercept)	<b>0.58</b>	<b>0.24</b>	<b>11.64</b>	<b>2.43</b>	<b>0.033</b>	<b>-85.32</b>	<b>7.73</b>	<b>7.84</b>	<b>-11.03</b>	<b>0.000</b>
Stiff. Diff.	<b>0.06</b>	<b>0.01</b>	<b>537.30</b>	<b>11.45</b>	<b>0.000</b>	<b>3.35</b>	<b>0.17</b>	<b>529.07</b>	<b>19.87</b>	<b>0.000</b>
PO: Back-to-front	0.14	0.35	11.24	0.39	0.707	-10.84	8.81	10.26	-1.23	0.246

articulators involved in making the constrictions needed for the two consonants in a C1C2 cluster: In short, when the articulator making the C1 constriction is more rapid than the articulator making the C2 constriction, there is more overlap than when the order of the rapidity of the articulators is the reverse. Jun (2004) assumes that the tongue tip has the highest inherent velocity, the tongue body has the lowest, and the lower lip is intermediate between the two lingual articulators. In the present study, we tested two hypotheses that are assumed in the proposal made by Jun (2004). The first was that articulators have ‘inherent velocities.’ The second, which is predicated on the first being true, was that overlap in a cluster will be modulated systematically by the specific combinations of articulators involved in making the consonantal constrictions in the cluster. We also tested a third hypothesis, as an alternative to the second, that overlap is modulated by the relative rapidity of the articulations involved, but that this relative rapidity is not determined by the specific articulators.

Support for the first hypothesis was limited, and discernible only when the rapidity of the articulators was quantified as measured stiffness. The only consistent, significant, articulator-related difference was that tongue body articulations had significantly lower measured stiffness (but not peak velocity) than lower lip and tongue tip articulations in C1 position and across cluster positions. This finding is consistent with the proposal by Jun (2004). However, there were no significant differences between the lower lip and tongue tip within or across cluster positions. This lack of difference is problematic for the proposal by Jun (2004), in which it is crucially assumed that tongue tip articulations are more rapid than those of both of the other articulators.

Given that there was only limited support for the first hypothesis, it was not surprising that overlap was not predicted correctly by the second hypothesis. The most relevant prediction of this hypothesis for the account of Jun (2004)—that two tongue tip-initial cluster types should show the most overlap—was undermined by significant results in the opposite direction. While the rest of the overlap patterns (i.e., those within and between lower lip-initial and tongue body-initial clusters) were more or less consistent with the second hypothesis, the exception of the tongue tip-initial clusters is deeply problematic for the account proposed by Jun (2004).

We found significant support for the third hypothesis we tested, i.e., the amount of overlap in a cluster was predicted significantly by the difference in rapidity of the specific articulations in a given token. This was true regardless of whether overlap was indexed by Relative Overlap or Onset Lag. However, this was only true when the rapidity of an articulation was quantified by stiffness. When the rapidity of an articulation was quantified by peak velocity, the difference in peak velocity did not predict either measure of overlap.

Accounting for the variation in overlap with Stiffness Difference (see **Figure 6C–D**)—which has continuous values—can explain much more of the data than accounts that appeal to categorical predictors like cluster type (i.e., the specific combination of articulators involved). Firstly, although there were some significant differences in overlap based on cluster type (see **Figure 5** and **Table 5**, in addition to the different intercepts by cluster type visible in **Figure 6**), the relationship between Stiffness Difference and both indexes of overlap held across cluster types (although the significance of this observation could not be tested statistically, as noted in Section 3.3). While an account based on cluster type could possibly be constructed to explain the significant differences found by cluster type, such an account would be unable to explain the fact that the relationship between Stiffness Difference and overlap holds for tokens within the same cluster type. Another aspect of the relationship between Stiffness Difference and overlap is that it captures the fact that it is the stiffness of C1 relative to that of C2 that is important, not the categorical

distinction of whether C1 has higher measured stiffness than C2 or vice versa. This can be illustrated by looking at the lower lip-tongue body data in **Figure 6C–D** (represented by the highest regression lines in both sub-figures). The Stiffness Differences of all tokens for this cluster type were positive, meaning that C1 always had higher stiffness than C2, and the relationship between Stiffness Difference and overlap was the same as for cluster types whose tokens encompassed both positive and negative values. An analysis that relied on categorizing clusters based on stiffness ‘order’ (i.e., indicating whether C1 or C2 has the higher value of measured stiffness) would not be able to capture this fact. Lastly, the fact that Peak Velocity Differences and Stiffness Differences within each cluster type spanned a wide range of negative and positive values further undermines the first hypothesis tested here. If the rapidity of a gesture were determined predominantly by its main oral articulator, then it would be expected that the difference in values within a given cluster type would be predominantly on one side or the other of the zero point for peak velocity and/or stiffness difference. However, only lower lip-tongue body clusters patterned this way, and only for Stiffness Difference. The other five cluster types did not pattern this way for either difference measure.

The lack of a result for the Peak Velocity Difference may at first seem at odds with the significant result for Stiffness Difference, but these results are compatible with each other. Looking at **Figure 1C–D**, Jun’s (2004) proposal predicts less overlap in C compared to D due to a ‘more rapid’ C2 articulator movement in C. However, these schematic drawings do not take articulator displacement into account. As discussed above, the peak velocity of an articulator is known to covary with the articulator’s displacement (e.g., Guenther, 1995; Kent & Moll, 1976; Munhall et al., 1985). In other words, the peak velocity of the articulator increases as the distance that the articulator has to travel increases. If we think of the schemas in **Figure 1C–D** as representing articulator movements through space as well as time with the C2 movement in **Figure 1C** traveling less far than the one in **Figure 1D**, we expect that the articulator movement in **Figure 1C** should have a lower peak velocity than the one in **Figure 1D**. Viewed this way, the schemas in **Figure 1C** and **D** are then consistent with the results of this study since the clusters where C2 had higher peak velocity than C1 (**Figure 1D**) had more overlap than when C2 had lower velocity than C1 (**Figure 1C**). At the same time, a C2 gesture like the one shown in **Figure 1C** could have a higher measured stiffness value than the C2 gesture in **Figure 1D** and yet still have a lower peak velocity than the C2 gesture in **Figure 1D**. If the C2 gesture in **Figure 1C** does not have to go as far as the C2 gesture in **Figure 1D** to achieve its target, the measured stiffness values factor out the confounding influence of displacement.

## 5. Conclusion

The results from the present study have several implications for the proposal of Jun (2004). One key insight of that proposal is that the amount of overlap in a consonant cluster is a function of the relative rapidity of the articulations of those two consonants. Our results provide strong support for this aspect of Jun’s proposal, but only when rapidity was indexed by the more abstract dynamical parameter of stiffness of the closing gestures of the two articulations in the cluster. Overlap was not predicted in the same clusters when the difference in the peak velocities themselves was used. The peak velocities of the articulations seemed to be a function of both the stiffness settings of the gestures and the distance that the articulators needed to travel to achieve their targets. We conclude that the settings of the stiffness control parameter associated with each articulatory gesture in a cluster, not the peak velocity of each articulator movement, contribute to differences in overlap.

The other aspect of the proposal by Jun (2004) is that typological patterns of the triggers and targets of regressive place assimilation can be explained by inherent differences in

rapidity across articulators. However, attributing differences in stiffness (or peak velocity) in a cluster based on the articulators involved seems implausible given the present results. We found limited support for systematic differences in stiffness values or peak velocities of articulations based on primary oral articulator. The amount of overlap in a cluster was also not predicted well by the specific combination of articulators in the cluster, and in fact patterned opposite to expectation in some crucial cases.

The present results do not change the typological generalizations noted by Jun (2004), but they do undermine the phonetic explanation that Jun (2004) proposes for them. The present results should be interpreted with some caution, since they come from one language. Nevertheless, the assumptions of Jun (2004) concerning both the ‘inherent velocities’ of articulators and the proposed effect that those inherent velocities have on overlap are cast by Jun (2004) as language-independent, which they would have to be in order to provide an adequate account of typological patterns. It is therefore reasonable to test those assumptions in any language. As we discussed in the Introduction, Moroccan Arabic is very useful in this regard given the rich combinations of consonants in this language. Differences in overlap may ultimately be part of an explanation for observed typological patterns of assimilation. Our results strongly suggest that further research with the goal of providing such an explanation should include the role of the dynamical control parameter of stiffness on overlap, since stiffness seems to stand at the right level of abstraction from the surface kinematics, which are highly context-dependent.

### **Additional Files**

The additional files for this article can be found as follows:

- **Supplementary Material 1.** Data file (CSV) containing all of the labeled tokens, as extracted from the EMA files. DOI: <https://doi.org/10.5334/labphon.272.s1>
- **Supplementary Material 2.** R script used to generate Figures 4–6, as well as to fit all statistical models. DOI: <https://doi.org/10.5334/labphon.272.s2>

### **Acknowledgements**

The authors thank Lisa Davidson, Maria Gouskova, Douglas H. Whalen, and the members of the Speech Production, Acoustics, and Perception Lab of the CUNY Graduate Center, who provided extensive feedback on initial drafts of this study. We are also indebted to two anonymous reviewers, whose comments resulted in material improvement to this study. All remaining errors are our own.

### **Funding Information**

KDR has been supported by NIH Grant DC-002717 to Haskins Laboratories and the City University of New York. PH has been supported by German Research Council Grant HO3271/3-1. AIG has been supported by the European Research Council (AdG 249440) and the Deutsche Forschungsgemeinschaft (SFB 1287, Project C04).

### **Competing Interests**

The authors have no competing interests to declare.

### **Author Contributions**

Kevin Roon designed the present experiment, labeled the data from the first two corpora to identify the consonantal gestures, created all of the data figures and statistical models, and was the primary author of all sections of the manuscript. Philip Hoole ran all of the equipment that was required for data acquisition, post-processed all of the articulatory and acoustic data to prepare it for analyses, and was involved in editing the manuscript.

Chakir Zeroual, as a native-speaker of Moroccan Arabic, was the primary consultant for all questions concerning the language, especially in designing the stimuli for the experiment. He was Speaker 1. Shihao Du labeled the data that identified the consonantal gestures from the three speakers in corpus 3. Adamantios Gafos designed the stimuli for the experiments, assisted in the collection of the data, devised the logic underlying the third hypothesis tested here, and was extensively involved in the writing of the manuscript. This study evolved from extensive discussions between KDR and AIG.

## References

- 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, H., ... 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 Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech Perception, Production and Linguistic Structure* (pp. 356–375). Tokyo: OHM Publishing Co.
- Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252. DOI: <https://doi.org/10.1017/S0952675700000658>
- Browman, C. P., & Goldstein, L. M. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201–251. DOI: <https://doi.org/10.1017/S0952675700001019>
- Browman, C. P., & Goldstein, L. M. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, *18*, 299–320. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30376-6](https://doi.org/10.1016/S0095-4470(19)30376-6)
- Byrd, D. (1992). Perception of assimilation in consonant clusters: A gestural model. *Phonetica*, *49*, 1–24. DOI: <https://doi.org/10.1159/000261900>
- 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., & 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>
- Chitoran, I., & Goldstein, L. M. (2006). Testing the phonological status of perceptual recoverability: Articulatory evidence from Georgian. *Paper presented at the Laboratory Phonology X*. Paris, France.
- Chitoran, I., Goldstein, L. M., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 419–447). Berlin/New York: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110197105.2.419>
- Cooke, J. D. (1980). The organization of simple, skilled movements. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North-Holland. DOI: [https://doi.org/10.1016/S0166-4115\(08\)61946-9](https://doi.org/10.1016/S0166-4115(08)61946-9)
- Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer. DOI: <https://doi.org/10.1007/978-94-010-0279-0>
- Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, *89*(1), 369–382. DOI: <https://doi.org/10.1121/1.400674>

- Fuchs, S., Perrier, P., & Hartinger, M. (2011). A critical evaluation of gestural stiffness estimations in speech production based on a linear second-order model. *Journal of Speech, Language, and Hearing Research*, 54(4), 1067–1076. DOI: [https://doi.org/10.1044/1092-4388\(2010/10-0131\)](https://doi.org/10.1044/1092-4388(2010/10-0131))
- Gafos, A. I. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory*, 20(2), 269–337. DOI: <https://doi.org/10.1023/A:1014942312445>
- Gafos, A. I., Hoole, P., Roon, K. D., & Zeroual, C. (2010). Variation in timing and phonological grammar in Moroccan Arabic clusters. In C. Fougeron, B. Kühnert, M. D’Imperio, & N. Vallée (Eds.), *Laboratory Phonology 10* (pp. 657–698). Berlin/New York: Mouton de Gruyter.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge/New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511790942>
- 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](https://doi.org/10.1016/0167-6393(85)90051-2)
- Hardcastle, W., & Roach, P. (1979). An instrumental investigation of coarticulation in stop consonant sequences. In H. Hollien & P. A. Hollien (Eds.), *Current Issues in the Phonetic Sciences* (pp. 531–540). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/cilt.9.56har>
- Heath, J. (1987). *Ablaut and Ambiguity: Phonology of a Moroccan Arabic Dialect*. Albany, NY: State University of New York Press.
- Hoole, P., & Zierdt, A. (2010). Five-dimensional articulography. In B. Maasen & P. H. H. M. van Lieshout (Eds.), *Speech Motor Control: New developments in basic and applied research* (pp. 331–349). Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199235797.003.0020>
- Hoole, P., Zierdt, A., & Geng, C. (2003). Beyond 2D in articulatory data acquisition and analysis. *Paper presented at the 15th International Congress of Phonetic Sciences (ICPhS XV)*. Barcelona, Spain.
- Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach*. (Doctoral dissertation), Los Angeles: University of California.
- Jun, J. (1996). Place assimilation is not the result of gestural overlap: Evidence from Korean and English. *Phonology*, 13, 377–407. DOI: <https://doi.org/10.1017/S0952675700002682>
- Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically Based Phonology* (pp. 58–86). Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486401.003>
- Kelso, J. A. S., Vatikiotis-Bateson, E., Saltzman, E. L., & Kay, B. A. (1985). A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling. *Journal of the Acoustical Society of America*, 77, 266–280. DOI: <https://doi.org/10.1121/1.392268>
- Kent, R. D., & Moll, K. L. (1976). Cineflourographic analyses of selected lingual consonants. *Journal of Speech and Hearing Research*, 15, 453–473. DOI: <https://doi.org/10.1044/jshr.1503.453>
- 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, 303–320. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31257-4](https://doi.org/10.1016/S0095-4470(19)31257-4)
- Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and C-center in selected French consonant clusters. *Paper presented at the 7th International Seminar on Speech Production (ISSP)*. Ubatuba, Brazil.
- 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>
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1–2), 71–102. DOI: [https://doi.org/10.1016/0010-0277\(87\)90005-9](https://doi.org/10.1016/0010-0277(87)90005-9)
- Monahan, K. P. (1993). Fields of attraction in phonology. In J. A. Goldsmith (Ed.), *The last phonological rule: Reflections on constraints and derivations* (pp. 61–116). Chicago: University of Chicago Press.
- 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>
- Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics*, 46(2), 135–147. DOI: <https://doi.org/10.1007/BF00339982>
- Oliveira, C., & Teixeira, A. (2007). On gestures timing in European Portuguese nasals. *Paper presented at the 16th International Congress of Phonetic Sciences (ICPhS XVI)*. Saarbrücken, Germany.
- 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. *Journal of the Acoustical Society of America*, 77, 640–648. DOI: <https://doi.org/10.1121/1.391882>
- 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>
- Redford, M. A., & Diehl, R. L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, 106(3), 1555–1565. DOI: <https://doi.org/10.1121/1.427152>
- 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](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>
- Son, M. (2008). *The nature of Korean place assimilation: Gestural overlap and gestural reduction*. (Doctoral dissertation), Yale University, UMI/ProQuest.
- Sorensen, T., & Gafos, A. I. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology*, 28(4), 188–215. DOI: <https://doi.org/10.1080/10407413.2016.1230368>
- Surprenant, A. M., & Goldstein, L. M. (1998). The perception of speech gestures. *Journal of the Acoustical Society of America*, 104(1), 518–529. DOI: <https://doi.org/10.1121/1.423253>

- Vatikiotis-Bateson, E., & Kelso, J. A. S. (1990). Linguistic structure and articulatory dynamics: A cross language study. *Haskins Laboratories Status Report on Speech Research, SR-103/104*, 67–94.
- Vatikiotis-Bateson, E., & Kelso, J. A. S. (1993). Rhythm type and articulatory dynamics in English, French and Japanese. *Journal of Phonetics*, 21, 231–265. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31338-5](https://doi.org/10.1016/S0095-4470(19)31338-5)
- Wright, R. A. (1996). *Consonant cluster and cue preservation in Tsou*. (Doctoral dissertation), Los Angeles: University of California.
- Zeroual, C., Hoole, P., Gafos, A. I., & Esling, J. (2014). Gestural overlap within word medial stop-stop sequences in Moroccan Arabic. *Paper presented at the 10th International Seminar on Speech Production (ISSP)*. Cologne, Germany.
- Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, 22, 121–140. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30189-5](https://doi.org/10.1016/S0095-4470(19)30189-5)

**How to cite this article:** 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, pp. 1–23. DOI: <https://doi.org/10.5334/labphon.272>

**Submitted:** 27 March 2020    **Accepted:** 18 February 2021    **Published:** 08 April 2021

**Copyright:** © 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 