Phonology **26** (2009) 187–215. © Cambridge University Press 2009
doi:10.1017/S0952675709001754 Printed in the United Kingdom

*Syllabification in Moroccan Arabic : evidence from patterns of temporal stability in articulation**

Jason Shaw
Adamantios I. Gafos
New York University and Haskins Laboratories
Philip Hoole
Ludwig Maximilians University, Munich
Chakir Zeroual
Faculté Polydisciplinaire de Taza, Morocco and Laboratoire de Phonétique et Phonologie (CNRS/Université Sorbonne Nouvelle, Paris)

Competing proposals on the syllabification of initial consonants in Moroccan Arabic are evaluated using a combination of experimental and modelling techniques. The proposed model interprets an input syllable structure as a set of articulatory landmarks coordinated in time. This enables the simulation of temporal patterns associated with the input syllable structure under different noise conditions. Patterns of stability between landmarks simulated by the model are matched to patterns in data collected with Electromagnetic Articulometry experiments. The results implicate a heterosyllabic parse of initial clusters so that strings like /sbu/ comprise two syllables, [s.bu]. Beyond this specific result for Moroccan Arabic, the model reveals the range of validity of certain stability-based indexes of syllable structure and generates predictions that allow evaluation of a syllabic parse even when stability-based heuristics break down. Overall, the paper provides support for the broad hypothesis that syllable structure is reflected in patterns of temporal stability and contributes analytical tools to evaluate competing theories on the basis of these patterns.

* We would like to thank the editors of the special issue, Andries Coetzee, René Kager and Joe Pater, as well as Dani Byrd and one anonymous reviewer for comments that greatly improved the paper. Thanks are also due to the NYU ph-lab group, especially Tuuli Adams, Adam Buchwald, Lisa Davidson, Maria Gouskova and Kevin Roon for comments on this work at various stages of development. Parts of the data were first presented at the CUNY conference on the syllable. We'd like to thank participants at that conference for their feedback, particularly François Dell, Paul Kiparsky, Patricia Shaw and Donca Steriade. Remaining errors are solely the responsibility of the authors. This research was supported in part by an Alexander von Humboldt research grant awarded to the second author.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

1 Introduction

There is a growing body of evidence in support of the hypothesis that temporal patterns in speech production are characteristic of phonological organisation (e.g. Browman & Goldstein 1988, Krakow 1989, 1999, Sproat & Fujimura 1993, Byrd 1995, 1996, Honorof & Browman 1995, Goldstein *et al*. 2007). The main aim of this paper is to employ these patterns in evaluating competing proposals on syllable structure. To this end, we develop rigorous tools for evaluating the relation between syllable structure and patterns of temporal stability in articulatory data. These include statistical procedures for evaluating stability and a model which generates temporal data given a syllabic parse as input. The statistical methods allow us to quantify and assess the effect of structural organisation on temporal stability. The model allows us to demonstrate the gradient range of temporal manifestations of an invariant phonological structure. Together, these tools provide the means to rigorously evaluate hypotheses about syllabification using experimental data.

Our empirical domain is initial consonant clusters in Moroccan Arabic, henceforth MA. Previous theoretical accounts of MA syllables can be divided into two broad classes, based on how they parse initial clusters. Those which admit complex onsets, henceforth collectively the 'complex onset hypothesis', organise strings such as /kra/ 'rent' and /skru/ 'his ploughshares' into one syllable (Benhallam 1980: 78, 1990, Keegan 1986: 214, Heath 1987: 266, 271–272, Benkirane 1998: 346). Those which ban complex onsets, henceforth collectively the 'simplex onset hypothesis', divide such strings into two syllables ([k.ra] and [sk.ru]) (Boudlal 2001: 62, Kiparsky 2003: 159–160, Dell & Elmedlaoui 2002: 292).[1] Each of these proposals unifies the description of the same set of core facts using different syllabic representations (at least if we restrict ourselves to 'internal evidence' in the sense of Kenstowicz & Kisseberth 1979: 139, 160–161).[2] For example, several authors adopt the complex onset hypothesis to account for the environments in which the *e* vocoid appears in descriptive grammars of MA (Harrell 1962, Harrell & Sobelman 1966, Harrell & Brunot 2004). The phonetic material produced in these environments has been analysed variably as either a short vowel (Harris 1942, Benhallam 1980, 1990, Keegan 1986, Heath 1987, Boudlal 2001) or as a transition between consonants (Dell & Elmedlaoui 2002, Gafos 2002). The key fact

[1] The '.' indicates a syllable boundary. Not all versions of the simplex onset hypothesis agree on the syllabic role of the initial consonant in a form like [k.ra]. The evidence presented in this paper bears on the presence or absence of a boundary between initial consonants, but not on the precise status of the initial consonant in terms of prosodic role, e.g. mora, syllable nucleus, etc.

[2] As an example of external evidence, Dell & Elmedlaoui (2002) analyse the orthometrics of sung verse in a particular genre of Moroccan Arabic song. The patterns in these data constitute impressive evidence for the simplex onset hypothesis.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

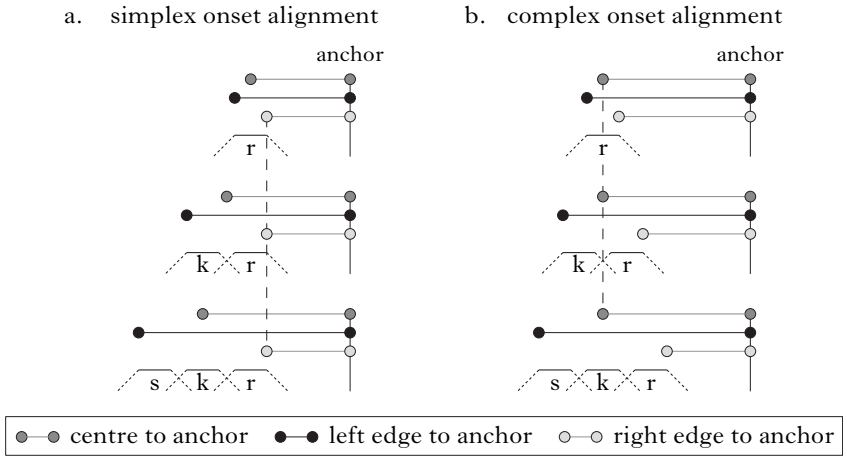


Figure 1

Schematic representation of three intervals delineated by points in an initial consonant cluster and a common anchor. The alignment schema on the left represents temporal predictions of the simplex onset hypothesis. The alignment schema on the right represents predictions of the complex onset hypothesis.

for our purposes is that, regardless of the nature of this vocoid, authors agree that it does not occur between consonants in initial clusters, e.g. #CCV. On the complex onset hypothesis, this is because the vocoid in question is prevented in complex syllable onsets. On the simplex onset hypothesis, it is because the vocoid is prevented at syllable boundaries.

In this paper we seek to distinguish between the simplex and complex onset hypotheses on the basis of temporal stability patterns in articulatory data. We begin by employing phonetic heuristics from past articulatory work on English (Browman & Goldstein 1988, Byrd 1995, Honorof & Browman 1995).[3] These heuristics are illustrated with the temporal alignment schemas in Fig. 1, corresponding to simplex (a) and complex (b) onsets. In these schemas, the temporal structure of segments is represented by three connected lines: a dotted line corresponding to movement towards constriction, a solid line corresponding to constriction duration and a second dotted line corresponding to movement away from constriction. For both alignment schemas, three words differing in the number of initial consonants, *r*, *kr*, *skr*, are shown. In addition, the figure shows three intervals for each word type. The intervals are left-delimited by the left edge, right edge and centre of the consonant cluster, and right-delimited by a common anchor point. The relevant pattern is in how the

[3] Patterns of simplex onset alignment can be found in English sequences such as CVCC#CV, e.g. [bakt#kab] (Byrd 1995).

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

duration of different intervals changes across words with increasing numbers of initial consonants, e.g. *laid*, *played*, *splayed* or *rue*, *crew*, *screw*. The most stable interval in each alignment schema is denoted by a vertical dashed line running from top to bottom, across all words within that schema. Simplex onset alignment corresponds to a pattern whereby the RIGHT EDGE TO ANCHOR interval is more stable than the CENTRE TO ANCHOR interval or the LEFT EDGE TO ANCHOR interval (Byrd 1995). Complex onset alignment corresponds to a different pattern, whereby the centre to anchor interval is more stable across words than the left edge to anchor interval and the right edge to anchor interval (Browman & Goldstein 1988, Honorof & Browman 1995).

When applied to our MA data, we find that the phonetic heuristics schematised above provide a good first diagnostic of syllable structure. However, under certain conditions these heuristics break down. That is, even though by and large the phonetic evidence in our data points in the direction of simplex onsets, we do find some evidence that appears to be more consistent with complex onsets. The conditions under which such evidence surfaces appear to be systematic. This brings us to an important and heretofore unaddressed issue in the relation between experimental data and syllable structure: how reliably do stability metrics reflect the underlying organisation of phonological form in terms of syllables? To address this question, the paper develops a model that allows us to study the relation between patterns of interval stability and posited phonological organisation. The model exposes the range of validity of the stability heuristics, and makes new testable predictions about the experimental data. These predictions allow us to evaluate the fit of a syllable parse to experimental data even in those cases where phonetic heuristics break down. Using the model as an analytical tool we are therefore able to verify that the simplex onset hypothesis is consistent with all of the data examined herein.

The remainder of this paper is organised as follows. After describing our methods in §2, §3 takes a first pass at using stability-based heuristics to evaluate syllable structure. This section analyses the stability of the three interval types schematised in Fig. 1. The analysis is conducted using two different anchor points. In §3.1, we find patterns of stability consistent only with the simplex onset hypothesis. In §3.2, using a different anchor point, we find patterns which appear at first to be inconsistent with the simplex onset hypothesis. To better understand the mixed results in §3, we constructed a model that encodes the simplex onset hypothesis as a series of articulatory landmarks linked by temporal relations. We report simulations based on this model in §4. The simulations matched the complete range of stability patterns in our data, including those which at first appeared to be inconsistent with the simplex onset hypothesis. In addition, the simulations generated new predictions which allowed us to further verify the simplex onset hypothesis even under those conditions in which the phonetic heuristics break down. In §5 we offer brief concluding remarks and indicate some directions for future work.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

2 Data collection and analysis

2.1 Data collection

2.1.1 *Electromagnetic Articulometry*. Electromagnetic Articulometry systems are highly appropriate tools for the collection of real-time lingual articulatory movement data. In Electromagnetic Articulometry (EMA; Perkell *et al*. 1992), an electromagnetic field is used to track movements of small receiver coils, about 2 mm in diameter, attached to the speech articulators using special adhesive. The transmitter coils produce alternating magnetic fields at different frequencies in the range of about 10 kHz. Receiver coils passing through this field generate an electric signal. The voltage of the signal is directly related to the distance and orientation of the receivers relative to the transmitter coils. This relationship is used to calculate the position of the receivers as a function of time.

Our data were collected from one male speaker of the Oujda dialect of MA using the three-dimensional system at the *Institut für Phonetik und Sprachverarbeitung*, Munich (Hoole *et al*. 2003).[4] Six transmitter coils were fixed on a cubic plastic apparatus surrounding the speaker's head. EMA receivers were placed on the speaker's jaw, tongue tip, lower lip, upper lip and tongue back.[5] Audio data was collected concurrently with a directional microphone at a sampling rate of 24 kHz.

2.1.2 *Materials*. Two sets of stimuli were constructed and presented to the same speaker in different recording sessions. The first set included three or four repetitions of each of the C ~ CC minimal pairs in Table I, for a total of 36 tokens. The words in each pair differed only in the number of initial consonants in the word. Target words were produced in the carrier phrase /galha__tɨlt mɨrrat/ 'he said to her__three times'.

C		CC	
bal	'to urinate'	dbal	'to fade'
tab	'to repent'	ktab	'book'
lih	'for him'	glih	'to grill'
bati	'to spend the night'	sbati	'belt'
bula	'urine'	sbula	'thorn'

Table I
Stimuli for first recording.

[4] One of the advantages of the 3D system is that it does not require fixating the speaker's head, a cumbersome necessity in 2D EMA systems, where a stabilisation apparatus must be used to maintain alignment between the device's measurement plane and the speaker's midsagittal plane (Hoole *et al*. 2003).

[5] The voltages in the receiver coils were captured at a sampling rate of 200 Hz. Voltage-to-distance conversions were conducted with a filter cut-off of 40 Hz for tongue tip and 20 Hz for all other articulators.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

In the second recording session, two sets of triads, given in Table II, differing only in the number of initial consonants, were produced ten times each for a total of 60 tokens. One set consisted of only real words; the other consisted of nonce words.[6] Target stimuli were produced in the carrier phrase /ʒibi__hnaja/ 'bring__here'. For both sessions, words in the target lists were separated by fillers, which consisted of unrelated words included for a separate experiment.

C	CC	CC
bulha 'her urine' dulha (nonce)	sbulha 'her ear (of grain)' kdulha (nonce)	ksbulha 'they owned it for her' bkdulha (nonce)

Table II
Stimuli for second recording.

2.2 Data analysis

The articulatory dataset produced by the EMA recordings was analysed using a MATLAB-based program, developed at Haskins Laboratories by Mark Tiede and adapted to our data by us. The program displays the positional signal acquired from the EMA data in synchrony with the acoustic waveform, spectrogram and velocity trajectories, which were computed by differentiating the positional signal of each receiver.[7]

The EMA receiver used to delineate movements associated with a consonant was the one corresponding to the consonant's primary oral articulator – tongue tip: [d t], lower lip: [b], tongue back: [g k]. The tongue-back receiver was also used to identify the vowels: [u a i]. Figure 2 displays the data collected for the /bulha ~ sbulha ~ ksbulha/ triad. Each word corresponds to a set of panels. Figure 2a corresponds to /bulha/, Fig. 2b to /sbulha/ and Fig. 2c to /ksbulha/. For each word, the positional signal in the *y*-dimension is shown.[8] Each panel shows ten trajectories (grey lines) corresponding to individual repetitions of the word along with a highlighted ensemble average (black line).

To facilitate discussion, three vertical lines are drawn for each word. The repetitions in Fig. 2 are aligned with respect to the maximal vertical

[6] Nonce words allowed us to test sequences of one, two and three stops. Although these sequences are phonotactically licit in MA, we were unable to find matched triads differing only in the number of initial stops amongst attested words in the lexicon.

[7] Additionally, the velocity signal obtained by calculating the first derivative of the positional signal was subjected to a five-point rectangular window moving average filter.

[8] The *y*-dimension is shown here for simplicity in presentation. As described below, the analysis was based on information in the vertical and horizontal dimensions combined.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

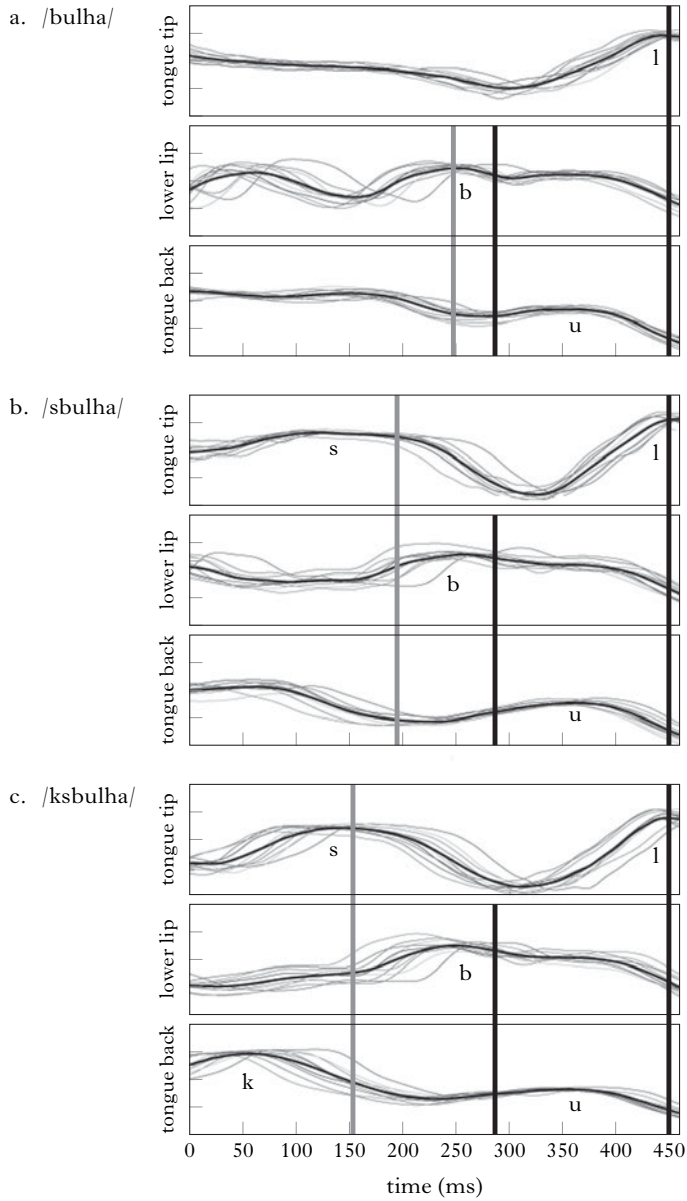


Figure 2

Positional signal in the *y*-dimension for 3 different receivers (tongue tip, lower lip and tongue back) for 10 repetitions each of (a) /bulha/, (b) /sbulha/ and (c) /ksbulha/. The leftmost vertical line demarcates the midpoint of the initial consonant cluster. The middle vertical line demarcates the release of the prevocalic consonant [b]. The rightmost vertical line demarcates the point of maximum constriction in the postvocalic consonant [l]. Distances between tickmarks on the *y*-axis correspond to 5 mm.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

displacement of the tongue-tip movement corresponding to [l]. The rightmost vertical line for each word is drawn at this anchor point. A second vertical line is drawn at the release of the lower lip constriction for the [b]. Another vertical line is drawn at the centre of the word-initial consonant cluster.

Inspecting these data leads to a number of impressionistic observations. First, the timing of the release of [b] relative to the vowel does not seem to change much across word types. Although we will quantify this stability more rigorously, for now it suffices to indicate it by the alignment of the vertical black lines across different words. Second, in contrast to what is observed for the release of [b], the centre of the consonant cluster gets farther away from the anchor point with each consonant added to the string. This is indicated by the progressive leftward shift of the vertical grey lines from /bulha/ to /sbulha/ to /ksbulha/.

Evaluating syllable-structure hypotheses from measures of interval stability requires two additional steps. First, we must extract the relevant intervals from continuous data. Second, we must evaluate the variability of these intervals across repetitions and across words. To quantify the impressionistic observations stated above, we defined intervals on articulatory landmarks extracted from the data. We now define and exemplify those landmarks.

The temporal life of a linguistically relevant articulatory movement can be decomposed into a series of landmarks (Gafos 2002: 276). These landmarks include the onset of movement toward a target, the achievement of a target, the onset of movement away from that target and the offset of movement. In the remainder of the paper we reserve the term onset to refer to a syllable onset and refer to the onset of articulatory movement toward a target as START. Other labels for articulatory landmarks are listed in (1).

(1) *Articulatory landmark labels*
   a. START: the onset of movement towards an articulatory target
   b. TARGET: achievement of an articulatory target
   c. RELEASE: the onset of movement away from an articulatory target
   d. END: the offset of controlled movement away from an articulatory target

For each segment, the landmarks in (1) were identified by referencing the velocity signal of the relevant articulator. Figure 3 shows the positional (top) and velocity (bottom) signals for the tongue-back receiver during the production of [k] in /ksbulha/. The landmarks defined above are labelled on the positional signal. Velocity peaks are labelled on the velocity signal. The peaks in tangential velocity correspond to movements to and away from linguistic targets. Troughs in the velocity signal correspond to the hold phase of a segment's articulation, also known as the plateau. The timestamps of START and TARGET were obtained by referencing the velocity

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

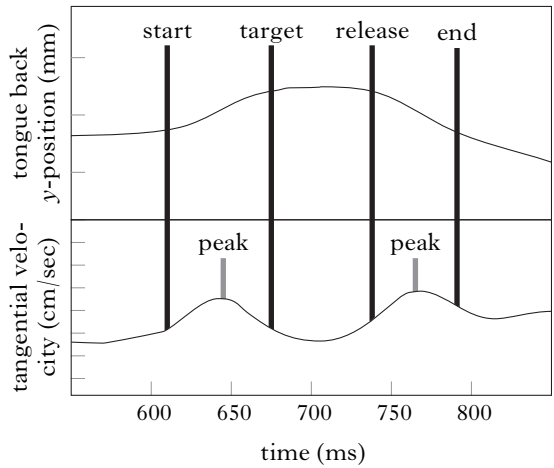


Figure 3

An example of measurements from the [k] portion of /ksbulha/. The top panel shows the positional signal in the vertical dimension of the tongue-back receiver (*y*-axis in mm) as a function of time (*x*-axis in msec); distances between tickmarks on the *y*-axis correspond to 5 mm. The bottom panel shows the instantaneous magnitude of the tangential velocity of the same receiver (in cm/sec); distances between tickmarks on the *y*-axis correspond to 5 cm/sec.

peak associated with movement towards a target constriction. The timestamps of RELEASE and END were obtained by referencing the velocity peak associated with movement away from constriction. The above landmarks were identified automatically by an algorithm. The algorithm locates the timestamp at which instantaneous velocity exceeds, in the case of START and RELEASE, or falls below, in the case of TARGET and END, a set percentage of the velocity peak associated with movement toward or away from an articulatory target.

Parsing articulatory landmarks from the continuous signal using the above procedure yields a series of timestamps intended to index specific articulatory events. In the following sections, these timestamps are used to quantify patterns of temporal organisation corresponding to different syllable-structure hypotheses.

3 Stability analysis

In this section, we use temporal stability to diagnose syllable structure. This involves comparing the duration of temporal intervals across different cluster sizes. Competing hypotheses on syllable structure implicate different patterns of interval stability. These differing predictions allow us to test for the presence of a syllable boundary between the consonants of initial clusters in MA.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

In order to evaluate which pattern of stability best characterises our MA data, we measured intervals from the left edge, the centre and the right edge of consonant clusters to a common anchor. We conducted this analysis twice, once using an anchor defined by a landmark in the post-vocalic consonant (§3.1) and once using an anchor defined by a landmark in the vowel (§3.2). Both of these analyses used the same consonant-cluster landmarks to delimit the left edge of intervals. That is, regardless of the anchor, the left edge of the consonant cluster was defined as the TARGET landmark of the first consonant in that cluster, e.g. b^{target} for /bulha/, s^{target} for /sbulha/, k^{target} for /ksbulha/. The right edge was defined as the RELEASE landmark of the immediately prevocalic consonant, e.g. b^{release} in /bulha/, /sbulha/ and /ksbulha/. The centre of the cluster was calculated by taking the mean of the midpoints of each consonant in the cluster, where consonant midpoint refers to the point equidistant from the TARGET and RELEASE landmarks in each consonant. This way of identifying the centre of the cluster involves a contribution from each prevocalic consonant. It therefore provides a global index of the entire consonant cluster (Browman & Goldstein 1988).

For each interval, a measure of its stability must be employed. The standard deviation (SD) of an interval's duration is one such widely used measure. However, in our case the intervals we wish to compare have intrinsically different means (i.e. the mean duration of the left edge to anchor interval and centre to anchor interval will always be larger than the mean duration of the right edge to anchor interval). Therefore, the relative standard deviation (RSD) or 'coefficient of variance', as calculated in (2), provides a better measure of stability for our purposes (Frank & Althoen 1995: 58–59).[9]

(2) *Relative standard deviation* (RSD)

$$\text{RSD} = 100 \times \text{standard deviation} / \text{mean}$$

To quantify the statistical reliability of differences in interval stability, we make use of the repeated measures ANOVA model. This model allows for multiple ('repeated') measurements of the same dependent variable under different conditions. In our specific case, the dependent variable is interval duration. The model evaluates whether the effect of phonetic context, i.e. cluster size and segment identity, is uniform across INTERVAL TYPE (left edge to anchor, centre to anchor, right edge to anchor). A significant interaction between INTERVAL TYPE and CLUSTER SIZE would indicate that some intervals are more stable across contexts than others. Both proposals on syllable structure predict such interaction. But, crucially, the complex and simplex onset hypotheses predict different sources

[9] For more on our decision to use RSD to index stability, see the methodological remarks in §3.3.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

	left edge to anchor			centre to anchor			right edge to anchor		
	mean	SD	RSD	mean	SD	RSD	mean	SD	RSD
bulha~sbulha ~ksbulha	333	82	24·6%	261	41	15·9%	200	22	11·2%
dulha~kdulha ~bkdulha	344	77	22·2%	258	46	17·7%	185	20	10·7%
bal~dbal	409	84	20·5%	337	33	9·7%	284	15	5·1%
tab~ktab	394	27	6·8%	315	18	5·7%	257	14	5·5%
bati~sbati	368	77	20·9%	294	27	9·1%	243	14	5·8%
bula~sbula	381	84	22·0%	302	34	11·1%	251	19	7·3%
set average	371	72	20%	295	33	12%	237	17	8%

Table III

Mean, SD and RSD of intervals delineated by three points in the consonantal cluster and a consonantal anchor for word sets containing dyads or triads differing in the number of initial consonants.

of interaction. The simplex onset hypothesis predicts an effect of cluster size on the left edge to anchor interval and the centre to anchor interval, but not on the right edge to anchor interval. The complex onset hypothesis predicts an effect of CLUSTER SIZE on the left edge to anchor interval and the right edge to anchor interval but not on the centre to anchor interval.

3.1 Consonantal anchor

We first report on intervals delimited by a consonantal anchor. This anchor was defined by the point of maximum constriction of the postvocalic consonant for each word set. For the set /bulha ~ sbulha ~ ksbulha/ the postvocalic consonant is [l], for /bati ~ sbati/ the postvocalic consonant is [t] and so on. Since there is no oral constriction in the postvocalic consonant of /lih ~ glih/, this set did not contribute any measurements.

Table III breaks the data down by stimulus set. We report the mean, SD and RSD of the measured intervals. The interval with the lowest RSD in each set is shaded. The table shows that for each word set (row), the right edge to anchor is the interval with the lowest RSD across tokens. In stark contrast, Browman & Goldstein (1988) find that for English data, the SD of the interval delineated by the right edge to anchor is roughly twice the SD of the centre to anchor interval. In our MA data, however, the SD is lower for the right edge to anchor interval in all cases. Thus, unlike in English, the centre to anchor interval has no clear stability advantage over other measured intervals.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

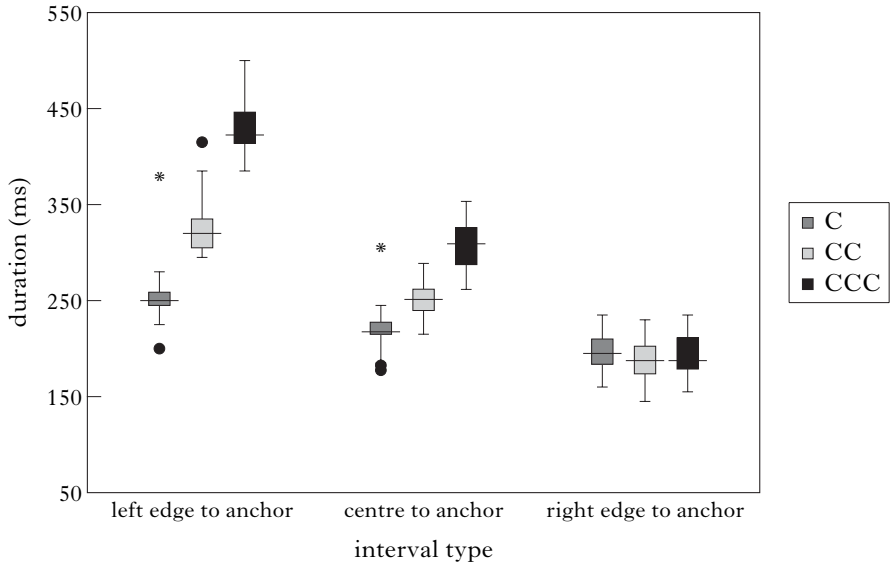


Figure 4
Intervals delineated by the left edge, right edge and centre of initial consonant clusters and a consonantal anchor.

Although Table III shows that right edge to anchor has the lowest RSD for all comparisons in our corpus, we would like to evaluate the statistical reliability of this difference, both in our corpus and in the general case. The RSD of the centre to anchor interval for the /bulha ~ sbulha ~ ksbulha/ triad, for example, was 15·9% compared with 11·2% for right edge to anchor. Is this a large enough difference to declare that the right edge is more stable?

To provide a statistical test of differences in interval variability, we subjected the set of interval durations for tokens comprising the /bulha ~ sbulha ~ ksbulha/ and /dulha ~ kdulha ~ bkdulha/ sets ($n = 60$) to a repeated measures ANOVA, with Greenhouse-Geisser correction. These sets were chosen because they have the largest number of repetitions per word and are matched for vowel quality and anchor-consonant identity. The repeated measures ANOVA model treats the different intervals (left edge to anchor, centre to anchor, right edge to anchor) as a within-token variable INTERVAL TYPE, and the number of consonants in the initial cluster (C, CC, CCC) and the segmental context (/ksb/, /bkd/) as across-token variables, CLUSTER SIZE and SEGMENT IDENTITY respectively. The box plots in Fig. 4 collapse across segmental contexts to illustrate the main trends in our data with respect to INTERVAL TYPE and CLUSTER SIZE. As the number of consonants in the initial cluster increases, the median duration of intervals measured from the left edge to anchor and centre to anchor

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

increase sharply. In contrast, the right edge to anchor interval varies little as a function of cluster size.

The main within-token effect of INTERVAL TYPE ($F(1{\cdot}08, 58{\cdot}07) = 2351{\cdot}11$, $p < 0{\cdot}001$) and the main across-token effect of CLUSTER SIZE ($F(2, 54) = 67{\cdot}61$, $p < 0{\cdot}001$) were significant. The main across-token effect of SEGMENT IDENTITY ($F(1, 54) = 2{\cdot}63$, $p = 0{\cdot}111$) and the interaction between CLUSTER SIZE and SEGMENT IDENTITY ($F(2, 54) = 1{\cdot}03$, $p = 0{\cdot}363$) were not significant. There was a significant interaction between INTERVAL TYPE and CLUSTER SIZE ($F(2{\cdot}15, 58{\cdot}07) = 296{\cdot}68$, $p < 0{\cdot}001$). This indicates that the effect of CLUSTER SIZE is not uniform across all intervals. Separate ANOVAs were conducted post hoc for each level of INTERVAL TYPE, with CLUSTER SIZE as an independent variable. Significant results were obtained for left edge to anchor ($F(2, 54) = 167{\cdot}22$, $p < 0{\cdot}001$) and for centre to anchor ($F(2, 54) = 69{\cdot}38$, $p < 0{\cdot}001$), indicating that the addition of a consonant reliably affected the timing of these intervals. Crucially, the effect of CLUSTER SIZE on the right edge to anchor interval was not significant ($F(2, 54) = 0{\cdot}688$, $p = 0{\cdot}507$). This shows that dividing the right edge to anchor interval measurements into groups based on the number of consonants in a cluster does not account for more variance than treating all right edge to anchor intervals as a homogeneous group. This result is in complete conformity with the simplex onset-alignment schema.

Post hoc tests were conducted to determine the source of the significant effect of CLUSTER SIZE on the left edge to anchor and centre to anchor intervals. A significant difference between #CV and #CCV for the centre to anchor interval would argue against the complex onset hypothesis. This would indicate that the centre to anchor interval is not stable across contexts.[10] The post hoc test results were that for both intervals, all comparisons (C *vs.* CC; CC *vs.* CCC; C *vs.* CCC) show significant differences at the $p < 0{\cdot}001$ criterion.

[10] Resyllabification of initial consonants in target words into the coda position of the preceding word in the carrier phrase, e.g. /ʒibi **k**sbulha hnaja → ʒibi**k** sbulha hnaja/, is plausible since all target words were preceded by a vowel. Although resyllabification would not affect how we calculate the RIGHT EDGE TO ANCHOR interval, it would affect calculation of the CENTRE TO ANCHOR interval. The complex onset hypothesis predicts that the interval left-delimited by the centre of the onset cluster is the most stable interval. If, however, the initial consonant in biconsonantal and triconsonantal clusters was syllabified with the preceding word in the carrier phrase, then the CENTRE TO ANCHOR interval evaluated above does not faithfully test the predictions of the complex onset hypothesis. We therefore recalculated the CENTRE TO ANCHOR interval on the basis of the assumption that initial consonants are resyllabified in #(C)CCV words. Under a resyllabification scenario, the two and three consonant clusters would be reduced to one and two consonants clusters respectively. The CENTRE TO ANCHOR interval calculated on the basis of these consonants is also not stable across words. We conducted an ANOVA on this version of the CENTRE TO ANCHOR interval with CLUSTER SIZE and SEGMENT IDENTITY as independent variables. As above, there was a significant main effect of CLUSTER SIZE ($F(2, 54) = 12{\cdot}20$, $p < 0{\cdot}001$). The interaction between CLUSTER SIZE and SEGMENT IDENTITY was not significant ($F(2, 54) = 1{\cdot}16$, $p = 0{\cdot}322$). These results verify that, under the resyllabification scenario as well, the pattern of stability predicted by the complex onset hypothesis is not apparent in our data.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

In sum, each consonant added before #CV, i.e. #((C)C)CV, significantly disrupts the duration of the left edge to anchor and centre to anchor intervals but not the right edge to anchor interval. Together with the comparison of RSDs above, these results provide support for the simplex onset hypothesis.[11] The phonetic heuristic for simplex onsets is right edge to anchor interval stability. We found that across clusters of different sizes, the right edge to anchor interval is more stable than the left edge to anchor or centre to anchor intervals.

3.2 Vocalic anchor

In this section we report a stability analysis using the END landmark of the vowel, V^{end}, as the anchor point. Table IV reports the mean, SD and RSD of the three interval types for each stimulus set.

Overall, intervals delineated by the vocalic anchor had higher SD across all measurements and word types than intervals delineated by the consonantal anchor. This difference is particularly salient for closed syllables. For several of the sets in Table IV, both the SD and RSD values are higher than in Table III for every measured interval. For example, in the /bal ~ dbal/ set from Table III, the SDs of the centre to anchor and right edge to anchor intervals were 33 ms and 15 ms respectively. When V^{end} was used as the anchor, the SD of these durations soared to 68 ms and 63 ms, as shown in Table IV. In the presence of this overall increase in interval variability, we see that the right edge to anchor interval no longer has a smaller RSD than the centre to anchor interval for all sets. For four out of the six word sets represented in both Table III and Table IV, the right edge to anchor interval lost its stability advantage. For /bulha ~ sbulha ~ ksbulha/, /bula ~ sbula/, /bal ~ dbal/ and /tab ~ ktab/, the centre to anchor interval has a lower RSD than the right edge to anchor interval. This is the pattern predicted by the complex onset hypothesis.

To determine whether the use of a vocalic anchor affected the qualitative results of the stability analysis, we again employ the repeated measures ANOVA, with Greenhouse-Geisser correction, on interval

[11] An anonymous reviewer points out that the distribution of laryngeal features offers another set of facts that can be used to evaluate syllable structure. These facts appear to support the simplex onset hypothesis in MA. A prohibition on mixed voicing onsets has been proposed as a linguistic universal (Kehrein & Golston 2004; see also Greenberg 1978, Lombardi 1995). MA permits a rather complex set of consonant combinations in initial position – all combinations of labial, coronal and dorsal places of articulations are attested, e.g. [gd dg kb bk tb bt] are possible. In addition, the laryngeal specifications of these clusters seem unrestricted, e.g. /kbaʃ/ 'sheep (PL)', /bkat/ 'she cries'. This distinguishes the initial clusters of MA from, for example, those of Georgian (Vogt 1971). Georgian allows comparable combinatorial freedom in place of articulation specifications for consonants in initial position but disallows mixed voicing, e.g. [bk] and [bt] are not attested. If MA has simplex onsets and Georgian has complex onsets (as claimed by Vogt 1971), the distribution of voicing features in these languages remains consistent with the cross-linguistic generalisation regarding laryngeal features.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

	left edge to anchor			centre to anchor			right edge to anchor		
	mean	SD	RSD	mean	SD	RSD	mean	SD	RSD
bulha~sbulha ~ksbulha	359	86	23·9%	287	51	17·8%	227	41	18·2%
dulha~kdulha ~bkdulha	323	92	28·5%	246	55	22·3%	172	35	20·3%
bal~dbal	373	103	27·5%	301	68	22·7%	249	63	25·3%
tab~ktab	394	48	12·2%	316	24	7·7%	258	26	10·0%
bati~sbati	379	73	19·3%	304	22	7·1%	254	13	5·2%
bula~sbula	498	73	14·6%	419	27	6·5%	369	26	6·9%
lih~glih	448	83	18·5%	355	38	10·7%	282	8	2·7%
set average	396	80	20·6%	318	41	13·6%	259	30	12·7%

Table IV

Mean, SD and RSD of intervals delineated by three points in the consonantal cluster and a vocalic anchor for word sets containing dyads or triads differing in the number of initial consonants.

durations. The independent variables were the same as the repeated measures ANOVA for the consonantal anchor reported above.

Figure 5 provides box plots, parallel to Fig. 4, of the intervals from consonant cluster landmarks to the V^{end} anchor by cluster size. Again, the intervals are collapsed across segment types. The general pattern of increasing left edge to anchor and centre to anchor intervals and fairly stable right edge to anchor intervals is the same as in Fig. 4. The most prominent difference between Figs 4 and 5 is that the durations in Fig. 5 are dispersed more widely about the median than those in Fig. 4. This indicates that the variance of intervals delineated by vocalic anchors is higher than the variance of intervals delineated by consonantal anchors. We now examine how this difference in overall variability affects the statistical results.

The results of the ANOVA show significant main effects of the within-token variable INTERVAL TYPE ($F(1·05, 56·47) = 2048·74$, $p < 0·001$) and a main effect of the across-token variables of CLUSTER SIZE ($F(2, 54) = 28·80$, $p < 0·001$) and SEGMENT IDENTITY ($F(1, 54) = 20·55$, $p < 0·001$). The interaction between CLUSTER SIZE and SEGMENT IDENTITY ($F(2, 54) = 1·96$, $p = 0·151$) was not significant. The significant effect of SEGMENT IDENTITY indicates that the absolute durations of the intervals are different depending on the segmental identity of the initial consonants. This factor did not reach significance for consonantal anchors above. However, the lack of interactions between SEGMENT IDENTITY and CLUSTER SIZE indicates that the effect of SEGMENT IDENTITY is of no theoretical importance with respect to the hypotheses at stake here.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

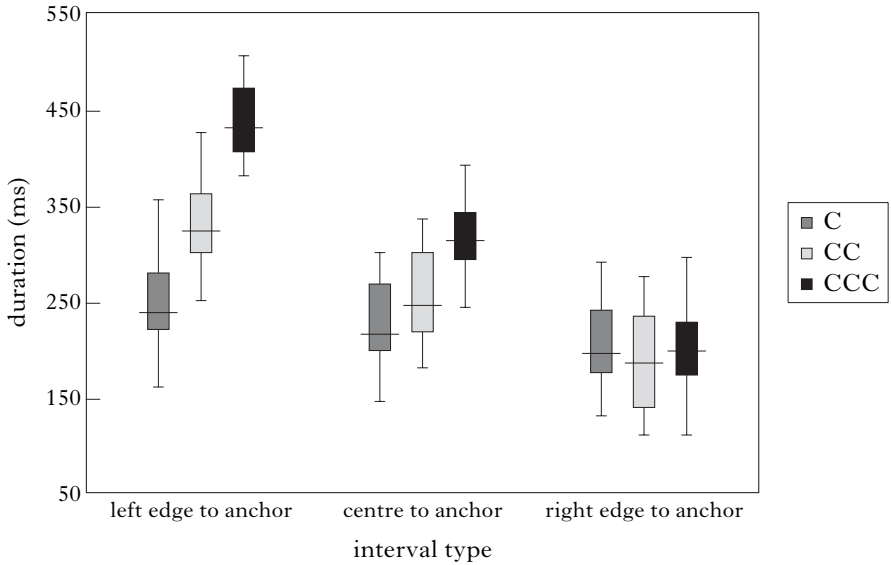


Figure 5
Intervals delineated by the left edge, right edge and centre of initial consonant clusters and a vocalic anchor.

As with the consonantal anchor intervals above, there was a significant interaction between INTERVAL TYPE and CLUSTER SIZE ($F(2{\cdot}09, 58{\cdot}07) = 318{\cdot}30$, $p < 0{\cdot}001$). To further investigate this interaction, one-way ANOVAs were conducted for each level of INTERVAL TYPE, with CLUSTER SIZE as the independent variable. Results showed a significant effect of CLUSTER SIZE on the duration of the left edge to anchor interval ($F(2, 54) = 99{\cdot}20$, $p < 0{\cdot}001$) and on the duration of the centre to anchor interval ($F(2, 54) = 27{\cdot}79$, $p < 0{\cdot}001$). The effect of CLUSTER SIZE on the right edge to anchor interval ($F(2, 54) = 1{\cdot}15$, $p = 0{\cdot}323$) was not significant. This indicates that, with V^{end} as the anchor as well, adding consonants to the initial cluster disrupts the stability of the centre to anchor interval and the left edge to anchor interval, but not the right edge to anchor interval. Thus the main effects obtained for the consonantal anchor above are duplicated with the vocalic anchor.

Post hoc tests with left edge to anchor and centre to anchor intervals as dependent variables allow us to determine the locus of the significant effect of CLUSTER SIZE. For the left edge to anchor interval, all levels of CLUSTER SIZE, C, CC and CCC, were significantly different from each other ($p < 0{\cdot}001$). For the centre to anchor interval, Tukey's post hoc shows that the significant effect of CLUSTER SIZE is attributable to a difference between the C/CC and CCC ($p < 0{\cdot}001$) contexts and that there is no significant difference between C and CC ($p = 0{\cdot}14$). This is a departure from the results obtained above with the consonantal anchor. Lack of

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

significant differences for the centre to anchor interval between C and CC contexts is consistent with the stability pattern characteristic of complex onset alignment. This result was not obtained for consonantal anchor intervals. On the basis of stability-based phonetic heuristics, then, this specific result would be consistent with a tautosyllabic parse of biconsonantal clusters (#CCV) and a heterosyllabic parse of triconsonantal clusters (#C.CCV or #CC.CV). Thus some evidence in our data is compatible with the complex onset hypothesis.

3.3 Summary and methodological remarks

The stability analyses in the preceding sections aim at evaluating competing theoretical proposals on syllable structure. In this section, we summarise their main results and highlight certain methodological points. Given the relative paucity of work in this area, a clarification of our methodological choices seems useful for future work. Our methodology builds on past work, but is also different in important ways.

First, as a measure of interval stability, we proposed the use of RSD as opposed to SD. It is an established property of human motor behaviour that the variance of an interval between two (timed) events increases with increasing mean temporal duration between the two events (Wing & Kristofferson 1973, Schöner 2002). This means that longer intervals will have larger SDs. In our data, robust correlations were obtained between the SD (the square root of the variance) and the mean of each interval type delimited by a consonantal anchor ($r(16) = 0{\cdot}534$, $p < 0{\cdot}05$). Thus, using raw SD to index stability constitutes a bias towards finding better stability for the shorter right edge to anchor interval than the longer centre to anchor and left edge to anchor intervals. To illustrate this, consider the SD of intervals delineated by the vocalic anchor. For the /bulha ~ sbulha ~ ksbulha/ triad, the SD of the right edge to anchor interval (41 ms) was lower than the SD of the centre to anchor interval (51 ms), yet the RSD of the right edge to anchor interval (18·2 %) was slightly higher than the centre to anchor interval (17·8 %). Likewise, for /bal ~ dbal/, the right edge to anchor interval had the lowest SD, but the centre to anchor interval had the lowest RSD. Thus different measures of stability would lead to different interpretations of the same data. In our case, the RSD measure of stability is more conservative than the SD measure because it factors out the effect of interval length on stability.

Another way in which our analysis departs from past work is in the use of articulatorily defined anchors. Past studies have used anchors defined by salient acoustic events. These studies used the achievement of closure of a voiceless (Browman & Goldstein 1988, Byrd 1995) or voiced plosive (Honorof & Browman 1995), obtained by visual inspection of the acoustic waveform. In the first study of this nature, Browman & Goldstein (1988) motivate this choice primarily from practical concerns. For plosives, the acoustic record provides a clear, consistent landmark, which offers an alternative to identifying a vowel in the continuous

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

articulatory signal. Byrd (1995: 290), however, points out that the validity of this anchor point rests on the assumption that the 'the vocalic gesture stands in a consistent relationship with its acoustic offset'. We conducted stability analyses with consonantal and vocalic anchors defined independently and articulatorily. By using articulatory landmarks as anchors we eliminate the restriction that the stability analysis be conducted on syllables closed with a plosive. This is advantageous when testing hypotheses about syllable structure. Typically, such hypotheses hold across a wide range of segmental identities. In addition, we avoid assumptions regarding the consistency of the relation between the movement of articulators and the acoustic offset of the vowel.

Using the methods described above led to clear results when intervals were delimited by a consonantal anchor. In these cases, the right edge to anchor was the most stable interval across cluster sizes for all dyads and triads in our data. Stability-based heuristics for syllable structure point clearly to the simplex onset hypothesis. When the V^{end} landmark was used as an anchor, then the variability of all intervals increased. For these more variable intervals, we obtained apparently conflicting results. For some dyads and triads, we saw the reverse pattern, whereby the centre to anchor interval had a lower RSD than the right edge to anchor interval. Additionally, the effect of CLUSTER SIZE on the centre to vocalic anchor interval was not significant for comparisons between #CV and #CCV. Both of these results are consistent with the complex onset hypothesis.

In what follows, we seek to better understand these seemingly contradictory results by investigating how stability-based indexes of syllable structure change under different conditions. To anticipate the main result, the analytic tools developed in the next section reveal that both patterns of stability obtained for vocalic anchors are consistent with a single invariant phonological structure.

4 A computational model of interval stability

To make sense of the apparently conflicting evidence in our data, we resort to a model that allows us to study how patterns of temporal stability respond to increases in variability. The model generates articulatory landmarks from a given syllable structure and therefore makes explicit predictions about how that structure affects the stability of temporal intervals. In addition to syllable structure, however, there are other factors (speech rate, lexical statistics of the target word, measurement error, etc.) that may affect the stability of temporal intervals. Our present focus is not the precise nature of the variability source, but rather the way in which variability affects stability-based indexes of syllable structure.

To study this, we introduce random noise into all intervals uniformly and study how stability-based indexes of syllable structure change under conditions of gradually increasing variability. This methodology leads to a novel result regarding the relation between syllable structure and stability

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

patterns. The pattern of stability whereby the centre to anchor interval is more stable than the left edge to anchor and right edge to anchor intervals has been seen as a reflex of complex syllable onsets (Browman & Goldstein 1988: 153). This does not mean, however, that this stability pattern necessarily implicates a complex syllable onset. The model developed in this section allows us to evaluate the full range of quantitative instantiations of a phonological structure. Simulations based on the model reveal that even when vowels are timed locally to the prevocalic consonant, as in the simplex onset alignment schema, the centre to anchor interval has a lower RSD than the right edge to anchor interval under certain noise conditions. We refer to such cases as (stability) reversals. In light of these results, we re-examine our experimental data and find that they are in full accord with a model that embodies the predictions of the simplex onset hypothesis. This enables a clear interpretation of the mixed stability results in §3. Overall, the model allows us to make sense of seemingly inconsistent results by illuminating how stability-based indexes are modulated under different conditions of variability.

4.1 Model parameters

The patterns of interval stability in our data by and large provided evidence for the simplex onset hypothesis. In some corners of the data, however, we found patterns of stability consistent with the complex onset hypothesis. These patterns never occurred for intervals delimited by consonantal anchors. However, they did arise for some intervals delimited by vocalic anchors. The latter intervals showed higher SD than the former for every interval type. It seems likely that this higher variability is related to the stability index in a way that may lead to an understanding of these results. The basic proposal is as follows. Consider two intervals of different mean durations. As the SD of these intervals is uniformly increased, the RSD increases at a slower rate for the longer interval than for the shorter interval. This is because RSD divides SD by a larger mean for the longer interval. Thus, the effect of variability on the RSD of an interval is modulated by the mean size of that interval. The longer interval is less sensitive to uniform increases in SD than the shorter one. This indicates that interval stability as a heuristic for syllable structure may be valid only within certain ranges of variability. At high levels of variability, the longer centre to anchor interval (as well as the left edge to anchor interval) may show better stability than the shorter right edge to anchor interval, regardless of phonological structure. Such stability reversals obscure the canonical stability pattern imparted by simplex onsets, where the right edge to anchor interval is expected to be the most stable interval. But as the above logic reveals, such a patterning is in fact consistent with the simplex onset hypothesis. The model described below tests this logic. Our modelling approach is related to earlier work in Gafos (2002), who models the acoustic consequences of posited coordination relations, and also Smith (1995), who pursues the modelling of the

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

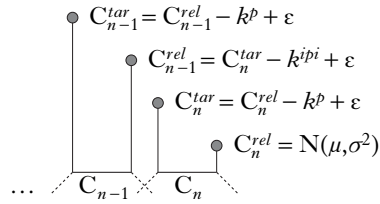


Figure 6
Generating landmark timestamps for #((C)C)CV words.

relation between rhythmic structure and articulatory data in Japanese and Italian.

Given a set of word types, #CV, #CCV, #CCCV, the model generates articulatory landmarks defining the plateau of each constituent segment. These landmarks are generated from stochastic versions of local timing relations between consonants and vowels (following Gafos 2002). The algorithm for generating landmarks is summarised in Fig. 6. Landmark generation proceeds by first selecting the timestamp of the RELEASE landmark of the immediately prevocalic consonant, C^{rel}_n, from a Gaussian distribution. The immediately preceding landmark, the TARGET of that consonant, C^{tar}_n, is then generated by subtracting consonant plateau duration, k^p, from C^{rel}_n and adding a noise term. These two landmarks, C^{tar}_n and C^{rel}_n, define the plateau of the immediately prevocalic consonant. For words with two or more initial consonants, the RELEASE landmark of the preceding consonant C_{n-1} (C_1 in #C_1C_2V words and C_2 in #$C_1C_2C_3$V words), was generated by reference to C^{tar}_n. The inter-plateau interval, k^{ipi}, was subtracted from C^{tar}_n, and a noise term was added. As before, C^{tar}_{n-1} was generated by subtracting plateau duration from C^{rel}_{n-1} and adding a noise term. Landmarks for the initial consonant in #CCCV words were generated following the same procedure.

In the simulation reported below, k^{ipi} and k^p were set to values reflecting averages across our data, 40 ms and 30 ms respectively. The precise values of these constants, however, are not crucial in determining the overall behaviour of the model. The model produces the same qualitative results with a range of k^p and k^{ipi} values. The noise term, ε, was generated from a normal distribution of Gaussian noise, with a mean of 0 ms and a SD of 20 ms. The C^{rel} of the immediately prevocalic consonant was selected from a Gaussian distribution, with a mean of 500 ms and a variance of 400 ms. The value for this mean is arbitrary. We choose 500 ms for clarity in presentation, as it ensures that temporally preceding timestamps will be positive. The value for the variance was set to match the noise term associated with other landmarks in the simulation.

The alignment of the vowel to the consonant cluster was in accord with the hypothesis that MA has simplex onsets. Thus, regardless of the number of consonants, the vowel was timed locally and directly to the prevocalic consonant. The anchor point was generated for each consonant

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

cluster by adding a constant, k^v, representing vowel length, to the time-stamp of the midpoint of the prevocalic consonant. The value of k^v used for the simulation results reported below was 250 ms. However, as with the other constants, the same qualitative result was obtained for a wide range of k^v values. The specific value of k^v chosen reflects the mean vowel length in our data, measured from the right edge of the consonant cluster to the end of the vowel.

To introduce variability into the intervals, a noise term with a mean of 0 ms and gradually increasing standard deviation was added to the anchor. There were a total of 20 stepwise increases. This yielded 20 anchors for each cluster differing only in the level of noise associated with the anchor. For the first anchor (anchor 1), noise had a standard deviation of 20 ms. For each subsequent anchor, the standard deviation increased by steps of 5 ms, so that the last anchor point (anchor 20) had a standard deviation of 120 ms. Introducing noise by systematically changing anchor variability is one way to ensure that variability is uniformly present across all intervals right-delimited by that anchor, e.g. centre to anchor and right edge to anchor.[12] Since in this way variability is uniformly present across all intervals of interest, we refer to it as 'overall variability'.

4.2 Simulations

The model described above was used to generate sets of 30 tokens. Each set contained ten instances of each #CV, #CCV and #CCCV word. Using the generated data, we then calculated the relative standard deviation (RSD) of the three interval types, left edge to anchor, centre to anchor and right edge to anchor. Since the anchor points were generated by adding a constant to the immediately prevocalic consonant, the right edge to anchor interval is expected to be more stable than the other intervals, at least under conditions of low overall variability. However, at higher levels of variability, the RSD of the centre to anchor interval should be lower than the right edge to anchor interval.

The result of the stability analysis for 1000 iterations of 30 simulated tokens is given in Fig. 7. The *y*-axis shows the average RSD of the three intervals, left edge to anchor, centre to anchor and right edge to anchor for each anchor (anchor 1 to anchor 20), shown on the *x*-axis. As expected, the right edge to anchor interval is the most stable interval at low levels of variability (anchor 1 to anchor 6). It can also be seen, however, that as interval variability increases, the RSD of the right edge to anchor interval increases at a faster rate than the RSD of the centre to anchor interval.

[12] Introducing variability through the standard deviation of the anchors is merely a convenient stand-in for other sources of variability that may exist in realistic data. As pointed out earlier, at issue is not the nature of the variability source, but rather if and how different variability conditions, regardless of their source, change quantitative indexes of syllabic constituency such as stability measures.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

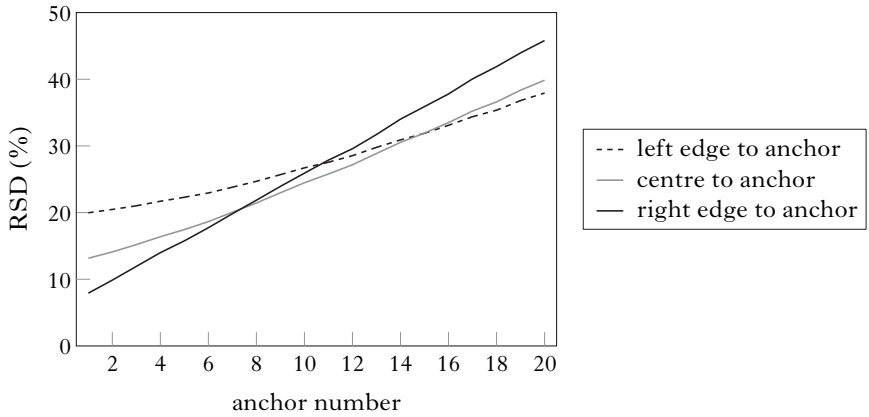


Figure 7

The relative standard deviation (*y*-axis) of three cluster to anchor intervals (left edge to anchor, centre to anchor, right edge to anchor) by anchors of increasing variability (*x*-axis).

Ultimately, when the SD of the anchor increases beyond that of anchor 7, the centre to anchor interval shows a lower RSD than the right edge to anchor interval. Thus, at high levels of overall variability, simplex onset alignment gives rise to the pattern whereby the centre to anchor interval is the most stable interval type. Although previous heuristics suggest that centre to anchor stability indicates a complex onset parse (Browman & Goldstein 1988: 153), the model qualifies this reasoning by exposing its full range of validity. Under conditions of high overall variability, simplex onset alignment may also generate patterns of centre to anchor interval stability. The simulation results thus confirm that simplex onset alignment can give rise to different patterns of stability under different conditions of variability.

Because the model is an explicit link between abstract phonological structure and its continuous indexes, it can serve as an analytical tool for reasoning about the former from the latter. Specifically, the model allows us to state precise predictions of the simplex onset hypothesis for patterns of interval stability in relation to overall variability. Although the model gives rise to both stability patterns attested in our data, it predicts stringent conditions on the levels of overall variability at with each pattern may occur. These predictions derive from the main trend shown in Fig. 7. As overall variability increases, stability declines faster for shorter intervals than for longer intervals. Thus, the simplex onset hypothesis, as embodied in the model, makes two related predictions about our data. The first concerns an overall trend. This prediction is that the stability advantage imparted to the right edge to anchor interval gradually diminishes in the presence of noise. The second prediction concerns the individual cases of stability reversals in our data. For dyads

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

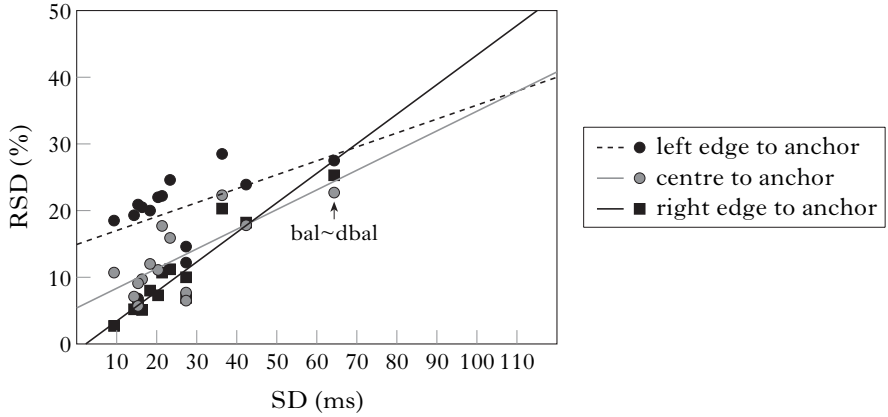


Figure 8

The relative standard deviation (*y*-axis) of three cluster to anchor intervals (left edge to anchor, centre to anchor, right edge to anchor) as a function of the SD of the right edge to anchor interval (*x*-axis). The best fitting line is shown for each interval type.

and triads that show reversals, the simplex onset hypothesis predicts an implicational relation between the stability pattern and the level of overall variability present in the intervals. We first evaluate the broad prediction of the model by comparing overall trends in the data to the simulation results in Fig. 7. We then evaluate the predictions for specific dyads and triads by zeroing in on just those word sets which showed stability reversals.

To see how stability patterns in our data change as variability changes, we can select an index of overall variability and plot the RSD for each interval in our data against this index. Figure 8 does this by redisplaying the RSD (*y*-axis) for each interval reported in Table III (consonantal anchors) and Table IV (vocalic anchors). RSD values are plotted as a function of an overall variability index (*x*-axis). This index is the SD of the right edge to anchor interval.[13] Each data point in the figure represents the RSD value for a specific interval at the level of variability present in the pairing of anchor and word set (dyads or triads) from which the interval was taken. For example, the overall level of variability in /bal ~ dbal/ intervals delineated by the vocalic anchor was 63 ms. At that value on the *x*-axis, three shapes are plotted. Each shape corresponds to an interval type, and its *y*-coordinate indicates relative stability in terms of RSD.

[13] This is a good index of overall variability, because its value is highly correlated with variability, in the rest of the intervals. That is, the SD of the right edge to anchor interval is highly correlated with the SD of the centre to anchor interval ($r(12) = 0{\cdot}935$, $p < 0{\cdot}001$), and with the SD of the left edge to anchor interval ($r(12) = 0{\cdot}675$, $p < 0{\cdot}05$).

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

To directly illustrate how variability influences stability patterns, regression lines were fit to each interval type. In effect, these lines show how RSD changes as a function of the variability present in the data. The lines were derived by fitting all of the data points in Fig. 8, collapsing across word sets and anchor types. The variability in the data comes from differences in anchors used to right-delimit intervals and also from differences between word sets. The lines show how the RSDs of different interval types behave at different levels of overall variability. The slopes of the lines can therefore be compared to the output of the model in Fig. 7. After looking first at the broad similarities between the data and the model, we then move on to examine the specific word sets that showed stability reversals in our data.

The key similarity between the simulated data in Fig. 7 and the experimental data in Fig. 8 is as follows. The lines in Fig. 7 produced by the model replicate the regression lines fit to the data in Fig. 8. Just as we saw in Fig. 7, the RSDs of the three interval types in the data in Fig. 8 increase at different rates as overall variability increases. More precisely, across the two figures, the RSD of the shortest interval type, right edge to anchor, increases at a faster rate than the RSD of the longer centre to anchor interval. Likewise, the centre to anchor interval increases at a faster rate than the longer left edge to anchor interval. This means that the stability advantage of the right edge to anchor interval in the data is gradiently linked to overall variability. This fact verifies a prediction of the model. As overall variability increases, the difference in RSD between the right edge to anchor and centre to anchor intervals decreases. A significant negative correlation ($r(26) = -0{\cdot}465$, $p < 0{\cdot}01$) between these variables (overall SD *vs.* difference in RSD between the right edge to anchor and centre to anchor intervals) validates the trend observable in Fig. 8.

We now zoom in from broad trends in the data to examine individual cases of stability reversals. For these cases, the model makes specific predictions. Given a word set (dyad or triad) and two sets of intervals delimited by different anchors extracted from that word set, the model embodying the simplex onset hypothesis predicts the following implicational relationship. If one set of intervals shows centre to anchor stability and the other set shows right edge to anchor stability, then the former set of intervals must have higher overall variability than the later. Further, the opposite relationship is precluded. This prediction allows us to evaluate whether cases of stability reversals are the product of the simplex onset hypothesis, as embodied in the model.

There are four word sets which showed a stability reversal in our data (one triad, three dyads). These are repeated in Table V. The table compares the RSD of the right edge to anchor interval with that of the centre to anchor interval. The left half of the table compares the RSDs of these intervals at low levels of variability (values from Table III, intervals right-delimited by a consonantal anchor). The right half of the table compares the RSDs of these intervals at high levels of variability (values

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

	lower variability		higher variability	
	SD	RSD comparison	SD	RSD comparison
bulha~sbulha ~ksbulha	22	right edge (11·2%) < centre (15·9%)	41	centre (17·8%) < right edge (18·2%)
bal~dbal	15	right edge (5·1%) < centre (9·7%)	63	centre (22·7%) < right edge (25·3%)
tab~ktab	14	right edge (7·7%) < centre (10·0%)	26	centre (7·7%) < right edge (10·0%)
bula~sbula	19	right edge (6·5%) < centre (6·9%)	26	centre (6·5%) < right edge (6·9%)

Table V

Comparison of the RSD of the right edge to anchor interval and the centre to anchor interval for the four word sets that showed stability reversals. The index of overall variability (SD of the right edge to anchor interval) is also given.

from Table IV, intervals right-delimited by a vocalic anchor). The SD of the right edge to anchor interval, the index of overall variability used above, is provided for each comparison.

The results in Table V show the same pattern for each stability reversal in our data. At the lower level of overall variability, the right edge to anchor interval is more stable than the centre to anchor interval. At the higher level of overall variability, the stability pattern is reversed. In just these cases, the centre to anchor interval has a lower RSD than the right edge to anchor interval. This pattern adheres to the implicational relationship for stability reversals derived from the simplex onset hypothesis. The predicted relationship holds for each stability reversal in our data. Verification of this prediction constitutes evidence in support of the simplex onset hypothesis. This support comes from precisely those cases in which the stability patterns are apparently consistent with the complex onset hypothesis. Thus, the model as an analytical tool also allows us to evaluate the predictions of a syllabic parse in cases where the phonetic heuristics break down.

In evaluating the fit between data and model we have so far concentrated our discussion primarily on intervals left-delimited by the centre and right edge landmarks. These are the intervals that are most crucial for evaluating competing syllabic parses of initial clusters. That is, no syllabic parse predicts that the left edge to anchor interval should show greater stability than other intervals. Our model, however, also makes a prediction about how the left edge to anchor interval is affected by overall variability: for the same reason that the RSD of the centre to anchor interval increases at a slower rate (as overall variability increases) than the right edge to anchor interval, the left edge to anchor interval should have an even slower rate of

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

RSD increase than the other two intervals. This prediction is born out in the data. At low levels of overall variability, the left edge to anchor interval shows the highest RSD of the three intervals. As variability increases, the RSD of the left edge to anchor interval increases at a slower rate than the other intervals. There are no shifts to left edge to anchor interval stability, simply because there are no intervals in our data variable enough to yield this reversal. The predicted trend, however, is apparent in our data. Across our corpus, the jump in RSD from the consonantal anchor to the vocalic anchor is largest for the right edge to anchor interval (5%), followed by the centre to anchor interval (2%), followed by the left edge to anchor interval (1%). This provides additional confirmation of our account of the differential effect of overall variability on the RSD of temporal intervals.

In short, we have seen that stability-based heuristics for syllable structure break down as reliable indexes of such structure at high levels of overall variability. This does not mean, however, that patterns of stability cannot inform phonological structure under these conditions. When heuristics fail, the model as an analytical tool allows us to make sense of the data. Simulations based on the simplex onset hypothesis reveal a gradient relationship between the stability of the right edge to anchor interval and the level of overall variability in the set of intervals. This gradient effect of variability constitutes the basis for stating further predictions. For cases of stability reversal, interval sets which show right edge to anchor interval stability are predicted to have lower overall variability than interval sets for which the centre to anchor interval is the most stable. Thus the model continues to make testable predictions based on the simplex onset hypothesis even for those cases when heuristics point in the direction of the complex onset hypothesis. Based on these finer predictions, we verified that each instance of stability reversal in our data is in accord with the simplex onset hypothesis.

With respect to MA, the core conclusion is that all patterns of interval stability in our data can be seen as consequences of the simplex onset hypothesis. Those results that at first appeared to be at odds with the broader picture from our data turn out to be necessary consequences of the simplex onset hypothesis at levels of high overall variability.

5 Conclusions and directions for future work

In recent years, there has been an increasing awareness that the temporal organisation of phonological form provides a rich and potentially highly informative area in the study of the relation between phonological theory and experimental data. In this paper, we have focused on the relation between certain abstract claims of syllabic organisation in Moroccan Arabic and temporal patterns in experimental data obtained using 3D Electromagnetic Articulometry.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

To evaluate competing proposals on initial cluster syllabification, we have looked into the temporal patterns exhibited in our articulatory data. In an analysis of interval stability, we provided evidence for simplex onsets except when overall variability was high. In these latter cases, biconsonantal clusters also showed stability characteristics consistent with the complex onset hypothesis. To make sense of the apparently conflicting evidence in our data, we introduced a computational model. This model served as an explicit link between theoretical claims on syllable structure and temporal patterns in our data. Segments were encoded as a series of articulatory landmarks coordinated in time. Relations between segments were encoded as temporal relations between these landmarks. The temporal relations embodying the simplex onset hypothesis were then used to produce simulated data under different noise conditions. The main result demonstrated was that the temporal organisation of simplex onsets can reproduce the complete range of stability patterns reported in our data. Though simple, the model captured patterns in our data under conditions of both high and low variability. In doing so, it provided evidence for simplex onsets even in those cases when phonetic heuristics broke down.

As the prosodic affiliation of initial consonant clusters is still a matter of debate in a number of languages, e.g. Bella Coola (Bagemihl 1991), Semai, Temiar, Kammu (Shaw 1993), Piro (Lin 1997) and Italian (Bertinetto 2004), there is immediate utility to the analytical techniques developed in this paper. Future work will seek to facilitate application of the tools developed herein to these and other open debates in phonological theory. First, we aim to generalise the methodology to evaluate phonological structure at levels of representation above and below the syllable. In particular, the class of theoretical proposals consistent with the simplex onset hypothesis disagree as to the prosodic status of C_1 in $\#C_1C_2V$, e.g. minor syllable (Boudlal 2001), mora (Kiparsky 2003) or syllable nucleus (Dell & Elmedlaoui 2002). Although we have provided evidence for the simplex onset hypothesis, we have not attempted to distinguish among the different proposals that adhere to this broad hypothesis. Future work will develop temporal predictions which distinguish between different syllabic roles and evaluate these predictions in our data. Second, the experimental data evaluated here was collected using Electromagnetic Articulometry. Although the number of languages for which articulatory data are available is steadily increasing, it is in principle possible to conduct an analysis of interval stability based on the acoustic correlates of articulatory landmarks. This would enable stability-based analyses of structure in languages for which it may be difficult to collect articulatory data. Lastly, our evidence in support of the simplex onset hypothesis for MA came from a model implementing that hypothesis. This model captured the entire range of stability patterns in our data. We have not, however, demonstrated that the patterns of stability in our data are inconsistent with the complex onset hypothesis with comparable rigour. Future work will expand the model to quantitatively evaluate the fit of competing syllable parses to a set of experimental data.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

REFERENCES

Bagemihl, Bruce (1991). Syllable structure in Bella Coola. *LI* **22**. 589–646.
Benhallam, Abderrafi (1980). *Syllable structure and rule types in Arabic*. PhD dissertation, University of Florida.
Benhallam, Abderrafi (1990). Moroccan Arabic syllable structure. *Langues et littératures* **8**. 177–191.
Benkirane, Thami (1997). Intonation in Western Arabic (Morocco). In Daniel Hirst & Albert di Cristo (eds.) *Intonation systems : a survey of twenty languages*. Cambridge: Cambridge University Press. 345–359.
Bertinetto, Pier Marco (2004). On the undecidable syllabification of /sC/ clusters in Italian: converging experimental evidence. *Italian Journal of Linguistics* **16**. 349–372.
Boudlal, Abdulaziz (2001). *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. PhD dissertation, Mohammed V University, Rabat.
Browman, Catherine P. & Louis Goldstein (1988). Some notes on syllable structure in articulatory phonology. *Phonetica* **45**. 140–155.
Byrd, Dani (1995). C-centers revisited. *Phonetica* **52**. 285–306.
Byrd, Dani (1996). Influences on articulatory timing in consonant sequences. *JPh* **24**. 209–244.
Dell, François & Mohamed Elmedlaoui (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer.
Frank, Harry & Steven C. Althoen (1995). *Statistics : concepts and applications*. Cambridge: Cambridge University Press.
Gafos, Adamantios I. (2002). A grammar of gestural coordination. *NLLT* **20**. 269–337.
Goldstein, Louis, Ioana Chitoran & Elizabeth O. Selkirk (2007). Syllable structure as coupled oscillator modes: evidence from Georgian *vs*. Tashlhiyt Berber. In Jürgen Trouvain & William J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*. Dudweiler, Saarbrücken: Pirrot. 241–244.
Greenberg, Joseph H. (1978). Some generalizations concerning initial and final consonant clusters. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.) *Universals of human language*. Vol. 2: *Phonology*. Stanford: Stanford University Press. 243–279.
Harrell, Richard S. (1962). *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.
Harrell, Richard S. & Louis Brunot (2004). *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.
Harrell, Richard S. & Harvey Sobelman (1966). *A dictionary of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.
Harris, Zellig S. (1942). The phonemes of Moroccan Arabic. *Journal of the American Oriental Society* **62**. 309–318.
Heath, Jeffrey (1987). *Ablaut and ambiguity : phonology of a Moroccan Arabic dialect*. Albany: SUNY Press.
Honorof, Douglas & Catherine P. Browman (1995). The centre or edge: how are consonant clusters organised with respect to the vowel? In Kjell Elenius & Peter Branderud (eds.) *Proceedings of the 13th International Congress of the Phonetic Sciences*. Vol. 4. Stockholm: KTH & Stockholm University. 552–555.
Hoole, Philip, Andreas Zierdt & Christian Geng (2003). Beyond 2D in articulatory data acquisition and analysis. In M. J. Solé, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions. 265–268.
Keegan, John M. (1986). The role of syllable structure in the phonology of Moroccan Arabic. In Gerrit J. Dimmendaal (ed.) *Current approaches to African linguistics*. Vol. 3. Dordrecht: Foris. 209–226.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press

Kehrein, Wolfgang & Chris Golston (2004). A prosodic theory of laryngeal contrasts. *Phonology* **21**. 325–357.
Kenstowicz, Michael & Charles Kisseberth (1979). *Generative phonology : description and theory*. New York: Academic Press.
Kiparsky, Paul (2003). Syllables and moras in Arabic. In Caroline Féry & Ruben van de Vijver (eds.) *The syllable in Optimality Theory*. Cambridge: Cambridge University Press. 147–182.
Krakow, Rena A. (1989). *The articulatory organization of syllables : a kinematic analysis of labial and velic gestures*. PhD dissertation, Yale University.
Krakow, Rena A. (1999). Physiological organization of syllables: a review. *JPh* **27**. 23–54.
Lin, Yen-Hwei (1997). Syllabic and moraic structures in Piro. *Phonology* **14**. 403–436.
Lombardi, Linda (1995). Laryngeal neutralization and syllable wellformedness. *NLLT* **13**. 39–74.
Perkell, Joseph S., Marc H. Cohen, Mario A. Svirsky, Melanie L. Matthies, Iñaki Garabieta & Michael T. T. Jackson (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *JASA* **92**. 3078–3096.
Schöner, Gregor (2002). Timing, clocks, and dynamical systems. *Brain and Cognition* **48**. 31–51.
Shaw, Patricia A. (1993). The prosodic constituency of minor syllables. *WCCFL* **12**. 117–132.
Smith, C. L. (1995). Prosodic patterns in the coordination of consonant and vowel gestures. In Bruce Connell & Amalia Arvaniti (eds.) *Papers in laboratory phonology IV : phonology and phonetic evidence*. Cambridge: Cambridge University Press. 205–222.
Sproat, Richard & Osamu Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *JPh* **21**. 291–311.
Vogt, H. (1971). *Grammaire de la langue géorgienne*. Oslo: Universitetsforlaget.
Wing, Alan M. & Alfred B. Kristofferson (1973). Response delays and the timing of discrete motor responses. *Perception and Psychophysics* **14**. 4–12.

https://doi.org/10.1017/S0952675709001754 Published online by Cambridge University Press