Temporal Coordination and Sonority of Jazani Arabic Word-Initial Clusters

Mohammed Ruthan, Karthik Durvasula, and Yen-Hwei Lin
Michigan State University

1   Introduction

There is a need for more techniques/methodologies to establish syllable structure, especially in cases where traditional analytic techniques do not lead to a clear parse. There has been some headway made with some evidence for syllabic organization coming from a correlation between the syllabic parse of consonant sequences and temporal coordination. Recent studies have shown that there is a correlation between the syllabic organization and the timing of the consonants and the vowels (Browman & Goldstein, 1988; Byrd, 1995; Goldstein et al., 2007). Our paper will expand upon the temporal coordination literature by presenting new data from an understudied dialect of Arabic, Jazani Arabic, and by showing the usefulness of acoustic methods to probe temporal stability patterns to understand syllabic organization.

In a pioneering study on the temporal organization in English syllables, Browman & Goldstein (1988) analyzed articulatory data from the Tokyo x-ray microbeam database that consisted of sets of nonsense words with shifted word boundaries, such as [splats] and [plats]. They measured the duration from the end of the vowel gesture (*anchor*) to three different points in the preceding consonant sequence: (a) the left-most consonant in the consonant sequence (*left-edge*), (b) the mean of the midpoints of all the consonants in the consonant sequence (*c-center*), (c) and the right-most consonant in the consonant sequence (*right-edge*). They found that the c-center to anchor interval duration was the most consistent (least variant) across different number of consonants in the syllable onset.

Extending this line of research, Shaw et al. (2011, 2009) study the temporal stability pattern for word-initial consonant sequences in Moroccan Arabic. While English has been standardly argued to have complex onsets, the syllabic organization of the word-initial consonant sequences in Moroccan Arabic has been a debatable issue. Benhallam (1980) argues that such sequences in Moroccan Arabic are complex onsets, thus strings such as /kra/ 'rent' and /skru/ 'his plowshares' are considered to be one syllable; whereas Boudlal (2001) argues that they are simplex onsets, and suggests that the strings are syllabified as [k.ra] and [sk.ru].[1] Shaw et al. (2011, 2009) attempted to tease apart the two syllabic analyses by looking at the temporal stability patterns of the relevant consonant sequences, using electromagnetic-articulography (EMA). They recruited 4 Moroccan Arabic speakers to produce nine target words like /lan/ 'to become soft', /flan/ 'someone', /kflan/ 'nonce', in a carrier phrase /zibi_hnaja/ 'bring_here'. They compared the durations between the three crucial points in the consonant sequences (left-edge, c-center and right-edge) to an *anchor point* in the following vowel (end of the vowel). They hypothesized that a language with a simplex onset organization, where the consonant sequences do not form a complex onset, should have more stability for the right-edge to anchor duration, while a language with a complex onset organization should have more stability for the c-center to anchor duration, as is true of English. They showed that Moroccan Arabic exhibited more stability in the *right-edge* to anchor duration than in the *left-edge* to anchor or the c-*center* to anchor durations; which suggests that the word-initial consonants in Moroccan Arabic were not all part of a single onset cluster.

Additional research has shown that temporal coordination is modulated by the Sonority Sequence Principle (SSP) (Hermes et al., 2013). In their study, they also used articulatory data from EMA to investigate Italian word-initial consonant sequences with different sonority profiles. Four Italian speakers produced 18 target words such as /rima/ 'rhyme', /prima/ 'first', repeated 10 times in a carrier phrase *Per favore dimmi la____ di nuovo* 'please, say ___ again'. The results showed that word-initial consonant sequences with rising sonority (obstruent–liquid sequences) e.g. /pr/, showed a c-center stability, suggesting that they are complex onsets as in English. However, word-initial consonant sequences with falling sonority (sibilant–obstruent sequences) e.g. /sp/ showed right-edge stability, suggesting that they are simplex onsets as in Moroccan Arabic.

---

[1] "." marks a syllable boundary.

Most, if not all, previous related work employed gestural coordination through articulatory techniques. However, Selkirk and Durvasula (2013) replicated previous results through *acoustic* measurements by showing that English has more stability in the *c-center* to *anchor* interval, than for *right-edge* to anchor or left-edge to anchor intervals. They conducted a production experiment, where acoustic data were recorded. Twelve native American English speakers produced 12 repetitions of 16 English words [8 test, 8 filler; total items/subject = 192] in a carrier phrase ["Say ____ here"]. The test items consisted of words that varied in the number of onset consonants (C_1, C_1C_2, sC_1C_2) in two different rhyme contexts [example set: ram, cram, gram, scram]. The Penn Phonetics Lab Forced Aligner (Yuan & Liberman, 2008) was used to automate identification of phone boundaries, and a Praat script was used to extract relevant measurements. Their results suggested that, as with articulatory measurements, English seemed to show a c-center to anchor stability in acoustic measurements, suggesting that temporal stability due to syllabic affiliation can be reliably probed through acoustic techniques. The fully automated measurement technique presented by them afforded a substantial decrease in the time and cost involved in such research.

In the current study, we probe whether word-initial consonant sequences in Jazani Arabic are simplex or complex onsets. Jazani is a dialect of Arabic that is spoken in the south-western province of Saudi Arabia in the city of Jazan. What is most relevant is that the dialect has word-initial consonant sequences - this is unlike most other Arabic dialects which typically have at most one consonant in word-initial position. This raises the issue of whether the Jazani dialect truly has complex onsets unlike most other Arabic dialects. It is also possible that the word-initial consonants in such sequences are not actually syllabified in the same syllable, and therefore, the Jazani syllable structure is actually identical to other dialects, like Moroccan Arabic. We attempt to tease apart these two possibilities. Furthermore, we also attempt to probe whether word-initial consonant sequences that conform to or violate the Sonority Sequencing Principle (SSP) have different syllabic organizations. Using acoustic method, we show that word-initial consonant sequences in Jazani Arabic are simplex regardless of their sonority profiles.

2   Methods

The methodology used in our study follows mostly that of Shaw et al. (2011), a production task in which we measure the same three intervals; the left-edge to anchor, the right-edge to anchor, and the c-center to anchor. However, the aim of this study and the primary difference in our methodology is the use of the acoustic methods, following Selkirk and Durvasula (2013), instead of the articulatory measurements, to find out the stability pattern of word-initial consonant clusters in Jazani Arabic. The following subsection will briefly describe our methodology.

2.1   *Participants*   We recruited seven native speakers of the Jazani dialect (all male), between the ages of 30-35. The data was collected in Jazan, Saudi Arabia, in the summer of 2016. The participants and one of the authors, Ruthan, met at an agreed time to collect the data from each participant individually.

2.2   *Materials*   78 target words (34 real, 44 nonce) grouped into 39 pairs were divided into different categories according to the sonority hierarchy: rising, equal, and falling sonority. For instance, test words with [sm] consonant sequences, like [smaʕ] 'listen', were categorized as 'rising' sonority, test words with [mn], like [mnaʕ] 'prevent', were categorized as 'equal' sonority, and test words with [nf], like [nfax] 'blow out', were categorized as 'falling' sonority. To ensure that the measurements of the segment boundaries, and subsequent left-edge to anchor interval, c-center to anchor interval, and the right-edge to anchor interval were accurate, the crucial consonants always consisted of nasals and fricatives. this combination of the word-initial consonant sequences was chosen on purpose. That is, annotating the beginning and the end of the nasals and the fricative consonants is easier to determine than other consonants.

2.3   *Procedure*   Each participant was asked to take part in the study by providing the production data individually. Acoustic data were recorded using Audacity 2.0.3 (Audacity Team, 2015), using a microphone, which was connected to a Macbook Pro laptop, at a sampling rate of 44 kHz. Participants sat in a quiet room in front of the computer, while target words were displayed in standard Arabic script. The instructions were provided in the Jazani Arabic dialect deliberately to encourage participants to pronounce the words in their own dialect. Participants were given three sentences for practice; these trial sentences were not used in the analysis. Participants read the test words in a carrier phrase, [ʔɪnta ____ marah θanjah] 'You __ Again'. Every participant produced (78*6) 468 tokens, yielding a total of 3276 tokens.

3   Data Analysis

3.1   *Scripting*   Each sound file was manually annotated in Praat (Boersma & Weenink, 2016), then a Praat script was used to automate the measurements for participants' production data. The annotation marked the segment boundaries of the word-initial consonant(s) and the following vowel. The same definitions used by Shaw et al. (2011) were employed in this study to measure the right-edge to anchor interval, the c-center to anchor interval and the left-edge to anchor interval.

3.2   *Measurements*   Acoustic and spectrographic information together was used to manually mark the labels mentioned above. To illustrate, in a word with a consonant sequence like [msak] 'catch', the left-edge to anchor interval duration was calculated from midpoint of the target consonant [m] to the end of the vowel [a]. The duration from the midpoint of the target consonant [s] to the end of the vowel determined the right-edge to anchor interval duration. The c-center interval duration was determined by calculating the duration from the mean of the mid-points of the two consonants to the end of the vowel. In the case of words with just a single consonant, like [sak] 'nonce', the three interval durations were the same, as the intervals as defined earlier overlap perfectly.
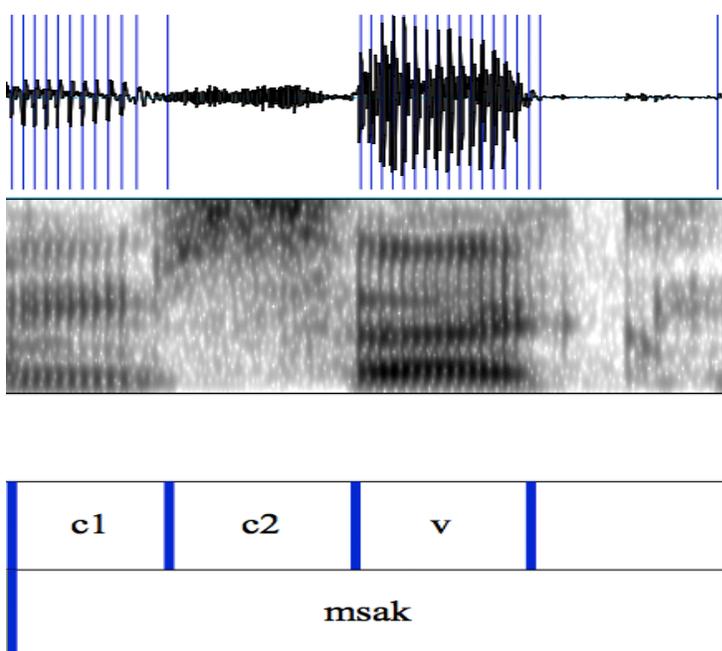
Figure 1: An example token. Durations are measured from the Left-edge (midpoint of [m]) the Right-edge (midpoint of [s]) and C-center (the mean of the midpoints of [m] and [s]) to the anchor point, defined as the end of the vowel.

4   Results and Discussion

4.1   *Stability Patterns in Jazani Arabic Data*   The first research question asks whether the syllabic organization of word-initial sequences in Jazani Arabic is a simplex or complex onset. As mentioned above, if temporal patterns show more stability for the c-center to anchor interval, it means that the word-initial consonant sequences in that language have a complex onset organization. On the other hand, if the temporal patterns show more stability for the right-edge, this means that the language or dialect has a simplex onset organization. Figure 2 shows boxplot, showing the number of consonants on the (x-axis), and the intervals duration on the (y-axis) in milliseconds, where 'CC' represents words that have a sequence of two consonants, while 'C' represents words that have only one consonant (Note: plots were made using the R package ggplot2 (Wickham, 2016). As the number of the consonants increases, from singleton to biconsonantal, the intervals changes behave similarly across the three intervals. The left-edge to anchor and the c-center to anchor interval durations change with the increase of the number of consonants (#CCVX, #CVX). On the contrary, the right-edge to anchor interval duration appears to be stable.
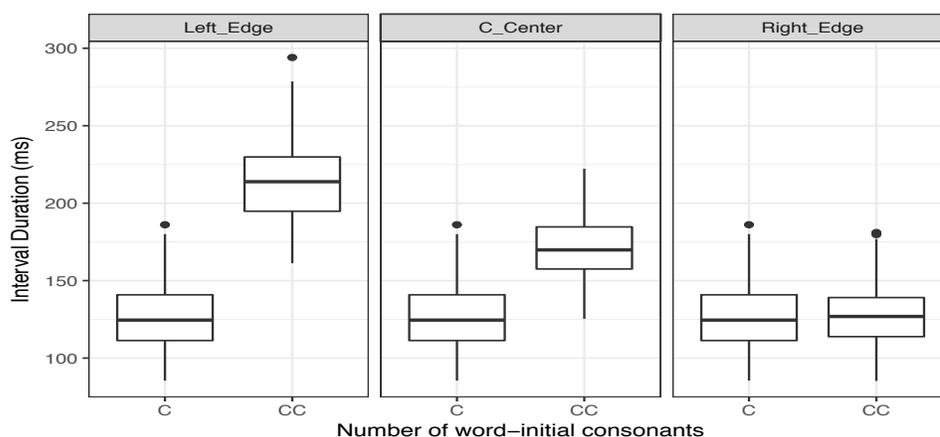
Figure 2: Boxplot of the left-edge, the c-center, and the right-edge interval durations.

The observed pattern based on raw durations of different intervals above is consistent with the simplex onset organization, in which the right-edge to anchor interval is the most stable interval. In order to further probe which interval is the most stable, we followed Shaw et al. (2011) and used (RSD), defined as *100*(standard deviation of the duration)/(mean duration)*, as an index for the interval stability rather than variance or standard deviation. Using variance or standard deviation does not control for the fact that shorter durations (right-edge and c-center) are likely associated with smaller variances. In contrast, using the RSD avoids the bias associated with such shorter intervals. **Error! Reference source not found.** presents the RSD values for some of the pairs across the seven speakers. The table shows that the right-edge always had the lowest RSD values compared to the other two intervals. The results support the simplex onset organization hypothesis for the syllable structure of Jazani Arabic.

Pairs	Left-edge	C-center	Right-edge
ħmad - mad	29.8	18.5	**10.7**
ħmaf - maf	30.5	19.6	**11.9**
nfad- fad	28.1	18.5	**11.5**
smaʕ - maʕ	31.6	21.2	**15.4**

Table 1: A sample set of the three intervals' RSD values across all seven speakers (the lowest RSD values in bold).

Figure 3 shows the interval types on the x-axis and the interval durations by their RSD values on the y-axis. The results of the RSD show that the right-edge to anchor interval was still the most stable interval compared to the other two intervals. In fact, the right-edge to anchor interval was associated with the lowest RSD values for all pairs tested.
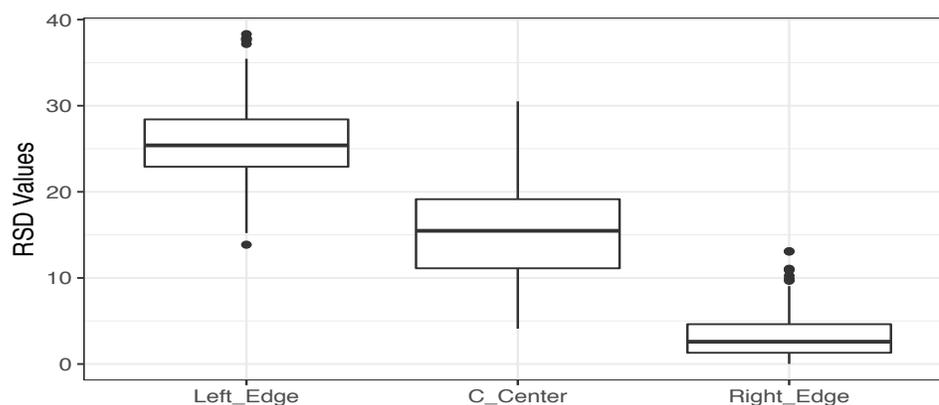

Figure 3: Boxplot for RSD values of the left-edge, the c-center, and the right-edge interval durations.

4

Figure 4 below shows the results for all seven participants. The figure shows the interval types on the x-axis and their RSD values on the y-axis. As can be seen the right-edge to anchor interval has the lowest RSD for all seven participants. That is, the right-edge to anchor interval has the most stability for all seven speakers.
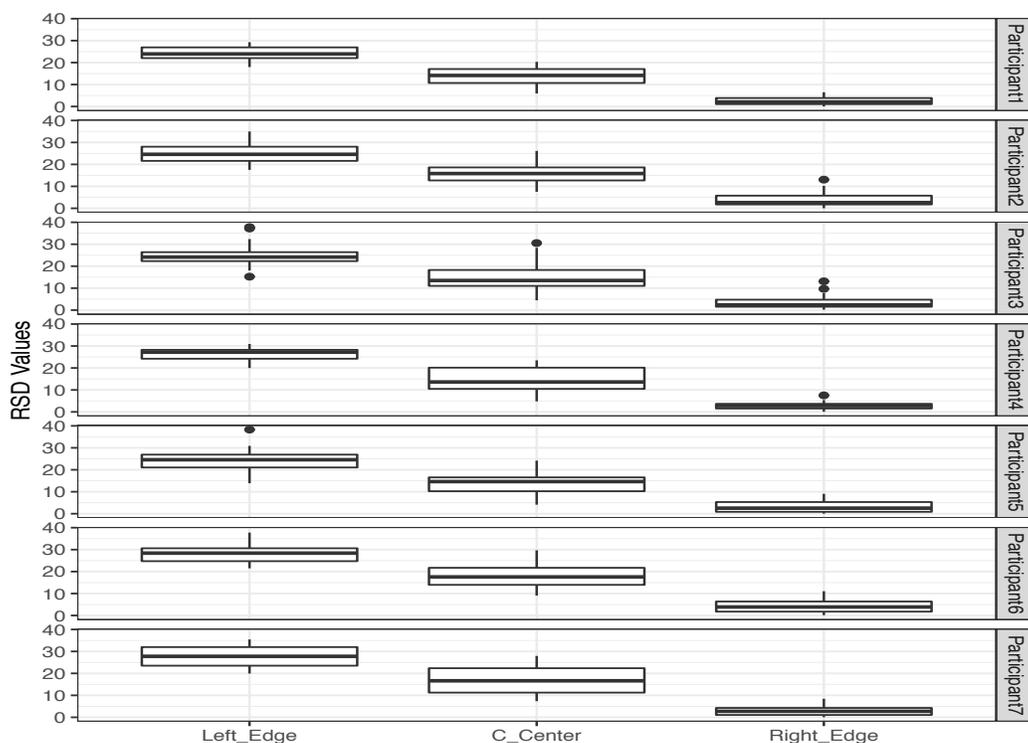

Figure 4: Boxplot for RSD values of the left-edge, the c-center, and the right-edge interval durations across seven speakers.

We fitted a linear fixed effects model to the RSD values in R (R Core Team, 2016). We used the packages lmer (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017). The random effects structure included a random slope of *Interval Type* (right-edge vs. c-center vs. left-edge) for *word pair* and *participants*. The best model included a fixed effect of *Interval Type* (

Table 2). As with the visual inspection, the model indicated that right-edge to anchor interval duration had the lowest RSD, suggesting again that the right-edge to anchor interval duration had the most stability.

	Estimate	Std. Error	df	t-value	Pr(>|t|)
(Intercept)	11	1.1	6.6	10.2	<0.001
C-center	7.8	0.8	9.4	9.9	<0.001
Left-edge	17.7	1.2	11.2	15.4	<0.001

Table 2: Linear mixed effects model with the RSD values as the dependent variable, and *Interval Type* as the independent variable [baseline: Right-edge].

In answer to the first question, the stability of the interval durations suggested that the temporal patterns showed the right-edge to anchor interval to have more stability in Jazani Arabic. This suggests that Jazani Arabic has a simplex organization for word-initial consonant sequences. That is, the consonants in a word-initial sequence are not all in the same onset, but rather are parsed as C.CVX; e.g., [smaʕ] 'listen' is parsed as [s.maʕ]. The pattern found in this study showed a similar pattern and comparable results to those observed in Moroccan Arabic by Shaw et al. (2011).

4.2 *The Stability Patterns and Different Sonority Profiles*   To answer the second question, which asks whether the syllabic organization differs depending on the different sonority profiles for the word-initial consonant sequences, it was necessary to categorize these word-initial consonant sequences into different groups based on the sonority profiles (rising, equal, and falling sonority). Based on these categories, we investigated whether the sonority profiles showed different temporal patterns that could indicate that some word-initial consonant sequences behave as simplex and others as complex onsets, as in the case of Italian.

Figure 5 shows the RSD values for words with different sonority profiles. The figure shows the interval types on the x-axis, the interval duration RSD values on the y-axis, and a different sonority profile in each of the facets. The RSD values for the right-edge to anchor interval duration, c-center to anchor interval duration, and left-edge to anchor interval duration showed similar patterns across all three sonority profiles. Based on visual inspection, the lowest RSD values seemed to be associated with the right-edge to anchor interval duration, suggesting that initial consonant sequences with different sonority profiles behave similarly.
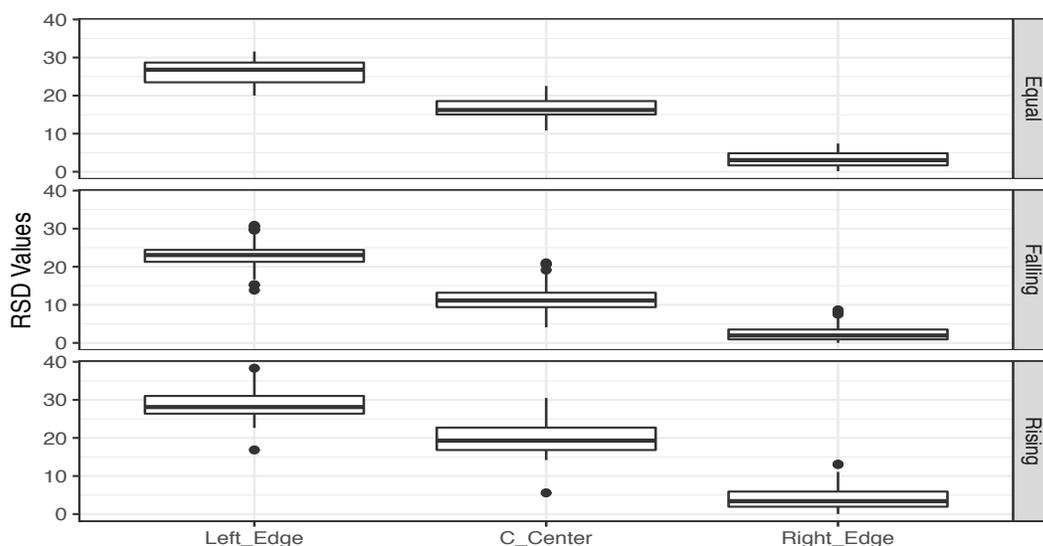

Figure 5: Boxplot for RSD values of the three intervals across the sonority profiles.

As before, we fitted a linear mixed effects models to the data. The RSD values were the dependent variable, and *Sonority Profile* (Equal vs. Falling vs. Rising) and *Interval Type* as the independent variables considered. The random effects structure included a random slope of *Interval Type* (right-edge vs. c-center vs. left-edge) for *word pair* and *participants*. The best model included separate fixed effects of *Interval Type* and *Sonority Profile* (**Table 3**). The results of the modelling suggest that the pattern of minimal RSD values being associated with the right-edge to anchor interval was similar for all sonority profiles.

	Estimate	Std. Error	df	t-value	Pr(>|t|)
(Intercept)	9.9	1.1	7.5	8.8	<0.001
C-center	7.8	0.8	9.3	9.9	<0.001
Left-edge	17.7	1.1	11.2	15.3	<0.001
Rising Sonority	1.0	0.7	33.1	1.4	0.17
Falling Sonority	2.2	0.5	33.3	4.0	<0.001

Table 3: Linear mixed effects model with the RSD values as the dependent variable, and *Interval Type* and *Sonority Profile* as the independent variables [baseline: Right-edge, Equal sonority].

5   Conclusion

In this paper, we investigated the syllabic organization of word-initial consonant sequences in Jazani Arabic and presented evidence from the patterns of temporal stability that showed that Jazani Arabic word-initial consonant sequences are simplex onsets rather than complex onsets. That is, the right-edge to anchor interval duration had the lowest RSD values, and thus was the most stable interval. This pattern was consistent with the simplex onset organization hypothesis in Shaw et al.'s (2011) study, as in the case of Moroccan Arabic. We also showed that regardless of the sonority profiles, the temporal pattern of the Jazani dialect remained the same, confirming the behavior of word-initial consonant sequences as simplex onsets across the board, unlike the case of Italian. Finally, the study also shows the effectiveness of acoustic measurements as a tool to understand syllabic organization, through the study temporal co-ordination patterns. We hope this paper will help other scholars, and lead to an expansion in studying more languages' syllable structures, due to the accessibility of the technique employed.

References

Audacity Team (2015). Audacity(R): Free Audio Editor and Recorder [Computer program]. Retrieved from audacity.sourceforge.net.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01.

Benhallam, A. (1980). *Syllable structure and rule types in Arabic*. University of Florida.

Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Retrieved from http://www.praat.org/.

Boudlal, A. (2001). *Constraint Interaction in the Phonology and Morphology of Casablanca Moroccan Arabic*. Ph.D Dissertation, Mohammed V University, Rabat, Morocco.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2–4), 140–155. http://doi.org/10.1159/000261823.

Byrd, D. (1995). C-Centers Revisited. *Phonetica*, *52*(4), 285–306. http://doi.org/10.1159/000262183.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007). Syllable Structure as Coupled Oscillator Modes: Evidence from Georgian vs. Tashlhiyt Berber. *Proceedings of the XVI International Congress of Phonetic Sciences*, 241–244.

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*(13), 1–26. http://doi.org/10.18637/jss.v082.i13.

R Core Team (2016). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.r-project.org/.

Selkirk, E. & Durvasula, K. (2013). Acoustic correlates of consonant gesture timing in English. *The Journal of the Acoustical Society of America*, *134*(5), 4202–4202.

Shaw, J. a., Gafos, A. I., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure: a case study of Moroccan Arabic consonant clusters. *Phonology*, *28(3), 455-490*. http://doi.org/10.1017/S0952675711000224.

Shaw, J., Gafos, A. I., Hoole, P., & Zeroual, C. (2009). Syllabification in Moroccan Arabic: evidence from patterns of temporal stability in articulation. *Phonology*, *26(1), 187-215*. Retrieved from http://www.journals.cambridge.org/abstract_S0952675709001754.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York,.

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America*, *123*(5), 3878–3878. http://doi.org/10.1121/1.2935783.