



Acoustics and articulation of medial versus final coronal stop gemination contrasts in Moroccan Arabic

Mohamed Yassine Frej¹, Christopher Carignan¹, Catherine T. Best^{1,2}

¹The MARCS Institute, Western Sydney University, Australia

²Haskins Laboratories, New Haven, Connecticut, USA

y.frej@westernsydney.edu.au, C.Carignan@westernsydney.edu.au C.Best@westernsydney.edu.au

Abstract

This paper presents results of a simultaneous acoustic and articulatory investigation of word-medial and word-final geminate/singleton coronal stop contrasts in Moroccan Arabic (MA). The acoustic analysis revealed that, only for the word-medial contrast, the two MA speakers adopted comparable strategies in contrasting geminates with singletons, mainly by significantly lengthening closure duration in geminates, relative to singletons. In word-final position, two speaker-specific contrasting patterns emerged. While one speaker also lengthened the closure duration for final geminates, the other speaker instead lengthened only the release duration for final geminates, relative to singletons. Consonant closure and preceding vowel were significantly longer for the geminate only in medial position, not in final position. These temporal differences were even more clearly delineated in the articulatory signal, captured via ultrasound, to which we applied the novel approach of using TRACTUS [Temporally Resolved Articulatory Configuration Tracking of UltraSound: 15] to index temporal properties of closure gestures for these geminate/singleton contrasts.

Index Terms: word-medial gemination, word-final gemination, duration, acoustics, articulation, speech dynamics, ultrasound, Moroccan Arabic.

1. Introduction

Some varieties of Arabic, including Moroccan Arabic (MA), have gemination occurring in all positions of the word (word-initially, -medially and -finally), yet most research conducted on gemination in Arabic and other languages has focused on medial gemination. The poor perceptual cues to gemination in word-final position have been said to lead to diachronic neutralization of the contrast [1] whereby words bearing a final geminated consonant are produced exactly as their non-geminated counterparts. Findings from some studies seem to support this claim. Word-final geminates are neutralized through the process of degemination in Hungarian, which affects glides and liquids, but this process is optional and some distinction may be carried by a lengthening of the preceding vowel for final geminates [2]. Similarly, geminates in Russian can freely be degeminated word-finally [3]. However, the contrastiveness of word-final gemination in Arabic has been debated among researchers. Ghalib [4] holds that geminates occurring word-finally are non-distinctive, i.e., neutralized, in Arabic, consistent with degemination. This view is in line with other studies [5,6]. However, El Saaran [7] argues against this claim by providing examples of geminate/singleton word pairs that attest to the distinctiveness of final gemination. These

studies, nonetheless, remain impressionistic in their nature and have not been verified with any instrumental data. To our knowledge, there are only two experimental studies on final gemination in other varieties of Arabic. While it was confirmed that final gemination is contrastive in Jordanian Arabic [8], word-final gemination was argued to be disappearing in Rural Jordanian Arabic because speakers are not always able to distinguish between the contrasts [9].

A variety of kinematic techniques such as MRI [10], X-ray [11] and electromagnetic articulography [12] have been employed to determine the articulatory characteristics of medial gemination in Arabic and other languages. However, articulatory research on word-final gemination is scarce.

Given the controversy over the contrastiveness or neutralization of final gemination in Arabic and the dearth of articulatory studies on it, the primary aim of this paper is to determine whether temporal cues in the acoustic signal, such as closure duration and durations of the flanking vowels are indicative of a length contrast between the MA coronals /t,d/ and /tt,dd/. If so, whether this length contrast is supported and further delineated by articulatory analysis. Moreover, we compare word-final gemination with word-medial gemination to investigate whether speakers use different or similar acoustic and articulatory features in producing these contrasts. These aims were broadly motivated for understanding general principles of how final gemination is achieved and how it compares to medial gemination.

Another novelty in this paper is the use of ultrasound to index articulatory durational properties of gemination. Ultrasound is an ideal technique because of its affordability, safety, and smaller effect on speech articulation and acoustics, compared to other kinematic techniques such as electromagnetic articulometry (EMA) and real-time magnetic resonance imaging (MRI). However, the traditional approach of fitting contours to tongue surfaces in ultrasound images at single time points would be inefficient for studying contrasts that are based primarily on duration differences (such as gemination), due to the often time-consuming nature of creating tongue surface contours for every single frame in the ultrasound video. More recent techniques have allowed for time-varying measurements of speech articulation created directly from ultrasound video [13], [14] and [15]. One such method, Temporally Resolved Articulatory Configuration Tracking of UltraSound (TRACTUS; [15]), is a reliable function that allows temporal articulatory analysis without the need for ultrasound image tracing.

2. Method and material

One female (44 years) (S1) and one male (40 years) (S2) MA

speaker produced, 10 times in a carrier phrase and 10 times in isolated form, four real disyllabic minimal pair MA words contrasting /t/ and /d/ with /tt/ and /dd/, respectively. Words were selected to allow for the comparison between geminates and their singleton counterparts word-medially and -finally. They were of the structure: /C₁aXaC₂/ (word-medial) and /C₁aC₂aX/ (word-final), with X being either a singleton or a geminate.

Acoustic data were co-registered with ultrasound data using the GE LOGIQ e ultrasound system with a GE 8C-RS probe, operating at 8 MHz and 8-10cm depths. Ultrasound video was extracted from the VGA output in real time using an Epiphan VGA2USB Pro video grabber. FFMPEG software was used to record a continuous .AVI file with ultrasound video at 30 fps with synchronous acoustic data, which were collected through USB using a Sennheiser MKH 416 microphone connected to a Sound Devices LLC USB Pre 2. A non-metallic head-mounted ultrasound probe holder, designed by Derrick and colleagues [16], was used to stabilize the ultrasound probe relative to the head within acceptable parameters for rotational and translational slippage, based on HOCUS measurements [Haskins Optically-Corrected UltraSound: 17].

3. Acoustic analysis and results

To perform the acoustic analysis, the data were imported into Praat, where segmentation was carried out manually following standard procedures. The following measurements were logged using a Praat script: closure duration of the target singleton and geminate consonants, the release duration as well as the duration of the preceding vowel.

The preceding vowel /a/ was segmented from the onset of clearly repeating periodicity and formant structure until the stop closure. The closure interval was marked from the end of regular periodicity of the vowel accompanied by noticeable drop of amplitude up until the burst. The burst of the stop release was identified by a clear brief spike of noisy energy in the spectrogram and measured until the start of the vowel in medial position and until the cessation of friction in the word-final position. This was possible because MA stops are produced with aspiration even in the word-final position. Results for each measurement are presented in Figures 1 and 2.

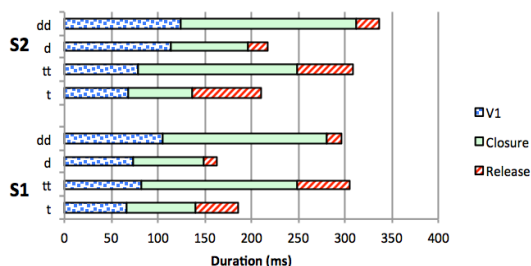


Figure 1: means of preceding vowel duration (V1), closure duration (closure) and release duration (release) in word-medial singletons and geminates for S1 and S2.

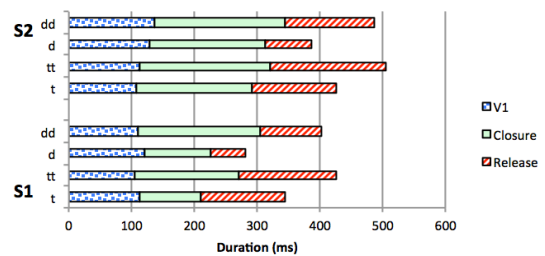


Figure 2: Means of preceding vowel duration (V1), closure duration (closure) and release duration (release) of word-final singletons and geminates for S1 and S2.

Each of the preceding vowel, closure and release durations were analyzed using a 2x2x2x2 repeated measures analysis of variance (ANOVA) with speaker (S1 vs S2) as a between-subject factor, and length (geminate vs singleton), position (medial vs final) and voicing (voiced vs voiceless) as within-subject factors. Unexpectedly, the results for the preceding vowel durations indicated no main effect of length, but there was a significant interaction between position and length [$F(3,48) = 24.5, p < 0.001$]. Post hoc pairwise *t*-test comparisons with a Bonferroni correction ($\alpha = 0.016$) indicated that the preceding vowel duration was longer significantly and contrastively before geminates than before singletons only word-medially [$F(1,12) = 17.7, p < 0.001$].

The results for the closure duration ANOVA manifested a significant main effect of speaker [$F(1,12) = 10.6, p < 0.01$] and position [$F(1,12) = 18.7, p < 0.001$] along with a significant interaction between length and position [$F(3,48) = 7.2, p < 0.05$] and one between speaker and length [$F(3,48) = 8.8, p < 0.05$]. It was revealed after breaking down the length-position interaction using corrected *t*-tests that geminates were produced with longer closure durations than singletons, but only word-medially [$F(3,48) = 9.8, p < 0.05$].

The results on the release duration ANOVA showed a main effect of voicing [$F(1,12)=28.1, p < 0.01$] and speaker [$F(1,12)=9.7, p < 0.01$] as well as a significant interaction between speaker and position [$F(3,48) = 11.8, p < 0.001$]. Since we found significant interactions involving speaker for both release and closure, we decided to run separate 3 way repeated measures ANOVAs for each speaker on each of these measurements (release and closure) with length (geminate vs singleton), position (medial vs final) and voicing (voiced vs voiceless) being within-subject factors.

For S1, the results of the closure ANOVA revealed a significant main effect of length [$F(1,6) = 13.7, p < 0.001$], i.e., S1 used closure contrastively to differentiate between geminates and singletons by producing geminates with longer durations than singletons. There was also a significant main effect of position [$F(1,6) = 20.6, p < 0.001$]. Pairwise *t*-tests demonstrated that her closure durations were significantly longer in the word-medial position than the word-final position [$F(1,6) = 26.5, p < 0.000$]. For S1's release duration ANOVA, there was no main effect of length, i.e., release was not used as a cue to contrasting a geminate with a singleton, but there was a significant effect of position [$F(1,6) = 11.3, p < 0.005$]. Pairwise comparisons showed that release durations were significantly higher in the word-final position than word-medial position [$F(1,6) = 10.3, p < 0.001$]. There was also a significant interaction between voicing and position [$F(3,18) = 7.5, p < 0.01$]. Bonferroni-corrected pairwise *t*-tests showed

that release durations were lengthened for geminates /dd/ but not /tt/ word-finally only [$F(1,6) = 14.6, p < 0.01$].

For S2, surprisingly, there was no significant main effect of length for the closure duration ANOVA, which means that closure didn't vary contrastively between a geminate and a singleton. Similar to S1, there was, however, a significant interaction between position and length [$F(3,18) = 9.9, p < 0.05$], from which we deduced, using pairwise *t*-test comparisons, that S1 used closure duration contrastively only in the medial position [$F(1,6) = 23.9, p < 0.001$]. S2, however, seemed to compensate for this by lengthening the release duration in final position, because there was a significant main effect of position [$F(1,6) = 13.5, p < 0.001$], and a significant interaction between position and length [$F(3,18) = 15.8, p < 0.001$]. Pairwise *t*-tests confirmed that S2's release durations were significantly longer word-medially than word-finally [$F(1,6) = 25.7, p < 0.000$]. More importantly, S2 used release duration as a cue to contrasting geminates with singletons in word-final position [$F(1,6) = 16.8, p < 0.001$], but not in word-medial position. The presence or absence of final gemination was previously reported to affect the duration of the release in some languages, such as Tashlhiyt Berber [18], but this effect was mostly attested for voiced stops (as was the case with S1 word-finally) with the exception of Cypriot Greek whose inventory lacks voiced stops [19].

To summarize the main findings, while both speakers employ closure duration as a cue to differentiate geminates from singletons word-medially, which is a common scenario for geminate/singleton contrasts, in the word-final position, the two speakers seem to be adopting different strategies. S1 continues to contrast geminates with singletons word-finally through closure duration, whereas S2 does so by differentiating release duration instead.

Next, we investigated whether these temporal differences are reflected in the articulatory signal, as well as whether the articulatory analysis would pinpoint additional information that might not be clearly/reliably evident in the acoustic measures.

4. Articulatory analysis and results

Rather than using the more frequently-applied but rather tedious and time-consuming task of manually tracing tongue contours, our ultrasound data were analyzed using TRACTUS [15]), which is a suite of Matlab functions designed to filter ultrasound images and submit the resulting pixel intensities to a Principal Components Analysis (PCA) model in order to identify a set of principal components (PCs) that explain the greatest amount of variance throughout the video recording. For this paper, 104 (S1) and 94 (S2) PCs were retained for further analysis, which each account for 80.3% of the total variance. A Linear Discriminant Analysis (LDA) model was then trained on the PC scores for ultrasound frames associated with the target coronal consonant closures across all the recording. The cumulative products of the PC scores and PC loadings were used to generate composite "heatmaps" for both singletons and geminates, which can be interpreted as average models of the lingual configuration during the closure period for the two consonant types. Having noticed that the heatmaps representing each position looked quite similar, we decided to combine them across the two word positions, and due to space limitations, we present here only the heatmaps from S2 (Figure 3 and 4), because S2's ultrasound image came out

clearer than S1's. The brighter a region is in the image, the more strongly that region is associated with the articulation of stop closure in the brightened region of the vocal tract. The LDA model was then used to predict linear scores for the PC scores across all of the words in the ultrasound recording. The result is an articulatory signal that can be interpreted as the likelihood that a given ultrasound frame resembles the articulatory configuration associated with a closure (Figures 5-8).

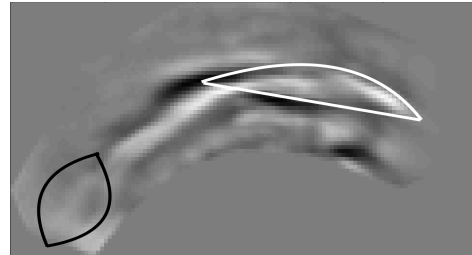


Figure 3: LDA Heatmap for singletons

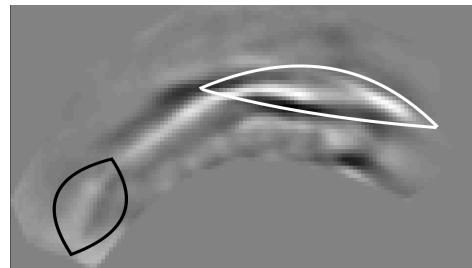


Figure 4: LDA Heatmap for geminates

A close observation of both heatmaps suggests that the gestures responsible for the production of both contrasts are quite comparable, with the exception of possibly greater pharyngeal and tongue blade constriction/approximation for geminates (see areas marked in black, in Figure 4). Using a Real-time MRI analysis, the release of geminated coronal stops was reported to be accompanied with tongue root retraction in Miyakojima Ikema [21]. S2 might well be using a similar mechanism in anticipation of the geminate release, based on the configurations observed in these heatmaps. Another interesting observation is that constriction location of the coronal stops in MA appears to be laminal for both singletons and geminates, contrary to the apical configuration of tongue tip that has been generally assumed for coronal singletons [22] (see area marked in white in Figure 3-4). These findings will, however, be substantiated with additional data from a larger number of speakers.

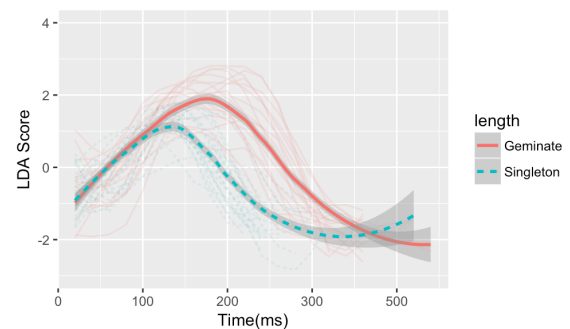


Figure 5: The LDA class scores over time for geminates and singletons in word-medial position for S1.

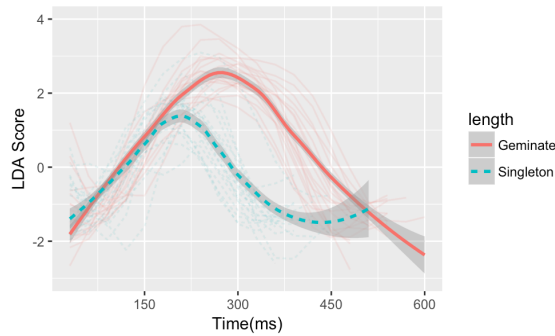


Figure 6: *The LDA class scores over time for geminates and singletons in word-medial position for S2.*

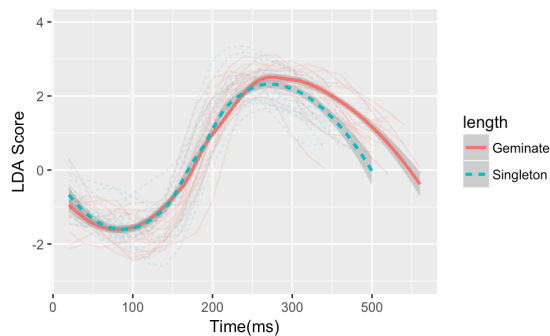


Figure 7: *The LDA class scores over time for geminates and singletons in word-final position for S1*

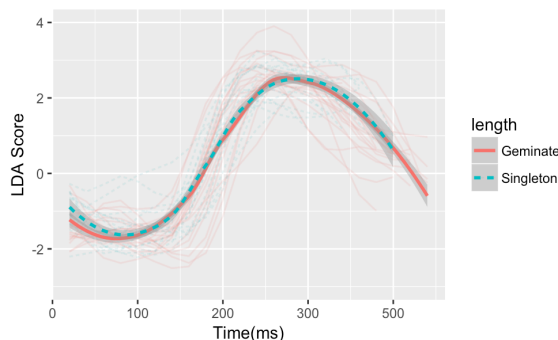


Figure 8: *The LDA class scores over time for geminates and singletons in word-final position for S2.*

Figures 5-8 display the LDA scores of the articulatory signals for individual tokens, as well as category means and 95% confidence intervals generated using t-based approximation (“loess” smoothing in the *ggplot2* package in R) [23]. The higher the score, the greater the likelihood that the tongue resembles the models in Figures 3 and 4. As such, the peak of the geminate and singleton lines would correspond to the closure. The difference between the geminate and the singleton lines is clearly more pronounced in the medial position (Figures 5 and 6). In word-final position, the contrast of the lines at the peak is slightly manifested for S1 only (Figure 7), which is supported by the acoustic analysis. Another interesting observation is that the onset of the geminate and singleton signals is identical for both speakers in both positions (the lines are unified). The offset, however, is

what sets S1 and S2 apart. Notably, despite a remarkable unified distribution, S2’s geminate line continues on longer than that of singleton in final position (Figure 8). This is reflective of the longer duration of the release that characterized S2’s production of word-final geminates compared with singletons. Moreover, at first sight, the line of the LDA articulatory signal for word-final gemination for both speakers seems not to be highlighting the full stop closure, because it does not go back down to baseline. However, this might not necessarily be the case. One possible explanation for this appearance in the LDA score figures is that the absence of a vowel after the release of the stop in the word-final position means that the tongue is not required to lower as much from the roof of the oral cavity with consonant release as it does for a following /a/ in the medial cases. Speakers might well be resting their tongue against or close to the roof right after the stop release in final position. This possibility will be taken up in detail in future investigations.

5. Discussion and conclusion

Given that closure duration is often reported to be the most consistent and robust correlate of the geminate/singleton contrast, it was expected that the participants in this study would distinguish geminates from singletons through means of lengthening the closure duration in both positions of the word. However, results show that this is the case for both speakers only for medial gemination and significant only for S1 in the final position. S2 did still produce higher durations for final geminates, but this was established through lengthening the release duration rather than the closure duration of geminates in this position. The fact that this speaker maintained this length contrast across both voicing types suggests that for this speaker, release duration is a strong correlate of final gemination. Nonetheless, given the scarcity of research on languages with final gemination and on MA in particular, we are not sure of the reliability of this release-duration technique of contrasting geminates with singletons word-finally. Given this and the closure duration distinction produced by S1, the question of whether final gemination is contrastive in MA is tentatively supported but whether it is reliable across speakers, and especially whether other speakers may use release duration rather than closure duration in this position remains inconclusive. An obvious next step would be to replicate the experiment with more participants to see if other speakers use the rather unusual pattern of release-duration contrast for final gemination. It would be also useful to test the relevance of closure versus release length contrasts to the perception of the final geminate/singleton contrast by MA native speakers.

The articulatory signals generated by TRACTUS proved to be capable of yielding temporal information about closure gestures in ultrasound images. The same analysis will be performed at a later stage on word-initial gemination contrasts in MA in order to demonstrate whether, even in the absence of clear acoustic information, in the case of voiceless stops, temporal analysis of gemination in this position of the word would be revealed in articulatory duration differences, as measured through ultrasound and TRACTUS.

6. References

- [1] J. Blevins, *Evolutionary phonology: The emergence of sound patterns*, Cambridge University Press, 2004.

- [2] E. K. J. Curtis, *Geminate Weight: Case Studies and Formal Models* (Doctoral thesis), University of Washington, 2003.
- [3] L. L. Kasatkin, and M. C. Choj, Dolgota/kratkost' Soglasnogo na Meste Sochetanij Dvuh Bukv v Sovremennom Russkom Literatunom Jazike. Dialog-MGU, Moskva. [Duration of the doubly-spelled consonants in the Contemporary Standard Russian], 1999.
- [4] G. Ghalib, *An experimental study of consonant gemination in Iraqi Colloquial Arabic* (Unpublished doctoral dissertation). University of Leeds, 1984.
- [5] M. W. Cowell, *A Reference Grammar of Syrian Arabic with Audio CD: Based on the Dialect of Damascus*, Georgetown University Press, 2005.
- [6] T. M. Johnstone, "Aspects of syllabication in the spoken Arabic of Anaiza," *Bulletin of the School of Oriental and African Studies*, 1967, 30(01), 1-16.
- [7] M. El Saaran, *A critical study of the phonetic observations of the Arab Grammarians* (Unpublished doctoral dissertation), University of London, 1951.
- [8] F. Al-Tamimi, K. Abu-Abbas, & R. Tarawnah, "Jordanian Arabic final geminates: An experimental clinical phonetic study," *Poznań Studies in Contemporary Linguistics*, 2010, 46(2), 111-125.
- [9] M. Al-Deaibes and N. Rosen, "Geminate-singleton contrast in Rural Jordanian Arabic," Paper presented at the ASAL 30, New York, 2016.
- [10] C. Hagedorn, M. I. Proctor, and L. Goldstein, "Automatic Analysis of Singleton and Geminate Consonant Articulation Using Real-Time Magnetic Resonance Imaging," In *INTERSPEECH*, 2011, pp. 409-412.
- [11] F. Bouarourou, B. Vaxelaire, R. Ridouane, F. Hirsch, and R. Sock, "Gemination in Tarifit Berber: X-ray and acoustic data," In *Proceedings of the 8th international seminar on speech production*, 2008, pp. 117-120.
- [12] C. Zeroual, D. Gafos, P. Hoole, & J. Esling, "Physiological and Acoustic Study of Word Initial Post-lexical Gemination in Moroccan Arabic," In *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.
- [13] T. Hueber, G. Aversano, G. Chollet, B. Denby, G. Dreyfus, Y. Oussar, P. Roussel, and M. Stone, Eigentongue, "Feature extraction for an ultrasound-based silent speech interface," *Proceedings of the Eighth International Seminar on Speech Production*, 2007, pp. 365-369.
- [14] S. R. Moisik, H. Lin, and J. H. Esling, "A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS)," *Journal of the International Phonetic Association*, 2014, 44, pp. 21-58.
- [15] C. Carignan, TRACTUS (Temporally Resolved Articulatory Configuration Tracking of UltraSound) software suite, URL: <http://christophercarignan.github.io/TRACTUS>, 2014.
- [16] D. Derrick, C. T. Best, & R. Fiasson, "Non-metallic ultrasound probe holder for co-collection and co-registration with EMA," In the *Proceeding of the 18th International Congress of Phonetic Sciences*, 2015.
- [17] D. H. Whalen, K. Iskarous, M. K. Tiede, The Haskins Optically Corrected Ultrasound System (HOCUS), 2005.
- [18] R. Ridouane, "Geminates at the junction of phonetics and phonology," *Papers in laboratory phonology*, 2010, 10, pp. 61-90.
- [19] G. Tserdanelis, & A. Arvaniti, "The acoustic characteristics of geminate consonants in Cypriot Greek," In *Proceedings of the fourth international conference on Greek linguistics*, 2001, pp. 29-36.
- [20] C. Zeroual, P. Hoole, & A. Gafos, "Spatio-temporal and kinematic study of Moroccan Arabic coronal geminate plosives," *Proc. 8th ISSP*, 2008, pp. 135-138.
- [21] M. Fujimoto, & S. Shinohara, "Articulation of voiced geminate and word-initial geminate obstruents in the Miyakojima Ikema dialect of Ryukyuan: A Real-time MRI analysis," *Poster presented at GemCon2015*, Glasgow, 2015.
- [22] C. Zeroual, P. Hoole, S. Fuchs, & H. Esling, "EMA study of the coronal emphatic and non-emphatic plosive consonants of Moroccan Arabic," In *Proc. 16th ICPHS*, 2007.
- [23] H. Wickham, *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. URL: <http://ggplot2.org>, 2009.