

# **Tongue movements and syllable onset complexity: Ultrasound study**

Tanja Kocjancic<sup>1</sup>

<sup>1</sup>Speech and Hearing Sciences, Queen Margaret University, UK

<sup>1</sup>Centre for Speech Technology Research, University of Edinburgh, UK

## **Abstract**

In this study ultrasound was used to investigate tongue movements in syllables with different number and type of onset consonants. Ultrasound recordings provided the information of the distance the tongue travels over a target, and audio recordings of the time needed. The speed of the tongue's travel was calculated from the two measurements. Results of ten speakers have shown that both duration and distance travelled increase with an increased number of onset segments, but that distance travelled is additionally influenced by the type of the segment, as is speed. Duration also seemed to be the least speaker-dependant of the three parameters.

Key words: tongue movement, ultrasound, syllable onset.

## **1. Introduction**

Ultrasound is a safe and non-invasive method which enables visualisation of tongue movement by depicting the upper tongue surface (Stone 2005). To obtain an ultrasound image of the tongue, the probe is placed under the speaker's chin and the inside of the mouth is scanned by emitting high frequency ultrasound waves. These waves are reflected at a boundary between mediums of different density ( tongue/air and tongue/bone) and detected by the probe. After the detection the point of reflection is calculated and an image is created at that point.

Characteristics of differently structured syllable onsets have been investigated in several studies (Greenberg et al. 2003; Crystal and House 1990). The findings generally suggest that syllable duration is mainly affected by the syllable's stress and the number of segments. Syllable duration increases with an increasing number of segments in both stressed and unstressed condition. Stressed syllables and their segments have greater duration than their unstressed counterparts.

The aim of this study was to investigate how tongue movements are affected by changing the type and number of syllable onset segments. Tongue movements are described by the distance travelled by the tongue over an utterance, the duration of an utterance, and the tongue speed. It was hypothesized that both measurements will increase with the increasing number of onset segments, and that speed will vary, depending on the type of segments.

## 2. Methodology

Ultrasound data of ten native English female speakers, aged between 19 and 30 years, was analysed in this study. Speech material consisted of six monosyllabic real English words: pay, say, lay, play, slay, splay.

Midsagittal view of the tongue was recorded with Concept M6 (Dynamic Imaging) ultrasound with a frame rate of 30fps, and a special helmet was used to fix the probe under the speaker's chin. Participants repeated each of the words five times in a frame sentence "a [word] today". Both ultrasound and audio signals were recorded at the same time using Articulate Assistant Advanced, which allows temporal synchronisation of the two signals.

Articulate Assistant Advanced was used for annotating and tracing tongue contours on the recorded ultrasound images. All the reported results are for the "a [word]" part of the recording. The travelled distance of a target utterance was calculated as the sum of average nearest neighbour distances (aNND) between every pair of consecutive tongue contours of the utterance. aNND is an average of all the nearest neighbour distances measured between the points on the two contours of a pair. Duration was measured from audio signal, and the tongue speed was calculated as the distance travelled over duration, to give information about the relationship between the two measurements.

## 3. Results

### 3.1. Duration

The results showed that duration increases with the increasing number of syllable segments (Figure 1a) and that it is statistically significant between all pairs of targets (Figure 1d, solid lines). Additionally, most of the speakers showed the same pattern of increasing duration: single onsets shorter than two consonants clusters, which were also shorter than the three consonants one.

### 3.2. Distance travelled by the tongue

Overall, the distance travelled by the tongue did increase with the increased number of onset segments (Figure 1b) but the increase was not stat. sig. between all of the syllable pairs (Figure 1d, dashed lines). Distribution of the measurements was also more variable than in the case of duration, and individual speakers showed less similar patterns of increasing the distance travelled by the tongue over a syllable. Not all speakers had greater measurements for targets with clustered onset than for those with single onset.

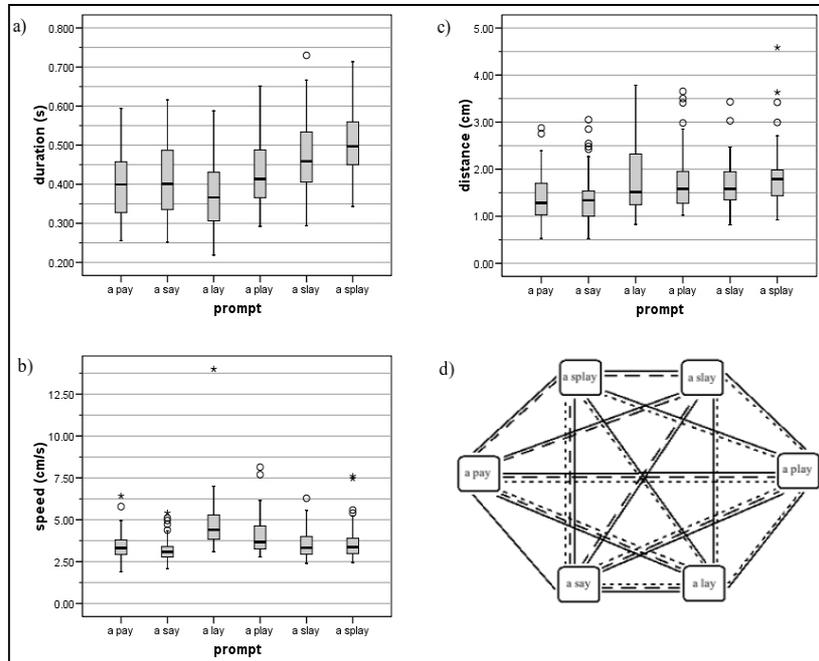


Figure 1. Duration (a), tongue’s distance travelled (b) tongue speed (c) and stat. sig. (Wilcoxon Signed Rank Test,  $p < 0.05$ ) between pairs of targets (d; dashed line = distance travelled, solid line = duration, dotted line = tongue speed.)

### 3.3. Speed of the tongue’s travelling

Speed of the tongue’s travelling is presented in Figure 1c. The tongue travelled the fastest over “a lay” and the slowest over “a say” with most of the target pairs being stat. sig. different from each other (Figure 1d, dotted lines). Number of onset segments did not influence tongue speed, and individual speakers showed very different patterns of increasing tongue speed over targets, although eight out of ten speakers had the fastest tongue speed over “a lay”.

## 4. Discussion and conclusion

Analysis has shown that duration increases with the addition of segments to syllable onset and thus confirmed the results of previous studies (Greenberg et al. 2003; Crystal and House 1990). Results of distance travelled measurements also showed the influence of the number of onset segments, as single onset targets were shorter than clustered ones. However, they were also influenced by the type of the segment. As expected, targets with single onset

/p/ and /s/ were shorter than those with single /l/, which had a distance travelled more similar to the targets with clustered onsets. /p/ is not a lingual consonant and does not contribute any tongue movement to the distance travelled, and /s/ has less tongue movement than /l/. In contrast, /p/ does contribute to the distance travelled in case of clustered onsets as it seems to make movement for /l/ more prominent and less restricted than in case of /sl/ clusters. Tongue speed, on the other hand, does not seem to depend on the number of segments, but mainly on the type.

Based on the results of this study it can also be concluded that duration of spoken target is less speaker-dependant than distance travelled. The latter depends on the size of individual speech organs and can not be adapted as duration can. Consequently, tongue speed is the most speaker-dependant and is appropriately adapted depending on the demands of the space over which the tongue has to travel, and the demands of appropriate speech timing.

This study has demonstrated that ultrasound is sensitive enough to describe continuous tongue movement during speech although it has some limitations. The most important is that it does not produce an image of raised tongue tip with air pocket below it, and thus we miss information about that part of the tongue. Measurements could be also affected by the probe which could potentially restrict jaw movement. Mooshammer et al. (2003) have shown that jaw position is low for /l/ and high for /s/. However, this effect is expected to affect all participants.

### **Acknowledgements**

Many thanks to my supervisor, Dr Nigel Hewlett. Financial support provided by the Marie Curie EdSST programme.

### **References**

- Articulate Instruments Ltd. 2007. Articulate Assistant Advanced User Guide: Version 2.07. Edinburgh, UK: Articulate Instruments Ltd.
- Crystal, T. H. and House, A. S. 1990. Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America* 88, 101-112.
- Greenberg, S., Carvey, H., Hitchcock, L. and Chang, S. Y. 2003. Temporal properties of spontaneous speech - a syllable-centric perspective. *Journal of Phonetics* 31, 465-485.
- Mooshammer, C., Geumann, A., Hoole, P., Alfonso, P., van Lieshout, P. and Fuchs, S. 2003. Coordination of lingual and mandibular gestures for different manners of articulation. In *Proceedings of the 15th ICPHS*, 81-84, Barcelona, Spain.
- Stone, M. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics & Phonetics* 19, 455-501.