Tongue articulation of front close vowels in Stockholm, Gothenburg and Malmöhus Swedish

Frid, Johan; Schötz, Susanne; Gustafsson, Lars; Löfqvist, Anders

Published in:
Proceedings of ICPhS 2015

2015

TONGUE ARTICULATION OF FRONT CLOSE VOWELS IN STOCKHOLM, GOTHENBURG AND MALMÖHUS SWEDISH

Johan Frid[1], Susanne Schötz[1], Lars Gustafsson[1] & Anders Löfqvist[2]

[1]Lund University Humanities Lab, Sweden
[2]Dept. of Logopedics, Phoniatrics and Audiology, Lund University, Sweden
johan.frid@ling.lu.se, susanne.schotz@ling.lu.se, lars.gustafsson@humlab.lu.se & anders.lofqvist@med.lu.se

ABSTRACT

Articulatory data were collected for the Swedish vowels /iː, yˌ, ʉ̟ː/ from nine speakers each of Stockholm, Gothenburg, and Malmöhus Swedish, and the tongue positions and their dynamics analysed using Functional Data Analysis (FDA). Results showed that the general tongue positions for /iː/ and /yː/ are similar and clearly different from /ʉ̟ː/ in all three dialects. Variation within the Stockholm and Gothenburg groups led to a subdivision into two types, where the tongue positions of type 1 resembled Malmöhus Swedish more. Several differences in tongue articulation between types 1 and 2 were observed, possibly explained by the presence of Viby-coloured /iː/ and /yː/ in type 2.

Keywords: articulography, Swedish, vowels, FDA

1. INTRODUCTION

In the Swedish vowel system, there are three contrastive long front, close vowels /iː, yː, ʉ̟ː/, characterised by a relatively small acoustic and perceptual distance. The magnitude of the lip opening is regarded as the major distinctive feature: unrounded /iː/, out-rounded /yː/, and inrounded /ʉ̟ː/ [3, 7]. Specifically the contrast between /yː/ and /ʉ̟ː/ is considered highly unusual among the world's languages. The tongue articulation is assumed to be basically identical, but the documentation of this is incomplete, especially for the articulatory dynamics [7]. To maintain the distinctions between these vowels, they are often characterised by a slight diphthongisation or consonantal off-glide at the end.

Another fairly common realisation of /iː/ and /yː/ in Swedish is as [ɨː] and [ʉ̟ː]. i.e. with a "damped" quality often referred to as Viby-colouring [2, 7]. There is disagreement in the Swedish phonetics literature if the major constriction for the damped /iː/ and /yː/ is further front compared to their regular counterparts, and basically alveolar, or instead further back and rather central [1, 4]. However, as adequate articulatory data seem to be lacking, these views are at best intelligent speculations.

[12] investigated the articulatory dynamics of /iː, yː, ʉ̟ː/ in Gothenburg Swedish (GS) and Malmöhus Swedish (MS), spoken in and near Gothenburg and Malmö. In MS, they found that the position of the tongue body was significantly lower for /ʉ̟ː/ than for /iː/ and /yː/. In GS, the speakers could be subdivided into two different types according to their articulation patterns; type GS1 resembled MS, while type GS2 had higher tongue body for /ʉ̟ː/.

The purpose of this study was to extend these findings by including Stockholm Swedish (SS), spoken in and near Stockholm, and compare the tongue articulation of /iː, yː, ʉ̟ː/ of this dialect to those of GS and MS. Our aim was to find out how SS relates to our findings for MS and GS. Based on the results of [12], we expected the tongue positions in the dimensions open–close and front–back to be different for /ʉ̟ː/ than for /iː/, /yː/ in all three dialects. Furthermore, we expected to find regional differences in the articulation of /iː/ and /yː/, as Viby-colouring is more common in SS and GS than in MS [2]. We also expected to find a subdivision into two types in both GS and SS.

2. MATERIAL AND METHOD

Nine speakers each of SS (3 females, 6 males, age: 21 – 63), GS (5 f, 4 m, age: 20 – 47), and MS (4 f, 5 m, age: 23 – 62) were recorded by means of electromagnetic articulography along with a microphone signal using an AG 500 (Carstens Medizinelektronik). Figure 1 shows the sensor positions and one subject with sensors attached. The speech material consisted of 15–20 repetitions by each speaker of /iː, yː, ʉ̟ː/ in carrier sentences of the

Figure 1: The twelve sensor positions and a speaker with the sensors attached.

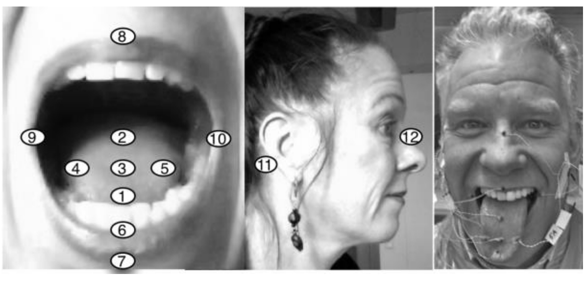


Figure 2: Mean tongue body height (z-score) as a function of normalised time for /iː, yː, ʉ̞ː/ in Malmö (MS) and two types of Gothenburg (GS1, GS2) and Stockholm (SS1, SS2) Swedish.

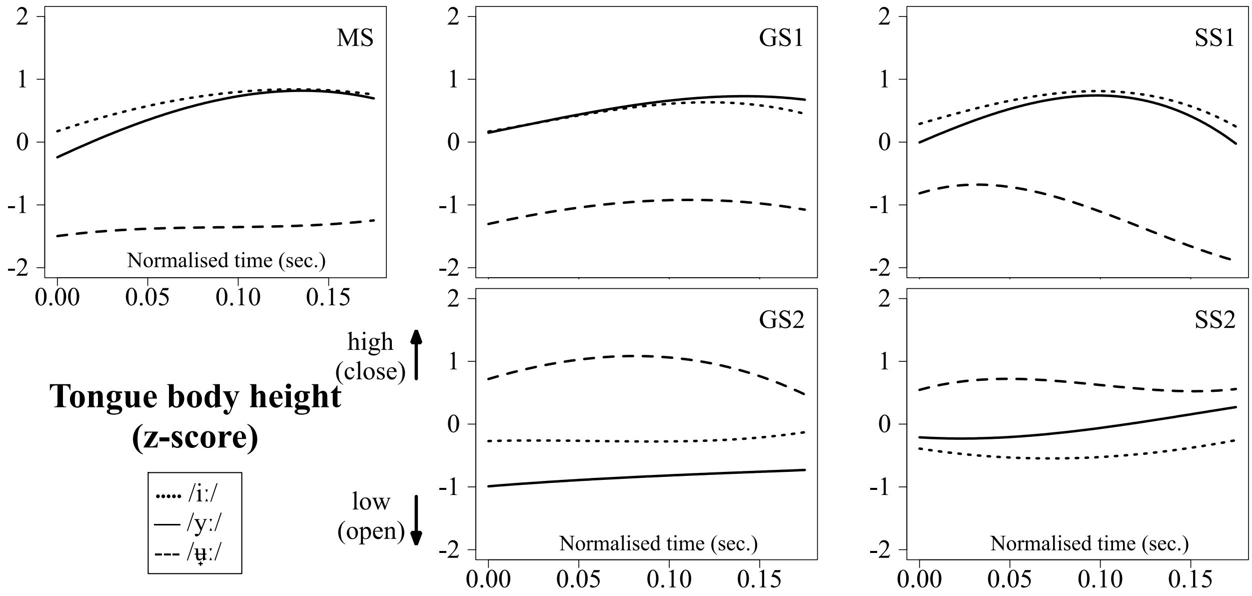
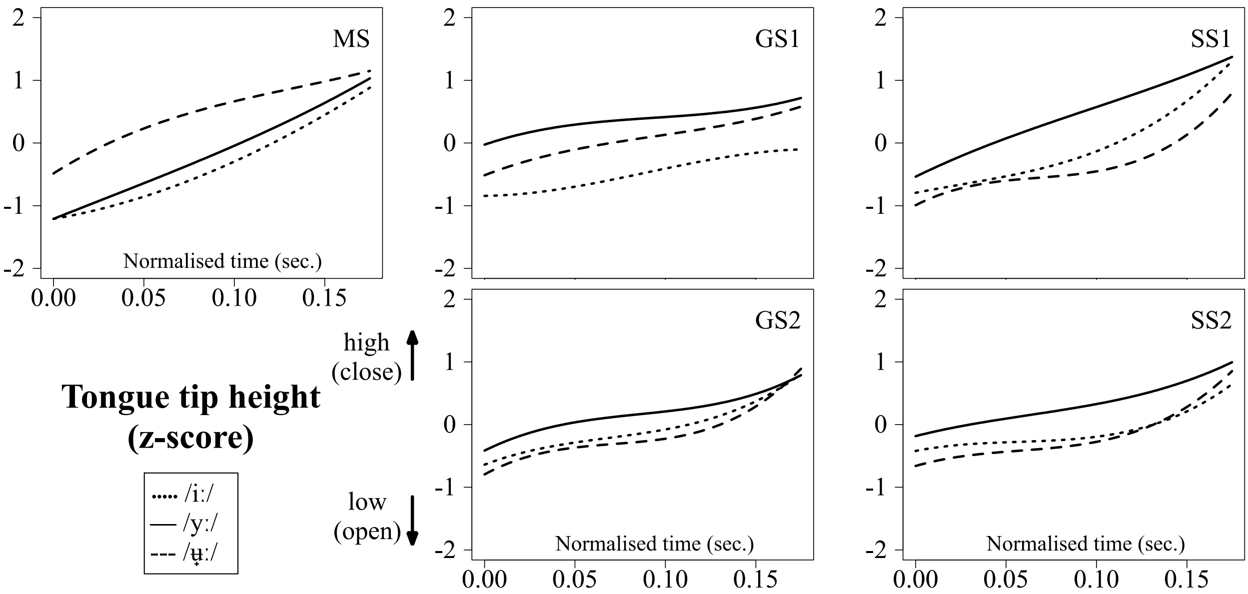


Figure 3: Mean tongue tip height (z-score) as a function of normalised time for /iː, yː, ʉ̞ː/ in Malmö (MS) and two types of Gothenburg (GS1, GS2) and Stockholm (SS1, SS2) Swedish.



type "De va inte hVt utan hVt ja sa" (It was not hVt, but hVt I said), where the target words containing the vowels were stressed. The sentences were displayed in random order on a computer screen, and the speakers were instructed to read each sentence in their own dialect at a comfortable speech rate.

2.1. Error detection and speaker normalisation

Noise and measurement errors are not uncommon in articulatory data due to quick head movements, sensors moving too close to each other, sensors breaking or falling off, or calculation errors. Errors were detected and excluded using the procedure described in [12]. Moreover, to compensate for differences in oral anatomy between speakers, data was normalized using z-score transformation.

2.2. FDA smoothing and aligning

Functional Data Analysis (FDA) is a technique for timewarping and aligning a set of signals to examine differences between them. FDA techniques and applications to speech analysis were first introduced by [11], and further developed in [8, 9, 5]. In this study, FDA was used to smooth the sensor traces, and to standardise the time to facilitate comparisons between repetitions. All FDA processing was done using the R package 'fda' (see [12] for details).

2.3. Analysis of tongue articulation

Sensors 1 and 2 were selected to represent the tongue tip and body (see Figure 1). FDA processed contours were plotted for the tongue body and tip dynamics in height and frontness, and the positions and dynamics compared within as well as across the regional varieties. Statistical analysis was done with functional t-tests (see [10] for details).

3. RESULTS

Generally, the vowel /ʉ̞ː/ displays distinct patterns from /iː/ and /yː/, and /ʉ̞ː/ also varies the most between regions. Among the SS and GS speakers we found a subdivision between speakers (5 type SS1, 4 type GS1) who articulate the three vowels with similar tongue positions as the MS speakers, and speakers (4 type SS2, 5 type GS2) who generally have different tongue positions compared to the MS speakers.

3.1. Tongue body height

Tongue body height is shown in Figure 2. In MS, GS1 and SS1 the position of the tongue body is lower for /ʉ̞ː/ than for /iː/ and /yː/, while in GS2 and SS2 the position is higher for /ʉ̞ː/. We found significant differences between varieties (pairwise functional t-tests, $p<0.05$) throughout the vowel in /ʉ̞ː/ for MS-GS2, MS-SS2, GS1-GS2 and SS1-SS2. For MS-GS1 and MS-SS1 the difference is not significant throughout the whole vowel. The main difference between SS2 and GS2 is that /iː/ has the lowest tongue body in SS2 while /yː/ is lower in GS2. SS1 displays slightly more arched contours for all vowels compared to the other varieties, suggesting a higher degree of diphthongisation or coarticulation.

3.2. Tongue tip height

Figure 3 shows that the tongue tip height for /yː/ is higher than for /iː/ and /ʉ̞ː/ in all varieties except MS, where /ʉ̞ː/ has the highest contour. Between varieties there are significant differences (pairwise functional t-tests, $p<0.05$) in the central part of /ʉ̞ː/ between MS and all others varieties. For GS1-GS2 and SS1-SS2 the difference is not significant. The dynamics for all the vowels in all the varieties is represented by slightly rising contours, suggesting closing diphthongisations, although some individual variation can be observed.

3.3. Tongue body frontness

As shown in Figure 4, the tongue body is more protruded in /iː/ and /yː/ than in /ʉ̞ː/ in all varieties except GS2, which displays an opposite pattern except in the final part of the vowel. /iː/ and /yː/ have similar contours in all varieties, with the clearest overlap in SS2. The vowel contours are either rising slightly (GS1, MS), arch-shaped (e.g. SS1, SS2) or slightly falling (GS2), suggesting different diphthongisation strategies.

3.4. Tongue tip frontness

Tongue tip frontness is shown in Figure 5. In MS the tongue tip is further back in /iː/ and /yː/ compared to /ʉ̞ː/, while the opposite pattern is found for all the other varieties. Between varieties, we found significant differences (pairwise functional t-tests, $p<0.05$) in the middle of /ʉ̞ː/ for MS vs. all the others. We also note somewhat different vowel dynamics in the different vowels and varieties, suggesting different types of diphthongisation gestures. In SS1 all vowels show slight forward-backward movements, but with an earlier timing for /ʉ̞ː/ than for /iː/ and /yː/. All vowels in GS1 move lightly forward, while they move backward in GS2. In MS /iː/ and /yː/ show a forward motion, while the arch-shaped contour for /ʉ̞ː/ suggests a forward-backward-movement.

4. DISCUSSION

The results of this study indicate that the tongue articulation for /ʉ̞ː/ is significantly different from /iː/ and /yː/ in both Stockholm, Gothenburg and Malmöhus Swedish. Our hypothesis of different tongue articulation for /ʉ̞ː/ than for /iː/ and /yː/ was thus confirmed.

Considerable regional variation was observed in this study, not only for each vowel in the front–back and open–close dimensions, but also in the vowel dynamics (diphthongisation). MS often displayed different patterns than SS and GS, supporting our hypothesis of different articulation strategies in different regional varieties, at least in part.

The intra-regional variation found in SS and GS led to a subdivision into the four types SS1, SS2, GS1 and GS2. A closer look showed that the SS1 and GS1 speakers were more often from the outskirts of the Stockholm and Gothenburg areas than the SS2, GS2 speakers. Furthermore, most SS2

Figure 4: Mean tongue body frontness (z-score) as a function of normalised time for /iː, yː, ʉ̟ː/ in Malmö (MS) and two types of Gothenburg (GS1, GS2) and Stockholm (SS1, SS2) Swedish.

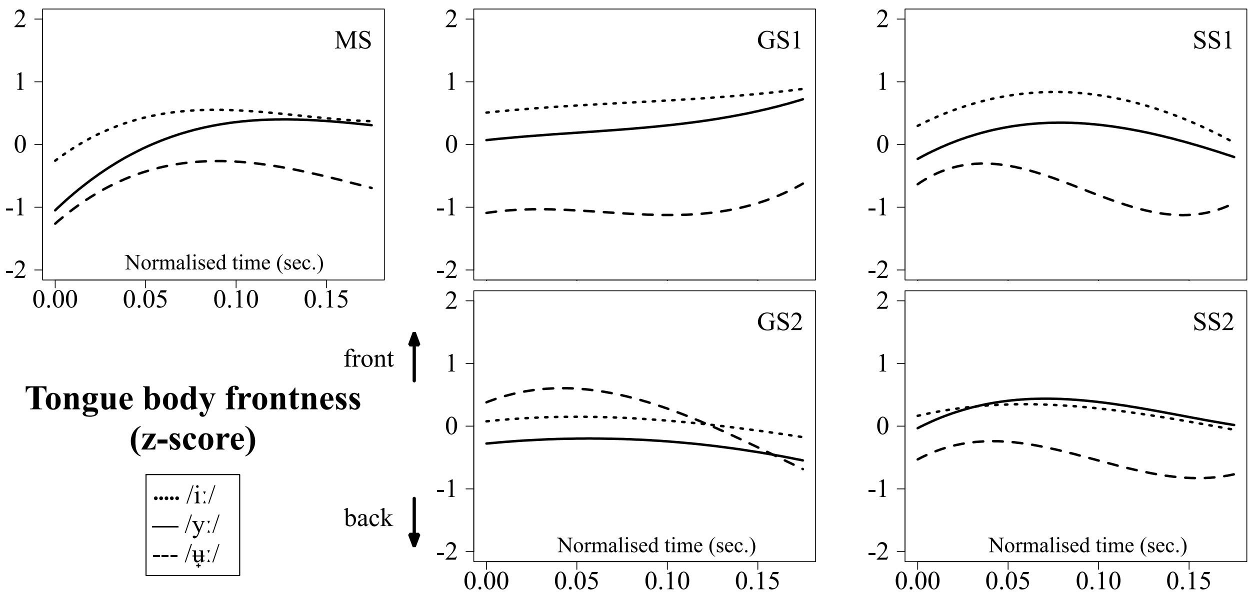
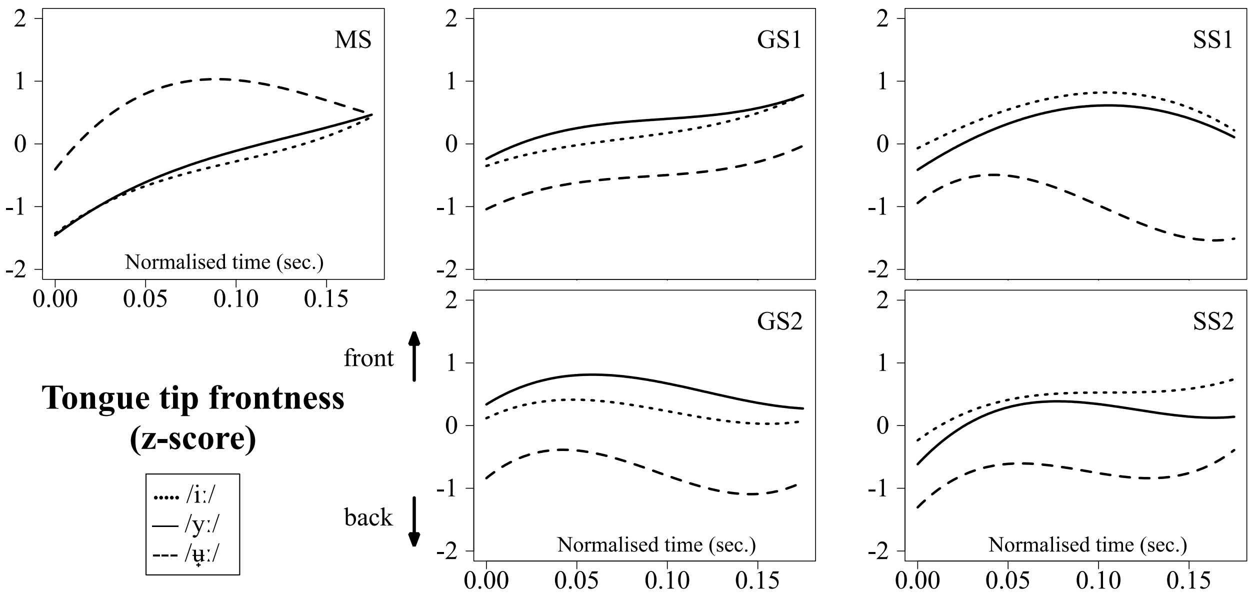


Figure 5: Mean tongue tip frontness (z-score) as a function of normalised time for /iː, yː, ʉ̟ː/ in Malmö (MS) and two types of Gothenburg (GS1, GS2) and Stockholm (SS1, SS2) Swedish.



and GS2 speakers had clear Viby-coloured /iː/ and /yː/, which was not the case for most of the SS1 and GS1 speakers. No MS speakers used Viby-colouring. The Viby-colouring may offer one explanation for the differences in tongue articulation. In future studies, we will investigate this further by comparing articulatory data and acoustic data, e.g. formant frequencies.

In this study we analysed only two discrete points and two dimensions of the tongue: tongue tip and body height and frontness, and used a standard z-score transformation for speaker normalisation. Although we did not look at lip rounding, traditionally regarded as the main difference between /iː/, /yː/ and /ʉ̟ː/, our results clearly show differences between these vowels in tongue body height as well. In future studies, we will compare tongue articulation to lip rounding and we also include a larger number of vowels, e.g. /eː/ and /øː/.

5. REFERENCES

[1] Björsten, S., Bruce, G. Elert, C-C, Engstrand, O., Eriksson, A. Strangert E. & Wretling, P. (1999). Svensk dialektologi och fonetik – tjänster och gentjänster. Svenska landsmål och svenskt folkliv: 7–23.

[2] Bruce, G. (2010). Vår fonetiska geografi. Lund: Studentlitteratur.

[3] Fant, G. (1959). Acoustic description and classification of phonetic units. Ericsson Technics 1. (Reprinted 1983 in Fant, G., Speech Sounds and Features. 32–83. Cambridge, MA: MIT Press).

[4] Engstrand, O., Björsten, S., Lindblom, B., Bruce, G. & Eriksson, A. (2000). Hur udda är Viby-i? Experimentella och typologiska observationer. Folkmålsstudier 39: 83–95. Helsingfors.

[5] Gubian, G., Cangemi, F., & Boves, L. (2011). Joint analysis of F0 and speech rate with functional data analysis. ICASSP, 4972–4975, Prague.

[6] Hadding, K., Hirose, H. & Harris, K. S. (1976). Facial muscle activity in the production of Swedish vowels: An electromyographic study. J. of Phonetics 4: 233–245.

[7] Ladefoged, P. & Maddieson, I. (1996). The Sounds of the World's Languages. Oxford: Blackwells.

[8] Lucero, J., Munhall, K., Gracco, V. & Ramsay, J. 1997. On the registration of time and the patterning of speech movements, J. Speech Lang. Hear Res. 4:1111–1117.

[9] Lucero, J. & Löfqvist, A. (2005). Measures of articulatory variability in VCV sequence, Acoust. Res. Lett. Online 6, 80–84.

[10] Ramsay, J., Hooker, G. & Graves, S. (2009). Functional Data Analysis with R and MATLAB. Springer.

[11] Ramsay, J. O., Munhall, K. G., Gracco, V. L. & Ostry, D. J. (1996). Functional data analysis of lip motion, J. Acoust. Soc. Am., 99, 3718– 3727.

[12] Schötz, S., Frid, J., Gustafsson, L, Löfqvist, A. (2013). Functional Data Analysis of Tongue Articulation in Palatal Vowels: Gothenburg and Malmöhus Swedish /iː, yː, ʉ̟ː/. In Proceedings of Interspeech 2013, Lyon.