An EPG and Ultrasound Study of Lingual Coarticulation in Vowel-Consonant Sequences

Natalia Zharkova

Queen Margaret University
E-mail: nzharkova@qmu.ac.uk

Abstract

While EPG registers the location and amount of tongue-palate contact, ultrasound can capture most of the tongue contour. Previous studies have not systematically quantified lingual coarticulation using EPG and ultrasound simultaneously. This study used both techniques for analysing vowel-on-consonant coarticulatory effects.

Four speakers of Scottish English produced /VC/ sequences with the consonants /p, f, t, s, l, r, k/ and the vowels /a, i/. The difference between each consonant in the two vowel contexts was computed using an EPG measure and an ultrasound measure. Additionally, temporal coarticulation was analysed, using EPG data.

A significant positive correlation was observed between the two measures, with labial consonants, followed by /r/, having the highest values. The two techniques also provided complementary data on lingual coarticulation. The velar stop was more coarticulated on the EPG measure than on the ultrasound measure, because EPG registered a shift in closure location across vowel contexts, while ultrasound captured the close proximity of the tongue root across the vowel contexts. The sibilant was more coarticulated on the ultrasound measure than on the EPG measure, because ultrasound, unlike EPG, registered vowel-dependent difference in the tongue root. Combined EPG and ultrasound data would be useful in future studies of coarticulation.

1 Introduction

Electropalatography (EPG) and ultrasound register tongue movements in different ways: EPG provides information on the location and amount of tongue-palate contact, while ultrasound captures differences between various tongue shapes.

EPG is an established technique for measuring lingual coarticulation. A review of EPG-based coarticulatory indices is presented in [4]. Ultrasound is an articulatory technique that captures most of the tongue contour, and is therefore capable of providing valuable data on lingual coarticulation ([8]). Ultrasound has recently been used for quantifying lingual coarticulation. Some studies quantify coarticulation extent in consonants based on the whole midsagittal curve data (e.g., [5, 2, 15]), others do not use whole curve data (e.g., [12, 13]).

There are very few studies using combined EPG, ultrasound and acoustic data for speech analysis. None of these studies used synchronised ultrasound and EPG data for systematic quantification of coarticulatory effects. In [9], cross-sectional tongue shapes based on coronal scans and linguopalatal contact were examined in several CVC utterances produced by one subject. The study focused on comparing tongue shapes and EPG patterns across phonemes, and on the relation between tongue shapes and palatal contact patterns within each phoneme. In [11], coronal ultrasound scans were collected in addition to EPG data and jaw displacement data, in order to investigate how the tongue adjusts its position to compensate for conflicting coarticulatory demands. For only one of three subjects EPG and ultrasound data were recorded simultaneously, and the ultrasound transducer was hand-held. In [10], EPG data were collected at a separate session from the ultrasound data, and used to reconstruct three-dimensional tongue surfaces. In [7], a methodological description was offered of the possible use of ultrasound and EPG together. Several studies (e.g., [1] and references cited there) used EPG and ultrasound for analysing vowel production in adolescents with hearing impairment. Quantification was based only on EPG, acoustic and transcription data; no ultrasound measurements were made.

The aim of this study was to use synchronised EPG, ultrasound and acoustic data for measuring vowel-on-consonant (V-on-C) coarticulatory effects in several English consonants. The difference between tokens of the same phoneme across two different environments was quantified separately from EPG and ultrasound data. This difference was taken as a measure of the environment influence upon the realisation of the phoneme.

The Coarticulation Index (CI, [3]) was used as an EPG measure of consonant coarticulation. This index represents the difference between overall tongue-palate contact of a given phonetic segment in two different contexts: for example, /t/ from /ata/ versus /t/ from /iti/. In order to be able to compare EPG and ultrasound results, a very similar measure needed to be taken using the ultrasound data. The difference between overall tongue shapes of a given consonant in two different vocalic contexts ([15]) was chosen to be the ultrasound measure of V-on-C coarticulation.

Additionally, as EPG data provide detailed information on timing of coarticulation, the onset of tongue movement towards the target consonant was measured and compared across lingual consonants.

2 Method

Simultaneous EPG, ultrasound and acoustic recordings were made for four adult native speakers of Scottish English, two male and two female, using the Queen Margaret University multi-channel system. A Concept M6 Digital Ultrasonic Diagnostic Imaging System was used, with an electronic endocavity transducer type 65EC10EA (120 degrees field of view). The transducer frequency was 6.5 MHz; the ultrasound frame rate was 30 Hz. A helmet was used for immobilising the head in relation to the transducer. WinEPG^{TM} was used; the EPG frame rate was 200 Hz. The software for data recording and analysis was "Articulate Assistant Advanced" Version 2.07 (Articulate Instruments Ltd, [14]).

The data were /$C_1V_1C_1V_1$/ sequences (with a word boundary after the first vowel: e.g., "Leigh leads") with the consonants /p, f, t, s, l, r, k/ and the vowels /a, i/, in real sentences, repeated five times. The target consonant for the analysis was the consonant following the word boundary.

For each token, an annotation was placed at the middle of the consonant (for stops, at the middle of the closure). The tongue-palate contact pattern and the tongue contour outline were captured at that time point. For each subject and for each consonant, two coarticulation measures were taken using 25 possible combinations of tokens from the two vowel contexts.

For each token, the number of activated electrodes in each row was expressed as a percentage of the total number of electrodes in that row. Then, for each row, the absolute value of the difference in the number of activated electrodes between the consonant in the context of /a/ and in the context of /i/ was computed. The CI was calculated by averaging the absolute difference values for all rows.

A cubic spline was fitted (automatically, with subsequent manual correction) to the tongue surface contour at each annotation point. Each spline was defined in terms of x-y values. For each subject and for each consonant, the Distance between tongue Curves (DC) in the two vowel contexts, in mm, was then computed in Matlab. Tongue curve comparison was carried out using the technique based on mean nearest neighbour distances between curves ([15]).

Univariate ANOVAs were run for cross-consonant comparison, separately for CI and DC. A Pearson's correlation was performed between CI and DC.

Temporal coarticulation was compared in five lingual consonants, using EPG data. For each subject and each VC token, total contact in each EPG palate row was measured at 25 equally spaced time points, the first point being at the middle of the preceding vowel, and the last point at the middle of the target consonant. For each token, it was recorded at which time point the steady increase of tongue-palate contact started in the posterior zone (the last four rows) for the velar consonant, and in the anterior zone (the first four rows) for the other consonants. Then, separately for each subject and each consonant, this time point value for each token from the context of /a/ was paired with the time point value for each token from the context of /i/, producing two sets of 25 values. Each pair was averaged; this resulted in 25 time point values for each consonant, representing the onset of tongue movement from the vowel towards the consonant (Ons). These values were compared across consonants, using a Univariate ANOVA. A Pearson's correlation was performed between Ons and CI.

3 Results

Table 1 contains DC and CI values, as well as Ons values (Ons is expressed as percentage from the transition onset), the data for all subjects are pooled together. A significant positive correlation between CI and DC was observed (r = 0.31; $N = 700$; $p < 0.01$). A significant effect of Consonant on both measures was reported (DC: $F = 71.56$; $df = 6$; $p < 0.001$; CI: $F = 83.09$; $df = 6$; $p < 0.001$). Labial consonants had the highest CI and DC values. The consonant /r/ had the highest values among lingual consonants on both measures, followed by /t/ and /l/. The consonants /k/ and /s/ had the smallest values of both measures, the velar having the lowest DC value, and the sibilant demonstrating the lowest CI value.

Table 1. *DC, CI and Ons values. Standard Deviations are in brackets.*

C	DC	CI	Ons
/p/	9.69 *(2.36)*	30.50 *(11.77)*	-
/f/	9.50 *(2.59)*	27.53 *(7.30)*	-
/t/	7.61 *(2.15)*	19.28 *(6.06)*	60.00 *(10.04)*
/s/	6.88 *(1.48)*	10.80 *(4.63)*	54.00 *(17.08)*
/l/	7.33 *(1.24)*	15.78 *(8.15)*	64.40 *(9.80)*
/r/	8.35 *(1.77)*	24.01 *(6.86)*	75.80 *(14.44)*
/k/	5.07 *(1.12)*	14.52 *(8.89)*	35.00 *(10.04)*

Figures 1 and 2 show tongue contours and tongue-palate contact patterns for /s/ and /k/, in subject S2. Five repetitions are presented as separate lines on tongue contour plots, and as averaged values on tongue-palate contact diagrams. Filled squares on the diagrams represent 100% contact in all repetitions; squares with numbers show percentage values; white squares represent no contact in any repetition.

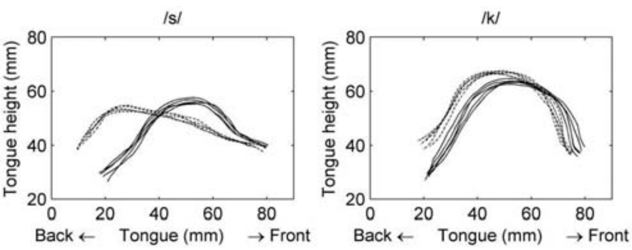

Figure 1. *Tongue contours for /s/ and /k/, subject S2. Solid lines – the context of /i/; dashed lines – the context of /a/.*

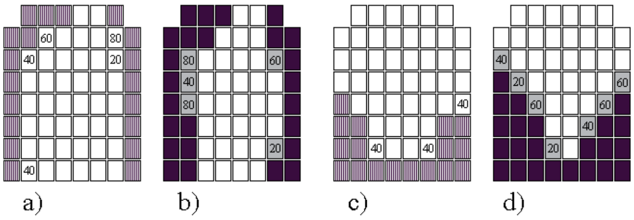

Figure 2. *EPG patterns for /s/ and /k/, subject S2: a) /s/ in the context of /a/; b) /s/ in the context of /i/; c) /k/ in the context of /a/; d) /k/ in the context of /i/.*

The results of the temporal coarticulation analysis showed a significant effect of Consonant ($F = 141.91$; $df = 4$; $p < 0.001$). The velar stop had the earliest onset of tongue movement towards the consonant (at around one third of the VC transition period); /r/ had the latest onset (at around three quarters of the transition). The consonants /l/ and /t/ had very similar onset times (just below two thirds of the transition). The sibilant was between the coronal stops and the velar stop, closer to the coronals.

A significant positive correlation was observed between CI and Ons ($r = 0.385$; $N = 500$; $p < 0.01$).

4 Discussion

Both EPG and ultrasound measures of spatial coarticulation showed that labial consonants were most affected by the vowels, and that /r/ was the most affected among the lingual consonants.

This study showed that the two techniques can provide complementary data on lingual coarticulation. EPG does not register any information on the behaviour of the part of the tongue that is further back than the most posterior place of tongue-palate contact. In /s/, ultrasound registered vowel-dependent difference in the tongue root, while EPG did not. This explains greater coarticulation in the sibilant on DC than on CI.

One more result due to differences between the two measurement techniques consisted in greater coarticulation of the velar stop on CI than on DC. EPG registered a shift in closure location in the front-back dimension, across vowel contexts, while ultrasound captured a relatively small change in tongue posture across the two vowel contexts, and this small change was reflected in a small DC value.

The positive correlation between temporal and spatial measures of lingual coarticulation suggests that the motion of the tongue region responsible for creating a constriction/closure towards the consonant target tends to start earlier in the consonants which are less affected by the preceding vowel. While this is generally the case for the coronal consonants analysed in this study, there is a noticeable difference between CI and Ons results between /k/ and /s/. The smallest Ons value was reported for /k/, however this consonant had a higher CI than /s/ did. This finding is consistent with X-ray dynamic data reported in [6]. The earlier onset of the active articulator motion towards the consonant for the velar might be explained as follows: "Because of differences between musculature and the larger mass which must be moved, this motion is somewhat slower and must begin earlier than the comparative tongue-tip gesture for postdental consonants" ([6], p. 20).

The results of this study suggest that using synchronised EPG and ultrasound data provides complementary information on lingual articulation, which would benefit future studies of coarticulation. Data from more speakers are needed, to establish the extent of possible inter-speaker variation.

References

[1]. P. Bacsfalvi, B.M. Bernhardt, & B. Gick. Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. *Advances in Speech-Language Pathology*, 9, 36-45, 2007.

[2] L. Davidson. Coarticulation in contrastive Russian stop sequences. In *Proceedings of the 16th International Congress of Phonetic Sciences*, 417-420, 2007.

[3]. E. Farnetani. V-C-V lingual coarticulation and its spatio-temporal domain. In W.J. Hardcastle and A. Marchal (Eds), *Speech Production and Speech Modelling*. Kluwer Academic, The Netherlands, 93-110, 1990.

[4]. F. Gibbon and K. Nicolaidis. Palatography. In W. Hardcastle & N. Hewlett (Eds), *Coarticulation: Theory, Data and Techniques.* Cambridge University Press, Cambridge, 229-245, 1999.

[5]. M. Gordon, R. Kennedy, D. Archangeli, & A. Baker. Distributed effects in coarticulation: an ultrasound study. [Oral presentation at Ultrafest IV, New York, USA, 28-29 September 2007.]

[6]. J.S. Perkell. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. MIT Press, Cambridge, MA, 1969.

[7]. J. Scobbie, S. Wood, & A. Wrench. Advances in EPG for treatment and research: an illustrative case study. *Clinical Linguistics and Phonetics*, 18**,** 373-389, 2004.

[8]. M. Stone. A guide to analyzing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, 19, 455-502, 2005.

[9]. M. Stone, A. Faber, L.J. Raphael, & T.H. Shawker. Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ], and [l]. *Journal of Phonetics*, 20, 253-270, 1992.

[10]. M. Stone & A. Lundberg. Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America*, 99, 3728-3737, 1996.

[11]. M. Stone & E. Vatikiotis-Bateson. Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics*, 23, 81-100, 1995.

[12]. Y. Vazquez Alvarez and N. Hewlett. The trough effect: an ultrasound study. *Phonetica*, 65:105-121, 2007.

[13]. S. Wodzinski and S. Frisch. Ultrasound study of velar-vowel coarticulation. *Journal of the Acoustical Society of America*, 120:3373-3374, 2006.

[14]. A. Wrench. Articulate Assistant Advanced: ultrasound module. [Oral presentation at Ultrafest IV, New York, USA, 28-29 September 2007.]

[15]. N. Zharkova. Quantification of coarticulatory effects in several Scottish English phonemes using ultrasound. *QMU Speech Science Research Centre Working Papers*, WP-13, 2007.

Acknowledgements

I am grateful to William Hardcastle, Fiona Gibbon and Marko Liker for advice, and to Alan Wrench and Steve Cowen for technical help. Supported by an ESRC postdoctoral fellowship, PTA-026-27-1268.