VARIABILITY OF LINGUAL STOPS IN ENGLISH: AN ELECTROPALATOGRAPHIC STUDY

D. Waters, K. Nicolaidis, W.J. Hardcastle and F. Gibbon
Queen Margaret College, Edinburgh, UK

ABSTRACT

This electropalatographic (EPG) study investigates aspects of normal articulatory variability in English lingual stops in VCV sequences. Variability due to the effects of vocalic context, voicing and type of speech material (real versus nonsense words) are examined. The methodology for data collection, segmentation and analysis is described. Results of the analysis will be available at the Stockholm Congress.

INTRODUCTION

Quantification of the extent of normal variability of articulatory movement during speech is of importance both in the development of robust automatic speech recognition systems and in the assessment and remediation of disordered speech. For the latter, it is essential to use baseline normative data against which to compare abnormal articulatory patterns. EPG investigation of tongue contacts with the hard palate during speech offers a means of making systematic investigations of aspects of normal articulatory variability. Previous studies have suggested various factors potentially affecting the variability of EPG patterns [1,2]. This study investigates variability associated with inter-speaker differences, vowel context, effects of voicing and of different types of speech material (real and nonsense words). It contributes towards establishing regions of relative invariance and variability in tongue contact patterns for lingual stops in English.

DATA

The speech data (from the EUR-ACCOR database, [3]) consisted of VCV sequences in English in both real and nonsense words, where V = /i, a/ and C = /t, d/. Five English speaking subjects each produced five repetitions of each VCV combination in nonsense and real words (80 tokens from each subject). Data were acquired with the mulitchannel system developed by Reading University and IBM. Electropalatographic (Reading EPG), laryngographic and acoustic data were recorded simultaneously for each item. [4, 5].

DATA SEGMENTATION PROCEDURE

A multi-level approach was adopted in the segmentation of the data. The placement of annotation points was based on information from the acoustic waveform, the electropalatographic data and the laryngographic signal. Twelve annotation points were marked and labelled for each VCV sequence. They were saved in separate annotation files for each utterance. These files were then used for further analysis of the data. Five of these annotation points are relevant for the current analysis of stops.

Annotation points

sce (stop closure) Taken at the first EPG frame showing complete constriction at the alveolar region. In cases of incomplete closure it was identified as the first frame of maximum constriction in the alveolar region.
sre (stop release) Taken at the last frame of complete constriction at the alveolar region. In cases of incomplete closure it was taken as the last frame of maximum constriction at the alveolar region.
sm (stop midpoint) Taken at the temporal midpoint between 'sce' and 'sre'.
mce (maximum contact from the EPG) Taken at the first frame of maximum constriction in the first four rows of electrodes during the stop closure.
lgp (last glottal pulse) Based on the laryngographic signal this point was taken at the end of periodic pulsing after the onset of closure for the consonant.

DATA ANALYSIS

Analyses were carried out using the Paradox 3 relational database and statistical analyses were performed using a PC running the SPSS statistical package. A number of quantitative and qualitative analysis procedures were undertaken to answer specific questions concerning the nature of articulatory variability of the two alveolar plosives.

Preliminary analysis of the data had indicated a number of cases where the stops were realised as fricatives with evidence of turbulent noise in the acoustic waveform trace. Such items were excluded from the quantitative analysis, however they were annotated separately and included in the qualitative analysis. The following analyses were carried out.

Contact totals

The total number of contacted electrodes in the four front and four back rows of the EPG palate were calculated separately. These measures were made for each /t/ and /d/ token involving symmetrical vowel environments. The measurements were made at the three consonantal points 'sce' (onset of constriction), 'mce' (first frame of maximum constriction in the first four rows) and 'sre' (end of constriction). This analysis aimed to identify any differences in the amount of contact between voiced and voiceless plosives. Because of the aerodynamic requirements for the production of voicing during stops the prediction would be that /d/ would involve less lingual-palatal contact than /t/.

Variability Index

EPG prototypical frames displaying frequency of electrode contact in five repetitions were computed. Based on these a variability index was calculated following Farnetani & Provaglio [6]. This index was calculated for consonants in symmetrical sequences at point 'sm' and used to examine differences in variability between the two consonants, between the two utterance types and among the five subjects.

Coarticulatory index

Contextual variability was quantified using a coarticulatory index (CI) [7]. This index quantifies the amount of tongue-palate contact in different vocalic environments. CIs calculated for symmetrical environments show global effects of the vocalic environment on the consonant. Contact for symmetrical sequences can then be compared with asymmetrical environments to determine possible anticipatory and carryover effects at the beginning (sce) and end (sre) of the consonant respectively.

Voicing

A descriptive analysis was made of the percentage of voicing during the closure phases for the consonants /t/ and /d/. Comparisons were made among subjects and between the two kinds of speech material (real and nonsense words). Results are displayed in the form of histograms.

Qualitative Analysis

Further descriptive analysis was carried out in order to display examples of contact patterns for /t/ and /d/ at the extremes of variation (maximum and minimum contact required for the

production of /t/ and /d/) for each subject.

STATISTICAL ANALYSIS

Various statistical analyses are currently being explored. The factors to be examined are:

Subject. Five subjects.
Consonant. /t/, /d/.
Vowel 1. / i /, /a/.
Vowel 2. / i /, /a/.
Speech Material. Real word, nonsense word.

RESULTS

At the time of writing the above analyses are in progress. Results will be displayed as a poster at the Stockholm Congress and will be made available in printed form to delegates, together with a discussion of the results, as a supplement to this paper.

REFERENCES

[1] Butcher, A., & Weiher, E. (1976), "An electropalatographic investigation of coarticulation in VCV sequences." *Journal of Phonetics,* vol. 4, No.1, pp. 59-74.

[2] Dagenais, P.A., Lorendo, L.C. & McCutcheon, M.J. (1994), "A study of voicing and context effects upon consonant linguapalatal contact patterns." *Journal of Phonetics,* vol. 22, pp. 225-238.

[3] Hardcastle, W.J. & Marchal, A. (1990), "EUR-ACCOR: a multi-lingual articulatory and acoustic database." *ICSLP 90, Proceedings of the International Conference on Spoken Language Processing, Kobe, Japan, November 18th - 22nd,* vol. 2, pp.1293-1296.

[4] Hardcastle, W., Gibbon, F. and Nicolaidis, K. (1991), "EPG data reduction methods and their implications for studies of lingual coarticulation." *Journal of Phonetics,* vol. 19, No. 3/4, pp. 251-266.

[5] Hardcastle, W.J., Jones, W., Knight, C. Trudgeon, A. and Calder, G. (1989), "New developments in electropalatography: a state-of-the-art report. *Clinical Linguistics & Phonetics,* 3, pp. 1-38.

[6] Farnetani, E., & Provaglio, A. (1991), "Assessing variability in lingual consonants in Italian." *Quaderni del Centro di Studio per le Ricerche di Fonetica del C.N.R.,* vol. X, pp. 117-145.

[7] Farnetani, E., Hardcastle, W. J. & Marchal, A., (1989), " Cross-language investigation of lingual coarticulatory processes using EPG." In Tubach, J.P. & Mariani, J.J. (Eds) *Eurospeech 1989, European Conference on Speech Communication and Technology, Paris,* pp. 429-432.