Contour-based analysis of EGG data from words in isolation and connected speech

Stephen D. Bier, Catherine I. Watson, Clare M. McCann

University of Auckland

s.bier@auckland.ac.nz, c.watson@auckland.ac.nz, c.mccann@auckland.ac.nz

Index Terms: Electroglottography, contact quotient, F0, contour analysis

1. Introduction

When examining changes in voice quality, it is common to use averaged measures from selected sections of speech recordings, giving static measures, as seen in [1, 2]. These include means and standard deviations of F0, contact quotient (Qx), and perturbation measures such as jitter and shimmer. Standard deviations and perturbation measures provide meaningful information when phonation is held in a constant state, such as with sustained vowels, but provide less information about speech stability in connected speech, where prosodic variation may disguise the variations caused by decreased stability.

2. Methodology

Electroglottographic (EGG) recordings of 15 young male speakers (ages 20-26) were taken. Each speaker was recorded reading the Rainbow passage and reading five randomised lists of 11 different hVd words [2]. An algorithm was developed in R (www.r-project.org/) to calculate the F0 and Qx contours from the EGG waveforms. This was adapted from previously developed algorithm to allow examination of the contours [2]. The examination of Qx contours is a novel approach to analysis of speech quality. F0 and Qx contours were calculated for the vowels in the hVd words and for the vowel portion from eight instances of the syllable ‘rain’ taken from the rainbow passage. The vowel from ‘rain’ was selected as it was repeated in stressed conditions. 874 hVd word and 116 ‘rain’ tokens were used in the analysis. Initial analysis of the shape of the contours was performed using the discrete cosine transform (DCT) [3]. The first coefficient (DCT1) was used to gain an indication of the mean for each contour, and the second coefficient (DCT2) was used to gain an indication of the slope of each contour. The contours for the vowel from ’rain’ were classified into upward sloped and downward sloped based on the first DCT coefficient for the F0 contour for further analysis.

3. Results and Discussion

Time normalised, averaged contours for the hVd vowels and the ‘rain’ vowels can be seen in Figure 1. The hVd results are split by vowel and the ‘rain vowels are split by F0 slope. Repeated measure ANOVAs were used to analyse DCT1 and DCT2 of the F0 and Qx contours for both the hVd vowels and the ‘rain’ vowels. Greenhouse-Geisser corrections were used where necessary. For the hVd words, there were significant interactions with vowel for F0 DCT1 ($F_{(3.8,102.6)}$=32, p<0.001), F0 DCT2 ($F_{(5.2,140.4)}$=3.8, p<0.01), Qx DCT1 ($F_{(4.4,118.8)}$=12.2, p<0.001), and Qx DCT2 ($F_{(4.6,124.2)}$=8.1, p<0.001). There were no significant interactions with slope for the ‘rain vowels’ other than for DCT2 of the F0 contour, which was used to generate the grouping.

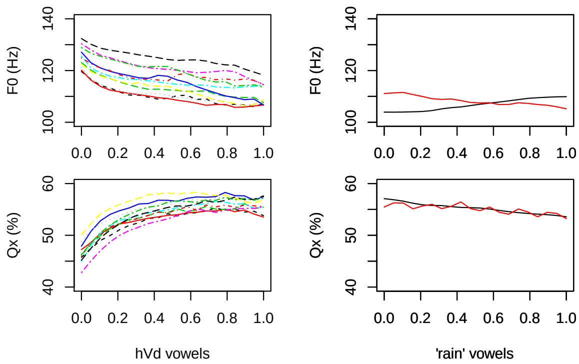

Figure 1: *Averaged time normalised F0 and Qx contours for the hVd vowels and ‘rain’ vowels. For the hVd vowels, each line is the average for a different vowel. For the ‘rain’ vowels, the lines correspond to upward sloping and downward sloping vowels.*

For hVd words the Qx contour has a large initial rise. This may be a coarticulation effect related to the use of an hVd frame. There was no obvious coarticulation effect apparent in the shape of the Qx contours for the ‘rain’ vowels. The ‘rain’ vowels show little difference in the Qx contour whether the F0 contour was upward sloping or downward sloping, suggesting that the Qx behaviour is independent of pitch. The Qx contour provides further insight into dynamic measures of voice behaviour, and has promising applications for studying connected speech. The Qx contour could be used in investigation of different coarticulation effects on vocal fold behaviour. It also allows quantification of dynamic vocal fold behaviour that could be used for comparisons between speaker groups, such as in a study of the aging voice.

4. References

[1] Schötz, S., “Acoustic analysis of adult speaker age”, Speaker Classification I, 2007.

[2] Bier, S.D., Watson C.I. and McCann, C., “Using the perturbation of the contact quotient of the EGG waveform to analyse age differences in adult speech”, Journal of Voice, 2014.

[3] Teutenberg, J., Watson, C.I. and Riddle, P., “Modelling and Synthesising F0 Contours with the Discrete Cosine Transform”, Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, 2008.