Quantifying vocal fold activity: two new methods for analysing electroglottographic data

This paper presents two new methods for assessing vocal fold activity using electroglottography and discusses the advantages that these methods have over the use of the contact quotient [2]. The position of the vocal folds within the larynx makes it difficult to measure glottal activity. Electroglottography (EGG) is a technique for studying vocal fold activity in a non-invasive manner, by estimation of vocal fold contact area from the admittance of an electrical current that passes through the neck [1]. From the EGG signal, it is possible to derive the contact quotient (CQ), i.e. the duration of the closed phase as a proportion of glottal cycle duration [2], and the wavegram, a technique for visualising overall vocal fold activity through time [3]. In this paper, I introduce a less resource-intensive visualisation method called tracegram, and I discuss a method for assessing wavegram data statistically using generalised additive models (GAMs).

Four native English speakers with phonetic training (1 F, 3 M) were recorded while uttering 10 tokens of the vowel [ɑ] with modal voicing and 10 tokens with breathy voicing. The first derivative (dEGG) of the filtered (EGG) signal was calculated and smoothed using a weighted moving average filter [5]. The times of the maximum and minimum peaks of the individual dEGG cycles—roughly corresponding to the beginning of the contacting and decontacting phase respectively—were extracted from a 500 millisecond portion of each vowel token (centred around the mid point of the token), and normalised as a proportion of the local dEGG cycle time. The CQ for each cycle was calculated as the difference between the normalised time values of the dEGG minimum and maximum [4]. A lower CQ indicates a breathier phonation, while higher CQ values correspond to creakier voice, with values for modal voicing lying around the middle. Wavegram data were also extracted following the method described in Herbst et al. [3], which provides a general overview of vocal fold activity by extracting measurements from the whole dEGG signal rather than just the peaks in the signal. Finally, the dEGG maximum and minimum were employed for visual assessment using dEGG tracegrams. A dEGG tracegram is a time-series graph in which the normalised time values of the dEGG maximum and minimum are plotted on the y-axis (where 0 and 1 represent the beginning and end of the cycle respectively). The two curves of a tracegram, one for the dEGG maximum and one for the minimum, represent vocal fold activity, and changes in the curves correspond to changes in fold activity.

As a general pattern, 3 of the 4 participants recorded for this study show lower CQ values for breathy tokens than for modal tokens (speakers 1, 2, 4; see Figure 1). Contrary to expectation, the CQ for modal and breathy tokens are almost identical in speaker 3, and slightly higher CQ values are in fact registered for the breathy tokens. This result is striking given that for speaker 3 the two phonation types are impressionistically distinguishable upon listening. A linear mixed effects regression model indicates that phonation type does not have an effect on average CQ ($\chi^2(1) = 3.159, p > 0.05$). However, a look at the tracegram of the individual participants reveals differences in fold activity not only in speakers 1, 2, 4, but also in speaker 3, although the magnitude of the difference is smaller in the latter (Figure 2). The GAM analysis of the wavegram data from the four participants indicates an overall significant difference in fold activity between modal and breathy phonation ($\chi^2(8) = 18015.72, p < 0.001$). This indicates that the wavegram GAM analysis is sufficiently sensitive even to small differences in vocal fold activity which might not be captured by the CQ. To conclude, wavegram GAMs are the recommended method for assessing vocal fold activity in running speech, while tracegram plots can offer a less resource-intensive method for visualising fold activity.


