

Characterization of features observed on the derivative of electroglottographic signal by the use of high-speed cinematography

Nathalie Henrich¹, Cédric Gendrot², Götz Schade³, Frank Müller³, Robert Expert¹

(1) Laboratoire d'Acoustique Musicale (CNRS, UPMC, Ministère de la Culture)

(2) Laboratoire de Phonétique et Phonologie (CNRS UMR 7018, Paris 3)

(3) Poliklinik für Hör-, Stimm- und Sprachheilkunde, Universitätsklinikum Hamburg-Eppendorf
henrich@lam.jussieu.fr, cgendrot@univ-paris3.fr

Abstract

The electroglottography (EGG) is a well-known and commonly used non-invasive method for measuring vocal folds contact area. However, few studies have paid attention to its derivative, although the differentiated EGG signal (DEGG) helps to enhance sudden changes in the vocal folds contact during the opening and closing phases. During the production of voiced sounds, the DEGG signal presents strong peaks, which can accurately be related to the glottal closing instants defined as the instants of termination of glottal area variation, and weak peaks of opposite sign, which can be related to the glottal opening instants defined as the instants of initialization of glottal area variation (Childers *et al.*, 1990). As glottal closing is usually abrupt, the closing peaks are often very strong and precise, whereas the opening peaks are often weaker and less precise, a fact that leads to some reservations on their usefulness for glottal opening instant detection (Baken, 1992). The DEGG peaks can be either single, or double, or imprecise (Henrich *et al.*, 2004). In particular, the double peak feature is not uncommon, for opening as well as for closing, and it may be found in specific cases: for a given singer, it can be consistently associated with either soft or loud production, and/or with either low or high pitches (Henrich *et al.*, 2004). These peaks may offer visual clues to some characteristic features of the vocal folds vibratory movement. As an example, Henrich *et al.* observed that the double peak feature could occur during a laryngeal mechanism transition, indicating that there may be some slower adjustments in the vocal fold contact process, even when the transition appears to be sudden when looking at the EGG amplitude change.

These observations call for research combining electroglottography with some kind of visualization. The purpose of the present study is to characterize some DEGG features using high-speed cinematography. The questions addressed by the study are the following ones: is the peak doubling feature related to a typical vibratory movement? Can the imprecision sometimes observed in the case of DEGG opening peaks be explained in terms of glottal abduction over the length of the vocal folds?

So as to answer these questions, high-speed images and EGG signals have been recorded simultaneously in the case of various voiced production. The recording session took place in the Universitätsklinikum Hamburg-Eppendorf (Phoniatrie und Pädaudiologie, Pr. Dr. Hess). Two French male subjects participated in the study: a male speaker (26 years old) and a professional counter tenor singer (40 years old). Both subjects

were asked to produce sustained sounds with different voice qualities (modal, creaky, breathy, tensed, harsh). Prior to each production, a sound example was played to the subject and the pitch was kept constant when possible. In addition, the professional singer was asked to sing sustained voiced sound on various pitches and using different voice registers (modal, produced in laryngeal mechanism M1; falsetto, produced in laryngeal mechanism M2; « voix mixte », produced either in M1 or M2), and glissandos with or without a noticeable voice break. The high-speed images and the EGG signal (Laryngograph Ltd London) were recorded simultaneously on a computer, using the high-speed camera unit Wolf HS Endocom 5560. The high-speed images of size 265*256 pixels were sampled at 4000 Hz and stored in binary files. The EGG signal was sampled at 44170 Hz, coded on 8 bits and stored in wav files. The data were processed using Matlab. The EGG signal was differentiated, and both EGG and DEGG signals were synchronized with the high-speed images. The observed oscillation patterns were characterized with the technique of multiline kymography (Neubauer *et al.*, 2001) and measurement of glottal area was obtained by applying an image processing algorithm based on the detection of contrast (light energy threshold).

The relation between the DEGG closing and opening peaks and the instants of initialization and termination of glottal area variation is explored to assess the validity of the previous findings on this extended database. Especially, much attention is given to the opening peaks, which seem to be dependent on the horizontal localization of initial vocal fold opening (anterior, midmembranous, or posterior).

References

- [1] Childers D.G., Hicks D.M., Moore G.P., Eskenazi L. and Lalwani A.L. (1990) *Electroglottography and vocal fold physiology*, J. Speech Hear. Res., vol. 33, pp. 245-254.
- [2] Baken R.J. (1992) *Electroglottography*, Journal of Voice, vol.6 (2), pp. 98-110.
- [3] Henrich N., d'Alessandro C., Castellengo M. and Doval B. (2004) *On the use of the derivative of electroglottographic signals for characterization of non-pathological voice phonation*, J. Acous. Soc. Am., vol. 115(3), pp. 1321-1332.
- [4] Neubauer J., Mergell P., Eysholdt U. and Herzel H. (2001) *Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes*. J. Acous. Soc. Am., vol. 110(6), pp. 3179-3192.