Assessing voice features of Greek speakers with hearing loss

Anna Sfakianaki & George Kafentzis
Department of Computer Science, University of Crete

Hearing loss compromises various aspects of normal speech production, effecting deviant articulatory configuration (McGarr & Campbell, 1995) but also problematic control of respiration and phonation (Stevens, Nickerson, & Rollins, 1983). Inappropriate management of glottal air flow, vocal fold abduction/adduction gestures, tension and mass, and velopharyngeal movement leads to faulty interarticulator coordination (Osberger & McGarr, 1982) resulting in severe difficulties in articulation and prosody. Hearing-impaired (HI) voice quality has been characterized as breathy, tense, harsh, throaty, flat, and common voice problems include inappropriate pitch and loudness, strain, roughness and vocal fatigue (Calvert, 1962; Thomas-Kersting & Casteel, 1989; Coehlo et al., 2015).

In order to study the vocal function of speakers with hearing loss, an assessment of glottal aerodynamics is required, as it would provide useful information about vocal fold movement and glottal air flow during speech (Baken & Orlikoff, 2000). Glottal characteristics have been examined in past and recent research in order to obtain suitable measures for the detection of HI voice deviations and for the examination of differences in vocal adjustments of speakers with HI in comparison with those of their normal hearing (NH) counterparts (e.g., Arends et al., 1990; Mahshie & Öster, 1991; Mora et al., 2012; Jaganathan & Kanagarah, 2016). Electrolaryngography (ELG) and electroglottography (EGG) have been utilized in the past so as to study the vocal function of speakers with hearing loss. Several problems in voice quality, such as breathiness, roughness or hoarseness have been associated with variations in fundamental frequency and amplitude, i.e., jitter and shimmer respectively (e.g., Monsen, 1983; Wolfe & Steinfatt, 1987; Jaganathan & Kanagarah, 2016), and other temporal and frequency metrics of the glottal source. For example, close-to-open phase ratio in a vocal fold vibration cycle, and steepness of glottal closure have been associated with breathiness (Maasen & Povel, 1987), while measures reflecting the extent of abduction of the vocal folds (open quotient) and glottal efficiency (speed quotient) have been found deviant from normal in HI speech suggesting reduced vocal fold mobility and oscillatory efficiency (Mahshie & Öster, 1991).

Besides ELG and EGG signal analysis, vocal fold movement can also be observed and measured using visual techniques (e.g., stroboscopy, kymography, etc.); however all these methods are both invasive and expensive. Alternatively, measurements directly from the glottal volume velocity signal of recorded speech can be made using glottal inverse filtering (GIF). GIF is based on the idea of inversion according to which, the effects of vocal tract and lip radiation are cancelled from the speech signal (Alku, 2011). Thus, by analyzing the speech signal (output) we estimate the glottal excitation (input).

The present paper examines voice features of Greek speakers with NH and with prelingual profound HI using the glottal inverse filtering program of Aalto Aparat (Pohjalainen, Airaksinen, Airas & Alku, 2015). To the best of our knowledge, there are no studies on HI glottal source features using Aalto Aparat and no studies using inverse filtering on Greek NH or HI voices. Estimations of the voice source signal were obtained from recordings of symmetrical /pVpV/ disyllables with the corner vowels /i, a, u/ produced by five speakers, two men and three women, with prelingual profound HI (> 90 db HL) and with different levels of speech intelligibility, and five speakers with NH, matched for age and sex. Both time- and frequency-domain parametrization methods were used in order to capture the most important features of the glottal source waveforms (Alku, 2011). Time-domain parametrization methods included the open quotient (OQ), the closing quotient (ClQ), the speed quotient (SQ) and the normalized amplitude quotient (NAQ), while frequency-domain parametrization methods involved the harmonic richness factor (HRF), the difference between the amplitudes of the fundamental and the second harmonic (H1-H2) and the parabolic spectral parameter (PSP). Also, additional measures such as the average fundamental frequency (F0), jitter, shimmer, noise-to-harmonics ratio (NHR) were computed using well-known digital signal processing (DSP) algorithms. Statistical analyses were performed for each measure so as to investigate significant differences between HI and NH speakers. Findings are examined in light of possible differences in the phonatory mechanism of speakers with HI. Moreover, the paper discusses the clinical value of inverse filtering in atypical voice research as well as the advantages and limitations of the application of freely available tools and algorithms in HI voice assessment.

REFERENCES

Alku, P. (2011). Glottal inverse filtering analysis of human voice production – A review of estimation and parameterization methods of the glottal excitation and their applications. *Sadhana, 36*(5), 623-649.

Arends, N., Povel, D-J., Van Os, E., & Speth, L. (1990). Predicting voice quality of deaf speakers on the basis of glottal characteristics. *Journal of Speech and Hearing Research, 33*, 116-122.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice.* 2nd ed. San Diego, California: Singular Publishing Group.

Calvert, D. R. (1962). Deaf voice quality: a preliminary investigation. *The Volta Review, 64,* 402-403.

Coehlo, A. C., Medved, D. M., & Brasolotto, A. G. (2015). Hearing loss and the voice. In F. Bahmad (Ed.). *InTech: Update on Hearing loss*, chapter 6. DOI: 10.5772/61217.

Jaganathan, N., & Kanagarah, B. (2016). Analysis of deaf speakers' speech signal for understanding the acoustic characteristics by territory specific utterances. *Circuits and Systems, 7,* 1709-1721.

Mahshie, J., & Öster, A-M. (1991). Electroglottograph and glottal air flow measurements for deaf and normal-hearing speakers. *STL-QPSR, 32*(2-3), 19-27.

Maasen, B., & Povel, D-J. (1987). Glottal determinants of deaf voice quality. *Proceedings of the XIth International Congress of Phonetic Sciences*, pp. 377-380. Tallinn, USSR: Academy of Sciences of the Estonian S.S.R.

McGarr, N. S., & Campbell, M. (1995). Opening inroads to understanding articulatory organization in persons with hearing impairment. In K. S. Harris, F. Bell-Berti & L. J. Raphael (Eds.), *Producing speech: Contemporary issues. For Katherine Harris,* 433-454. New York: AIP Press.

Monsen, R. B. (1983). Voice quality and speech intelligibility among deaf children. *American Annals of the Deaf, 128,* 12-19.

Mora, R., Crippa, B., Cervoni, E., Santomauro, V., & Guastini, L. (2012). Acoustic features of voice in patients with severe hearing loss. *Journal of Otolaryngology-Head & Neck Surgery, 41*(1), 8-13.

Osberger, M. J., & McGarr, N. (1982). Speech production characteristics of the hearing impaired. *Status Report on Speech Research, SR-69*, 227-288. Haskins Laboratories.

Pohjalainen, H., Airaksinen, M., Airas, M., & Alku, P. (2015). *Aalto Aparat, Manual v. 2.0,* Dept. of Signal Processing and Acoustics, Aalto University.

Thomas-Kersting, C., & Casteel, R. L. (1989). Harsh voice: Vocal effort perceptual ratings and spectral noise levels of hearing-impaired children. *Journal of Communication Disorders, 22,* 125-135.

Wolfe, V. I., & Steinfatt, T. M. (1987). Prediction of vocal severity within and across voice types. *Journal of Speech and Hearing Research, 30,* 230-240.