Forensic voice comparison using glottal parameters in twins and non-twin siblings

Eugenia San Segundo Fernández[1], *Pedro Gómez-Vilda*[2]
[1]*Universidad Internacional Menéndez Pelayo (UIMP), Madrid, Spain*
eugeniasansegundo@gmail.com
[2]*Center for Biomedical Technology, Universidad Politécnica de Madrid, Spain.*
pedro@fi.upm.es

Forensic phoneticians have traditionally relied on the information found in the vocal folds for speaker identification: from the analysis of classical distortion parameters like jitter and shimmer (Künzel & Köster, 1992) and other laryngeal features (Jessen, 1997) to the automatic approaches exploring the usefulness of the combined use of vocal source and vocal tract information in order to improve speaker-recognition systems (Zheng, 2005; Farrús, 2008). Based on previous voice-pathology investigations (Gómez et al., 2007), other studies by these authors have more recently shown that their voice-analysis methodology based in the decoupling of vocal tract from glottal source estimates can also be useful for the biometric characterization of speakers (Gómez et al., 2009, 2010). Following these studies, San Segundo (2012) was a pilot experiment with a relatively small sample of MZ and DZ twins (12 and 8, respectively), and using only some of the glottal parameters provided by a specific software package implementing vocal tract inversion and glottal source parameterization (www.biometrosoft.com).

For the current study 54 speakers were recruited: 24 MZ twins, 10 DZ twins, 8 non-twin brothers and 12 reference-population speakers. In a first step, as a follow-up of San Segundo (2012), some reexaminations and in-depth voice analyses were carried out for all the 20 speakers' voices already analyzed in the above-mentioned proof of concept: 1) anamnesis reexamination to discard possible voice-related pathologies; 2) reexamination of the parameter values extracted, since the analysis in the pilot experiment was carried out in a batch-mode and this kind of processing may entail certain evaluation software artifacts (ESA); and 3) new voice analysis and back annotation with the aim of visually inspecting the glottal waveform of the speakers' voices and checking their fitting to usual normophonic thresholds. Besides, if deemed necessary, a DNA test was carried out to confirm the twins' zygosity. In a later step, the naturally-sustained [e:] fillers of all the 54 speakers (2 sessions per speaker) were extracted and analyzed with the same software creating a vector of 68 parameters from each vowel segment, comprising: 1) f0 and distortion parameters; 2) cepstral coefficients of the glottal source power spectral density (PSD); 3) singularities of the glottal source PSD; 4) biomechanical estimates of vocal fold mass, tension and losses; 5) time-based glottal source coefficients; 6) glottal gap (closure) coefficients; and 7) tremor (cyclic) coefficients. Finally, a forensic comparison was carried out using the methodology described in Gómez et al. (2012). The results suggest that the parameters analyzed are somehow genetically related, as more similarity is found in MZ twins than in DZ twins or non-twin siblings. Besides, the between-speaker comparisons for unrelated speakers yield LLRs homogeneously around -10, indicating a very good performance of the system.

References

Farrús, M. (2008). Fusing prosodic and acoustic information for speaker recognition. PhD Thesis, Universitat Politècnica de Catalunya.

Gómez, P., Fernández-Baillo, R., Nieto, A., Díaz, F., Fernández-Camacho, F.J., Rodellar, V., Alvarez, A. and R. Martínez. (2007). Evaluation of voice pathology based on the estimation of vocal fold biomechanical parameters. *Journal of Voice*, **21** (4), 450-476.

Gómez, P., Fernández-Baillo, R., Rodellar, V., Nieto V., Álvarez-Marquina, A., Mazaira-Fernández, L.M., Martínez, R. and J.I. Godino. (2009). Glottal Source biometrical signature for voice pathology detection. Speech Communication, **51** (9), 759-781.

Gómez, P., Álvarez-Marquina, A., Mazaira-Fernández, L.M., Fernández-Baillo, R., Rodellar, V. and V. Nieto. (2010). Glottal Biometric Features: Are Pathological Voice Studies appliable to Voice Biometry? *WTM-IP Workshop Tecnologías Multibiométricas para la Identificación de Personas*, Universidad de Las Palmas de Gran Canaria, 4-6 July 2010.

Gómez, P., Mazaira-Fernández, L.M., Martínez, R., Álvarez-Marquina, A., Hierro, J.A., and R. Nieto. (2012). Distance Metric in Forensic Voice Evidence Evaluation using Dysphonia-relevant Features. *VI Jornadas de Reconocimiento Biométrico de Personas (JRBP)*, Las Palmas de Gran Canaria, 26-27 January 2012.

Jessen, M. (1997). Speaker-specific information in voice quality parameters. *Forensic Linguistics* **4**(1): 84-103.

Künzel, H.J. and J-P. Köster. (1992). Measuring Vocal Jitter in Forensic Speaker Recognition. *Proceedings 44th Annual Meeting American Academy of Forensic Sciences*, New Orleans, 113-114.

San Segundo, E. (2012). Glottal source parameters for forensic voice comparison: An approach to voice quality in twins' voices, *21th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA)*, Santander, 6-8 August 2012.

Zheng, N. (2005). Speaker recognition using complementary information form vocal source and vocal tract. PhD Thesis, The Chinese University of Hong Kong.